

Utvärdera effekter av offentligt stöd till näringslivet

Förslag till analysverktyg

Lars Bager-Sjögren, Federica Calidoni, Håkan Gadd,
A.M.M. Shahiduzzaman Quoreshi

Regleringsbrevsuppdrag nr 3, 2007

Dnr 1-010-2007/0023

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

För ytterligare information kontakta Federica Calidoni
Telefon 063-166 617
E-post federica.calidoni@itps.se

Förord

I 2007 års regleringsbrev fick ITPS i uppdrag av Näringsdepartementet att undersöka möjligheten att skapa en modell eller flera olika modeller för utvärdering av statligt stöd i Sverige. Denna rapport fokuserar på utvärderingsfrågor ur två olika perspektiv.

Å ena sidan presenteras erfarenheter och god praxis i andra länder (Holland, Irland och USA) för att belysa den så kallade *utvärderingskulturen*. I denna rapport föreslås två möjligheter för att effektivisera utvärdering: (1) att utse en extern myndighet som utför både ex ante- och ex post-utvärdering eller (2) att utse en enda myndighet som ansvarar för genomförande, analys och utvärdering av åtgärder.

Å andra sidan fokuserar rapporten på utvärderingsmetoder på såväl mikro- som makronivå. Efter en kortfattad litteraturgenomgång av matchningstekniker och rumslig ekonometri, föreslås att ITPS utvecklar en modell för att undersöka effekterna av statligt stöd på tre olika nivåer: företagsnivå, regional nivå och nationell nivå.

För att kunna utveckla en sådan modell krävs ett omfattande utvecklingsarbete inom ITPS dagliga verksamhet och att 6–8 personer uteslutande arbetar med detta i 3–4 arbetsår. Den totala kostnaden för att utveckla modellen uppskattas till ca 5 miljoner kr. per år eller 20 miljoner kr. under en fyraårsperiod.

ITPS vill rikta ett tack till den referensgrupp som följt projektet. Gruppen har bestått av: Runar Brännlund (Umeå Universitet), Jonas Månsson (Växjö Universitet), Richard Murray (Statskontoret), Magnus Sjöström (Konjunkturinstitutet), Anders Stenberg (SOFI).

Rapporten har tagits fram av en projektgrupp vid ITPS som består av Federica Calidoni (projektledare), Lars Bager-Sjögren, Håkan Gadd, A.M.M. Shahiduzzaman Quoreshi och Johan Lundberg (CERUM, Umeå Universitet)

Östersund, augusti 2007

Sture Öberg
Generaldirektör

Innehåll

Sammanfattning	7
1 Inledning	10
1.1 Statligt stöd.....	11
2 Utvärderingsprocess och utvärderingskultur	13
2.1 Utvärdering i olika länder	13
2.1.1 Nederländerna.....	14
2.1.2 Irland.....	19
2.1.3 USA.....	22
2.2 Lärdomar från andra länder	26
2.3 Utvärdering av näringspolitik i Sverige.....	27
2.3.1 Ramverk för utvärdering.....	27
2.3.2 Styra med mål	28
2.3.3 Vad är uppföljning och utvärdering?	30
2.3.4 Nuteks och Vinnovas utvärderingsstrategier	32
2.4 Utmaningar	35
3 Utvärderingsmetoder	39
3.1 Analys på mikronivå: icke experimentell utvärdering och matchande bedömning	39
3.2 Analys med rumslig ekonometri.....	42
3.2.1 Ett konkret exempel på tillämpning av rumslig ekonometri	45
3.2.2 Slutdiskussion och rekommendationer	48
3.3 Allmän jämviktsmodell.....	49
3.3.1 Befintliga modeller.....	51
4 Mikro-makro modellen	63
5 Sammanfattande kommentarer	69
Referenser	71
Bilaga 1. Diskussionsämnen för intervjuer	75
Bilaga 2. Några program och metoder från andra länder	77
A2.1. Nederländerna.....	77
A2.2. Irland.....	80
Bilaga 3. Olika ekonometriska program för rumslig analys	82

Sammanfattning

År 2007 fick ITPS i uppdrag av Näringsdepartementet att undersöka möjligheten att skapa och genomföra en modell eller flera olika modeller för utvärdering av statligt stöd i Sverige. Denna rapport fokuserar på utvärderingsfrågor ur två olika perspektiv. Å ena sidan presenteras erfarenheter och god praxis i andra länder (Holland, Irland och USA) för att belysa den så kallade *utvärderingskulturen*. Å andra sidan fokuserar rapporten på utvärderingsmetoder på såväl mikro- som makronivå. Efter en kortfattad litteraturgenomgång av matchningstekniker ges en skiss av en möjlig jämviktsmodell (CGE) som skulle kunna fungera som allmän ram för utvärdering av politiska effekter på både nationell och regional nivå.

Arbetet är indelat i två delar. I den första delen av arbetet (kapitel 1 och 2) är de viktigaste frågorna: Vad är utvärdering? Varför bör en myndighet utveckla en utvärderingskultur? Vad krävs för att utveckla den? Det är tydligt att utvärdering är ett komplext koncept som både syftar till att övervaka vad som sker i en organisation och att dra lärdomar av det som sker. Utvärdering kan vara ett ansvarsområde för en liten grupp specialister inom eller utanför institutionen eller en del av dess allmänna verksamhet. Under alla omständigheter uppfattas utvärdering som ett användbart och viktigt koncept och därför måste sättet att uppnå, införa och hantera utvärdering vara noggrant genomtänkt.

Detta är en stor utmaning för komplexa åtgärder om exempelvis statligt stöd och ekonomiska bidrag, delvis beroende på att utvärderingar naturligtvis är politiskt känsliga. För att utvärderingen ska bli effektiv måste institutionerna vara beredda att godkänna och publicera rapporter som innehåller negativa eller obehagliga slutsatser. För detta krävs en djupt rotad kultur med demokratisk utvärdering på alla administrativa nivåer.

I kapitel 1 beskrivs hur offentlig intervention på marknaden motiveras i den teoretiska litteraturen.

I kapitel 2 beskrivs utvärderingens roll i tre olika länder: Nederländerna, USA och Irland. Genom utvecklingen av VBTB-programmet ”*From Policy Budget to Policy Accountability*” år 1999 har Nederländerna fått en ledande roll inom den europeiska debatten om utvärdering. VBTB har medfört att Nederländerna omvärderat utvärderingsfunktionen och infört en mer systematisk och strukturerad strategi när det gäller utvärdering. Dessutom skall man, enligt den holländska lagstiftningen, för varje nytt eller anpassat politiskt mål eller verktyg överväga om ex ante utvärdering är användbar. Vidare skall varje befintlig stödåtgärd utvärderas minst en gång var femte år. De irländska erfarenheterna skiljer sig mycket från de holländska. Intressanta utvärderingsfunktioner när det gäller den ökade uppmärksamheten för utvärdering på EU-nivå framhävs. Irland utvecklade ett offentligt utvärderingssystem till följd av kraven för anslutning till Europeiska unionen, vilket är ett tydligt exempel på hur starka yttre påtryckningar kopplat till tillgång på viktiga resurser kan vara en avgörande katalysator när det gäller att främja utvärdering. Slutligen är erfarenheten från USA intressant på grund av den karakteristiska politiska strukturen och det stora antalet aktörer som är engagerade på både federal och statlig nivå och eftersom den är ett viktigt exempel på ledarskap vid utvärdering inom Advanced Technology Program (ATP). ATP har engagerat nationellt framstående utvärderare som tillämpar nya och befintliga metoder för att skapa en analytisk och empirisk grund för ATP:s verksamhet och resultat.

I kapitel 2 dras också slutsatser om lärdomarna av andra länders erfarenheter och undersökningen om den svenska situationen. Den holländska erfarenheten visar hur viktigt det är med en bra och genomtänkt utformning av åtgärderna som blir en utmärkt utgångspunkt för utvärdering och hur viktigt det är med detaljerad datasammanställning. Av den irländska erfarenheten kan vi lära oss att program som är begränsade och öppna när det gäller syfte är lättare att utvärdera. Slutligen visar erfarenheten från USA att det finns ett behov av att öka antalet retrospektiva analyser, att införliva både direkt och indirekt stiganalys i lönsamhetsundersökningar, att övervaka och uppdatera information, att kontinuerligt utveckla ny utvärderingsteknik, att upptäcka och ta itu med nya frågor som uppkommer till följd av utvärderingen och att i högre grad utnyttja utvärderingsresultaten i beslutsfattandeprocessen.

I analysen av den svenska utvärderingskulturen betonas behovet av att fördjupa förståelsen för effekterna av olika åtgärder och ta hänsyn till externa experters resultat. Med tanke på att det inte räcker med övervakning för att få relevanta insikter om åtgärdens effektivitet måste de olika organen som ansvarar för utvärdering ta ytterligare ett steg framåt mot utvärdering. Det är det enda sättet att förse lagstiftarna med de verktyg som behövs för att utforma nya åtgärder och program. Sverige måste röra sig mot en mer komplex utvärderingskultur för att kunna starta en inlärningsprocess där man genom utvärderingar får information och nödvändig kunskap för att bättre kunna utforma åtgärderna. Det finns två alternativa vägar som den svenska regeringen kan följa för att uppnå effektivare utvärderingar. (1) Att utse en myndighet som utför både ex ante och ex post utvärderingar, men som inte genomför politiken. Genom att tilldela olika ansvar (genomförande och utvärdering) till de två myndigheterna kan departementet öka oberoendet och tillförlitligheten i utvärderingsresultaten och minska eventuell partiskhet hos utvärderaren. (2) Att utse en enda myndighet som ansvarar för att genomföra, analysera och utvärdera politiken. Med detta alternativ finns det större risk för att myndigheten är partisk, i synnerhet när den utför ex ante utvärderingar, för att öka sin budget.

Den andra delen av arbetet (kapitel 3 och 4) är inriktad på utvärderingsmetoder och innehåller ett förslag på en modell som ITPS skulle kunna utveckla i framtiden. Den viktigaste frågan när det gäller utvärdering är *identifiering* av en kontrafaktisk situation och behovet av att svara på det kontrafaktiska dilemman: "Vad skulle ha hänt om åtgärden inte genomförts?"

I kapitel 4 undersöks hur den kontrafaktiska situationen kan analyseras på både mikro- och makronivå, beroende på målsättningen med åtgärden, hur stora effekterna förväntas bli och datatillgängligheten. Med hjälp av mikroanalys jämför utvärderaren en så kallad "experimentell" grupp där åtgärden tillämpas med en annan "kontrollgrupp" som motsvarar den grupp företag som inte tagit emot offentliga medel. I kapitlet beskrivs de ekonometriska metoder som används för att bilda en kontrollgrupp som påminner så mycket som möjligt om den experimentella gruppen, åtminstone när det gäller observerade egenskaper. Eftersom analysen på mikronivå inte alltid är genomförbar kan utvärderaren framhäva de allmänna effekterna på den nationella (eller regionala) ekonomin med hjälp av en allmän eller partiell jämviktsanalys där de två scenarierna beskrivs vid en bestämd tidpunkt (den jämvikt som uppnås av ekonomin utan intervention och den jämvikt som uppnås efter genomförande av en åtgärd) och ange åtgärden som orsak till skillnaden mellan dem. Analys på makronivå utgör ett värdefullt alternativ till mikroanalys vid brist på data eller egenartad utformning av åtgärder.

Kapitlet 4 fokuserar också på den roll som rumsliga villkor och placering spelar vid undersökning av effekter på ekonomipolitiken och besvaras följande frågor: Hur kan rumslig ekonometri användas vid utvärdering av nationella åtgärder och specifika regionala åtgärder? Hur användbart kan det vara? Är det värdefullt att inkludera rumslig analys i någon typ av utvärdering av statliga stödprogram? Hur ska rumsliga effekter redovisas inom ”traditionella” ekonometriprogram (Stata, SAS, E-views)? Är det möjligt? Hur är det med specifika program (GIS, GeoDa, SpaceStat)? Vilka resurser skulle ITPS behöva för att skaffa den kompetens och kunskap som behövs för att arbeta med rumsliga ekonometri modeller?

I kapitlet sammanfattas de befintliga modellerna i den svenska ekonomin och deras för- och nackdelar diskuteras, framför allt: EMEC (Environmental Medium Term Economic Model), FIMO (finansiell modell), LUMOD (Long term dynamic simulation), KIMOD, ISMOD, SAMGODS, RAMSES (Riksbankens Aggregerande Makromodell för Studier av Ekonomin i Sverige) och Sveriges Konkurrenskraft – att förstå och mäta nationell konkurrens.

I kapitel 4 beskrivs även den så kallade ”Mikro-Makro modellen” som ITPS skulle kunna utveckla i framtiden för att studera effekterna av statligt stöd på tre olika nivåer: företagsnivå, regional nivå och nationell nivå. Det idealiska sättet att undersöka effekterna på de tre olika nivåerna vore att börja med en företagsspecifik analys av varje företag i ekonomin, gruppera dem på regional nivå för att undersöka regionala effekter och slutligen sammanställa uppgifterna från de olika regionerna för att studera effekterna på hela ekonomin. Med tanke på att ett sådant förfarande på grund av det stora antalet företag inom ekonomin skulle bli både tidskrävande och tekniskt omöjligt föreslår vi att ett flexibelt system, i likhet med RAMSES/BVAR som utvecklats av Riksbanken, utvecklas på regionala nivåer, där vi kan undersöka effekterna av olika åtgärder på mindre grupper av företag.

För att kunna utveckla en sådan modell krävs ett omfattande utvecklingsarbete inom ITPS dagliga verksamhet och att 6–8 personer uteslutande arbetar med detta i 3–4 arbetsår. Den totala kostnaden för att utveckla modellen uppskattas till ca 5 miljoner kr. per år eller 20 miljoner kr. under en fyraårsperiod.

I kapitel 5 dras slutsatser.

1 Inledning

I 2007 års regleringsbrev fick ITPS i uppdrag av Näringsdepartementet att undersöka möjligheten att skapa en modell eller flera olika modeller för utvärdering av statligt stöd i Sverige. Denna rapport fokuserar på utvärderingsfrågor ur två olika perspektiv. Å ena sidan presenteras erfarenheter och god praxis i andra länder (Holland, Irland och USA) för att belysa den så kallade *utvärderingskulturen*.¹ Å andra sidan fokuserar rapporten på utvärderingsmetoder på såväl mikro- som makronivå. Efter en kortfattad litteraturgenomgång av matchningstekniker ges en skiss av en möjlig jämviktsmodell som kan användas som en allmän ram för utvärdering av politiska förändringar på regional nivå.

Enligt beräkningar gjorda av Institutet för tillväxtpolitiska studier (ITPS S2007:007) satsade staten under perioden 2003 till 2006 årligen mellan 25 och 31 miljarder SEK på olika typer av företagsstöd. Här ingår bland annat anställningsstöd, starta-eget-bidrag, regionalstöd i form av nedsatta socialavgifter, energistöd, transportstöd och regionala utvecklingsbidrag, samt stöd till små företag.

Summan av alla stöd i ITPS rapporten ökade 2006 till 30,7 miljarder, en ökning med 2,3 miljarder eller 8 procent. Merparten av det senaste årets ökning förklaras av att energiskatten har höjts för hushåll och serviceföretag, utan motsvarande höjning för industriföretagen. Det ger avtryck i statistiken genom att skatteutgifterna vid användning av energi ökat med nästan 502 miljoner eller med 2,6 procent till 20,2 miljarder. Inom de andra stödkategorierna hade de arbetsmarknadspolitiska stöden ökat, med 16 procent 2005 och minskade med 1 procent 2006. Energistöden ökade år 2006 med 51 procent. En minskning av FoU-stöden 2005 med 9 procent hade delvis bokföringsmässig förklaring. Emellertid ökade FoU-stödet med 14 procent 2006. Kategorin övrigt stöd minskade 2005 med 75 procent och ökade 2006 med 99,4 procent till 1629 miljoner. Det senare berodde delvis på en ökning av Boverkets stöd till mindre hyres- och studentbostäder från 373 till 560 miljoner mellan 2005 och 2006. Större genomslag hade dock det nya stödet för efterbehandling av förorenade områden, som uppgick till 588 miljoner 2006. De andra stödkategorierna hade mindre variationer över tiden. (ITPS S2007:007)

Ett syfte med den här typen av stödåtgärder är att öka sysselsättningen och på så sätt även främja en god ekonomisk tillväxt. För att utvärdera effekterna av statliga stödåtgärder är det av stor vikt att för ändamålet relevanta statistiska metoder och teoretiska modeller tillämpas.

Traditionellt är kärnområdena inom ramen för ITPS analysverksamhet ekonomisk tillväxt, produktivitet, sysselsättning och regionala frågor. Med detta i åtanke undersöker vi både möjligheten att anpassa befintliga modeller och möjligheten att skapa en ny modell. Resten av analysen är därför inriktad på utvärderingsmodeller, med särskild tonvikt på regionala frågor.

¹ När vi undersökte utvärderingskulturen i andra länder bestämde vi oss för att träffa de personer som ansvarar för utvärdering och reste därför till den Haag, Washington D.C. och Dublin för att göra intervjuer. Intervjuerna byggde på en rad diskussionsämnen som presenteras i bilaga 1.

1.1 Statligt stöd

I de flesta fall sker myndighetsinterventioner när marknaderna inte fungerar optimalt och fördelningen av resurser på en marknad eller inom en industri är suboptimal eller för att stimulera landets konkurrenskraft. Enkelt uttryckt kanske marknaden inte alltid fördelar knappa resurser effektivt eller inte konkurrerar effektivt på den internationella marknaden. Myndigheterna behöver därför motivera sin intervention med att den ligger i allmänhetens intresse.

Enligt ekonomisk litteratur har marknadsmisslyckanden haft negativa effekter på ekonomin eftersom en optimal fördelning av resurser inte kan nås. Med andra ord minimeras inte de sociala kostnaderna för att producera varan eller tjänsten och detta medför slöseri med vissa resurser. Problemet med marknadsmisslyckanden och hur de bör hanteras har gett till diskussioner mellan olika ekonomisk skolor.

Om ett visst resultat är paretoeffektivt, så betraktas det enligt det neoklassiska perspektivet inte som ett marknadsmisslyckande oavsett om det ligger i allmänhetens intresse. Den keynesiska och nykeynesiska skolan kopplar misslyckandet med att automatiskt uppnå ett fullständigt utnyttjande av resurser till ett marknadsmisslyckande. Framför allt betonar den nykeynesiska skolan det faktum att priser och löner inte anpassats. Vidare hävdar den så kallade Public Choice-skolan och förespråkare för laissez faire kapitalism att konceptet marknadsmisslyckande överhuvudtaget inte finns. De hävdar att myndigheterna inte nödvändigtvis bör försöka åtgärda marknadsmisslyckanden eftersom kostnaderna skulle bli högre om myndigheterna misslyckas än vad kostnaderna för det marknadsmisslyckande de försöker åtgärda är. Andra, såsom socialdemokrater och New Deal-anhängare, ser marknadsmisslyckanden som ett mycket vanligt problem i oreglerade marknadssystem och förordar därför omfattande statlig intervention inom ekonomin.

Om statlig intervention å ena sidan är nödvändig för att uppnå paretoeffektivitet kan den å andra sidan vara skadlig för konkurrensen och detta har lett till omfattande lagstiftning om statligt stöd.

I EG-fördraget² definieras statligt stöd som ekonomiskt stöd till verksamhet med fyra egenskaper:

- Det ges av en medlemsstat och med hjälp av statliga medel.
- Det gynnar vissa företag eller viss produktion.
- Det snedvrider eller hotar att snedvrیدا konkurrensen.
- Det påverkar handeln mellan medlemsstaterna.

Några exempel på statligt stöd inom ramen för fördraget är: statliga bidrag, ränteavdrag, skattelättnad, skatteavdrag, statliga garantier eller aktieinnehav, statligt tillhandahållande av varor och tjänster med särskilda villkor, direkta subventioner, skattebefrielser, förmånliga räntesatser, lånegarantier med särskilt gynnsamma villkor, förvärv av mark eller byggnader som kostnadsfritt eller fås på gynnsamma villkor, tillhandahållande av varor och tjänster på gynnsamma villkor, säkerhet mot verksamhetsförlust, återbetalning av kostnader vid framgångsrikt resultat, direkta eller indirekta statliga garantier för att kreditera verksamhetens diskonton, utdelningsgarantier, förmånlig emission på marknaden av aktier eller obligationer, minskning av eller befrielse från avgifter eller skatter, inklusive

² www.europa.eu.int/comm/competition/state_aid/legislation/aid3.html - A

degressiv avskrivning och minskning av sociala avgifter, uppskjutna skatteavgifter eller sociala avgifter, stöd som finansieras med hjälp av särskilda avgifter, kapitalöverföringar, vissa statliga innehav när det gäller företags kapital.

Mindre uppenbara exempel där statligt stöd kan bli aktuellt är: konsultrådgivning, förmåner genom verksamhet på området stadsförnyelse, stöd till företag att investera i miljöprojekt, stöd till statliga företag för att förbereda sig för privatisering, lagstiftning för att skydda eller garantera marknadsandelar, offentligprivata partnerskap och kontrakt som inte är öppna för anbudstävlan, medel för avfallsskatt som delvis går till miljöprojekt.

Sammanfattningsvis finns det tre huvudkrav för statligt stöd:

- Stödet inriktas på en väldefinierad marknadsbrist.
- Stödet måste vara inriktat på ett lämpligt område: statligt stöd måste vara ett lämpligt instrument, stödet måste ha en incitamentseffekt och stå i proportion till det problem som ska hanteras.
- Snedvridningen av konkurrensen och handeln som stödåtgärden ger upphov till måste vara begränsad så att den på det hela taget kan kallas förenlig.

2 Utvärderingsprocess och utvärderingskultur

Under de senaste decennierna har utvärdering i allt högre grad blivit en självständig vetenskap med rötter i flera discipliner och ett användbart verktyg för att tolka och genomföra politiska studier, resultatbedömning, teknisk design, investeringsportföljer osv. På den politiska arenan har utvärderingsmetoder fått ökad betydelse och många länder har börjat lagstifta om obligatorisk utvärdering av olika program.

Olika länder har därför olika sätt att se på utvärderingens roll i den politiska processen. Den vanliga uppfattningen är att utvärderare strävar efter att bidra till sociala förbättringar och att detta endast kan uppnås om utvärderingsresultaten återförs för att informera programadministratörer, politiskt ansvariga och andra aktörer så att programstruktur och verksamhet förbättras.

2.1 Utvärdering i olika länder

I följande avsnitt beskrivs vilken roll utvärdering spelar i tre olika länder: Nederländerna, USA och Irland. För att samla in specifik information om dessa länder har besök gjorts i the Haag, Washington D.C. och Dublin för att träffa några utvärderingsansvariga personer. Intervjuerna baserades på en rad diskussionsämnen som presenteras i Bilaga 1.

Skälen för dessa länder är variationerna i utvärderingskultur. I Nederländerna har mycket gjorts sedan mitten av 1980-talet för att förbättra legitimiteten och effektiviteten i användningen av offentliga medel. I mitten av 1980-talet fokuserades den politiska uppmärksamheten framför allt på statsutgifternas användning. Till följd av detta koncentrerades uppmärksamheten på granskning och kontroll. Under 1990-talet ändrades fokus och frågor om effektivitet och verkan fick utrymme i debatten. Den viktigaste förklaringen bakom denna utveckling var önskan om att förbättra styrningen av statliga organisationer som borde inriktas mer på prestationer och, i synnerhet, utfall. Att förbättra resultatstyrningen sågs som den främsta målsättningen, genom att klargöra förhållandet mellan utnyttjade resurser, produkter och tjänster och de resultat som ska uppnås samt att använda detta som utgångspunkt för att 1) utforma, 2) genomföra och 3) utvärdera politiken. Resultatet är en övergång från enbart finansiell redovisning till policyredovisning.

Men det är framför allt genom att utveckla ett program för resultatbaserad budgetering, *"From Policy Budget to Policy Accountability"* (VBTB), som Nederländerna har fått en ledande roll i den europeiska utvärderingsdebatten. VBTB³ har lett till att Nederländerna har omprövat utvärderingens funktion och till att introducera ett mer systematiskt och strukturerat förhållningssätt till utvärdering.

³ De viktigaste slutsatserna i VBTB-programmet kan sammanfattas enligt följande:

- För all ny policyutveckling som introduceras bör en noggrann och seriös bedömning göras av behoven av en *ex ante* utvärdering.
- De viktigaste policymålen bör utvärderas regelbundet.
- Kvaliteten i utvärderingsinstrumenten bör säkerställas.
- De politiskt och administrativt ansvariga bör informeras om resultaten av utvärderingar.
- Inom departementen bör det finnas en tydlig ansvarsfördelning för hur utvärderingarna ska genomföras.

I Nederländerna, liksom i andra europeiska länder⁴, har utvärderingar av program för statligt stöd gett värdefulla insikter i hur befintliga program kan förbättras och hur ny identifierad kunskap kan tillämpas i nya stödprogram. I denna bemärkelse kan kostnaden för utvärderingsstudierna återvinnas genom vinster i form av verkan och effektivitet. Varje befintlig stödåtgärd ska utvärderas åtminstone vart femte år.

De irländska erfarenheterna skiljer sig helt från de holländska och visar på intressanta utvärderingsfunktioner i förhållande till det växande intresset för utvärderingsmetoder på EU-nivå. Irland utvecklade sitt statliga utvärderingssystem för att uppfylla kraven för EU-anslutning. Systemet har därefter stärkts av interna orsaker som hänför sig till regeringens strävan att få bättre utdelning på satsade medel inom alla områden av offentliga utgifter. Detta återspeglas i regeringens projekt för utgiftsstyrning, Expenditure Review Initiative.

Slutligen är erfarenheterna från USA intressanta på grund av den speciella politiska strukturen och den stora mängd aktörer som berörs på både federal och delstatlig nivå. På federal nivå förkroppsligas utvärderingskulturen i USA av den federala revisionsmyndigheten General Accounting Office (GAO)⁵, vars ansvar är att studera den federala regeringens verksamhetsområden och utgifter samt bistå kongressen och förvaltningsmyndigheternas chefer med råd om hur de ska styra verksamheten effektivare. GAO utvärderar federala program, granskar federala utgifter och lämnar rättsliga utlåtanden. Samtidigt har varje delstat sin egen statliga revisionsmyndighet som gör motsvarande studier på delstatlig nivå.

Det andra skälet som motiverar vårt intresse av USA är att utvärdering är en väsentlig del av offentligt finansierade FoU-program, både som stöd för programstyrning och för att uppnå politiska mål. Ett viktigt exempel på ledarskap i utvärdering är programmet för avancerad teknik, Advanced Technology Program (ATP), där framstående utvärderare anlitas som tillämpar nya och befintliga metoder för att bygga upp en analytisk och empirisk bas för ATP-programmets verksamhet och resultat.

Efter en kort sammanfattning av hur utvärdering leder till lagstiftning och genomförande samt vilken roll berörda aktörer har, kommer vi att fokusera på några intressanta metoder och viktiga egenskaper som kan visa vägen för hur utvärderingsarbetet kan utvecklas i Sverige.

2.1.1 Nederländerna

Sedan 1999 har en holländsk lag om resultatbaserad budgetering, VBTB "*From Policy Budget to Policy Accountability*", föreskrivit att policyprogram ska utvärderas i detalj varje gång allmänna medel används för att granska om resurserna har använts på ett bra sätt.

Vi har genom intervjuer försökt förstå hur utvärderingsprocessen utvecklats i Nederländerna. Med hänsyn till begränsad tid har fokus varit på FoU- och innovationsprogram. Möten och intervjuer har skett vid nederländska ekonomidepartementet Ministry of Economic Affairs, SenterNovem (nederländsk myndighet som stöder hållbara energi- och miljöinitiativ) och CPB Netherlands Bureau for Economic Policy Analysis (Nederländernas byrå för ekonomisk-politisk analys).

⁴ Till exempel Österrike, Belgien, Tjeckien, Finland, Irland, Norge, Portugal, Slovakien och Storbritannien.

⁵ GAO har dessutom samarbetat med ITPS vid många tillfällen.

Tack vare ett möte med Luuk Klomp och Arjan Wolters på Directorate General for Enterprise and Innovation Strategy (Generaldirektoratet för näringsliv och innovationsstrategi) respektive Research and International Affairs Department (Avdelningen för forskning och internationella affärer) inom Ministry of Economic Affairs (Ekonomidepartementet) har vi skapat oss en bild av den aktuella situationen i Nederländerna.

De program som används idag kan sägas tillhöra två olika kategorier: tematiska program och grundläggande instrument. Inom varje tematiskt program finns ett antal instrument, medan de grundläggande instrumenten mer eller mindre utgörs av grundläggande principer. De enskilda instrumenten representerar den lägsta nivån, som enligt lagen måste utvärderas minst en gång vart femte år.⁶

Sedan den första januari 2002 regleras policyplanering och utvärdering av ett regeringsbeslut om resultatmätning och utvärdering (Regeling Prestatiegegevens en Evaluatieonderzoek, RPE). Beslutet lägger fast ett antal krav på förberedelser (ex ante utvärdering), uppföljning och ex post-utvärdering av politiska program. Kraven gäller:

- 1 Användningen av utvärderingsinstrument
- 2 Skyldigheten att ta hänsyn till ex ante utvärdering då övervägs nya instrument. Maktgivarna måste överväga olika policyalternativ om det inte finns goda skäl för att inte göra det (brådska, begränsade finansiella eller samhällsliga risker, tillräcklig information från ex post-utvärderingar, enligt föreskrifter i internationella överenskommelser)
- 3 Frekvensen och omfattningen av ex post-utvärderingar. Varje instrument måste utvärderas.
- 4 Utvärderingsinstrumentens kvalitet
- 5 Information till minister, departementschef och riksdag om utvärderingsresultaten
- 6 Ansvarsfördelning inom departementet avseende verkställandet av detta påbud.

För närvarande går Nederländerna igenom en övergångsfas mot färre instrument, högre flexibilitet och fler möjligheter att kombinera olika instrument. I ett framtida scenario kommer större program att presenteras beroende på kombinationen av olika instrument och ökad utvärdering av de olika instrumentsammansättningarna.

I själva verket ansvarar departementet för att utforma de olika instrumenten och kombinationerna av dem i tematiska program, medan SenterNovem ansvarar för genomförandet av samtliga instrument som avser innovation och miljöfrågor. Dessutom håller SenterNovem kontakten med mottagare och ansvarar för uppföljning och datainsamling för de program som de genomför. Slutligen kommer externa utvärderare in i bilden för att genomföra den obligatoriska utvärderingen när instrumentet har använts i fem år.

⁶ 2006 ändrades lagen delvis och 5-årsperioden är inte längre så strikt, de 5 åren är en flexibel gräns som huvudsakligen betraktas som en riktlinje, men utvärderingar kan göras efter längre eller kortare perioder av politiska orsaker eller i enlighet med politiska cykler. Det finns emellertid fortfarande en förpliktelse att utvärdera varje instrument.

Även om gällande lag inte innehåller någon skyldighet att använda ex ante utvärdering, betraktas sådana undersökningar som nödvändiga för att jämföra olika alternativ och för att lösa föreliggande problem. Ex ante utvärderingar genomförs därför mycket ofta, men i motsats till ex post-undersökningar, som alltid läggs ut på externa konsultbyråer, utförs de internt för att verifiera vilka instrument som ger den bästa lösningen på problemet, normalt med hjälp av befintliga utvärderingar. För det mesta genomförs en nollbasanalys innan ett program införs för att senare ska kunna genomföras en ex post-utvärdering. Nollbasanalysen ger en bild av situationen innan införandet och görs i huvudsak för tematiska program.

Ekonomidepartementet har det finansiella ansvaret för de utvärderingar som utförs och är involverat under hela processen. Det är departementet som bestämmer vilka ramar som gäller för utvärdering och resultatindikatorer. Departementet ansvarar för att organisera effektutvärderingen, även om en oberoende panel övervakar utvärderingen, som utförs av privata konsulter. Detta innebär en skillnad mellan utvärderingen (konsulten) och uppföljningen (organisationen som är engagerad i genomförandet – SenterNovem).

Ex post-utvärderingar utförs därför av externa oberoende privata konsulter, men interna kontroller och avvägningar garanterar att den privata konsulten tar upp rätt frågor och det finns en oberoende panel som ansvarar för kvaliteten och oberoendet i forskningen.

Olika utvärderingstekniker och modeller används, beroende på data tillgänglighet och vilket ämne som avhandlas och

ibland kommer utvärderingarna i exakt rätt tid i förhållande till cyklerna för beslutsfattande, men det finns en spänning mellan bra utvärderingar och policycyklerna eftersom den fulla effekten av olika program endast är synlig långt bortom ansvariga politikernas tidshorisont.

Under intervjun framkom att för närvarande diskuteras definitioner och syften, i synnerhet för processen att ange mål och lämpliga indikatorer och vilka resurser som är tillgängliga för dem samt hur målen skånås. Faktum är att departementet definierar syftena och resultatindikatorerna både för tematiska program och grundläggande instrument, men målen i sig diskuteras tillsammans med näringslivet och forskare för att göra dem så realistiska som möjligt. Departementet utarbetar delmål och huvudmål i samspel med den verkställande myndigheten. Enhetschefen i Generaldirektoratet för näringsliv och innovation Luuk Klomp menar att

Målen är alltid tydligt mätbara och tidsavgränsade. De är dock inte alltid realistiska. Till exempel: Lissabonstrategin satte upp målet att 3 procent av FoU-investeringarna skulle finansieras av den privata sektorn. Detta har antagits som en målsättning i den holländska politiken, men är inte realistiskt i Nederländerna där privata satsningar på FoU har legat stilla på 2 procent i 10 år. Målet är alltså inte realistiskt även om det är relevant, mätbart och tidsavgränsat. Problemet är att även om målet faktiskt specificeras så beror allt på mycket mer än politiken och programmet. Om målet inte uppnås betyder det inte nödvändigtvis att politiken inte fungerar, men orsakerna är svåra att fastställa.

Dessutom

Om målen utarbetas externt är det betydligt svårare att definiera rätt mängd resurser som bör tilldelas. Om departementet själv definierar målen är det lättare att veta var

man står och hur långt man kan nå. Mål som angetts av departementet är mer realistiska och tilldelade resurser bättre avpassade.

Ett annat aktuellt samtalsämne är tillgången till data för utvärderingssyften, vilket påverkar de metoder som kan användas. De data som samlas in under genomförandeprocessen består normalt av grundläggande information och efter fem år, då utvärderingen ska göras, samlas mer detaljerad information.

De olika departementen besvarar varje år den nederländska centralregeringens frågeformulär om utvärderingar i den s.k. Central Government Evaluation Survey (EOR) för att informera revisionsmyndigheten och andra organisationer om vilka utvärderingsstudier som planeras, är på gång eller har genomförts och om deras aktuella resultatdatasystem. EOR är ett användbart instrument som bland annat kan hjälpa departementen att skapa en bild av i vilken utsträckning deras verksamhet är föremål för utvärderingsstudier och ska redovisas i resultatdatasystem samt ge externa parter information om de utvärderingsinstrument som används.

SenterNovem är den myndighet under nederländska ekonomidepartementet som ansvarar för genomförandet av politiska beslut avseende FoU och innovationer. Som anges på deras webbplats är deras kärnkompetens att ”göra verklighet av regeringspolitiken” och på den nederländska regeringens vägnar verkställer de politiken avseende innovation, energi- och klimatförändringar samt miljö- och utrymmesplanering.

Nedan återges en del av intervjun med Hendrik Blanksma från Innovation Intelligence and Coordination, Directorate of Innovation (avdelning som informerar om och samordnar innovationsverksamheten inom Direktoratet för innovation).

Under senare år har myndigheterna börjat få större inflytande på själva programmen och för närvarande har en del av utformningen av den ekonomiska politiken förts över till SenterNovem för att uppnå närmare samverkan mellan dem som drar upp riktlinjerna för politiken och dem som verkställer den. Nu sysslar SenterNovem mer aktivt med att följa upp och rapportera om effekterna av programmen samt med utvärdering.

SenterNovem deltar i diskussionen om att definiera målen och på vilket sätt de ska mätas. Vi använder SMART-kriterierna och försöker tillämpa målen i de långsiktiga analyserna. Vi tittar på resultaten av aktiviteterna, projekten och programmen och måste alltid referera till de mål som definierats av departementet. När departementen rapporterar till riksdagen talar de ofta om resultaten på fältet, vare sig det är resultat av programmet eller inte. Ofta frågar riksdagen: hur fungerar samarbetet mellan näringslivet och universiteten nu jämfört med för 10 år sedan? Och så redogör de för skillnaden. Nästa fråga blir då: hur mycket bidrog programmet med till den förändrade bilden? Men detta är inte alltid den viktigaste frågan eller den viktigaste delen av rapporten. För departementet är det ofta den allmänna trenden som är viktig.

SenterNovem arbetar på tre nivåer: uppföljning, effektregistrering och utvärdering. Uppföljningen sker kontinuerligt för varje program. Den avser endast de företag som fått pengar: ingen uppföljning görs av kontrollgruppen. Sedan måste vi titta på effekterna av de olika projekten i programmet och SenterNovem hjälper de enskilda projekten att registrera sina effekter. Därefter görs en extern utvärdering av programmet, för vilken SenterNovem tar fram en del av uppgifterna.

Departementet anger inte tydligt hur utvärderingen budgeteras: i vissa program anges inte resurserna för utvärdering tydligt och inte heller resurserna för uppföljning och rapportering.

Netherlands Bureau for Economic Policy Analysis (CPB) (Nederländernas byrå för ekonomisk-politisk analys) är en av de konsulter som ansvarar för utvärderingen av den offentliga politiken. CBP Netherlands Bureau for Economic Policy Analysis är ett oberoende forskningsinstitut med avseende på innehållet, men formellt en del av den centrala regeringsmakten och som finansieras med allmänna medel. CPB genomför oberoende ekonomiska analyser för att informera såväl politiker och makthavare som samhällsorganisationer, vetenskapsvärlden och allmänheten.

Marc van der Steeg och Björn Vrommer, forskare vid Department for Sector Growth, Structure and Knowledge Economics CPB (departementet för sektortillväxt, struktur- och kunskapsekonomi), kommenterar utvärderingsverksamheten i Nederländerna.

CPB utför huvudsakligen ex post-utvärderingar, men ibland utförs även ex ante utvärderingar, till exempel ex ante analyser av kostnader och intäkter för politiska förslag om att investera i vissa infrastrukturprojekt, kunskapsprojekt, osv.

CPB genomför utvärderingar både på uppdrag av departementet och på eget initiativ. Vi gör mycket gärna experimentella utvärderingar med slumpvisa kontroller och skulle vilja att dessa utvärderingsmetoder skulle utökas och bli rutin då riktlinjerna för politiken dras upp.

I allmänhet engageras vi endast i slutskedet, när det är dags att utvärdera ett program, så många gånger måste vi förklara att vissa program är omöjliga att utvärdera på grund av att tillgängliga data är undermåliga eller så måste vi påpeka bristen på data och nackdelarna. Många gånger vill departementet att vi ska identifiera framgångs- eller misslyckandefaktorer och är inte intresserade av orsakssammanhang eller additionalitetseffekter. De tillskriver politiken resultatet trots att de inte har kontroll över andra faktorer och detta är ett problem.

Många utvärderingar genomförs på eget initiativ när vi ser en möjlighet att övertyga departementet om hur det kan fungera. Med de egna utvärderingarna försöker vi bevisa att om departementet utformar politiken på ett visst sätt, är det lättare att genomföra en bra utvärdering och säga mer om kausaliteten.

I allmänhet är målsättningarna inte särskilt tydliga, inte preciserade utan alltför generellt angivna. Målsättningarna behöver därför översättas till mätbara indikatorer.

Hur mäts beteendeadditionalitet? Det är ett nytt koncept och vi har försökt skapa ett frågeformulär och identifiera några hypotetiska konstruktioner som visar hur det kan uppfattas. Det är ett lovande utvärderingsområde, men inte så starkt som det experimentella upplägget. I FoU och nätverksstudier är beteendeadditionalitet ett bra sätt att analysera nätverk och socialt kapital, men reser problem när det gäller socialt önskvärda svar.

Utvärderingskultur: många människor är inte vana tänka i termer av utvärdering när de utformar politiken. I Nederländerna, i synnerhet vid utbildningsdepartementet, finns det personer som är övertygade om användbarheten av utvärderingar och tänker på det för nya program. Det sjunker allt mer in i människors medvetande.

De största problemen är politikutformningen, brist på data, brist på kontrollgrupper och bristen på studier innan programmet verkställs.

Politikutformningen är avgörande. Dessutom bör data samlas in både för kontroll- och experimentgrupp, men utan en bra struktur kan mängder med data samlas in och ändå blir inte utvärderingen lyckad.

Departementet måste göra någon slags utvärdering och inte nödvändigtvis en utvärdering med experiment- och kontrollgrupper. Att ha utvärderingen i åtanke när politiken utformas kan vara mycket praktiskt och det vore idealt, men Nederländerna är inte den ideala världen. Den grundläggande utvärderingskulturen håller på att förändras, men det finns fortfarande mycket att göra och nu försöker vi övertyga dem om att experimentell utvärdering är bättre.

Departementet bör undvika att bara plocka ut godbitarna när resultaten presenteras utan också ta hänsyn till alla effekter av politiken, inte bara dem som rör de viktigaste fastställda målsättningarna.

Mer detaljerad information om program och metoder presenteras i Bilaga 2.

2.1.2 Irland

Irland fick sitt oberoende 1922. Under 1950- och 1960- talen präglades Irland mycket av protektionism i syfte att skydda den inhemska industrin. Irland fick medlemskap i EU 1973 och det medförde större strukturella förändringar, men innebar också att ”kalla vindar” började blåsa eftersom Irland var tvungen att öppna upp sin ekonomi mot Europa. I mitten av 1980-talet förbättrades den makroekonomiska situationen på Irland. Regeringen samlade alla aktörer för att diskutera (unikt vid den tidpunkten på Irland) näringspolitiken och skapa enighet om nödvändiga åtgärder. Branschvisa investeringar och FDI ökade efter att Irland hade genomfört reformer på utbildningsområdet och inom arbetsmarknadspolitiken

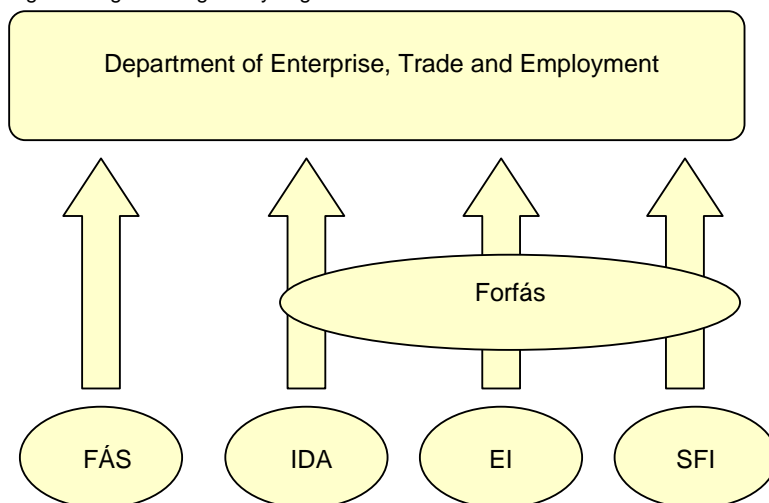
Irlands ekonomiska utveckling har genomgått en omvandling sedan mitten av 1980-talet och de senaste femton åren har den ekonomiska utvecklingen varit utan historiskt motstycke. Den ekonomiska tillväxten har i jämförelse med andra EU-länder varit rekordartad och den Keltiska tigern har varit ett begrepp. Inflödet av EU-medel har sedan slutet av 1980-talet varit stort och bidragit till Irlands utveckling. Uppföljning, utvärdering, övervakning, revision är alla delar i ett system för att undersöka, studera och analysera resultat och effekter av offentliga satsningar. Vilken roll har utvärdering inom närings- och regionalpolitik på Irland?

Information och intervjuer har insamlats/genomförts den 24–25 april i Dublin. Intervjuer har genomförts med personer vid Department of Enterprise, Trade and Employment; Margaret Ryan, state aid unit, Eamonn Culbert och Gerard Monks, Enterprise agencies unit. Vidare har intervjuer genomförts med personer vid Forfás (Forfás is the national policy and advisory board for enterprise, trade, science, technology and innovation), Jaqueline Allan och Seamus Bannon samt med Gerhard O’Brien vid Enterprise Ireland.

Närings- och regionalpolitik är sammanflätat med arbetsmarknads- och forskningspolitik och styrs av ett gemensamt departement, Department of Enterprise, Trade and Employment, och till departementet tillhörande myndigheter (se Figur 1). Näringspolitiken genomförs i huvudsak av Enterprise Ireland (EI), en myndighet under departementet, med en årsbudget på 2,6 miljarder kronor. Andra viktiga myndigheter är IDA med uppgift att skapa förutsättningar för utländska direktinvesteringar och Irländska forskningsrådet (SFI) som ska stimulera forskningssystemet. På utförarnivå finns även en myndighet med ansvar

för arbetsmarknadsfrågor för individer och företag (FÁS). En central myndighet under departementet är Forfás med uppgift att övervaka, forska, analysera och utvärdera politiken samt ge råd till departementet. Inom Forfás har utvärderingar bedrivits under ca 15 års tid. Forfás har även till uppgift att koordinera näringspolitiken vilket görs av dess styrelse tillsammans med cheferna för IDA och Enterprise Ireland⁷.

Figur 1. Organisering av myndigheter



Irlands utvärderingskapacitet har påverkats av två drivande krafter. Irlands medlemskap i EU har medfört ett externt tryck och regeringens ambitioner att sedan slutet av 1990-talet utveckla en utvärderingskultur för alla områden som omfattas av offentliga utgifter har inneburit ett internt tryck⁸.

Irland har varit mottagare av EU-stöd sedan strukturfondernas införande 1989. Fonderna finansierar delar av ett antal operativa program som är länkade till ett övergripande styrdokument (Community Support Framework Plan). Kravet på utvärderingar i fondernas regelverk medförde att Irlands regering 1989 inrättade två oberoende⁹ utvärderingsenheter som under perioden 1994–1999 utökades med en central utvärderingsenhet baserad i Finansdepartementet. Under strukturfondsperioden 1999–2006 minskade inflödet av strukturfondspengar då delar av Irland inte längre klassificerades som mål 1 område. I samband med denna förändring utökades kapaciteten för den centrala utvärderingsenheten, men de två oberoende enheterna lades ned. Vid ingången till 2007 består den offentliga utvärderingskapaciteten av två centrala utvärderingsenheter inom Finansdepartementet och myndigheter med ansvar att utvärdera. Den ena enheten arbetar med utvärderingar av alla operationella program på Irland och den andra enheten arbetar med analys/utvärdering av statliga program

⁷ 2003. *Review of Industrial performance and Policy 2003*. Department of Enterprise, Trade and Employment

⁸ Boyle, R. 2005. *Evaluation capacity development in the republic of Ireland*. ECD working paper series no 14, June 2005, Washington: World Bank

⁹ Enheterna var oberoende från genomförandeorganisationen för strukturfondsprogrammen.

Efter Australisk modell introducerades år 1996 på Irland ett program¹⁰ för att granska de offentliga utgifterna och som 1997 följdes av Expenditure Review Initiative (ERI) som syftade till att granska departementens utgifter vart tredje år. Två huvudsakliga syften skulle vägleda arbetet; att systematiskt analysera vad som har uppnåtts med offentlig finansiering och utarbeta ett underlag som underlättar beslut om prioriteringar inom och mellan program¹¹. Intentionen med ERI var att skapa en utvärderingsstrategi för alla departement och uppmuntra en utvärderingskultur inom departementen som fokuserade på produktivitet, effektivitet, relevans och prioritering av insatser. Ansvariga för övervakning av ERI systemet utgörs av en central styrgrupp (Expenditure Review Steering Committee ERCSC).

Efter en granskning av Irlands riksrevision etablerade ERCSC ett nätverk mellan personer involverade i ERI-systemet, införde ett system för kvalitetssäkring av granskningar och ett system för att följa upp resultaten av de granskningar som genomförts¹².

Huvudintrycket är dock att utvärderingskapaciteten är splittrad när det gäller de resurser som finns att tillgå för utvärdering/analys av näringspolitiken. Förändringar över tid av Kommissionens krav på genomförda utvärderingar och analyser av specifika ekonomiska faktorer som underlag för godkännande av nya stödprogram har medfört ett ökat medvetande om att utvärdering och övervakning är nödvändigt för nya eller förändrade stödprogram. Denna utveckling har även visst stöd i och med att näringspolitiken går igenom på makronivå vart tionde år.

Statsstöd på Irland styrs i stor utsträckning av ”kartor” godkända av EU och av anvisningar som reglerar hur stödet kan användas. Kartorna över statsstödsregioner har utarbetats av statsstödsenheten på Näringsdepartementet. Utöver notifierade stödprogram finns även ett antal stöd som faller under *de minimis* reglerna.

Ända sedan slutet av 1980-talet har mycket politik på Irland kanaliserats genom program. I början 1990-talet genomfördes en total genomgång av den Irländska näringspolitiken och genomförandet systematiserades i program. All policyformulering sker på nationell nivå.

Olika typer av stöd är en ingrediens i de flesta program. Program är i detta sammanhang något vidare än ett traditionellt stödprogram som ofta omfattar ett stöd. Programtanken genomsyrar makronivån för ett flertal politikområden. Under perioderna 2000–2006 och 2007–2013 har Nationella Utvecklingsplaner arbetats fram (NUP). Den nationella utvecklingsplanen är det primära styrdokumentet för hur Irland ska utvecklas. Under NUP finns Irlands och EUs styrdokument om hur Irland ska använda EUs strukturstöd. Kopplat till den nationella utvecklingsplanen finns operativ program och till dessa fastlagda nationella strategier¹³. I den nationella strategin för näringspolitiken anges den nationella

¹⁰ Programmet benämndes ”*Delivering better Government*”

¹¹ *Department of Enterprise, Trade and Employment. 2004. ”Expenditure review of IDA Ireland’s property programme”, sid 13*

¹² *Boyle R. 2005. Evaluation capacity development in the republic of Ireland. ECD working paper series no 14, june 2005, Washington: World Bank, sid 16*

¹³ *Statement of Strategy 2005-2007 lägger fast mål och identifierar åtgärder för närings- och arbetsmarknadspolitiken. För respektive politikområde finns en särskild strategi. För näringspolitiken fastställs en strategi vart tredje år. 2004 kom ”Ahead of the curve – Ireland’s place in the global economy” och sommaren 2007 presenteras en ny strategi. Strategy for Science,*

ambitionen för olika områden med mål (goals), delmål (objectives), åtgärder (strategic actions) och indikatorer (performance indicators). (Se Bilaga 2)

Forfås har till uppgift att utvärdera näringspolitikens uppgifter i en vidare kontext, medan Enterprise Irelands fokus är att utvärdera de program organisationen driver och ansvarar för. Utvärderingar i egen regi genomförs framförallt av Forfås i övrigt används universitet och fristående konsultbolag samt revisionsföretag. Av de program Enterprise Ireland driver ansvarar en "programansvarig" för genomförande och utvärdering och rapporterar till Enterprise Irelands styrelse som i sin tur ansvarar för rapportering till departementet.

Mycket av det som görs på nationell nivå (departementsnivå) är genomgångar av statens utgifter på olika områden. Från departementets perspektiv är detta exempel på utvärdering. I övrigt hänvisas till motsvarande Riksrevisionen som genomför utgiftsgenomgångar i enlighet med "public service management act".

En genomgång av ett antal utvärderingar utförda av Forfås visar att utvärderingarna fokuseras till analyser av input och output från programmet. Någon koppling till regionalekonomiska eller makroekonomiska sammanhang görs inte. Genomgången i Bilaga 2 visar att kvalitativa metoder dominerar utvärderingarna.

Utvärderingar som diskuterar dödvikt är "egentligen" inte möjliga då de flesta företag på Irland deltar eller har deltagit i någon form av stödprogram. Det är därför mer relevant att diskutera partiell dödvikt vilket egentligen är effektiviteten av ett stöd mätt för ett antal output-variabler i förhållande till en annan grupp.

Ett skäl till att utvärderingarna inte i någon större utsträckning laborerar med kontrollgrupper är tillgången till data. På Irland finns inte samma möjligheter att samla in data i jämförelse med Sverige. Vid ITPS finns individ och företagsdatabasen (IFDB) vilken möjliggör en sammankoppling av företag med individer vilket ger betydande analysmöjligheter. På Irland finns inga sådana möjligheter. Ett exempel på detta problem framkom vid intervjuerna. En myndighet planerar en undersökning som syftar till att följa nyantagna forskarstuderande för att förstå vilka faktorer som styr examinationsfrekvensen. Bakgrunden är att den Irländska regeringen beslutat att antalet disputerade ska öka. Data om de forskarstuderande kräver att universiteten frivilligt ställer dessa data till förfogande. I Sverige är denna typ av data offentlig genom SCB.

Det finns inget motsvarande till ITPS databaser på Irland. En avgörande skillnad är avsaknaden av lagstiftning rörande insamling av statistik. En annan orsak är också de resurser som satsas på utvärderingar. Forfås kan jämföras med ITPS, men avsätter bara en tiondel av vad ITPS avsätter till utvärderingar varje år. Denna jämförelse ger en indikation om förhållandena, men säger naturligtvis inget om den faktiska situationen då inga sådana data har kunnat ställas samman.

Boyle menar att de hinder som behöver övervinnas för att utveckla utvärderings- kapacitet och kultur på Irland är bristen på reglering som gör att utvärdering kan integreras med beslutsfattande.

2.1.3 USA

Intervjuer genomfördes den 19 april 2007 med Bob Boerner från National Conference of State Legislatures (NCSL) (organisation för beslutsfattarna i delstaterna) och Gary

Technology and Innovation 2006-2013 presenterar mål, åtgärder och förväntade effekter av politiken inom området.

VanLandingham från Florida Legislature Office of Program Policy Analysis and Government Accountability (OPPAGA) (en enhet i den lagstiftande församlingen som granskar myndigheter och program för att förbättra servicen och minska kostnader). Har en tydlig bild av hur politiken för statligt stöd är organiserad i USA byggdes.

Det första vi vill peka på är att det förekommer två överlappande lagstiftningsnivåer: federal och delstatlig nivå. Utformning, genomförande och utvärdering av politiken har olika kännetecken beroende på om det är på federal eller delstatlig nivå. När det gäller utvärdering finns det absolut inget samband mellan den federala revisionsmyndigheten Government Accountability Office (GAO) och enskilda staters olika revisionsmyndigheter, utom att delstaterna måste följa riktlinjerna i ”Den gula boken” för finansiell revision. De enskilda staterna är helt skilda från den federala regeringen. Ungefär hälften av staterna genomför utvärderingar enligt den Gula boken, andra stater följer utvärderingsnormer hos den amerikanska utvärderingsföreningen American Evaluation Association (AEA) eller iakttar god utvärderingssed. Till skillnad från den federala nivån där utvärderingar styrs av riktlinjerna i GAO:s Gula bok, så finns det inga krav på riktlinjer på delstatsnivå.

Dessutom har statliga revisionsmyndigheter som OPPAGA inga befogenheter när det gäller att utforma eller verkställa politiken, de ansvarar endast för utvärderingen och kan ge råd till den lagstiftande församlingen. Kommunikationsnivån med övriga berörda aktörer beror på vilka myndigheter som staten har att göra med. Uppföljnings- och datainsamlingsprocessen sker i de myndigheter som ansvarar för genomförandet och datakvaliteten beror i hög utsträckning på myndigheternas arbete.

Ett intressant förhållande är autonomi för de olika federala programmen och den varierande betydelse som utvärdering har i respektive program beroende på hur politiken är utformad.

Ett märkligt exempel är NIST:s (National Institute of Standards – Nationella institutet för standarder och teknik) program för avancerad teknik, Advanced Technology Program (ATP). ATP är ett federalt program vars målsättning är att hjälpa industrin att investera i långsiktig högriskforskning med avkastningsmöjligheter som vida överskrider den privata vinsten. Genom att dela kostnaden med företagen påskyndar ATP utvecklingen av innovativ teknik på ett tidigt stadium och hjälper industrin att öka sin konkurrenspotential.

Kritisk utvärdering av ATP:s effekt på ekonomin är en viktig del av programmet. För att mäta de långsiktiga ekonomiska effekterna av FoU-medel har ATP tagit fram rutiner för ekonomisk analys som ytterligare flyttar fram positionerna när det gäller utvärdering av långsiktiga utfall av en FoU-investering.

Från 1993 finns en lag, Government Performance and Results Act (GPRA) som gör federala myndigheter ansvariga för att använda resurser på ett klokt sätt och uppnå resultat med programmen.

GPRA kräver att myndigheterna utarbetar planer för vad de avser uppnå, mäter hur de lyckas, fattar lämpliga beslut utifrån den information de samlat in och informerar kongressen och allmänheten om sina resultat. GPRA föreskriver att myndigheter utarbetar:

- en femårig strategisk plan, som innehåller en programförklaring och ställer upp långsiktiga mål och syften,
- årliga resultatplaner, som anger årliga resultatåtaganden för att uppnå de mål och syften som presenteras i den strategiska planen,

-
- årliga resultatrapporter, som utvärderar en myndighets framsteg med att uppnå utlovade resultat.

Det åligger därför myndigheter att planera för att uppnå mål och syften; se till att det finns resurser tillgängliga för att verkställa planerna; mäta och bedöma framsteg och koppla de resurser som faktiskt använts till uppnådda resultat samt att rapportera om uppnådda framsteg och effekter på framtida insatser.

GPRA anger därför ett nytt ramverk för utvärdering eftersom lagen är avsedd är att fokusera uppmärksamheten på utfallen av dess program.

Som Susan E. Cozzens¹⁴ framhåller:

GPRA krockar med det tidigare ramverket [...] på flera sätt. Den första inkonsekvensen är mellan utvärderingen av indata till forskningsprocessen och utvärderingen av dess utfall. [...] Den andra krocken är den autonoma kulturen kontra planeringskulturen, vilken återspeglats i kampen om tolkningen av resultatmålen. [...] Den tredje nya komponenten som GPRA har introducerat är samråd med berörda aktörer.

Sammantaget kan GPRA göra undersökande myndigheter, programansvariga och forskare till bättre strategiska tänkare, med en klarare insikt i hur deras strävanden kommer allmänheten till godo.

Advanced Technology Program (ATP)

Programmet för avancerad teknik, Advanced Technology Program, beskriver sig självt som att det “överbryggar klyftan mellan forskningslaboratoriet och marknaden”. Det skapades för att finansiera offentlig-privata partnerskap för att stödja utvecklingen av nya tekniker med potential att användas i hela den amerikanska ekonomin.

Sedan programstarten har NIST poängterat att utfallen måste bedömas och utvärderas och efterlyst en objektiv analys av partnerskapsprogrammets mål, verksamheter och resultat. De ekonomiska effekterna av ATP-projekt kan mätas på flera sätt, till exempel i form av ökad produktivitet, nya affärer, sysselsättningsfördelar och ökad BNP, inklusive både privat avkastning för det företag som deltar i projektet och olika spridningseffekter.

Vid utvärdering av effekterna av ATP-programmet har NIST varit mycket noggranna med att spåra programmets direkta och indirekta effekter. Den direkta effekten följer stödmottagarna och inkluderar privat avkastning till just de företag som varit direkt involverade i de ATP-finansierade projekten och spridningseffekterna till deras kunder. Den indirekta effekten handlar om att kunskapen som genererats av projektet utnyttjas av andra som står utanför projektet och som inte direkt bidragit till investeringskostnaden. Tillsammans representerar direkta och indirekta effekter det som ekonomerna kallar för projektets sociala avkastning.

För att spåra effekterna använder ATP olika verktyg, som:

¹⁴ Cozzens S. (2003). *Ramverk för att utvärdera vetenskaplig och teknisk politik i USA. I “Learning from Science and Technology Policy Evaluation - Experiences from the United States and Europe”, redigerad av Philip Shapira och Stefan Kuhlmann. Edward Elgar, Cheltenham UK och Northampton, USA.*

-
- Statistisk profilering av sökande, projekt, deltagare och tekniker
 - Uppföljning av alla projekts och deltagares framsteg
 - Statusrapporter för alla slutförda projekt
 - Detaljerade mikroekonomiska fallstudier av valda projekt och program
 - Ekonometriska och statistiska undersökningar av innovations-, produktivitets- och portföljeffekter
 - Makroekonomisk analys av valda projekt
 - Studier av särskilda frågor
 - Utveckling och testning av nya värderingsmodeller och -verktyg.

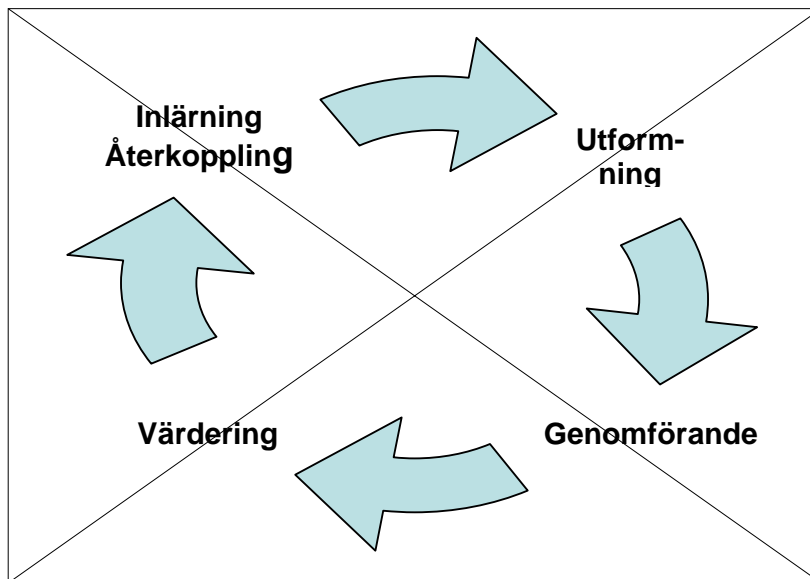
Som Irwing Feller från Pennsylvania State University förklarar:

ATP har gått längre än andra program som har försökt mäta direkta fördelar, genom att försöka mäta indirekta effekter eller spridningseffekter. Att mäta dessa effekter är en svår uppgift. ATP:s värderingstekniker ligger i framkanten av utvecklingen och har på många sätt flyttat fram positionerna ytterligare.

En utvärderingsinsats infördes i programmet av två skäl: 1) som ett styrverktyg, för att uppfylla programmets mål och förbättra dess effektivitet; 2) för att uppfylla de många externa önskemålen om ATP-programmets resultat. Utvärdering är ett kraftfullt verktyg när det integreras i programstyrningen och för att maximera effektiviteten behöver den göras i fyra steg:

- Utformning
- Genomförande
- Värdering
- Inläring och återkoppling

Figur 2. Programledningscykeln



Källa: www.atp.nist.gov

Men, som framgår i Figur 2, är detta inte enbart en linjär sekvens, utan en cykel i vilken bedömningar, inhämtade lärdomar och återkoppling bör återspeglas i lämplig modifiering av programmets utformning, genomförande, osv. Programutvärdering är därför fullständigt integrerat i programmets dynamiska struktur och bidrar till kontinuerliga förbättringar av programmet.

Rosalie Ruegg hos NIST påstår att:

Det finns några grundläggande principer att följa när man skapar ett utvärderingsprogram. Den grundläggande principen är att mäta i förhållande till uppdraget. En annan viktig princip är att koppla programmets aktiviteter till dess uppdrag på ett systematiskt sätt; prestationen till aktiviteterna samt kort- och långsiktiga utfall till prestationerna. Detta kallas för att "utveckla en utvärderingslogisk modell".

2.2 Lärdomar från andra länder

Det är tydligt att den tidigare beskrivningen av utvärderingsmetoder i andra länder föranleder en djupare diskussion. På vilken nivå bör utvärderingen genomföras? Hur viktigt är det med ex ante utvärdering? Vad är en rimlig tidsperiod för utvärdering? Hur ser situationen ut i Sverige när det gäller utvärdering? Räcker det att följa upp resultaten?

Den holländska erfarenheten erbjuder flera slutsatser avseende utvärderingsdelen av den politiska processen.

- 1 En väl genomtänkt utformning av politiken är en bra utgångspunkt för att uppnå övertygande bevis på orsakssambandet mellan det politiska instrumentet och dess resultat.
- 2 Det är nödvändigt med detaljerad datainsamling från både de företag som gynnas av politiken och sådana som inte gör det. Utan information om kontrollgruppen av icke-användare är det mycket svårt att hitta övertygande bevis på effektivitet. Och utan

tydliga och specifika frågor och begrepp kommer svarsfrekvensen hos de små och medelstora företag som deltar i undersökningen att vara låg och svaren svåra att tolka.

- 3 Effekter som uppträder på längre sikt kan endast observeras på lång sikt.
- 4 Vissa politiska instrument kanske också påverkar beteendet hos icke-mottagande företag vilket man bör ta hänsyn till för att undvika att över- eller underskatta effektiviteten av olika politiska program.

I Irland har programtänkandet införts systematiskt på flera politiska områden. Ur utvärderingssynpunkt har program både för- och nackdelar. Ett program som består av många olika områden och åtgärder är i allmänhet svårare att utvärdera. En fördel är dock att program är begränsade och genomskinliga i vad de avser att uppnå, vilket gör utvärdering enklare. Dessutom är Irlands system med "målsättningar på hög nivå" och ett antal makroindikatorer som systematiskt följs upp och granskas mer fördelaktigt för att bedöma ansvar. I de flesta utvärderingar har syftet inte varit att utvärdera effekterna utan metoderna för att förbättra stödprogrammet. I detta syfte har man använt "peer reviews" (bedömning av sakkunniga) för att få en djupare förståelse för hur stödet används. Utvärderingens roll som inlärningsfaktor saknar inte betydelse.

Slutligen kan de mest intressanta lärdomarna från USA sammanfattas med ett par intervjupersoners ord. Med Gary VanLandinghams ord (OPPAGA)

... kan den lagstiftande församlingen endast uppmärksamma ett visst antal frågor åt gången, så om man utvärderar allting kommer lagstiftaren inte att bry sig och regeringen att bortse från merparten av resultaten... Vi måste fråga lagstiftaren vad som är viktigt i ett visst program så att vi vet vilka frågor vi måste besvara och vilka som är utvärderingens syften. Lagstiftaren anger i regel inte syftena på ett särskilt specifikt sätt eftersom programmen normalt utformats som en politisk kompromiss.

Eller om vi använder Kathleen McTigues ord (NIST – ATP-programmet):

Vi kan lämna några rekommendationer: öka antalet retrospektiva analyser, införliva både direkta och indirekta analyser i lönsamhetsanalyser; kontinuerlig uppföljning och uppdatering av information, fortlöpande utveckling av utvärderingstekniker, identifiera och ta itu med nya frågor som uppstår vid utvärderingen, analysera både misslyckanden och framgångar, använd en effektiv blandning av interna och externa utvärderingsstudier och slutligen, dra större nytta av utvärderingsresultat i beslutsfattandet.

2.3 Utvärdering av näringspolitik i Sverige

I detta avsnitt återges det svenska regelverk som berör utvärdering samt exempel på två myndigheters arbete med att utveckla utvärdering inom sina områden.

2.3.1 Ramverk för utvärdering

I lagen om statsbudgeten (1996:1059) anges i de två första paragraferna att

I§ I statens verksamhet skall hög effektivitet eftersträvas och god hushållning iakttas.

Med statens verksamhet avses i denna lag sådan verksamhet som sköts av regeringen, domstolarna och de förvaltningsmyndigheter som lyder under regeringen.

2§ Regeringen skall för riksdagen redovisa de mål som åsyftas och de resultat som uppnåtts på olika verksamhetsområden.

Budgetlagen från 1996 innebar ett omfattande förändringsarbete inom offentlig förvaltning. Bland annat genomförde finansdepartementet ett projekt, Verktyg för Ekonomisk styrning inom Staten (VESTA) publicerad *Ekonomisk styrning – effektivitet och transparens* (Ds 2000:68) som en del i detta utvecklingsarbete.¹⁵

En utredning som också diskuterade konsekvenserna av budgetlagen var *Utvärdering och vidareutveckling av budgetprocessen* (SOU 2000:61). I den anges (sid 35) att med paragraf 2 skall förstås att budgetlagen ställer krav på regeringen att inför riksdagen redovisa de mål som ställts upp och de resultat som uppnåtts

Budgetlagen ställer i 2 § krav på att regeringen för riksdagen redovisar de mål som åsyftas och de resultat som har uppnåtts på olika verksamhetsområden. Bestämmelsen syftar enligt budgetlagspropositionen till att slå fast betydelsen av mål- och resultatstyrning i statlig verksamhet. Avsikten med resultatstyrningen anges vara dels att klargöra för mottagaren av statliga medel att denne måste ange vad som skall uppnås, dels att underlätta för riksdagens utskott att i ökad utsträckning ägna sig åt uppföljning och utvärdering. (SOU 2000:61 sid 35)

Riksdagsordningen ändrades året därpå (SFS 1974:153, kap 4 §18) och i denna stadgas 2001:

I utskottens beredning av ärendena skall ingå uppgiften att följa upp och utvärdera riksdagsbeslut inom de ämnesområden som för varje utskott anges i 4–6 §§ med tillhörande tilläggsbestämmelser. Lag (2003:180).

Till stöd för riksdagen och dess utskott bildades 2002 myndigheten Riksrevisionen som en del av riksdagens kontrollmakt. Skapandet av Riksrevisionen medförde dock att regeringens resurser för utvärdering minskade då riksrevisionsverket lades ned. Detta har delvis kompensrats av utökade utvärderingsresurser inom övriga myndigheter.

2.3.2 Styra med mål

Förhållandet mellan regeringen, myndigheterna och riksdagen regleras förutom i budgetlagen och i riksdagsförordningen även i verksförordningen (1995:1322) och förvaltningslagen (1986:223) samt i grundlagarna. Verksförordningen är mest intressant som styrdokument av myndighetens verksamhet från ett utvärderingsperspektiv.

I verksförordningen anges i §7 om myndighets chefens uppgifter att:

1) hushålla väl med statens medel

samt

2) fortlöpande följa upp och pröva den egna verksamheten och konsekvenserna av de författningsföreskrifter och särskilda beslut som rör verksamheten samt vidta de åtgärder som behövs

¹⁵ Se även regeringens skrivelse Skr 1000/01:151

Begreppet effektivitet finns med som en del i styrelsens ansvar att *pröva* (sic) om myndighetens arbete bedrivs effektivt. Begreppet mål finns även med men endast i kravet att de anställda skall vara införstådda med målet för myndighetens verksamhet.

I förordningens 31§ anges kraven på vad som skall ingå i en handling (dokumentation) för varje myndighetsbeslut. Här framgår att i svenska lagen endast kräver att detta dokument skall innehålla:

- Datum
- Beslutets innehåll
- Vem som har fattat beslutet
- Vem som har varit föredragande
- Vem/vilka som har deltagit i handläggningen av beslutet utan att delta i avgörandet

Varje specifik myndighet styrs utöver ovanstående lagar av en instruktion (i form av ett SFS), ett årligt regleringsbrev samt med särskilda uppdrag.

Med dessa instrument styrs verksamheten dels med avseende på de resurser som används i verksamheten så kallad finansiell styrning, dels med avseende på prestationer och effekter så kallad resultatstyrning. Den ekonomiska styrningen är verktyget för att skapa kontroll över statens finanser i syfte att dels åstadkomma den valda resursfördelningen bestämd av regeringens politiska prioriteringar, dels kunna återrapportera till riksdagen enligt budgetlagen. Konkurrerande fördelningsambitioner utifrån en begränsad budget leder till att krav på effektivitet i användandet av statens resurser men även krav på verifiering av effekter av genomförd verksamhet.

I den ovan nämnda departementsskriften (Ds 2000:68) diskuteras hur Riksdagen kan ställa krav på regeringens mål att de skall vara mätbara och uppföljningsbara samt att det skall vara meningsfullt att ställa prestationer mot kostnader. De så kallade SMART- kriterierna har förts fram som lämplig norm vid uppställandet av mål. Enligt dessa skall mål vara:

Tabell 1. SMART- kriterier

Specifika	De skall tydligt anges vad som skall uppnås
Mätbara	De skall vara möjliga att följa upp med hjälp av resultatindikatorer, nyckeltal eller liknande
Accepterade	De skall vara accepterade och uppfattas som relevanta av dem som ska genomföra den aktuella verksamheten
Realistiska	De skall vara möjliga att uppnå
Tidsatta	Tidpunkten då målen skall vara uppnådda skall anges

Källa: DS 2000:63 sid 54¹⁶

I skriften framförs att SMART-kriterierna kan tillämpas i större utsträckning på lägre nivåer av regeringens arbete dvs. under politikområdesnivå som verksamhetsområde och verksamhetsgrensnivå. Detta förslag är i linje med den utveckling som sker över hela världen i krav på ökad transparens och tydlighet i offentlig förvaltning. Förslaget har dock i Sverige inte fått något genomslag i lagtext, men regeringens formuleringar i de årliga

¹⁶ SMART-kriterierna finns även i engelsk version och EU-kommissionen har rekommenderat dessa sedan år 2002 se hänvisning i "SMART Innovation a practical guide to evaluation innovation programmes".

regleringsbrev och myndigheternas interna arbete (se nedan) utvecklas i linje med de mål som SMART-kriterierna anger.

Ett uppmärksammat problem med målstyrning är förhållandet mellan mål och de resurser som är satta till att nå målet helt eller delvis. Det ankommer naturligtvis på den valda regeringen att genomföra sin politik och sätta mål därefter. Ett problem är att dessa mål kan formuleras på olika nivåer. Vissa mål är mer visionära och svårigen kan kopplas till uppföljningsbara indikatorer. Den svenska IT-politiken hade till exempel en gång målet att ”bli först med ett informationssamhälle för alla”. Utan att någon del i satsen var tydliggjord i den av riksdagen antagna propositionen.¹⁷

Det andra problemet är att koppla visionära mål med myndigheters begränsade resurser. I den svenska lösningen ingår årliga samtal mellan regeringskansliet och myndigheterna i den så kallade mål- och resultatdialogen (något som diskuteras i SOU 2000:61 nämnd ovan).

Ett ytterligare problem med målstyrning är ju mer konkreta mål som regeringen beslutar om desto mer arbete får den att följa upp dessa. Arbetsfördelningen blir därför att myndigheterna måste tolka regeringens mål och omforma dessa till konkreta åtgärder inom myndighetens verksamhetsområde. Myndigheter behöver förankra detta hos regeringen och återkomma med resultat. Regeringen å sin sida måste bedöma myndigheternas ”tolkning” och utvärdera om detta är en framgångsrik väg att gå för att nå regeringens mål.

Ett lämpligt redskap för att anpassa SMART-kriterier till verksamheter och politikområden är att arbeta med så kallad *verksamhetslogik*. Ekonomistyrningsverket har beskrivit vad detta går ut på i en rapport (ESV 2001:16). Verksamhetslogik är en benämning bland flera som åsyftar metoden att analysera kopplingen mellan valda åtgärder och de mål man avser uppnå med åtgärderna. Verksamhetslogisk analys syftar till att identifiera tänkta förlopp och bakomliggande antaganden för valda åtgärder. Analysen leder till ett tydliggörande av troligheten att ett resultat uppstår och vilken del till detta som den genomförda åtgärden har.¹⁸

2.3.3 Vad är uppföljning och utvärdering?

I frånvaro av marknader behöver offentlig förvaltning ett instrument för sin effektivitetsrevision. Uppföljning och utvärdering är i allmänhet de verktyg som används. Emellertid så finns ingen gemensam standard som säger vad som är uppföljning/utvärdering och vad som inte är uppföljning/utvärdering. Förenklat kan sägas att uppföljning och utvärdering kan relateras till resultat som offentliga åtgärder producerar inom en viss tidsperiod. I Ds 2000:68 anges uppföljning och utvärdering i förhållande till uppställda mål vilket innebär att resultat sätts i relation till i förväg satta mål.¹⁹

Med uppföljning och utvärdering menas helt enkelt att beskriva och analysera hur dessa resultat eller frånvaron av resultat av offentliga åtgärder kan förstås. I syftet med

¹⁷ Se ITPS (2005) *En lärande IT-politik för tillväxt och välfärd ITPS 2003:015*

¹⁸ Se mer nedan om verksamhetslogik. Begreppet har en mängd liknande namn, programteori, händelsekedja, förändringsteori. Från företagsledningsmodeller är begreppet ”balanced scorecards” näraliggande.

¹⁹ Det finns en diskussion om standard se t.ex. www.fteval.at exempel på vad utvärdering bör innehålla. Inom utvärderingsvärlden finns det dock en konflikt mellan ”kvalitativ” och ”kvantitativ” metodologi (sic inte metod) som även svenska myndigheter möter vid förfrågningar om anbud på utvärdering.

uppföljning och utvärdering ingår sedan länge att forma ett underlag som kan avgöra om den offentliga åtgärden skall fortgå, kanske med en viss förändring i nivån, eller avslutas.

Effekt och effektivitet är nyckelord som också måste tas upp i detta sammanhang. Effektivitet kan endast relateras till någon form av jämförelsenorm. Är denna norm inte definierad går det inte att uttala sig effektiviteten. En offentlig åtgärd kan endast sägas ha effekt om det kan påvisas att åtgärden åstadkommer något som annars inte skulle realiseras inom den tidsperiod som åtgärder ska realiseras.

Effektivitet i bemärkelsen kostnadseffektivitet anger att offentliga åtgärder skall produceras med minsta möjliga kostnad.

Ur den uppsjö av litteratur om utvärdering är Vedung (1997)²⁰ ett internationellt standardverk.²¹ Vedung definierar utvärdering med:

Noggrann efterhandsbedömning av utfall, slutprestationer och förvaltning i offentlig verksamhet (Vedung 1997 s 33)

Vedung menar att utvärdering består av de två huvudkomponenter (s 32) *kvalificerad uppföljning* och *effektmätning*.

Inom EU delas utvärderingar upp i *ex ante*, *mid-term* och *ex post*. Dokument skall produceras över en offentlig åtgärds lämplighet före (*ex ante*), under utförande (*mid-term*) samt efter att åtgärden är genomförd (*ex-post*). Denna uppdelning är ett praktiskt tillvägagångssätt och förmodligen det som kommer att dominera framöver i än större utsträckning.

Vedungs terminologi skiljer sig från EU i två aspekter. Den ena är att Vedung betonar att utvärdering är något som sker efter att en offentlig åtgärd är utförd/igångsatt. *Ex ante*utvärdering vill Vedung ersätta med begreppet förhandsgranskning eller konsekvensanalys. Den andra aspekten är att Vedung menar att kvalificerad uppföljning är mer än bara uppföljning av enstaka resultatvariabler och inkluderar det vi nedan kommer att diskutera såsom verksamhetslogik. I EU:s *mid-term* begrepp omfattar det som även kallas ”process” eller en granskning av utförandeorganisationens lämplighet. Effektmätning är fokuserat på de kausala relationerna mellan offentliga utfall/resultat och önskvärda konsekvenser något som i EU:s terminologi är placerat i framför allt *ex post* utvärdering.

Någon terminologi som a priori är mest korrekt är inte möjligt att bestämma. Rent pragmatiskt är det rimligt att anpassa sig till EU:s konventioner, men sträva efter att utveckla kvaliteten i innehållet i respektive utvärdering oavsett om den kallas *ex ante*, *mid-term* eller *ex post*. Två begrepp som ITPS finner är av strategisk betydelse är verksamhetslogik och effekt då dessa fokuserar på de kausala kopplingarna som den offentliga åtgärden antar gäller. Diskussionen om dessa är fundamental under hela programmets livslängd men framför allt innan programmet igångsätts, *ex ante*, och naturligtvis i samband med *ex post*analysen. En *ex ante* utvärdering som inte utförligt granskar dessa kan inte ligga till grund för åtgärdens realiserande.

²⁰ Vedung E (1997) Utvärdering i politik och förvaltning, *Studentlitteratur*

²¹ Några aktuella rapporter är ”Review of methodologies to measure the effectiveness of state aid to SMEs” av Mosselman, Prince & Kemp samt ”SMART Innovation a practical guide to evaluation innovation programmes”.

2.3.4 Nuteks och Vinnovas utvärderingsstrategier

Utvärderingsambitionerna i närings- och innovationspolitiken har förändrats under de senaste åren. Vinnova har krav att rapportera effekter av sina (och sina föregångares) program. Nutek har förändrat sin organisation delvis för att ledningen skall få ett större inflytande över utformandet av nya åtgärder.

Var för sig har Nutek och Vinnova nyligen beskrivit i text hur de tänker sig hur uppföljning och utvärdering skall utvecklas på myndigheten i framtiden. Båda dokumenten är enligt ITPS steg i rätt riktning för att de skall kunna bestämma mer systemiskt huruvida de närings- och innovationspolitiska åtgärder de ansvarar för åstadkommer några samhällsekonomiska effekter.²²

Avsnitten nedan beskriver utvärderingsambitionerna i Sverige och kan jämföras med beskrivningarna från andra länder i denna rapport, men skall inte uppfattas som en utvärdering av dessa strategier.

Nutek

Nuteks häfte *Nytta* tar upp SMART-kriterierna ovan och understryker betydelsen av väl formulerade mål. Skriften tar även upp begreppet *förändringsteori* vilket är samma sak som begreppet verksamhetslogik. Man beskriver även skillnader mellan åtgärders aktivitet, resultat och effekt samt nämner att indikatorer för dessa bör identifieras som krav på kvaliteten i valda indikatorer. I skriften tas upp det betydelsefulla med att formulera en uppföljnings- och utvärderingsplan i vilka de olika momenten dokumenteras.

Organisatoriskt har Nuteks förändrats sedan 90-talet då analysverksamheten stod vid sidan av de operativa verksamheterna (som i Nuteks nya terminologi kallas leveransprocesser). Analys, som utgörs av analys, nyutveckling och utvärdering, är nu en del av ledningsprocessen för myndigheten och står därmed över de operativa verksamheterna. De senare har endast uppgiften att utföra specificerade uppdrag och utveckla nya åtgärder endast tillsammans med medarbetare i ledningsprocessen.

I skriften *Nytta* anges exempel hur Nutek önskar att det skall gå till:

Tabell 2. Nuteks organisation för planering och uppföljning av nya åtgärder

Moment	Aktivitet	När	Av vem
Planering	Målformulering	Samtliga insatser	Nyutvecklingsprocess
	Förändringsteori	Samtliga insatser	
	Definiera indikatorer	Samtliga insatser	
	Upprätta uppföljnings- och utvärderingsplan	Samtliga insatser	Upprättas av medarbetare i leveransprocess i samråd med medarbetare i inom utvärdering
	Referensmätning	Vid identifierat behov	
	Förhandsbedömning (ex ante)	Större och komplexa insatser bör ha en oberoende ex ante	

²² Nutek (2006) *Nytta – Nuteks system för uppföljning och utvärdering*

Vinnova (2007) *Vinnovas samlade strategi för uppföljning, utvärdering och effektanalys*” DNR 2006-04079. Vinnova har tidigare publicerat en skiss till strategi se Winqvist 2003 ”Strategies behind Vinnova’s evaluation policy” Fteval nr 19 samt en skiss till uppföljningssystem över forskningsdelen i innovationspolitiken Winqvist (2006) *Approach to a national system for monitoring university research in Sweden*” Fteval nr27

Uppföljning		Samtliga insatser	Genomförs av medarb i leveransprocess i samråd med medarb i delprocess utvärdering
Effektvärdering		Vid identifierat behov	Genomförs av utvärdering i samråd med medarb i leveransprocess
Utvärdering		Vid identifierat behov	Genomförs av utvärdering

Källa: Nutek (2007) s. 21

I ett schematiskt exempel visar Nutek på hur ett fiktivt program konkretiseras i utfallsmål, resultatmål och effektmål samt en beskrivning av (en rudimentär) förändringsteori.

Vinnova

Vinnovas inriktning på innovationspolitik dels som stöd till ren forskning inom vissa områden (sk behovsmotiverad forskning), framför allt i form av stöd till kompetenscentra, dels som stöd i syfte att stimulera till ökad forskning inom mindre företag, medför särskilda utmaningar för att identifiera resultat och effekter.

Vinnovas rapport har ett liknande innehåll som Nuteks rapport. Precis som Nutek placeras information från uppföljning och utvärdering i ett systemsammanhang. Informationen skall syfta till att beskriva bidraget av alla myndighetens åtgärder samt hur dessa samspelar med varandra. Det betonas att, eftersom många av Vinnovas åtgärder syftar till effekter på lång sikt att insamling av information kan ta nära 20 år. Även i denna rapport anges betydelsen av att

...planeringen av verksamheten dessutom innefattat en mycket tydlig effektlogik, dvs. en genomgång av vilka effekter som verksamheten på lång sikt ska leda till och hur vägen dit ser ut, eller vilka mekanismer som antas ha betydelse för utfallet. (sid 14)

Vinnova anger vidare hur viktigt det är att åtgärdens effektlogik bygger på rimliga antaganden om hur sambanden (mellan åtgärd och utfall) ser ut. Vinnova ägnar mer utrymme än Nutek att diskutera utveckling över tid. I ett diagram beskrivs att ackumuleringen av samhällsekonomiska effekter förväntas ta fart flera år efter projektavslut.

Till skillnad från Nutek anger inte Vinnovas utvärderingsstrategi några organisatoriska förändringar av myndigheten som kan befrämja genomförandet av utvärderingsstrategin. Man konstaterar att organisationen är "decentraliserad" med vilket menas i detta sammanhang att ansvaret för uppföljning och utvärdering av konkreta åtgärder ligger på den avdelning, enhet respektive handläggare som är ansvarig för genomförande.

I syfte att stärka uppföljning och utvärdering anför Vinnova istället utvecklandet av en norm (enhetligt arbetssätt s 31) för hantering av program, utlysningar och politikutveckling med så kallade piloter. Utöver detta menas att stöd i form av ett fungerande IT-system som medger flexibel inmatning, bearbetning och analys av projektdata skall medverka till effektivare uppföljning av enskilda projekt, men även jämförelse mellan olika programinsatser. Systematisk datafångst i alla FoU-program är en åtgärd som är angelägen att genomföra snarast. Stöd till avdelningarna finns även organisatoriskt genom utvärderingskompetens inom Vinnova, framförallt avdelningen för strategiutveckling. Rapporten anger även att det finns en särskild grupp, QA-gruppen, som ska stödja avdelningarna i formulerandet av programdokument med avseende på förväntade resultat

och effekter samt hur utlysningstexter skall formuleras. I rapporten anges att en ”tydlig organisation för utveckling av Vinnovas effektanalyser pågår” (s 21).

Vinnova deltar i flera internationella nätverk²³ i syfte att utveckla kompetensen inom utvärdering och analys.

Nutek och Vinnovas arbete med att tydliggöra verksamhetslogiken i de åtgärder de initierar är vällovlig. Sverige har troligtvis inte varit sämre än andra länder att granska sina offentliga åtgärder, men från den tidigare synen att analys är en stödverksamhet till att ”analys” är en strategisk verksamhet för myndigheten är det en förändring som möjliggör en utveckling mot effektivare program. Det är naturligtvis för tidigt att bedöma hur verksamheten kommer att utvecklas men utifrån dessa dokument anser ITPS att en diskussion bör utvecklas i åtminstone följande avseenden.

Fördjupa förståelsen av vad effekt kan vara

Det första är att myndigheternas ledning visar insikt över att myndighetens ”resultat” inte är detsamma som myndighetens ”effekt” på samhällsekonomin. Myndigheterna måste vinnlägga sig om att utveckla sin kompetens rörande vad som kan kallas effekter och vad som är resultat. För Nuteks del är valet av begrepp som ”effektvärdering” betydelsefullt. Man har dock låtit använda detta begrepp till stödmottagares egenvärdering av myndighetens stöd.²⁴ Detta tillvägagångssätt är vanligt förekommande vid beställda utvärdering från myndigheter. Emellertid är det otillräckligt som belägg för förekomst och storlek på effekt och kritiserat inom utvärderingslitteraturen.²⁵ Det är möjligt att Nutek menar att mer kausalt valida metoder kommer att utnyttjas under deras begrepp ”utvärdering” men detta återstår att se. Ett dilemma uppstår dock när Nutek hänvisar till att i regeringsuppdrag skall ”effektutvärdering” genomföras två år efter programslut (se sid 29). Frågan är vilken ”effekt” man talar om från respektive parter. Ett sätt att komplettera de subjektiva mätningarna som Nutek nämner är att etablera en mätning av utgångsläget. Något sådant refereras dock inte i effektvärderingen av produktutvecklingsprogrammet.

Ovan nämndes att begreppet ”effekt” uttrycker i det här sammanhanget en relation mot en specificerad norm eller med andra ord den kontrafaktiska situationen. Precis som alla andra kontrafaktiska utsagor betingas ”effekt” slutsatser utifrån mätningar av utgångsläget och hur väl specificerat det relevanta utgångsläget är.

Vinnovas rapport nämner EU:s begrepp för effekter, ”additionalitet” men nämner inte de distinktioner i form av ”input”, ”beteende” samt ”output” där det senaste är det mest önskvärda, men samtidigt det svåraste, vilket medfört ett intresse av att studera effekter på beteende och insatsstrategier.

Inkludera risker i verksamhetslogiken

Enligt ITPS bedömning består design och utfall av program av om myndigheten lyckas identifiera relevant förändringslogik. Ett dilemma i både Vinnovas och Nuteks rapport är

²³ Till exempel Taftie – The Association for Technology Implementation in Europe www.taftie.org

²⁴ Nutek (2006) ”Effektutvärdering av produktutvecklingsprogrammet” Dnr 70-2006-4350

²⁵ Storey, D.J. (2000). *Six Steps to Heaven: evaluating the impact of public policies to support small businesses in developed economies*, in Landstrom H and Sexton D.L. (eds.) *Handbook of Entrepreneurship*, Blackwells, Oxford, pp. 176-194.

Duignan P (2007) ”Seven possible outcome designs” <http://www.parkerduignan.com/oiiwa/model/>

att ingen av dessa tar upp risker och händelser som ligger utanför myndigheternas kontroll, men som kan ha inflytande på det myndigheten avser att påverka. Förmodligen är det realistiskt att anta att det finns inflytelserika faktorer som kan antingen ta udden av åtgärden eller dränka ”effekten”, t ex en kraftig konjunkturuppgång som medför en ökad tillgång på finansieringsmöjligheter av tidigare riskabla projekt. För att beskrivningar av verksamhetslogik skall stimulera till en utveckling av relevanta åtgärder bör sådana faktorer identifieras.²⁶

Diskutera populationer

Offentliga åtgärder inom näringspolitiken och innovationspolitiken i synnerhet i form av finansiella företagsstöd är relativt små till storleken om jämförelse med näringslivets totala investeringsvolym görs. Hur är det tänkt att dessa gör någon skillnad? Myndigheter och andra offentliga aktörer som Almi, Innovationsbron och Industrifonden bör relatera sina åtgärder till relevanta populationer. I den ovan nämnda effektvärderingen av produktutvecklingsprogrammet anges att Nutek år 2005 förmedlade 60 miljoner till 322 företag. Ingenstans i rapporten finns uppgifter om hur många företag som fanns i gruppen: legitima företag att söka stöd. Ej heller anges hur många av dessa som nått informationen.

I en verksamhetslogisk diskussion måste det rimligen finnas en diskussion om möjligheten att öka omfattningen av programmet under förutsättning att det lyckas. Men detta förutsätter en analys av representativitet dvs. vad är det för grupp av företag som ingår i programmet.

Bjuda in till diskussion av extern expertis

Den sista av ITPS synpunkter rör offentligheten av dokumentation. Av både Nutek och Vinnovas rapporter framgår inte i vilken omfattning man tänker delge externa intressenter dokumentation för diskussion. Framför allt är det av intresse att diskussion sker kring kausala länkar och eventuellt dolda antaganden innan åtgärder sätts. ITPS menar att det är en styrka att man utnyttjar sedvanligt referensgruppsförfarande även här. Vinnova har t ex nyligen presenterat ett förslag till strategi för ökad innovationskraft med inriktning på små och medelstora företag. I arbetet med strategin har en omfattande referensgrupp utnyttjats. När denna strategi operationaliseras bör även den härledda verksamhetslogiken diskuteras med externa experter.

2.4 Utmaningar

Behovet av ett gemensamt ramverk och möjligheter till ömsesidigt lärande genom att jämföra metoder i olika länder har inspirerat OECD att ta fram riktlinjer som möjliggör internationella jämförelser. Enligt detta ramverk bör program utvärderas med hänsyn till deras verkan inom ett antal områden, som t ex:

- Ändamålsenlighet: behandlar programmet en viktig målsättning?
- Överlägsenhet: är programmet effektivare än andra tillvägagångssätt, program eller instrument som skulle kunna uppnå samma mål?
- Systemeffektivitet: hur samverkar programmet med andra program och i vilken utsträckning beror dess verkan på förhållanden som skapats av andra statliga åtgärder?

²⁶ Se *ESV (2001) Verksamhetslogik* ESV 2001:16 se även *ESV 2006:7 Måluppfyllelseanalys* sid 27ff

-
- Egeneffektivitet: är programmet kostnadseffektivt?
 - Adaptiv effektivitet: i vilken utsträckning har resultaten från utvärderingarna återförts till utformningen och genomförandet av politiken? Säkerställer politikutformningen att det finns ett visst mått av flexibilitet när det gäller att bemöta oförutsägbara förändringar?

Dessutom betonar OECD att utvärderingsmetoden ska anges tydligt, att utvärderingsplanerna ska inkludera kostnads- och intäktsramar, att mätbara kriterier ska användas för att fastställa framgång när det är möjligt, och att ett kontrafaktiskt scenario ska användas. Enligt OECD:s rapporter har det sistnämnda varit den svagaste aspekten i många OECD-länders utvärderingsverksamhet.

När vi diskuterar befintliga utmaningar för utvärdering i Sverige måste vi komma ihåg OECD:s ramverk och fokusera på de aktörer som berörs och på aktuella rutiner för utvärdering i Sverige, både när det gäller ex ante- och ex post-utvärderingar.

OECD:s utvärderingsmodell är särskilt intressant för ex ante-undersökningar, i synnerhet avseende ändamålsenlighet och överlägsenhet. De viktigaste målen för en ex ante-utvärdering är:

- 1 att beskriva ett problem, som sannolikt beror på brister i marknaden, politiken eller systemet
- 2 att undersöka alternativa sätt att lösa problemet och föreslå en programlogik som gör att problemet kan lösas steg för steg.

Dessutom poängterar OECD-modellen att det är omöjligt att genomföra en utvärdering om det saknas tydligt angivna syften. Statsmakterna bör ange syften, och om det finns mer än ett syfte, ange vilket som har högst prioritet. När syftena formulerats måste de kvantitativa målen anges och endast därefter kan utvärderingen göras.

David Storey (2000) har gjort en analytisk ram för att klassificera en rad olika typer av analyser. Som framgår i Tabell 3 identifieras sex olika metoder där uppföljning visas i steg 1–3 och utvärdering i steg 4–6.

Tabell 3. De sex stegen.

UPPFÖLJNING	
Steg 1	Planerna tas i bruk
Steg 2	Mottagarens synpunkter
Steg 3	Mottagarens åsikt om vilken skillnad programmet medfört
UTVÄRDERING	
Steg 4	Jämförelse av mottagarnas resultatet med företag som inte mottagit stöd
Steg 5	Jämförelse med <i>matchade</i> företag
Steg 6	Hänsynstagande till urvalsskevheter

Källa: Storey (2000)

Mot bakgrund av denna enkla struktur är det nu intressant att ställa följande fråga: utför svenska myndigheter (t.ex. ITPS, Nutek och Vinnova) *utvärdering* enligt definitionen i Storeys steg 4–6? Det som Nutek kallar *effektivradering* är en slags subjektiv utvärdering där mottagande företag ombeds bedöma effektiviteten i en viss politik med hjälp av frågeformulär och en intervju. Detta är en vanlig metod, inte bara i Sverige, utan också i de flesta OECD-länder, både när det gäller utvärderingar som utförs av statliga myndigheter och av externa utvärderare. Men kan vi verkligen kalla detta för utvärdering?

Detta tillvägagångssätt har förvisso sina fördelar, särskilt ur kostnadssynpunkt. Uppföljningsprocessen som identifierar egenskaper och beskaffenhet hos mottagarna i ett projekt, liksom att samla in mottagarnas åsikter, är billiga och enkla procedurer som informerar om de mottagande företagens subjektiva bedömning och är allmän praxis i många länder. Men att följa upp med hjälp av frågeformulär och intervjuer ger ingen indikation om politikens effektivitet. Dessutom berättar de ingenting om måluppfyllelse eller additionalitet. De representerar endast delvis de företag som överlevt och troligen är svaren som ges socialt önskvärda svar.

Med bara uppföljning kan man därför inte få policyrelevanta insikter om politikens effektivitet och de olika myndigheter som ansvarar för utvärdering måste gå ett steg vidare. Detta är det enda sättet att ge lagstiftarna verktygen de behöver för att utforma nya åtgärder och program.

Steg 4 och 5 i Storeys klassificering är inte utan invändningar. Att jämföra mottagarna med typiska företag och matchande företag kan innebära problem med tanke på att de företag som får hjälp inte är typiska. Perfekt matchning kan vara mycket svårt och löper stor risk att bli föremål för snedvridet urval.

Oavsett dessa problem och behovet av att använda statistiska tekniker för att ta hänsyn till urvalsskevheter är det uppenbart att med hjälp av matchningsmetoder med mycket större sannolikhet kan få tillförlitliga resultat om de kausala effekterna av en viss politik. ITPS uppfattning är att Sverige behöver gå vidare mot en mer komplex utvärderingskultur för att sätta igång en inlärningsprocess där utvärdering ger information och skapar nödvändig kunskap för att möjliggöra bättre politikutformning.

Det finns två alternativa vägar som den svenska regeringen kan följa för att uppnå effektivare utvärderingar.

- 1 Det första alternativet är att utse en myndighet som utför både *ex ante* och *ex post* utvärderingar. En sådan myndighet skulle ansvara för datainsamling, men inte genomföra politiken. Därmed skulle tre aktörer beröras: departementet som ansvarar för politikutformningen, en myndighet (till exempel Nutek eller Vinnova) som

ansvarar för att genomföra politiken och en myndighet som ansvarar för att utvärdera den (t.ex. ITPS). Genom att tilldela olika ansvar (genomförande och utvärdering) till de två myndigheterna kan departementet öka oberoendet och tillförlitligheten i utvärderingsresultaten och minska eventuell partiskhet hos utvärderaren, men får samtidigt högre administrativa kostnader eftersom det finns två aktörer.

- 2 Det andra alternativet kan vara att utse en enda myndighet som ansvarar för att genomföra, analysera och utvärdera politiken. Berörda aktörer skulle därmed endast vara två: departementet och myndigheten och allt skulle göras *internt* enligt principerna i Nuteks system Nytt²⁷. Detta alternativ har fördelen att minska de administrativa kostnaderna genom att skapa eller utse en enda myndighet, men med en större risk för att myndigheten är partisk, i synnerhet när den utför ex ante utvärderingar, för att öka sin budget.

Detta är två alternativa vägar som Sverige kan välja, men det är viktigt att komma ihåg att, som visas i utvärderingsteori, det är bara genom oberoende utvärdering att vi kan undvika systematisk avvikelse. En utvärdering kallas oberoende när den utförs av enheter och personer som står fria från de ansvariga för utvecklingsinsatsens utformning och genomförande. En utvärderings trovärdighet beror delvis på i vilken mån den utförs på ett oberoende sätt. Oberoende förutsätter frihet från politisk påverkan och organisatorisk tvång. Oberoendet kännetecknas av obegränsad tillgång till information och fullständig självständighet vid utförandet av undersökningar och rapportering av fynd.

Oavsett vilken väg som väljs är det uppenbart att fokus behöver andras till en ny utvärderingskultur som riktar in sig på att angivna syften och kvantitativa mål är förståeliga; att man ser bortom enkla metoder för uppföljning och förbättrar instrumenten så att de tar hänsyn till urvalsskevhet och leder till en inlärningsprocess.

För att uppnå denna typ av kultur krävs därför ett långsamt förändrat beteende hos alla berörda aktörer. Det är endast när syftena är tydligt angivna, data är systematiskt insamlat och rollerna väl definierade som det är möjligt att koncentrera sig på de metoder som är bäst lämpade för att utföra utvärderingar.

Mot denna bakgrund fokuseras nästa kapitel på olika metoder, både på mikro- och makronivå, för att besvara frågan: "vad skulle ha hänt om politiken inte hade genomförts?"

²⁷ Nutek (2006) *Nytt* – Nuteks system för uppföljning och utvärdering

3 Utvärderingsmetoder

Nyckelfrågan i utvärderingen är *identifiering*. Vi behöver antaganden för att dra en provisorisk slutsats att en variabel påverkar en annan. Vi måste särskilt fokusera på det bästa sättet att identifiera en kontrafaktisk situation och försöka svara på det kontrafaktiska dilemmat: ”vad skulle ha hänt om politiken inte infördes?”

En sådan kontrafaktisk situation kan analyseras på både mikro- och makronivå, beroende på målet med politiken, storleken på de förväntade effekterna och tillgången till data.

När det gäller statligt stöd anses mikroanalysen vara det bästa sättet att hitta den kontrafaktiska situationen, eftersom den tillåter en jämförelse av en så kallad ”experimentell grupp” som tar emot politiska åtgärder med en annan ”kontrollgrupp” som representerar de företag som inte tog emot statliga medel. Ekonometriska metoder används för att konstruera en kontrollgrupp som är så lik behandlingsgruppen som möjligt, åtminstone när det gäller observerade kännetecknen.

Att slumpvis fördela programdeltagare i kontrollgrupper och experimentella grupper är ofta inte möjligt i fält och forskarens önskan om att utvärdera ett program med rigorös experimentell design är ofta oförenligt med målet att uppfylla de uttryckta behoven hos programdeltagarna. I sådana fall kan en analys på makronivå som belyser de allmänna effekterna på den nationella (eller regionala) ekonomin vara ett värdefullt alternativ. Genom en allmän eller partiell jämviktsanalys kan vi beskriva två scenarier vid en given tidpunkt – den jämvikt som uppnås i ekonomin utan ingripande och den jämvikt som uppnås efter införandet av politiska åtgärder, och vi kan hänföra skillnaden mellan dem till politiken. Analysen på makronivå har två huvudsakliga nackdelar: den kan bara användas för att studera effekterna av program som representerar en stor andel av de offentliga utgifterna och den ger ingen information om orsaksförhållanden. Å andra sidan representerar den dock ett värdefullt alternativ till mikroanalysen vid brist på data eller när programmet har en hög grad av komplexitet.

3.1 Analys på mikronivå: icke experimentell utvärdering och matchande bedömning

Nyckelfrågan vid en utvärdering är *identifiering* av en kontrafaktisk situation och behovet av att besvara det kontrafaktiska dilemmat: ”vad skulle ha hänt om politiken inte infördes?”

I experimentella studier fördelas deltagarna slumpmässigt ur en stor grupp av sökande som kan komma i fråga och genom en jämförelse mellan den stödmottagande gruppen och kontrollgruppen erhåller vi en opartisk bedömning av den genomsnittliga behandlingseffekten. Detta är dock inte fallet i icke-experimentella studier, eftersom de olika behandlingsgrupperna troligen kommer att skilja sig åt på ett icke slumpmässigt sätt och det är nödvändigt att konstruera en jämförelsegrupp som ligger så nära den experimentella kontrollgruppen som möjligt.

En rad olika ekonometriska metoder används för att konstruera en kontrollgrupp som är så lik behandlingsgruppen som möjligt, åtminstone när det gäller observerade kännetecken: slumpmässiga experiment, kvasiexperimentella utformningar, matchning av enskilda kännetecken som skiljer behandlings- och kontrollgrupper åt samt propensity score-matchning.

De första två metoderna används sällan i studier av ekonomiska effekter av offentlig politik av olika anledningar – slumpmässiga experiment är dyra och tidsödande och de är inte genomförbara när variablerna inte kan manipuleras, medan kvasiexperimentella utformningar främst kritiserats för sin stora partiskhet vid urvalet.

Problemet med partiskhet vid urvalet i konsekvensutvärderingar orsakas av att programdeltagarna skiljer sig från icke-deltagarna i kännetecken som inte kan observeras av utvärderaren vilket hänför sig till beslutet att delta i programmet och dess resultat.

Om utvärderaren med hjälp av slumpmässiga urval kan säkerställa att behandlings- och kontrollgrupperna är statistiskt likvärdiga i kvasiexperimentella och icke-experimentella utformningar används ekonometriska tekniker för att forma deltagande- och resultatprocesserna och för att komma fram till en opartisk bedömning av programmets effekt.

En metod som föreslagits för att lösa detta problem är matchning. Matchningsmetoden har utvecklats och använts mycket i den statistiska och medicinska litteraturen (Rubin 1978; Rosenbaum & Rubin 1983, 1985; Rubin & Thomas 1992), men är ganska nya för ekonomiska och politiska utvärderingar. Matchning kan anses vara en ”korrigeringsstrategi” som korregerar partiskheten vid urvalet och betyder att personer från olika behandlingsgrupper paras ihop med andra som liknar deras observerbara kännetecken.

Det faktum att matchning inte kräver en utarbetad resultatekvation gör den värdefull i utvärderingsstudier. Traditionella matchningsstudier parar ihop varje programdeltagare med en enda matchad icke-deltagare, men på senare år har flera studier använt ett viktat genomsnitt som innebär att en programdeltagare paras ihop med flera icke-deltagare.

Med Petra Todds (2006)²⁸ ord

antar vi att det finns två möjliga resultat, som kallas (Y_0 , Y_1), vilka representerar tillstånden utan och med behandling och en person kan bara vara i ett tillstånd åt gången, så bara ett av resultaten observeras. Det resultat som inte observeras kallas ett kontrafaktiskt resultat.

Behandlingseffekten för en enskild person är skillnaden mellan resultatet med behandling och resultatet utan behandling och det är inte direkt observerbart. Utvärderingsproblemet beror på data som saknas, för om vi kunde observera både Y_0 och Y_1 samtidigt skulle vi kunna konstruera en sådan skillnad för alla. En bedömning av effekten av ett program kräver därför att en slutsats dras om vilka resultat som skulle ha observerats i tillståndet utan program med tanke på att vi bara kan observera det totala resultatet på både den behandlade gruppen och den obehandlade.

Utvärderare måste därför använda utvärderingsparametrar som kan variera. På grund av brist på data är det vanligast att utvärderaren koncentrerar sig på den genomsnittliga effekten av åtgärder på de stödmottagande och försöker besvara frågan: hur mycket gynnades de personer som deltog i programmet i genomsnitt, jämfört med om de inte skulle ha deltagit?

En matchning bygger därför på ett antagande om att det finns en uppsättning observerbara kännetecken²⁹ Z och att resultaten är oberoende av programdeltagandet men beroende av Z eller, i Rosenbaum och Rubins (1983) terminologi, fördelningen av behandlingen är strikt

²⁸Todd, P. (2006). *Matching Estimators*. athena.sas.upenn.edu/~petra/papers/mpalgrave2.pdf

²⁹ Observerbara kännetecken kan exempelvis vara: kostnader, antal anställda och geografisk lokalisering.

ignorerbar med tanke på uppsättningen observerbara kännetecken. Det krävs vidare för alla Z att det finns en positiv sannolikhet för att antingen delta eller inte delta i programmet och att fördelningen av de matchande variablerna Z inte påverkas av huruvida behandlingen tas emot eller ej.

Todd (2006) belyser att matchande studier kan vara svåra att genomföra när uppsättningen villkorsvariabler Z är många och när Z är separata. Om Z är kontinuerliga kan konvergenstalen dessutom bli långsamma på grund av den så kallade *dimensionella förbannelsen*. Rosenbaum and Rubin (1983) diskuterar det dimensionella problemet och konstaterar att när resultatet Y_0 är oberoende av programdeltagande beroende på Z , är det även oberoende av deltagande, beroende på sannolikheten för deltagande. När matchningen på Z är giltig är därför propensity score-matchningen $P(Z)$ också giltig. Det är en av anledningarna till att det mesta av litteraturen om matchning är inriktad på propensity score-matchning.

Med *propensity score-metoder* bedömer utvärderaren propensity score med en parametrisk logit- eller probit-modell eller med en halvparametrisk metod³⁰ och sedan, i nästa steg, matchas enskilda personer på grundval av osannolikhet att delta.

Andra begränsningar med icke-experimentella metoder jämfört med experimentella metoder är att icke-experimentella metoder inte garanterar att stödet för jämförelsegruppen motsvarar stödet för den behandlade gruppen – de kommer sannolikt att kombinera två olika uppsättningar data för deltagare och icke-deltagare, vilka ofta samlas in med hjälp av olika metoder.

I litteraturen föreslås därför alternativa matchningsbedömare, exempelvis: Nearest Neighbor Matching, Interval Matching, Kernel och Local Linear Matching och Difference in difference Matching³¹.

Om matchning med propensity score och regressionsmetoder med flera variabler används, kontrolleras urvalet utifrån observerbara kännetecken som sker med hjälp av instrumentella variabelmetoder som utvärderaren kan kontrollera för, vilket utgår från urvalet av icke observerbara kännetecken. Instrumental Variables är en teknik som identifierar en faktor som avgör mottagandet av ett stöd, men som inte påverkar resultat av intresse. Denna faktor används sedan för att simulera vem som skulle ha varit i behandlingsgruppen och vem som skulle ha varit i kontrollgruppen om mottagandet av stödet baserades på denna faktor. Skillnaden i resultat mellan dessa simulerade behandlings- och kontrollgrupper är sedan effekten.

³⁰ *Propensity score kan inte bedömas med en icke-parametrisk modell, för då återuppstår det dimensionella problemet.*

³¹ *Skillnaden i ett givet resultat mellan mottagare av åtgärder (behandlingsgruppen) och en jämförelse- eller kontrollgrupp beräknas innan projektet genomförs. Denna skillnad kallas den "första skillnaden". Skillnaden i resultat mellan behandlings- och kontrollgrupper beräknas på nytt en tid efter projektets genomförande, och detta kallas den "andra skillnaden". I difference-in-difference-tekniken är projektets effekt den andra skillnaden minus den första skillnaden. Logiken är att åtgärds effekt är skillnaden i resultat för behandlings- och kontrollgrupper efter att studien har genomförts, netto efter redan existerande skillnader i resultat mellan behandlings- och kontrollgrupper som uppstår före studien.*

3.2 Analys med rumslig ekonometri *

Traditionellt har empiriska analyser av effekterna av olika typer av statligt stöd baserade på nationalekonomisk teori fokuserat på hur dessa stöd påverkat utvecklingen i just det företag eller den region de utbetalats till. Då det är rimligt att anta att det inte endast är statliga stöd som påverkar ett företags eller en regions utveckling tas ofta andra potentiellt viktiga förklaringsfaktorer med i analysen för att på så sätt kunna kontrollera, eller konstanthålla, för dessa. Exempel är tillgången på humankapital, råvaror och naturresurser, investeringar, den lokala offentliga servicen, lokala skattesatser, demografiska faktorer etc. Även tidsdynamiken, dvs. det faktum att det ofta tar viss tid innan effekten av en åtgärd blir synlig, har givits tämligen stort utrymme inom både den tillämpade empiriska litteraturen och den mer metodinriktade ekonometriska litteraturen. Här utgör utvecklingen av så kallade dynamiska paneldatametoder viktiga bidrag (se exempelvis Holtz-Eakin, m fl (1988), Arrelano och Bond (1991), och Ahn och Schmidt (1995), se även Baltagi (2001) för en översikt).

Även om det är en rimlig utgångspunkt att utvecklingen inom ett företag eller en region till stor del påverkas av faktorer inom företaget eller regionen, är det rimligt att anta att utvecklingen inom ett företag eller en region påverkas av utvecklingen inom närliggande företag eller regioner. Skälen till att så skulle vara fallet är flera. Exempelvis är ett rimligt antagande att ett statligt stöd till ett företag inom en region även kan påverka det företagets underleverantörer och/eller återförsäljare lokaliserade i närliggande regioner. Men stödet kan även påverka ett företags konkurrenter och därmed få totalt sett en negativ effekt på exempelvis sysselsättningen i regionen eller närliggande regioner. Det finns empirisk forskning baserad på svenska data som tyder på att om statligt regionalstöd har en negativ effekt på den regionala arbetslösheten, har detta i sin tur en positiv effekt på inflyttningen till regionen vilket är av stor regionalpolitisk betydelse (se Aronsson m fl (2001) och Lundberg (2003)). Andra studier visar att inflyttningen till attraktiva regioner även har en positiv effekt på inflyttningen till närliggande regioner (Lundberg (2006a)). En möjlig förklaring till detta resultat är att fastighetspriserna inom attraktiva regioner tenderar att vara högre än inom mindre attraktiva regioner vilket kan leda till att individer med lägre inkomster bosätter sig i närliggande regioner med lägre fastighetspriser.

För att fånga den här typen av effekter har allt fler studier under senare tid kommit att tillämpa så kallad rumslig ekonometri. Rumslig ekonometri är ett samlingsbegrepp för deskriptiva statistiska test och förklarande regressionsmodeller som tar hänsyn till att händelser som uppstår inom exempelvis ett företag eller en region kan påverka utvecklingen inom andra företag eller närliggande regioner. Just det faktum att det med rumslig ekonometri går att ta hänsyn till olika grad av beroende mellan företag och regioner gör att dess inbördes lokalisering modelleras. Från att tidigare nästan uteslutande varit en metod som används inom regional vetenskap, urban och regional ekonomi samt ekonomisk geografi, har rumslig ekonometri under senare tid blivit allt mer använd även inom traditionell empirisk nationalekonomi och teoretisk ekonometri. En förklaring till detta är att rumslig ekonometri rent principiellt är tillämplig så snart det finns anledning att tro att en agents beteende (region, individ, företag, myndighet, etc.) påverkar beteendet eller utvecklingen hos en annan agent. Som exempel på tillämpningsområden kan nämnas efterfrågeanalys (Case, 1991) internationell ekonomi (Aten, 1996) arbetsmarknadsekonomi

* av Johan Lundberg, Centrum för regionalvetenskap vid Umeå universitet (CERUM)

(Topa, 1996), offentlig ekonomi och lokal offentlig ekonomi (Case m fl, 1993; Murdoch m fl, 1993; Holtz-Eakin, 1994; Brueckner, 1998; och Lundberg, 2006b), skattekonkurrens mellan regioner (Besley och Case, 1995; Brett och Pinkse, 1997; Bivand och Szymanski, 1997; Revelli, 2002), regional tillväxt och migration (Armstrong, 1995; Rey och Montouri, 1999; Fingleton, 2001; Lundberg, 2006a).

I en allt mer internationaliserad värld där transportkostnaden tenderar att utgöra en allt mindre del av företagets totala produktionskostnad och där konsumenter i allt större utsträckning handlar och inhämtar information om olika produkter och priser via Internet är det motiverat att ställa frågan om lokaliseringen av företag och produktion längre har någon betydelse? Och spelar det någon roll för företagen hur de är geografiskt lokaliserade relativt varandra? För många slutkonsumenter är det heller inte av avgörande betydelse vart produkten är producerad.³² Samtidigt som det finns argument för att lokaliseringen av olika verksamheter inte längre skulle spela så stor roll finns tydliga tecken på att en mängd verksamheter tenderar att samlas till vissa specifika platser. För att ta Sverige som exempel så finns stark koncentration av glasbruk i Småland, telekommunikation och elektronik i Kista, finansmarknad i Stockholms innerstad, läkemedel i Lund etc. Orsakerna till detta är flera. Frånsett transportkostnaderna kan ett mindre företag som inte har resurser att genomföra en lokaliseringsanalys dra nytta av att lägga sig i närheten av en större konkurrent då närheten till ett större företag ofta ger närhet till vissa kringfunktioner så som speciell administrativ (exempelvis för branschen specifik juridisk hjälp) och teknisk kompetens, underleverantörer och/eller återförsäljare. Det kan även vara lättare att rekrytera personal med specifik kompetens inom en region där det redan finns ett antal företag inom samma bransch. Men även stora företag kan dra fördel av att lägga sig nära andra företag inom samma bransch. Informella möten och informationsläckage mellan företag är ofta vanligare om det geografiska avståndet mellan företagen är litet. Vid dessa möten kan för branschen viktig information om nya teknologiska framsteg, nya trender, information om medarbetares kompetens etc. spridas. Andra viktiga faktorer som påverkar företagets lokalisering är tillgången på naturresurser och råvaror (exempelvis gruvnäring och skogsindustri), transportleder och kommunikationer (tidigare skogsindustrins lokalisering vid älvmyningar). Ytterligare en kanske inte helt oväsentlig faktor är beslutsfattande chefers personliga preferenser för olika platser.

Ur ett fördelningspolitiskt perspektiv är lokaliseringen av produktionen av stor betydelse. Skälet är att med produktionen följer inkomster och arbetstillfällen och om produktionen blir ojämnt fördelad över regioner tenderar även inkomsterna att bli ojämnt fördelade. Samtidigt som produktionen tenderar att lokaliseras till storstadsregioner eller regioner med specifik kompetens eller tillgång till viktiga råvaror och naturresurser, finns också tecken som tyder på att de regionala skillnaderna i medborgarnas inkomster ökar. För att ta Sverige som ett konkret exempel kan nämnas att år 1981 var den genomsnittliga medelinkomsten i den svenska kommun med de högsta inkomsterna drygt dubbelt så hög som medelinkomsten i den svenska kommun som då hade de lägsta inkomsterna. År 1999 hade den siffran stigit till drygt 3 gånger. Dessutom ökar befolkningen inom storstadsområdena medan den minskar i mer glest befolkade regioner. Sammantaget kan

³² Detta naturligtvis betingat på att produktionen sker i enlighet med de värderingar som konsumenten har vad gäller arbetsförhållanden, politisk uppfattning och fördelning av inkomsterna av produktionen. Det ligger utanför ramen för denna rapport att utförligt diskutera dessa faktorer varför det antas att produktionsförhållandena är de samma oberoende vart produktionen sker.

en sådan utveckling resultera i stora skillnader i service nivå, offentlig så väl som privat, mellan olika regioner.

Detta avsnitt behandlar översiktligt på vilket sätt rumslig ekonometri kan användas för att analysera effekterna av statligt stöd, både allmänt stöd till olika regioner och mer specifika företagsstöd. Paragrafen gör på intet sätt inte anspråk på att vara en heltäckande sammanfattning av samtliga test, estimatorer eller tillämpningsområden för rumslig ekonometri utan ska ses som en kortfattad översikt.

Rumslig ekonometri skiljer sig från ”vanlig” ekonometri genom att hänsyn tas till att det som sker inom en region³³ även kan påverka vad som händer inom närliggande regioner. Detta utan att exkludera andra potentiellt viktiga förklaringsfaktorer till ett viss skeende. Det går även att inkludera och modellera tidsdynamiken i rumsliga modeller.³⁴

För att ta ett konkret exempel. Antag att staten beslutar att ge ett extra arbetsmarknads-politiskt stöd i form av starta-eget-bidrag. Stödet är i och för sig riktat direkt till det enskilda företaget men kommer sannolikt inte att vara jämnt fördelat över samtliga kommuner. Det faktiska syftet med bidraget antas i detta exempel vara att påverka sysselsättningen. Med hjälp av rumslig ekonometri går det att på statistisk väg testa om denna åtgärd även påverkar sysselsättningen i närliggande regioner. Exempelvis kan det nystartade företaget nyttja underleverantörer i andra regioner, bli underleverantör till företag lokaliserade i andra regioner eller anställa personal bosatt i andra regioner. Det går även att med rumslig ekonometri testa hur långt detta beroende sträcker sig. Är det endast till angränsande regioner eller även till regioner som ligger längre ifrån? Ytterligare en intressant frågeställning som går att analysera med rumsliga statistiska metoder är om det i exemplet ovan nämnda nystartarbidraget har effekt på överlevnaden för redan existerande företag inom samma bransch, och om så är fallet, hur långt denna påverkan sträcker sig?

De ovan angivna frågeställningarna är intressanta utifrån ett policyperspektiv då det ger kunskap om en åtgärds spridningseffekter. Om en åtgärd till ett visst specifikt företag eller region tenderar att påverka även andra företag och regioner än de som erhåller stöd bör detta tas hänsyn till vid utformningen av åtgärder och regelsystem för det specifika stödet. Därför är det viktigt att öka kunskapen kring olika typer av stöds spridningseffekter och ta med dessa vid analyser av stödets samlade effekter. Om det föreligger rumsliga effekter av en viss stödåtgärd och dessa inte tas med i analysen kommer de övriga parametrarna i modellen och dess signifikansgrad inte att bli korrekt estimerade vilket leder till felaktiga slutsatser.

³³ Som nämnts i inledningen behöver det inte röra sig om just regioner. Det kan vara ett beroende mellan företag, individer, etc. Region kommer dock genomgående att användas för att i texten kunna ge konkreta exempel.

³⁴ En genomgång av statistiska metoder för att fånga och modellera tidsdynamiken ligger utanför ramen för denna rapport. En bra genomgång av dynamiska paneldatametoder ges i Baltagi (2001).

3.2.1 Ett konkret exempel på tillämpning av rumslig ekonometri³⁵

För att åskådliggöra hur rumslig ekonometri praktiskt kan användas för den här typen av analyser baseras en del av den fortsatta framställningen och beskrivningen av olika rumsliga test och estimatorer på ett konkret men hypotetiskt exempel. Utgångspunkten är att staten under ett visst år beslutat ge ett extra arbetsmarknadspolitiskt stöd i form av starta-eget-bidrag. Stödet är riktat direkt till det enskilda företaget och kommer därmed inte att vara jämnt fördelat över samtliga kommuner. Det faktiska syftet med bidraget antas i exemplet vara att påverka sysselsättningen inom den region där företaget är lokaliserat, men det är även av intresse att fånga dess potentiella effekt på sysselsättningen inom närliggande regioner och, om så är fallet, hur långt denna påverkan sträcker sig.³⁶ Genomförandet löper över fyra steg:

- 1 Insamling av data och användning av Geographic Information System (GIS)
- 2 Viktmatriisen – definition av det rumsliga beroendet
- 3 Deskriptiva test för rumslig korrelation
- 4 Regressionsanalys

Insamling av data och användning av Geographic Information Systems (GIS)

För att genomföra den här typen av analys krävs information om företagets geografiska lokalisering, hur mycket stöd företaget erhållit, antalet anställda, samt sysselsättningen eller förändrad sysselsättning inom regionen från det att stödet betalats ut och exempelvis två eller tre år framåt i tiden. För att kontrollera för andra potentiellt viktiga förklaringsfaktorer till förändrad sysselsättning är det viktigt att ha information om exempelvis tillgången till annan offentlig service inom regionen, andra sysselsättningsåtgärder inom region, regionala demografiska uppgifter (åldersstruktur), tillgången på humankapital inom regionen, etc.³⁷ Vidare bör datamaterialet omfatta samtliga regioner och helst över en längre tidsperiod. Genom att ha information över en längre tidsperiod blir analysen inte betingad på valet av år, och därmed kan mer generella slutsatser dras.

GIS är ett datorbaserat system för att samla in och länka lägesbunden information till dess geografiska koordinater. Det kan exempelvis vara ett företags vinst, omsättning, antalet anställda, etc. Just det faktum att informationen är länkad till dess geografiska koordinater gör GIS mycket användbart vid rumsliga analyser. Informationen blir därmed relativt enkel att visualisera med hjälp av kartor och det är enkelt att beräkna det geografiska avståndet mellan olika observationer. I exemplet ovan är all information i data setet bunden till en specifik geografisk plats eller ett geografiskt område. Företaget är lokaliserat på ett visst ställe, det statliga stödet direkt kopplat till ett specifikt företag, åldersstrukturen och sysselsättningen bunden till en viss kommun etc.

³⁵ För en mer teknisk genomgång av de nedan beskrivna testen och estimatorerna hänvisas till Anselin (1988) vilket är den kanske mest heltäckande läroboken inom rumslig ekonometri och utgör en mycket bra introduktion till området. För senare referenser hänvisas till enskilda artiklar, framför allt av Anselin, Cressie, Florax, Haining, Kelejian, LeSage, Pace, Pinkse, Prucha, och Rey.

³⁶ Syftet skulle naturligtvis även exempelvis kunna vara att öka antalet företag. Det är då bara att byta ut region mot företag i texten. För att göra framställningen så enkel som möjligt ges bara ett exempel.

³⁷ Det finns naturligtvis andra potentiella förklaringsfaktorer till en förändrad sysselsättning inom en region. Vi bortser dock från dessa för att förenkla framställningen något. Vi bortser även från en mer komplicerad tidsdynamik.

Viktmatrisen – definition av det rumsliga beroendet

Efter att datainsamlingen är genomförd är det första steget i en rumslig analys att definiera det rumsliga beroendet. Med andra ord, graden av beroende mellan de olika regionerna måste definieras. Det vore önskvärt att tillsammans med övriga parametrar i modellen på statistisk väg estimeras graden av beroende mellan de olika regionerna. Oftast är detta dock inte möjligt eftersom antalet parametrar i en sådan estimation skulle överstiga antalet observationer.³⁸ Inte ens om det finns tillgång till en rad observationer över tiden är det i de flesta fall inte praktiskt genomförbart eftersom det skulle innebära att antalet tidsperioder översteg antalet individer (eller regioner) varför graden av beroende mellan de olika regionerna måste antas i förväg. Detta innebär i sin tur att resultaten från all rumslig ekonometri, deskriptiva test så väl som förklarande regressionsanalys, bygger på ett i förväg gjort antagande om hur det rumsliga beroendet, eller graden av påverkan mellan de olika regionerna, ser ut. Detta är viktigt att komma ihåg vid tolkningen av resultaten från dessa modeller.

Vanligast är att beroendet bygger på det geografiska avståndet mellan de olika regionerna eller om regionerna direkt angränsar till varandra. I det första fallet definieras ofta beroendet som avtagande med avståndet mellan regionerna vilket i många fall är rimligt. I det andra fallet är beroendet diskret, dvs. antingen finns ett beroende eller inte. Det är i flera fall av intresse att veta hur långt effekterna av en viss åtgärd sträcker sig. Exempelvis kan det vara intressant att studera om ett givet företagsstöd sprider sig till regioner inom ett visst avstånd från den region där företaget som erhållit bidraget är lokaliserat. Genom att konstruera en mängd viktmatriser med olika så kallade ”cut off” värden går det med hjälp av rumslig ekonometri testa hur långt effekterna sträcker sig. Med ”cut off” värde menas att beroendet antas vara inom ett visst avstånd, därefter antas inget beroende alls. Det finns även andra definitioner som exempelvis bygger på närhet i befolkningsstorlek, regionernas yta, graden av invandrare etc. Exempelvis kan det vara så att Göteborg i större utsträckning påverkas av saker som sker i Stockholm än i geografiskt mer närliggande kommuner. Skälet till detta skulle vara att Göteborg och Stockholm är Sveriges befolkningsmässigt största kommuner och därmed på olika sätt skiljer sig från landets övriga kommuner. Samtidigt kan det i vissa fall vara så att kranskommuner till Stockholm i större utsträckning påverkas och influeras av utvecklingen i kranskommuner som ligger i närheten av Göteborg.

Utifrån definitionen av hur beroendet mellan de olika regionerna ser ut skapas sedan en viktmatris där varje element i viktmatrisen definierar beroendet mellan två regioner. Det viktiga med definitionen av elementen i viktmatrisen är att de kan antas vara exogent givna. Om så inte är fallet uppstår metodmässiga problem. Givet att antagandet är exogent givet går det sedan att med hjälp av de test och regressionsmetoder som beskrivs nedan testa för olika typer av beroende för att se vilken definition som är mest tillämplig på den aktuella frågeställningen och datamaterialet. Som tidigare nämnts är det viktigt att hela tiden vara medveten om att samtliga resultat är betingade på det antagande som gjorts kring hur det rumsliga beroendet ser ut. Detta ställer höga krav på rimligheten i de definitioner som görs av elementen i viktmatrisen, något som bör ges stor uppmärksamhet

³⁸ *Antag en analys av sysselsättningen i Sveriges samtliga kommuner vilka idag uppgår till 290 stycken. Då skulle antalet tidsperioder behöva med god marginal överstiga 290 för att det rumsliga beroendet ska kunna estimeras, vilket i praktiken inte är möjligt att samla in data för. Dessutom har indelningen av Sveriges kommuner och kommunikationerna förändrats under den tidsperioden vilket skapar metodmässiga problem.*

och utrymme i presentationen av analysen och resultaten. Av estimations- och tolkningsmässiga skäl radstandardiseras ofta viktmatrisen vilket innebär att varje rad i matrisen summeras till 1.

Deskriptiva test för rumslig korrelation

Efter att det rumsliga beroendet mellan regionerna är definierat går det att genomföra både deskriptiva test för rumslig korrelation och inkorporera detta beroende i regressionsmodeller. Skillnaden mellan deskriptiv statistik och regressionsanalys är att den deskriptiva statistiken beskriver hur datamaterialet ser ut medan regressionsanalysen används för att försöka förklara vad som påverkar olika skeenden. I en regressionsanalys går det att ta hänsyn till andra potentiellt viktiga förklaringsfaktorer till ett visst skeende och även analysera olika faktorer inbördes relativa betydelse. Det finns olika typer av statistiska test som på ett systematiskt sätt söker säkerställa tendenser och mönster i data. Regionala skillnader i exempelvis sysselsättning eller förändrad sysselsättning kan i och för sig åskådliggöras visuellt i olika typer av geografiska kartor, men visuellt åskådliggjorda mönster får dock en större tyngd om de dessutom kan statistiskt säkerställas. Naturligtvis kan de nedan beskrivna testen även användas för att analysera regionala mönster vad gäller utbetalningen av själva bidraget.

Det finns två olika grupper av test för rumsligt beroende. Den ena gruppen kallas globala test vilka kan besvara frågan om det föreligger något allmänt beroende mellan samtliga regioner som ingår i datamaterialet. Exempel på globala test är Moran's I och Geary's C . Baserat på dessa två test, som tolkningsmässigt inte skiljer sig åt, går det att säkerställa förekomsten av kluster av höga eller låga värden, eller om data är organiserat så att höga och låga värden ligger nära varandra. Det går dock inte att utifrån testen uttala sig om det är antingen höga eller låga värden som är korrelerade med varandra, och inte heller vart denna korrelation är som starkast.

Den andra gruppen av test för rumslig korrelation kallas lokala test och med hjälp av dessa går det att upptäcka klusterbildningar inom datamaterialet. Det går alltså att utifrån dessa test att upptäcka observationsspecifika mönster. De lokala testen ger svar på frågan om en viss observation är omgiven av andra observationer med antingen lika eller olika värden, antingen jämfört eller oberoende av värdet i den egna regionen. Det går exempelvis att med hjälp av de lokala testen upptäcka om en viss region med hög sysselsättning är omgiven av regioner med höga eller låga sysselsättningsnivåer. De lokala testen används också ofta för att upptäcka extrema observationer som kan störa en senare regressionsanalys. Exempel på lokala test är G_i -statistikorna och LISA (Local Indicators of Spatial Association). Globala och lokala test används ofta i kombination med varandra för att på så sätt kunna upptäcka både generella mönster och klusterbildningar.

Regressionsanalys

Till skillnad från deskriptiva test går det med hjälp av regressionsanalys att ta hänsyn till andra potentiellt viktiga förklaringsfaktorer till ett visst skeende och vilken av dessa faktorer som har störst betydelse. Ur ett policy perspektiv är detta av stor vikt då en regressionsanalys kan ge svar på frågan om exempelvis en viss stödåtgärd har större eller mindre betydelse för den regionala utvecklingen jämfört med andra potentiellt viktiga förklaringsfaktorer.

Utifrån de ovan nämnda testen för rumslig korrelation går det att påvisa rumsligt beroende, men testen ger ingen vägledning kring hur detta beroende ska inkorporeras i en regressionsmodell. Inom rumslig ekonometri skiljs nämligen ofta mellan två olika typer av regressionsmodeller. Den ena, den så kallade lag-modellen, ingår det rumsliga beroendet som en extra variabel i regressionskvationen. Från den typen av regressionsanalys går det att besvara frågeställningen om det föreligger något direkt rumsligt beroende mellan regioner så som det definierats i viktmatrisen. Det går dessutom att besvara frågeställningen om exempelvis effekterna av en åtgärd inom en region har en positiv eller negativ effekt på utvecklingen inom närliggande regioner. Det går också att utifrån lag-modellen att estimeras så kallade reaktionsfunktioner, dvs. hur en viss region reagerar på någonting som händer inom en annan region. Reaktionsfunktioner är vanligt förekommande inom exempelvis litteraturen kring skattekonkurrens. Den andra typen av rumslig regressionsmodell är den så kallade error-modellen där det rumsliga beroendet istället uppträder i slumtermen i regressionskvationen. Tolkningen av ett signifikant rumsligt beroende i error-modellen är att en störning som införs i systemet sprider sig till samtliga regioner, inte endast till de närmast angränsande. Det finns statistiska test för att diskriminera mellan lag- och error-modellen, dvs. test som vägleder vilken av de två specifikationerna som passar det aktuella data setet bäst.

En källa till felaktigt specificerade rumsliga modeller och därmed felaktiga tolkningar av parameterestimatet kommer av förekomsten av så kallad strukturell instabilitet. Strukturell instabilitet kommer antingen av att de ”sanna” parameter estimaten inte är de samma för alla regioner eller att själva funktionsformen skiljer sig åt mellan olika regioner. Något förenklat innebär detta att exempelvis effekterna av en stödåtgärd inte är de samma för samtliga regioner vilket i många fall är realistiskt. Exempelvis är det rimligt att anta att effekterna av en stödåtgärd skiljer sig åt mellan norrländska inlandskommuner och kommuner i stockholmsområdet. Det går att i rumsliga modeller ta hänsyn till båda dessa orsaker till strukturell instabilitet.

3.2.2 Slutdiskussion och rekommendationer

Rumslig ekonometri har under de senaste åren kommit att i allt högre grad tillämpas för att analysera ekonomiska fenomen. Skälet till detta är metodens förmåga att fånga interaktionen mellan olika aktörer eller regioner. Senare tids empiriska forskning kring regional utveckling har i allt större omfattning kommit att tillämpa olika typer av rumslig ekonometri, både deskriptiva test och mer förklarande regressionsanalys. Även om det inte alltid kan statistiskt beläggas att det förekommer ett rumsligt beroende mellan regioner tyder empirisk forskning baserad på svenska data på att den typen av samband bör testas för innan de förkastas. Det finns heller inga tecken som tyder på att beroendet mellan regioner skulle avta i framtiden varför det är av stor vikt vid analyser och uppföljning av effekterna av olika typer av offentligt stöd till näringslivet och andra regionala stödåtgärder testa för rumslig korrelation. Dessutom finns många andra tillämpningsområden för rumslig ekonometri så som efterfrågeanalys, internationell ekonomi, arbetsmarknadsökonomi, offentlig ekonomi och lokal offentlig ekonomi, skattekonkurrens mellan regioner, regional tillväxt och migration vilket ytterligare motiverar en ökad användning av rumsliga ekonometriska metoder.

För att kunna tillämpa den här typen av metoder krävs mycket goda kunskaper i statistik och ekonometri samt förmåga att självständigt kunna programmera olika typer av estimatorer och test. Vad gäller valet av dataprogram rekommenderas ett kartritningsprogram för att visuellt presentera data, ett specialprogram för rumslig analys samt ett ”vanligt”

statistikprogram. Kartritningsprogrammet bör vara länkat till ett specialprogram för rumslig analys för att på så sätt underlätta skapandet av viktmatriser samt en visuell presentation av resultaten. Exempelvis är kartritningsprogrammet ArcView i kombination med SpaceStat och Stata en utmärkt utgångspunkt för att både kunna estimeras rumsliga modeller och andra typer av statistiska modeller. Även Matlab och GAUSS är mycket kompetenta programpaket för rumslig analys, men kräver mer av användaren. Ännu en god kombination är S-PLUS och ArcView. I det slutliga valet av program bör användarens tidigare erfarenhet av de olika programmen vägas in. En van användare av exempelvis Limdep kan eventuellt förorda och utträta mer med just Limdep framför exempelvis Matlab.

Slutligen bör nämnas att forskningen och utvecklingen av nya estimatorer och test inom rumslig ekonometri idag går oerhört snabbt vilket innebär att tidigare estimationsmetoder genomgår en noggrann granskning. Detta märks inte minst genom att allt fler ekonometriker och statistiker på senare tid kommit att intressera sig för området (se specialnumret av *Journal of Econometrics* (2007)). Estimatorer och test blir allt mer avancerade och samtidigt bättre i det avseendet att de på ett mer korrekt sätt fångar olika samband vilket leder till bättre analyser. För att kunna nyttja dessa framsteg krävs naturligtvis att analytikern har hög kompetens inom området, men även flexibla statistikprogram med stort utrymme för egen programmering.

3.3 Allmän jämviktsmodell

Allmän jämvikt syftar till att ge en förståelse av hela ekonomin med en bottom-up-strategi, som omfattar enskilda marknader och aktörer. Den allmänna jämviktsmodellen används traditionellt för att analysera effekterna av en förändring i den ekonomiska politiken och den bidrar till att förutse konsekvenserna av den föreslagna förändringen utifrån en rad olika ekonomiskt betydelsefulla variabler – priser, resultatnivåer, offentliga mottaganden och inkomstfördelningen mellan de konsumerande enheterna. En av de största fördelarna med den allmänna jämviktsmodellen är dess förmåga att spåra konsekvenserna av stora förändringar i en viss sektor genom hela ekonomin.

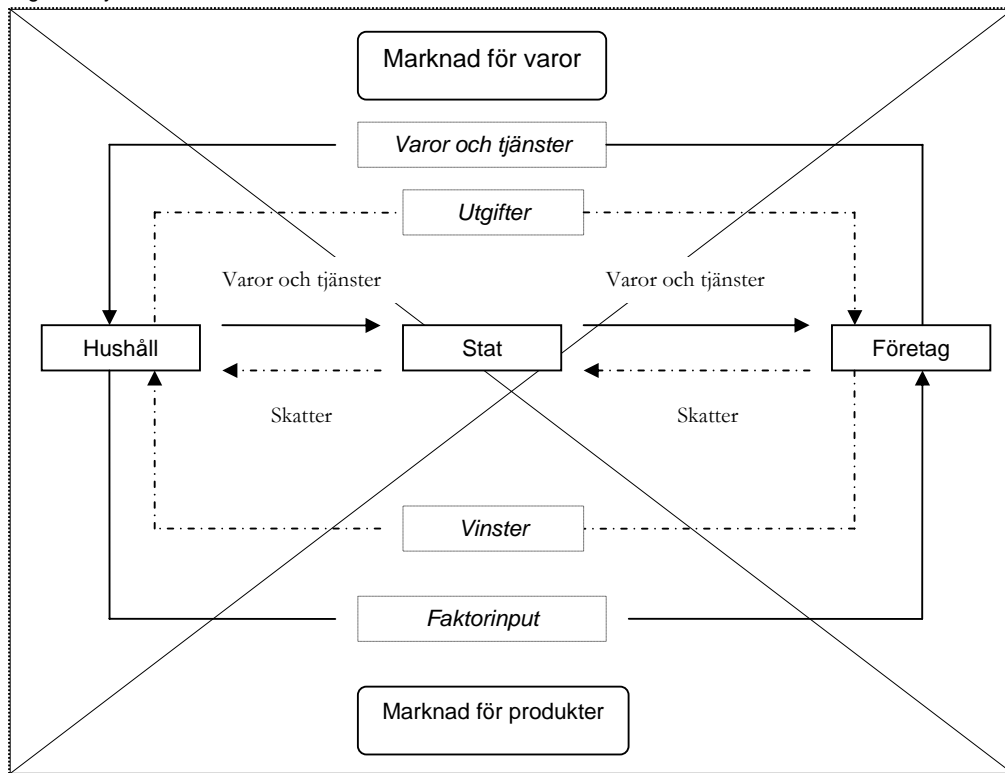
Begreppet *allmän jämviktsmodell* har både en metodisk och en teoretisk sida. Allmänna jämviktsmodeller betraktar ekonomin som ett slutet system med mellanliggande relationer där förändringar i en av variablerna leder till förändringar i alla andra variabler i systemet. Exempelvis sänker införandet av en skattesänkning eller ett stöd inte bara kostnaderna för företaget, utan det har även effekter på sysselsättningen och produktionen. En högre eller lägre produktion i ett företag kan leda till högre eller lägre produktion i andra företag, exempelvis hos kunder eller leverantörer. Det är därför viktigt att bedöma alla parametrar i systemet samtidigt för att fånga in både de direkta och de indirekta effekterna av en politik. Genom att bedöma alla parametrar kan vi därför minska antalet exogena variabler.

Den allmänna jämviktsteorin betraktar priser och kvantiteter med jämvikt i ett system med perfekt konkurrens, vilket innebär att den betraktar alla samspel mellan marknader som funktioner på den individuella marknaden³⁹.

³⁹ För en djupare förståelse av den allmänna jämviktsteorin rekommenderas Varian (1992) och Mas-Colell, Whinston and Green (1995).

I praktiken är det nästan omöjligt att genomföra allmänna jämviktsmodeller, eftersom det vanligen producerar icke-linjära variabler. Förekomsten av icke-linjära variabler i ett system gör systemet omöjligt att lösa analytiskt och det leder till användning av numeriska metoder som kallas CGE (Computable General Equilibrium Model), vilka använder en symmetrisk matris som kallas SAM (Social Accounting Matrix) för att kalibrera koefficienterna för beräkningarna i modellen.

Figur 3. Cykliskt flöde i en sluten ekonomi.



En CGE-modell bygger därför på den fråga som är för handen, det problem som ska lösas och tillgången till data. Det grundläggande begreppet för CGE-modeller är det cykliska flödet av varor i en sluten ekonomi, vilket beskrivs i Figur 3.⁴⁰ Hushåll och företag är huvudaktörerna i diagrammet och staten har vanligen en passiv roll i det cykliska flödet, även om staten beskrivs uttryckligen i många CGE-modeller. Hushållen äger produktionsfaktorerna och de konsumerar varor och tjänster som produceras av företagen. Företagen anlitar produktionsfaktorer för att producera varor och tjänster. Staten driver in skatter och fördelar intäkterna till hushåll och företag i form av bidrag, subventioner och överföringar. Flödet i Figur 3 visar att allt som produceras av företagen behålls eller konsumeras av hushållen. Företag absorberar å andra sidan hela tillförseln av faktorer. Det innebär att för hela aktiviteten i ekonomin bokförs utgifternas vägde mot inkomsternas

⁴⁰ Man kan förstås bygga en CGE modell för en öppen ekonomi.

värde. Det dubbla bokföringssystemet avspeglas i SAM-matrisen, där ekonomin delas upp i olika sektorer utifrån det aktuella problemet.⁴¹

När partiella allmänna jämviktsmodeller såsom modeller med efterfrågan på faktorer används flyttas fokus till ett begränsat antal marknader i stället för hela ekonomin. I praktiken är det nästan omöjligt att bedöma alla parametrar i ekonomin samtidigt och det är nödvändigt att begränsa antalet variabler som beaktas och bedöms.

Med hjälp av CGEM kan vi studera ekonomin som helhet, medan fördelen med partiella jämviktsmodeller är möjligheten att använda data på företagsnivå för att studera effekterna av statliga åtgärder.

3.3.1 Befintliga modeller

I Sverige har ett antal CGE-modeller utvecklats de senaste åren med olika mål. Här diskuterar vi fördelarna och nackdelarna med de befintliga modellerna, med tanke på att det inom ramen för det nuvarande arbetet även är viktigt att förstå vilka av de befintliga modellerna som skulle kunna användas och utvecklas för att utvärdera politiken när det gäller statligt stöd i Sverige.⁴²

EMEC – Environmental Medium Term Economic Model

Konjunkturinstitutet har utvecklat en statisk långtidsmodell som kallas EMEC (Environmental Medium Term Economic Model), vilken är särskilt utformad för analys av effekterna av den svenska miljöpolitiken för hushåll och företag. EMEC är en statisk CGE-modell med 26 industrier och 33 sammansatta varor och en offentlig sektor som producerar en enda artikel. Producerade varor och tjänster exporteras och används tillsammans med importen för att skapa sammansatta varor för inhemsk användning. Sammansatta varor används som input av industrier och för kapitalbildning. Dessutom konsumerar hushållen sammansatta varor och det finns 26 konsumentvaror. Produktionen kräver två typer av arbetskraft och kapital som huvudsakliga faktorer, samt input av material, transporter och energi. Det speciella med EMEC är den mycket detaljerade specifikationen av utsläpp från användning av både bränsleinputproduktion och hushållens bränsleförbrukning – modellen innefattar utsläpp av koldioxid, kolmonoxid, metangas, svaveldioxid och partikelämnen.

Företagens eller hushållens energianvändning omfattas av en energiskatt och föroreningskatter. Skatteundantag på grund av användningen av koldioxid eller av andra orsaker avspeglas i de beräknade skattesatserna. Konsumentvaror är också föremål för mervärdesskatt samt andra indirekta skatter. Användningen av arbetskraft är föremål för sociala avgifter och hushållen betalar skatt på arbetsinkomster. Företag och hushåll reagerar på priser, inbegripet skatter, och justerar sin blandning av input eller sina konsumentvaror genom att byta bort sin relativt dyraste input eller vara.

⁴¹ En detaljerad beskrivning av CGE-modeller finns i Shoven och Whalley (1984), Kehoe and Kehoe (1995).

⁴² I Sverige har ett antal mikrosimuleringsmodeller också utvecklats de senaste åren med olika mål. Mikrosimulering innebär att man bygger upp detaljerade modeller över, till exempel, Sveriges befolkning. Dessa modeller används för att köra simuleringar av vad som händer varje individ med avseende på boende, arbete, familjebildning, utbildning och andra parameter. Den första utvecklade modellen byggdes i 1970 av Riksdagsförsäkringsverket (RFV) för att simulera framtiden av den svenska pensionplan. En annan modell, MICROHUS, utvecklades i 1980 och en tredje modell SVERIGE byggdes som bearbetning från U.S. modellen CORSIM.

Det representativa företaget antas välja en optimal blandning av två typer av arbetskraft och en optimal blandning av energi i tre steg. Företaget beslutar sedan om blandningen av arbetskraft och fysiskt kapital i skapandet av mervärde samt blandningen av energi och material i skapandet av energi-materialinput. Olika output och input måste transporteras, och företaget väljer en optimal transportlösning i två steg. En optimal blandning av mervärde och energi-materialinput väljs på högsta nivå, för att skapa företagets output. En annan typ av byte hör samman med varor i samma klass. Inhemska varor är ickeperfekta substitut för utländska varor i såväl inhemsk som utländsk användning, dvs. i såväl import som export.

Som nämnt ovan är EMEC en modell som är utformad för analysen av miljöproblem och det skulle vara till nytta att utvärdera effekterna av energiskattesänkningar. På grund av dess uppbyggnad skulle det kräva vissa justeringar och det skulle i alla händelser vara till nytta att bara utvärdera vissa typer av politik.

För närvarande finns det i Sverige ett antal skattesänkningssystem som är igång – sänkning av den selektiva inköpsskatten för produktion av koldioxidneutrala bränslen, tillverkningsindustrins skattelättnad för koldioxid, tillverkningsindustrins skattelättnad på el, fullt undantag från energiskatt för energiintensiva företag.

Tack vare ett möte med Göran Östblom och Magnus Sjöström från Konjunkturinstitutet har vi försökt besvara följande frågor:

Kan man använda och anpassa EMEC för att studera effekterna av de skattesänkningssystem som nämns ovan? Hur lång tid skulle det i så fall ta?

En del av de angivna problemen kan analyseras med EMEC. Det enklaste är att studera skattelättnaden för koldioxid för tillverkningsindustrierna. För att beräkna hur mycket tid det skulle ta att göra regressionerna är det nödvändigt att på ett bättre sätt definiera vad som har gjorts och med vilka antaganden. I grova drag bedömer vi att det skulle kunna ta ett par månader för varje problem med tanke på att vissa förändringar av modellen är nödvändig.

Hur lång tid skulle det ta att utveckla en modell som EMEC?

Det beror på omständigheterna, särskilt på om det finns en liknande modell att börja med och vilka kompetenser som behövs. Utvecklingen av den nuvarande versionen av EMEC har tagit 3–4 arbetsår, utspritt på en period av 10 år, och 10 forskare.

Finns det en möjlighet för Konjunkturinstitutet att samarbeta med ITPS när det gäller utveckling av en ny modell eller anpassning av en modell som härrör från EMEC?

Utöver de uppdrag som KI tar emot från staten finns det utrymme för den miljöekonomiska avdelningen att ägna sig åt extern forskning. Det innebär att EMEC skulle kunna utvecklas och anpassas i enlighet med behoven. För externa uppdrag är kostnaden 900 kr/h. Det är inte möjligt att utveckla nya modeller för andra myndigheter, men ett samarbete mellan ITPS och KI för att gemensamt utveckla en modell är tänkbart, efter korrekta avtal. För externa uppdrag behöver KI långsiktig planering.

Kan EMEC användas för att diskutera regionala frågor? Eller är det bättre att utveckla en ny metod för att studera de regionala effekterna?

EMEC kopierar den svenska ekonomin på övergripande nivå och innefattar inte en regional nivå. Resultaten från simuleringen kan delas upp med hänsyn till regioner om det finns ett lämpligt distributionsverktyg. Detta är emellertid något helt annat än att simulera

regionala effekter med en modell med regional input. Det största problemet är inte att bygga en modell utan att ha tillgång till data som i de nationella räkenskaperna, vilka utgör grunden för simuleringen i EMEC. Om vi skulle dela upp EMEC:s resultat på regional nivå skulle vi definitivt tolka dem försiktigt.

Det är svårt för KI att säga om det är bättre att bygga upp en regional modell för att fånga in regionala effekter, med tanke på att det också beror på de tillgängliga resurserna och de frågor som ska besvaras.

Bedömer ni alla parametrar som används i EMEC? Hur är det med tillgången till data?

Ingen statistisk bedömning görs i EMEC. Grunden för simuleringen är en så kallad input-output-tabell (SAM) från de nationella räkenskaperna. Elasticitet och koefficienter kommer från olika källor och bara ett mindre antal av dem beräknas på KI. Grundscenarierna skapas från de antaganden som presenteras i LU.

FIMO – finansiell modell

FIMO är en beräknings- och simuleringsmodell som årligen beskriver finansiella flöden och nettolångivning fördelat på de institutionella sektorerna i ekonomin enligt definition i de nationella räkenskaperna. Modellen används bland annat av Finansdepartementet.

Sektorsuppdelningen i modellen omfattar den centrala regeringen, pensionssystem, kommuner, hushåll, företag och övriga världen. Modellen består av mer än 500 variabler på lägsta nivå, varav majoriteten beskriver transaktioner inom den allmänna regeringen och mellan den allmänna regeringen och hushållen. Modellen genererar ett stort antal summor, både på sektors- och transaktionsnivå. Exempel på de förstnämnda är totala intäkter och utgifter, nettolångivning och förändringar i tillgångar och skulder, medan summor av betalningsflöden mellan sektorer såsom totala skattebetalningar och överföringar till hushåll är exempel på det sistnämnda.

Modellen innehåller även en exogent fastställd total beskrivning av ekonomins verkliga sida, vilken de finansiella flödena är beroende av. Möjliga tillämpningar av modellen innefattar beräkningar av finansiella flöden och nettolångivning baserat på en given utveckling av ekonomins verkliga sida, samt olika typer av känslighetsanalyser.

Dess svaghet finns på modellens verkliga sida, som är baserad på ett stort antal antaganden och som fastställs exogent.

LUMOD – Långsiktig dynamisk simulering

LUMOD är en dynamisk modell som studerar sparande och kapitalbildning, utvecklad av Stockholms Universitet. Den kan betraktas som en tillväxtmodell, men den innefattar även den keynesianska konsumtionsfunktionen och multiplikatorn samt investerings-acceleratorn.

LUMOD innehåller både en tillgångs- och en efterfrågesida och priserna är kostnadsdrivna. Modellen tar hänsyn till produktivitet, vinstandelar, löner, inkomstskatter, överföringar, räntor och importpriser.

Växelkursen och tillväxttakten för utländska priser fastställs exogent. Med hjälp av en import-exportfunktion relateras den utländska handeln till den inhemska och externa efterfrågan och pristrenden. Den offentliga sektorn behandlas som en helhet. Å andra sidan studeras offentliga intäkter och utgifter i större detalj. På intäktssidan har vi inhemska skatter, inkomstskatter och indirekta skatter. På utgiftssidan har vi konsumtion,

investeringar, utländska överföringar och överföringar till hushåll och subventioner till företag.

LUMOD är en öppen modell med ett stort antal exogena variabler. LUMOD är i första hand ett instrument för att testa konsekvensen hos exogena antaganden och leder till en integrerad ram.

LUMOD:s beräkningar ger inga prognosresultat, utan illustrerar möjliga utvecklingsmönster för den svenska ekonomin med hänsyn till olika antaganden. Syftet är att öka förståelsen av sambandet mellan utvecklingsfaktorer samt att ge ett experimentverktyg för att analysera olika typer av politik.

KIMOD

KIMOD är en dynamisk allmän jämviktsmodell för den svenska ekonomin utvecklad av KI som ska användas för en total ekonomisk-politisk analys och makroekonomiska medellånga scenarier. Det är en starkt sammansatt makroekonomisk modell i vilken alla företag är lika, liksom alla hushåll. Dessutom är den allmänna regeringssektorn konsoliderad och därmed inte uppdelad i central regering, kommuner och distrikt. Modellen är mikrobaserad på så sätt att företag och hushåll fattar optimala beslut om output respektive konsumtion, med tanke på rationella förväntningar om andra aktörers beteende och om den sannolika framtida utvecklingen av modellen som helhet.

KIMOD är dynamisk, i betydelsen att investeringar och sparande under en period påverkar de framtida möjligheterna till output respektive konsumtion, och att alla beslutsfattare tar hänsyn till detta. Tiden är uppdelad i skilda perioder under ett år. De prognoser som genereras av modellen är därmed tidsserier med årlig frekvens och de ekonometriska ekvationerna i modellen beräknas på grundval av årsdata. De nationella räkenskaperna är den källa som föredras för inledande data och för att beräkna parametrar.

KIMOD är en jämviktsmodell i två avseenden. För det första fastställs priserna för varje period så att tillgången motsvarar efterfrågan på alla marknader utom arbetsmarknaden. Den sistnämnda formas i stället som en sökningsmarknad, där lönerna fastställs i förhandlingar mellan arbetsgivare och arbetstagare. Arbetslöshet uppstår till följd av både sökrelaterad friktion och konkurrens som inte är perfekt. För det andra närmar sig modellen på lång sikt ett stabilt tillstånd som är oberoende av ekonomins tillstånd i början. I detta stabila tillstånd är ekonomin på en balanserad tillväxtstig med en konstant relativ tillväxttakt.

KIMOD är därför avsedd att användas i makroekonomiska analyser och medellånga scenarier, där medellång avser en tidshorisont på mellan två och sex år. För andra tidshorisonter är den undre gränsen för modellens användbarhet en följd av att periodernas längd fastställs till ett år. Det innebär att det inte finns någon som helst säsongsdynamik, och att data för delar av innevarande år inte kan användas som modellens inledande tillstånd. För långsiktiga syften är modellens användbarhet begränsad av att demografin och andra strukturella utvecklingar över tiden inte har modellerats.

ISMOD

ISMOD utvecklades ursprungligen i början av 1980-talet och har sedan dess använts för långsiktiga ekonomiska prognoser för finansdepartementet. Under de senaste tio åren har modellen bara använts av Nutek och Sika. På senare år har modellen justerats och beräknats om med nya data vid Handelshögskolan i Jönköping.

ISMOD-modellen är formulerad för att generera lösningar som gäller i mellan 5 och 15 år framåt. Tidsperioden bör inte vara kortare på grund av investeringsprocesser och inte längre eftersom inputdata (teknologiskt alternativ) blir mindre tillfredsställande när tidsperioden är längre.

Modellen kan beskrivas enligt följande delar:

1. Produktionsstruktur: grunden är input-output-matriser, där varje sektors efterfrågan på leverans av input fastställs och inputkoefficienterna är specifika för varje teknisk kategori i sektorn. Kapaciteten är begränsad i modellen. Kapaciteten kan utvidgas under tidsperioden, men det kräver investeringar och att investeringarna genererar en efterfrågan på leveranser från andra sektorer (enligt en vektor av investeringskoefficienter). Samtidigt minskas kapaciteten i vissa av sektorerna. Vinstnivån i sektorn avgör reduktionshastigheten (och avgörs i sin tur av priser och lönenivå).
2. Tillgång till varor och tjänster: under tidsperioden förändras tillgången på grund av i) kapacitetsreduktion, ii) investeringar i ny kapacitet och iii) import. Alla dessa tre komponenter är beroende av jämviktspriserna under tidsperioden.
3. Efterfrågan på varor och tjänster: efterfrågesidan har följande fem beståndsdelar: i) aktuella köp för produktion, ii) privat konsumtion, iii) köp för den offentliga sektorn, iv) export och v) köp av investeringsvaror. Efterfrågans beståndsdelar är direkt och indirekt beroende av jämviktspriserna.

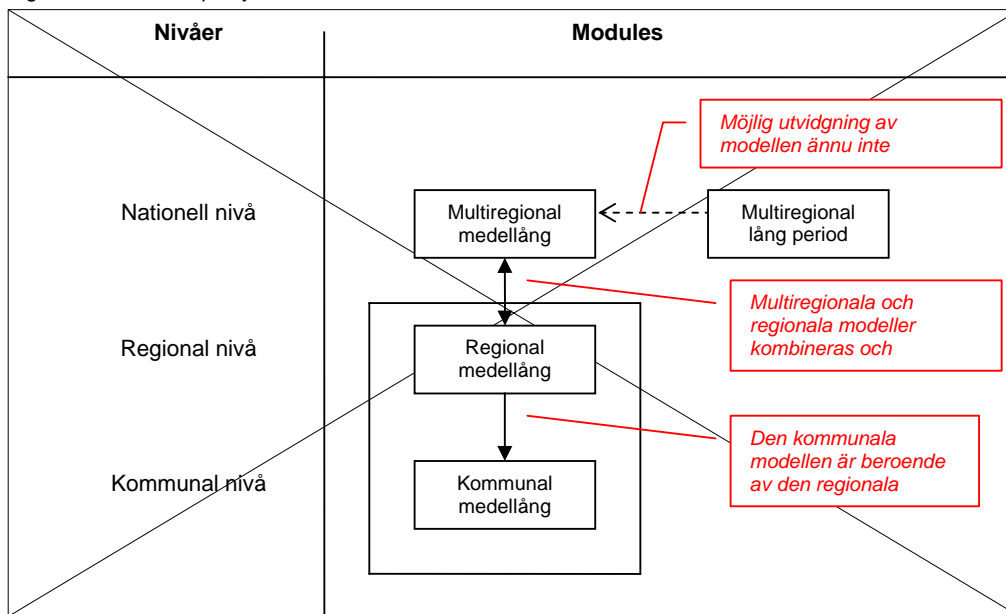
Ett problem med ISMOD är att den verkar med industrisektorer och inte varugrupper. Handel och transport hanteras på ett mycket förenklat sätt. Output från dessa sektorer antas uppgå till de totala handelsmarginalerna för de övriga sektorerna.

Erfarenheten visar att förändringar i transportkostnader och infrastrukturinvesteringar har en effekt på sektorsstrukturen och transporterade volymer och mängder. I ISMOD är detta inte möjligt. Vidare finns det ingen geografi i ISMOD och ett annat problem är tidshorizonten, eftersom analysen av infrastrukturinvesteringar kräver långa prognosperioder.

rAps

rAps är ett integrerat system för regionala analyser som bygger på en central och en lokal databas som relaterar till olika modeller. Det kan användas både för enkla analyser och för mer avancerade analyser såsom prognoser och scenarier. rAps är användbart för både regionala och centrala aktörer och bygger på ett system av modeller. Systemet består av en serie moduler med specifika syften, vilka tillsammans kan ge stor grad av flexibilitet och anpassningsbarhet till systemet.

Figur 4. Moduler i rAps-systemet



Källa: https://www.h.scb.se/raps/Pdf/raps_beskrivning.pdf

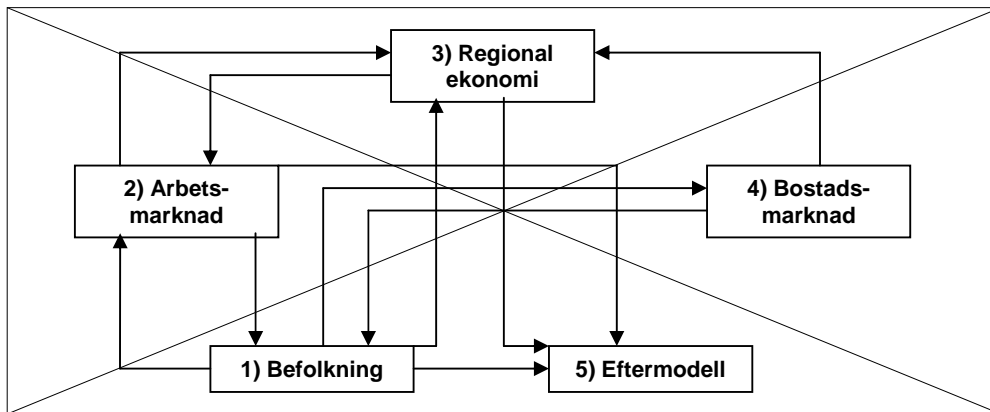
Såsom visas i Figur 4 är modulerna i rAps grupperade efter nivå och efter tidsperspektiv och kommun och de minsta enheter av vilka hela ekonomin är uppbyggd.

Den regionala modellen för den medellånga perioden (5–10 år) är en ekonomisk och demografisk modell som är avsedd för analys och prognos på regional och kommunal nivå. Regionerna är inte geografiskt statiska, eftersom de är uppbyggda som grupper av kommuner och kan variera utifrån den fråga som är under utredning.

Modellen byggs upp utifrån anslutningen på kommunal nivå och består av fem delmodeller: 1) befolkning, 2) arbetsmarknad, 3) regional ekonomi 4) bostadsmarknad och 5) eftermodell per kommun. Dessa fem delmodeller avgör sysselsättningen per kommun och utbytet, samt kommunens intäkter och utgifter.

Modellen är uppbyggd i syfte att skapa möjliga framtida scenarier och den används i verkligheten i princip för att analysera olika scenarier, parametrarna beräknas inte och modellens resultat beror på användarens antaganden.

Figur 5. De fem delmodellerna i den regionala modellen

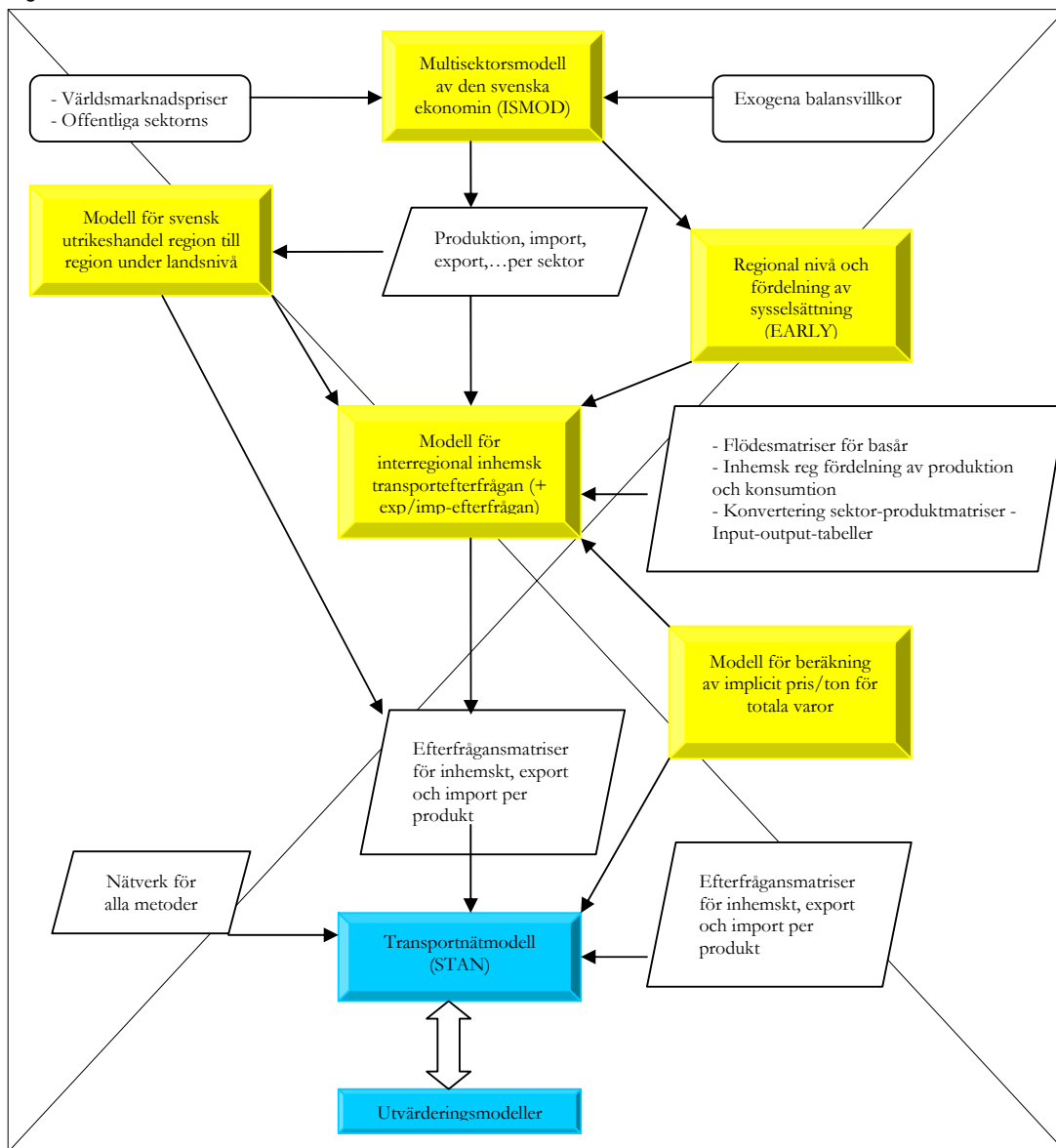


Källa: https://www.h.scb.se/raps/Pdf/raps_beskrivning.pdf

SAMGODS

SAMGODS består av ett antal sammankopplade moduler, som verkar på olika nivåer av rumslig detalj, såsom framgår av Figur 6.

Figur 6. SAMGODS-modellen



Källa: SAMPLAN 2004:1

De gula rutorna i figuren avser främst förändringar i efterfrågan över tiden, medan de blå rutorna handlar om effekten av särskild politik. Det finns fem olika modeller som rör prognoser om efterfrågan över tiden:

- Ett flersektorssystem av modeller av den svenska ekonomin (ISMOD)
- En modell för regional uppdelning av sektorssysselsättningen (EARLY) kopplad till ISMOD

-
- En modell för att forma den mellanregionala transportefterfrågan inom Sverige (VTI/TPR)
 - Modell för regionala prognoser av den svenska utrikeshandeln
 - Modeller för prognoser av implicita varuvärden för totala varor

SAMGODS-modellen har flera syften: att ge efterfrågansprognoser, information om politik och projektutvärdering, analys av effekter och konsekvenser av alternativa strategier. Vi är främst intresserade av användningen av modellen för analyser av allmänna transportpolitiska åtgärder samt politiska åtgärder som hör samman med specifik infrastruktur, t ex väg, järnväg osv., med hänsyn till många beslutsprocesser som skulle kunna använda information om fraktförfrågan.

Alla de modeller som beskrivs ovan är uppbyggda som simuleringsmodeller i syfte att skapa möjliga framtida scenarier. Det mesta av de data som används är baserad på en rad antaganden och det gör modellerna olämpliga för våra syften eftersom de inte kan ge en verklig bild av världen. De används i princip endast för att analysera olika scenarier, med oberäknade parametrarna och modellernas resultat beror på användarens antaganden.

Det finns dock två andra modeller som är intressanta på grund av deras kopplingar till och likheter med den ”Mikro-makro modell” som nämns i nästa kapitel.

RAMSES – Riksbankens aggregerade makromodell för studier av ekonomin i Sverige

RAMSES är en allmän modell som har till syfte att förklara hela ekonomin, inte bara en viss beståndsdel. Den består av många olika hushåll och företag som samspelar på marknader för varor, kapital och arbetskraft. Precis som i många andra nykeynesianska allmänna jämviktsmodeller antas marknaderna för varor och arbetskraft kännetecknas av monopolistisk konkurrens. Det innebär att företag och anställda, i stället för att ta priser och löner för givna, är medvetna om att de kan påverka dem genom sitt beteende. Eftersom priser och löner antas vara orörliga kan valutapolitiken dock påverka den verkliga ekonomin (t ex output och tillgång till arbetskraft) på kort sikt eftersom de nominella priserna och lönerna inte är fritt justerbara efter en förändring av den nominella räntan.

Modellen omfattar även en centralbank som fastställer den kortfristiga räntan och en offentlig sektor som antas finansiera sina konsumtionsutgifter genom att beskatta arbete och konsumtion. Det finns även en utländsk ekonomi som antas vara opåverkad av den inhemska ekonomiska utvecklingen i Sverige. I modellen importeras konsument- och investeringsvaror delvis från övriga världen och en del av den inhemska outputen exporteras.

Modellen kan formuleras som ett antal matematiska villkor som beskriver hur hushållen och företagen agerar, med hänsyn till antagandet om ett optimerande beteende med rationella förväntningar och med den kortfristiga nominella räntan som kontrolleras av Riksbanken. RAMSES innehåller 24 ekvationer för detta. Tillsammans med ekvationerna för villkor utomlands, bedrivandet av skattepolitik och ekonomiska chocker utgör de ett konsekvent matematiskt system av icke-linjära differentialekvationer som producerar en relativt acceptabel bild av hur ekonomin utvecklas över tiden.

I detta sammanhang är det relevant att understryka att de teoretiska modellvariablerna i RAMSES är knutna till mätekvationer till statistiska observationer av exempelvis output,

priser och räntor. Det är inte nödvändigt att inkludera observerade variabler för var och en av modellvariablerna. Det finns i princip ingen empirisk motvikt till vissa modellvariabler och det finns inget tillfredsställande sätt att mäta andra. Beräkningsförfarandet består av att använda RAMSES modellstruktur tillsammans med de utvalda observerade variablerna för att skapa en bild av de variabler för vilka mätningar inte är tillgängliga. Det är dock viktigt att uppsättningen observerade variabler är tillräckligt informativ för att identifiera modellens parametrar och de bakomliggande modellvariabler som inte kan observeras.

Beräkningar av RAMSES görs för närvarande med följande 15 makroekonomiska variabler: BNP-deflator, konsumtion, investering, reallöner, realränta, politisk ränta, arbetstid, BNP, export, import, UNDI1X, investeringsdeflator, utländsk BNP, utländsk inflation och utländsk ränta. RAMSES beräknas på data från 1986 Q1 fram till i dag. Valet av beräkningsperiod måste väga kvantitet mot kvalitet – mycket data behövs för att parameterberäkningarna ska bli korrekta, men dessa data bör även hänvisa till en period utan beräkningsbara strukturella förändringar.

Följande anges i Adolfson, Laséen, Lindé och Villani (2007):

Eftersom många aspekter på ekonomin formas alltför enkelt i RAMSES är denna modell inte lämplig för ett antal syften. Ett viktigt exempel är att RAMSES inte innehåller finansiella friktioner, vars betydelse för förståelsen av den valutapolitiska överföringsmekanismen som diskuteras omfattande i litteraturen. Ett annat exempel är skattepolitikens mycket begränsade roll i RAMSES. Arbetet med att utveckla modeller är dock en fortlöpande process på Riksbanken som har till syfte att lära av de brister som är en oundviklig del av detta område. I detta arbete är det naturligtvis också viktigt att dokumentera modellens empiriska egenskaper.

[...] vi vill betona att trots att vi inte tror att formella modeller som RAMSES kan ersätta det omfattande analytiska arbetet av sektorsexpert och andra har utvecklingen av den nya generationen jämviktsmodeller nu blivit så framgångsrik på olika sätt att dessa modeller har förtjänat en framstående plats i en centralbanks verktygslåda.

Sveriges konkurrenskraft – att förstå och mäta nationell konkurrenskraft

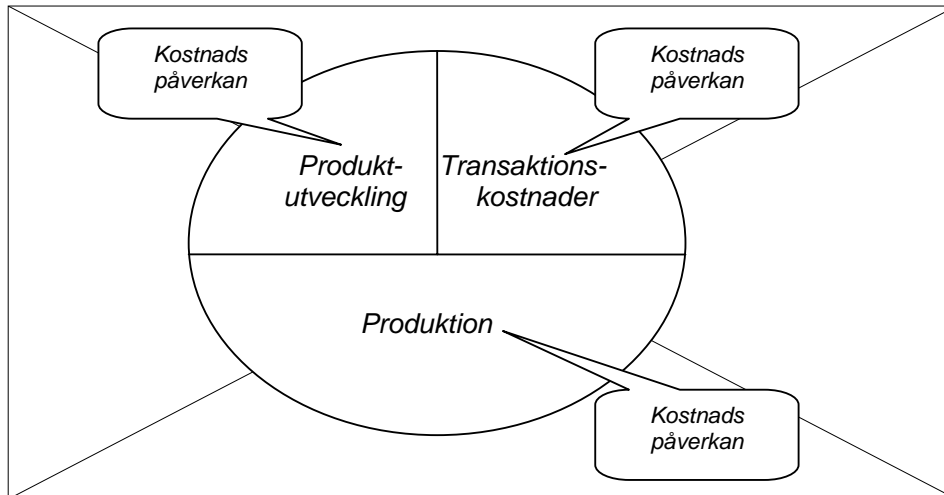
I en färsk rapport från ITPS (ITPS A2007:002) belyser Sandro Scocco tanken på att konkurrenskraft har fått allt större betydelse de senaste årtiondena. Författaren har därför utvecklat en otvetydig definition av konkurrenskraft som gör det lättare att mäta och förstå detta begrepp.

Ett lands konkurrenskraft betraktas i ITPS rapport som hela den arbetande befolkningens intäkter jämfört med motsvarigheten för en grupp av elva jämförbara OECD-länder. Måttet på Sveriges konkurrenskraft blir sedan förändringen i intäktsvärdet för den arbetande befolkningens produktion jämfört med motsvarande förändring för länder med liknande produktionsförutsättningar. Om värdet för Sverige ökar snabbare än jämförelsegruppen har den svenska konkurrenskraften stärkts.

Ett lands konkurrenskraft kan analyseras med utgångspunkt i mikronivån och med hänsyn till kostnadsfunktionen för enskilda företag. Konkurrenskraft är ett relativt mått som inte kan uttryckas i absoluta termer men som, såsom framgår av Figur 7, kan uttryckas som en funktion av företagets kostnader. Ett företag har tre olika typer av kostnader –

produktionskostnader, produktutvecklingskostnader och transaktionskostnader – och med dessa i åtanke anses ett företag vara konkurrenskraftigt när dessa kostnader är internationellt gångbara och den totala kostnaden inte överstiger kostnaden för konkurrenter med samma kvalitet och pris. Det är viktigt att understryka att dessa är kostnader som påverkar företagets mervärde och är därför inte kostnader som hör samman med mellanliggande produkter och tjänster såsom el eller råmaterial.

Figur 7. Företags mervärde



Källa: ITPS A2007:002

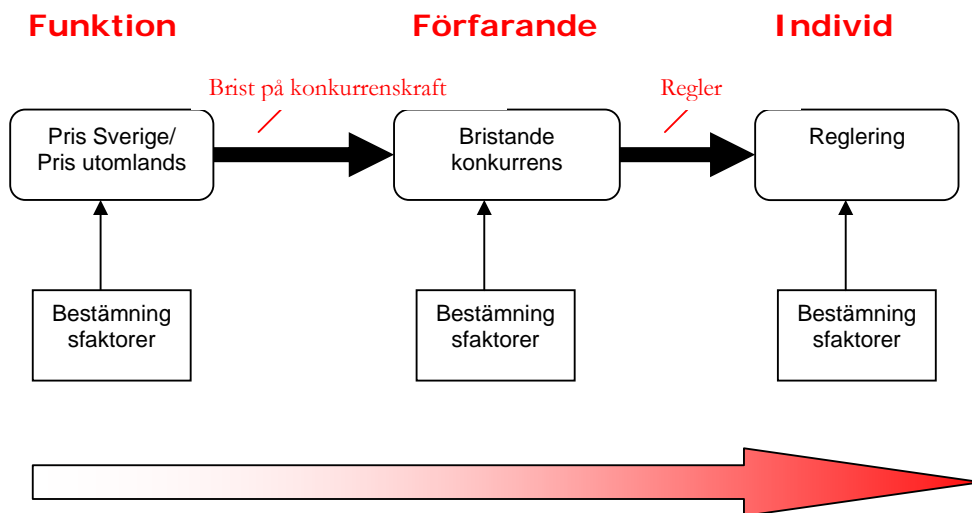
Med priser och kostnader som utgångspunkt är det därför möjligt att bygga upp en intressant och mätbar koppling mellan mikronivå (företagsnivå) och makronivå (nationella räkenskaper).

Det faktum att ITPS har utvecklat ett mätbart konkurrenskraftsbegrepp baserat på företagets kostnadsfunktion är mycket viktig för modellutveckling. Det är uppenbart att regeringen med hjälp av statligt stöd och bidrag påverkar företagets kostnader. Staten kan även påverka produktionsnivån genom resurser, regler och förordningar samt maktutövning.

Genom en trestegsanalys (funktionell nivå, förfarandenivå och individuell nivå) av de tre huvudsakliga kostnadsområdena (produktion, produktutveckling och transaktionskostnader) har staten därför ett verktyg för att identifiera relevanta faktorer och aktörer som påverkar kostnadsfunktionen och kan därefter rikta sin uppmärksamhet mot företagsnivån.

Genom att införa en trestegsanalys som visas i Figur 8 kan regeringen analysera de frågor som är ekonomiskt och politiskt relevanta. En sådan analys skulle därför leda till mer ekonomiskt lämpliga politiska åtgärder.

Figur 8. Trestegsanalys



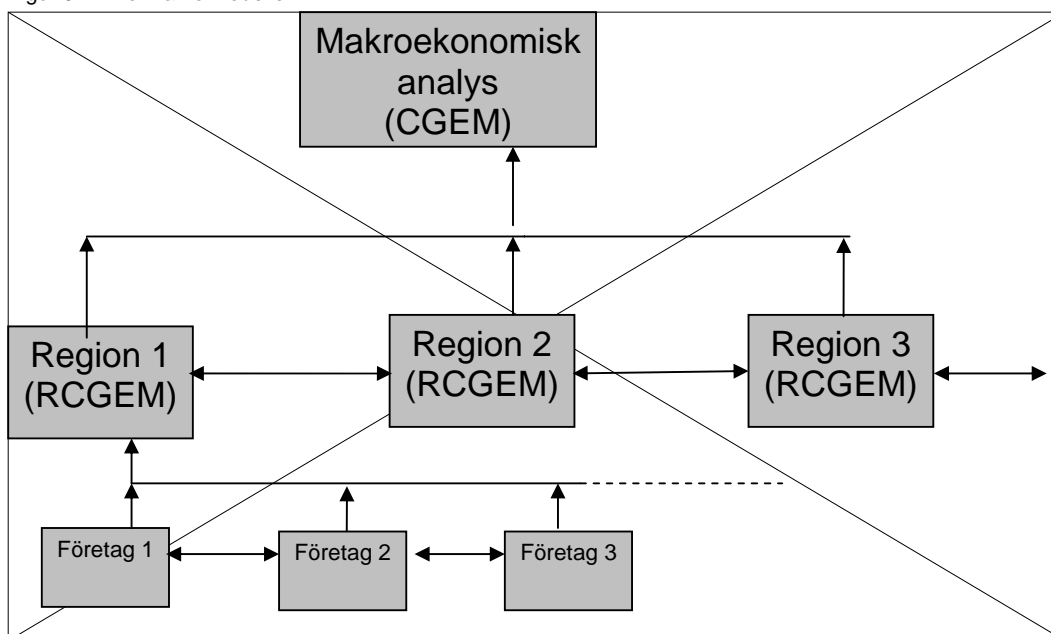
Källa: ITPS A2007:002

I nästa kapitel presenteras en utförlig modell. Den föreslagna modellen är uppbyggd kring Riksbankens RAMSES-modell och Soccos modell, och den har till syfte att ge det ekonometriska verktyg som behövs för att bygga upp en intressant och mätbar koppling mellan mikronivån (företagsnivå) och makronivån (nationella räkenskaper).

4 Mikro-makro modellen

Vid diskussioner om effekterna av offentlig stödpolitik är olika typer av analyser möjliga beroende på aggregeringsnivån och vilka stödmottagarna är. Effekter kan studeras på företagsnivå, regional nivå och makroekonomisk nivå. Införandet av en skattesänkning inom en viss bransch har till exempel inte bara effekter på de företag som tillhör en viss bransch utan även på andra företag såsom leverantörer och kunder. En utvärdering av effekterna av politiken behöver analysera företagsspecifika spillover-effekter och undersöka de inbördes relationerna mellan olika företag för att fånga effekterna på mottagare och icke-mottagare. Vi får inte heller bortse från mottagarnas geografiska placering och förutom de företagsspecifika spillover-effekterna⁴³ behöver även regionala spillover-effekter beaktas. Slutligen påverkar införandet av en skattesänkning den allmänna jämvikten i ekonomin på en aggregeringsnivå.

Figur 9. Mikro-makro modellen.



Såsom Figur 9 visar vore det idealiska sättet att undersöka effekterna på tre olika nivåer genom att börja med en företagsspecifik analys av varje företag i ekonomin, gruppera dem på regional nivå, för att kunna undersöka regionala effekter och slutligen sammanställa uppgifterna från de olika regionerna för att analysera effekterna på hela ekonomin.

Med tanke på det stora antalet företag inom ekonomin vore ett sådant förfarande emellertid både tidsödande och tekniskt omöjligt.

En möjlig lösning är att utveckla ett flexibelt system där vi kan undersöka effekterna av olika åtgärder på mindre grupper av företag. På så sätt skulle vi kunna undersöka hur förändringar påverkar ett enskilt företag, men inte alla företag på samma gång.

⁴³ En direkteffekt, antingen positiv eller negativ, av någons vinst eller välbefinnande som blir en biprodukt av en annan persons eller ett annat företags verksamhet.

Företagen kan grupperas i så homogena grupper som möjligt inom gruppen och så heterogena grupper som möjligt mellan grupperna. Om vi till exempel vill studera effekterna av en skattesänkning på tillverkningssektorn kan vi gruppera företagen enligt sektor, men om vi vill undersöka effekterna av en skattesänkning som är relaterad till emissionsrättigheterna kan företagen grupperas enligt emissionsnivån, oavsett vilken sektor de tillhör. När företagen har grupperats homogent kan vi välja en undergrupp av representativa företag för varje grupp.

Enligt Adolfson, Laséen, Lindé och Villani (2007) dominerades analys och prognoser om inflation och konjunkturcykler av Keynes modeller fram till början av 1980-talet. Enligt dessa modeller antas aktörerna inom ekonomin styras av olika tumregler och de tror inte att förväntningarna på framtiden skapas med rationella metoder. I stället används enkla beräkningar som bygger på tidigare datamönster. Användningen av Keynes modeller motiverades främst med bristen på tekniska verktyg som teorier och datorer som är nödvändiga för att lösa komplexa ekvationsproblem med långsiktiga förväntningar. Ett annat skäl var att de tidigaste versionerna av allmänna jämviktsmodeller för undersökning av makroekonomisk utveckling inte tycktes överensstämma med data.⁴⁴

Keynes modeller, där efterfrågans betydelse för att förstå konjunkturcyklerna betonades, upphörde emellertid med att beskriva data från 1970-talet och tidigare när stagflation⁴⁵ uppstod i västvärlden och dessa modeller misslyckades med att fånga ett antal strukturella skiften.

Under det senaste decenniet har omfattande akademisk forskning utvecklats en ny generation makroekonomiska allmänna jämviktsmodeller där tonvikten i den ekonomiska beskrivningen ligger på utbudssidan, men där efterfrågan på kort sikt även har en viktig roll genom förekomsten av olika marknadsbrister i kombination med nominella och reella rigiditeter. Med hänsyn till marknadsbristerna och rigiditeterna förekommer reaktionerna på de olika störningarna mer gradvis i modellen. Dessa nya makroekonomiska allmänna jämviktsmodeller kallas nykenynesiska modeller och har hållbara och väldokumenterade empiriska egenskaper.

Såväl Riksbankens RAMSES-modell som den modell som föreslås i denna rapport bygger på BVAR-utvärdering och visar tydligt att bristen på tekniska verktyg som teorier, datorer och data inte längre är ett problem.

De befintliga allmänna jämviktsmodellerna i Sverige (dvs. RaPS och EMEC) har än så länge inte använts för att undersöka regionala frågor. Detta beror på olika faktorer, men framför allt på följande: (1) modellerna skapades för att undersöka specifika frågor och med hänsyn till ekonomin på en aggregeringsnivå och (2) enligt modellutvecklarna skulle det krävas data om varuflöden mellan regioner för regional analys⁴⁶, till exempel de som finns på nationella konton och som är så gott som omöjliga att spåra.

⁴⁴ Till exempel Kydland & Prescotts (1982) grundläggande modell av den verkliga konjunkturcykeln där utbudsdelens betydelse för förståelsen av makroekonomisk utveckling kritiserades för bristen på empirisk realism. Ekonomin tycktes kännetecknas av mycket större rigiditet än vad som kunde förväntas om den styrdes av marknadsmekanismer och rationella hushåll och företag.

⁴⁵ En kombination av låg tillväxt och hög inflation.

⁴⁶ Den enda modellen som tar hänsyn till regionala varuflöden är RaPS, men den gör det på ett sätt som inte passar ITPS syfte.

Även om sådana data skulle bli tillgängliga skulle de befintliga modellerna inte vara lämpliga och det skulle krävas många ändringar med tanke på att ITPS är intresserade av effekterna av offentlig politik på sysselsättning och produktivitet på regional nivå.

För att kunna nå syftet behöver vi följande:

- En modell som skapats på regional basis.
- En modell som skapats utifrån datatillgänglighet.

Utifrån befintliga data och de frågor som ska besvaras kan vi skapa en modell som liknar Riksbankens RAMSES/BVAR, men på regionala nivåer, och på följande sätt.

Eftersom vi är intresserade av att utvärdera effekterna av åtgärder som inriktats mot företag eller grupper av företag på regional nivå behöver vi kunna framhäva olika effekter på mottagare och icke-mottagare. Vi kanske är intresserade av att studera spillover-effekter och kausala effekter av en planerad åtgärd. Detta kan vara användbart för beslutsfattare innan åtgärden beslutas och genomförs.

Eftersom ITPS har tillgängliga data om bokslut för företag kan dessa användas som utgångspunkt i undersökningen av mottagare och icke-mottagare. Med hjälp av paneldata – datauppsättningar med data om ett visst antal n företag under en tidsperiod – kan vi utvärdera tillgång och efterfrågan på olika sätt. En möjlig ansats är att tillämpa en ”poolad ansats” som bygger på mycket strikta antaganden, till exempel att alla företag som tillhör en sektor använder samma teknik, oavsett vad de tillverkar. Det innebär att marginaleffekterna av en viss åtgärd är densamma för alla företag i en bestämd sektor och blir heller inte möjlig att undersöka hur det påverkar olika företag.

En annan mindre restriktiv möjlighet är att tillåta olika tekniker i olika sektorer, vilket innebär att parametrarna är sektorsspecifika för den valda aggregeringsnivån och att man i praktiken utvärderar olika sektorsspecifika partiella system. Fördelen med denna modell jämfört med den ”poolade” ansatsen är att alla parametrar kan variera mellan sektorerna. Nackdelen är att den valda aggregeringsnivån inte nödvändigtvis motsvarar de riktiga skillnaderna när det gäller teknik.

Partiella jämviktsmodeller utvärderar egenpriselasticiteten och korspriselasticiteten för varor och undersöker sedan hur efterfrågan på en viss producentkapitalvara fluktuerar till följd av prisförändringar på andra varor, medan allt annat förblir oförändrat.⁴⁷ Att utveckla och bibehålla en partiell jämviktsmodell är inte lika tidsödande som en allmän jämviktsmodell och partiella jämviktsmodeller har därför använts mycket oftare vid undersökningar av åtgärders effekter.

I stället för paneldata kan tidsseriedata användas för att undersöka effekterna av en bestämd åtgärd. Vid faktorefterfrågemodeller används såsom nämnt ovan paneldata för att utvärdera effekterna på aggregerad nivå. Med tidsseriedata för varje företag blir det möjligt att undersöka effekterna av olika åtgärder på företagsnivå.

Det kan vara intressant att undersöka hur olika offentliga åtgärder påverkar produktionsnivån på olika företag och inom olika sektorer. Minskade kostnader på grund av statligt stöd leder sannolikt till högre produktionsnivåer, men företag inom samma sektor har inte nödvändigtvis samma kostnadsfunktion och effekterna av åtgärden kan vara

⁴⁷ Den här metoden har använts i många undersökningar. Se exempelvis Dargay (1983), Berndt (1991), Brännlund (1997) och Brännlund & Lundgren (2004, 2005).

olika för olika företag. Därför kan det vara intressant att undersöka externaliteter – undersöka hur produktionsnivån på ett företag påverkar nivån på andra företag inom samma sektor – och Granger-kausaltitet vid genomförande av vissa åtgärder. Om det går att ställa bättre prognoser om den aktuella produktionsnivån på företag 1 med hjälp av data om produktionsnivån på företag 2 föregående år, kan vi uttala oss om det föreligger Granger-kausaltitet mellan produktionen på företag 2 och produktionen på företag 1.

Om vi antar att det finns n företag inom en bransch är Y_{it} totalproduktionen på företag i , $i=1, \dots, n$, vid tidpunkten t och vi antar att Y_{it} beror på totalkostnaderna C_{it} , på företagets produktion vid tidpunkten $t-1$ och på produktionsnivån hos ett anslutet företag j inom samma bransch vid tidpunkten $t-1$, Y_{it-1} och Y_{jt-1} , där $i \neq j$. För enkelhetens skull antar vi också att det är ett loglinjärt samband mellan variablerna.

Vi kan till exempel skriva en tidsseriemodell, så kallad VAR, enligt följande:

$$\begin{aligned} \ln(Y_{1t}) &= \alpha_{10} + \alpha_{11} \ln(Y_{1t-1}) + \alpha_{21} \ln(Y_{2t-1}) + \dots + \alpha_{n1} \ln(Y_{nt-1}) + \beta_1 \ln(C_{1t}) + \varepsilon_{1t} \\ \ln(Y_{2t}) &= \alpha_{20} + \alpha_{21} \ln(Y_{1t-1}) + \alpha_{22} \ln(Y_{2t-1}) + \dots + \alpha_{2n} \ln(Y_{nt-1}) + \beta_2 \ln(C_{2t}) + \varepsilon_{2t} \\ &\vdots \\ \ln(Y_{nt}) &= \alpha_{n0} + \alpha_{n1} \ln(Y_{1t-1}) + \alpha_{n2} \ln(Y_{2t-1}) + \dots + \alpha_{nn} \ln(Y_{nt-1}) + \beta_n \ln(C_{nt}) + \varepsilon_{nt} \end{aligned}$$

Där ε_{it} är feltermen för ekvationen i och α_{ji} och β_i , $j=0,1,\dots,n$ och $i=1,2, \dots, n$ är parametrarna. α_{ji} , för $i=j$, och fångar hur produktionen på ett bestämt företag under en bestämd period påverkar produktionen på samma företag under följande period, medan α_{ji} , för $i \neq j$, visar hur produktionen för företag j under en bestämd period påverkas av produktionen för företag i föregående period.

Enligt ekonomisk teori kan vi förvänta oss att koefficienten β_i är negativ, vilket tyder på att om $\ln(C_{it})$ ökar med en enhet och allt annat är likadant, minskar $\ln(Y_{it})$ med β_i . Den modell som beskrivs är en enkel version av flera möjliga modeller.⁴⁸

Om statligt stöd leder till ökad produktion inom en viss bransch kan det samtidigt ha effekter på sysselsättningen i samma bransch, åtminstone på kort sikt. Effekterna på sysselsättning i en bestämd bransch kan utvärderas med en tidsseriemodell. Vidare är det möjligt att utveckla tidsseriemodeller för att fånga effekterna av en bestämd åtgärd på olika branscher.

Vi bör emellertid hålla i åtanke att även tidsseriemodeller har begränsningar, till exempel på grund av behovet av tillräckligt långa tidsserier och de praktiska svårigheterna med att utvärdera ett system med många variabler. En möjlig lösning är att gruppera homogena företag inom varje bransch med hjälp av matchningsmetoder (Rosenbaum och Rubin, 1983) och välja ett representativt företag eller en grupp av företag för varje grupp som inkluderas i systemet över ekvationer som skall utvärderas. Grupperingen kan göras med olika typer av beräkningar av matchning i enlighet med det syfte som beskrivs ovan (dvs. propensity score, difference in difference, Kernel etc.).

⁴⁸ För en bättre beskrivning, se även Brocwell och Davis (2002) och Hamilton (1994). Fabiani et al. (2000) har tillämpat en SVAR-modell för att förklara fluktuationen när det gäller arbetslösheten.

Med hjälp av en tidsseriemodell av typen VAR får vi därför parametrarna α_{ji} och β_i för de representativa företagen och vi använder dem för att undersöka effekterna av genomförande av åtgärder på företagsnivå, där företagsnivå betyder grupperna av homogena företag. Observera emellertid att VAR-modellerna gör det möjligt för ITPS att även undersöka ett visst företag, men inte alla företag inom ekonomin samtidigt på grund av det stora antalet företag inom ekonomin. Såsom nämnt ovan är ITPS även intresserade av att undersöka regionala effekter. Det enklaste sättet att göra detta är att tillämpa VAR-modellen på regional nivå på det sätt som tidigare beskrivits för företagsnivån. Gruppering på regional nivå eller företagsnivå beror på målsättningarna med åtgärder.

VAR-modellerna är lämpliga för prognoser på kort och medellång sikt och längden på prognosperioden beror på längden på tillgänglig data.

För att undersöka effekterna av en särskild åtgärd på hela ekonomin eller för att göra en prognos för en lång period kan vi skapa en SAM-matris som bygger på de beräknade parametrarna av VAR-modellen. Fördelen med den här modellen är att parametrarna beräknas direkt från historiska data i stället för att härledas från det teoretiska antagandet som finns i majoriteten av befintliga modeller.

Den föreslagna modellen är därför fullständig och möjliggör undersökning av ekonomin på olika nivåer. En sådan modell kan vara av intresse för de organisationer som sysslar med subventioner, utlåning eller till exempel policy som påverkar kostnader och vinster eftersom modellen kommer att byggas utifrån kostnads- och vinstfunktioner. Men är den genomförbar? Vilka resurser krävs i så fall? Vilken kompetens och kunskap behövs? Hur många människor skulle vara engagerade? Vilken typ av data behövs?

Det är svårt att ge ett entydigt svar på dessa frågor, men ITPS har diskuterat dessa frågor med personer som ansvarar för liknande modeller. Efter samtal med Göran Östblom och Magnus Sjöström från Konjunkturinstitutet kan slutsatsen dras att utvecklingen en modell som EMEC kräver minst 3–4 arbetsår och en grupp bestående av åtta till tio forskare/analytiker. EMEC är inte den enda modellen som under lång tid utvecklats, även LUMOD, ISMOD, FIMO och andra har utvecklats över en lång tidsperiod och vi förväntar oss inte att den föreslagna modellen är ett undantag från regeln när det gäller tidsåtgång.

Nedan förklaras hur mycket resurser som krävs för att utveckla modellen och vilken resursåtgång som krävs.

Enligt en preliminär beräkning, krävs det 6–8 forskare/analytiker på heltid. En grupp bör bestå av: tre disputerade ekonometriker, två disputerade nationalekonomer, en statistiker samt en programmerare som kan arbeta under 3 till 4 år. En del kompetenta personer finns redan på ITPS. För närvarande finns det en disputerade ekonometriker, två disputerade nationalekonomer och tre statistiker på ITPS som kan vara av intresse för projektet. Men nyrekrytering är nödvändig för att frigöra dessa från ordinarie arbetsuppgifter. Kostnaden för projektet har beräknats till ca 5 miljoner kronor per år, vilket är 20 miljoner under 4 årsperiod.⁴⁹

Givet att modellen skulle vara av intresse för andra organisationer som arbetar med subventioner och utlåning, förutom ITPS, finns en möjlighet att hitta olika organisatoriska

⁴⁹ *Kostnader täcker lönekostnader, resekostnader, datorer och program varor, men inte andra logistiska kostnader (t.ex. lokalhyra och administrativa kostnader). Kostnader har beräknats med en lönenivå på 40 000 kr för fem personer och 27 000 för två personer.*

lösningar till utvecklingsarbete. Modellen skulle därför kunna utvecklas i samarbete med andra organisationer, exempelvis regeringskansliet.

Om regeringen beslutar att avsätta resurser till modellutveckling kommer ITPS att arbeta på följande sätt:

- 1 att genomföra ett test inom sex månader för att kontrollera om modellen ger avsedda resultat, dvs. om modellen är lämplig
- 2 att genomföra en testutvärdering av en policyåtgärd inom 18 månader
- 3 inom 36–48 månaders beräknas modellen vara klar för att utvärdera olika policyåtgärder.

När det gäller data finns redan det mesta som behövs tillgängligt från MM Partners⁵⁰ och ITPS har redan köpt eller samlat större delen av de data som behövs för modellen.

⁵⁰ *MM Partners databas innehåller bokslutsdata för alla aktiebolag i Sverige. Andelen enskilda och handelsbolag är negligierbar.*

5 Sammanfattande kommentarer

I det här arbetet beskrivs möjligheter till bättre utveckling av utvärderingskulturen och utvärderingsmetoderna när det gäller utvärdering av effekterna av statliga stödprogram på ekonomin.

Den första delen av arbetet tar hänsyn till den befintliga utvärderingskulturen i olika länder: Holland, USA, Irland och Sverige. Vi kan sammanfatta lärdomarna på följande sätt:

- Betydelsen av en genomtänkt utformning av åtgärder är en utmärkt utgångspunkt för utvärdering
- Betydelsen av detaljerad och systematisk sammanställning av data
- Programmen bör vara begränsade
- Tydliga politiska mål och hur de uppnås
- Behovet av att öka antalet retrospektiva analyser
- Behovet av att införliva både direkt och indirekt stiganalys i lönsamhetsstudier
- Betydelsen av kontinuerlig övervakning
- Betydelsen av att utveckla nya utvärderingsmetoder
- Att upptäcka och ta itu med nya frågor som uppkommer till följd av utvärdering
- Att i högre grad utnyttja utvärderingsresultaten i beslutsprocesser.

I analysen av den svenska utvärderingskulturen betonas behovet av att fördjupa förståelsen för effekter av olika åtgärder och ta hänsyn till externa experters resultat. Eftersom det krävs mer än övervakning för att få relevanta insikter om åtgärdernas effektivitet måste ytterligare ett steg mot en mer vetenskaplig utvärdering tas för att förse lagstiftarna med de verktyg som behövs för att utforma nya åtgärder och program.

Sverige måste röra sig mot en mer komplex utvärderingskultur för att kunna starta en inlärningsprocess där man genom utvärderingar får information och nödvändig kunskap för att bättre kunna utforma åtgärderna. Det finns två alternativa vägar som den svenska regeringen kan följa för att uppnå effektivare utvärderingar.

- 1 Det första alternativet är att utse en myndighet som utför både ex ante och ex post utvärderingar. En sådan byrå skulle ansvara för datainsamling, men inte genomföra politiken. Därmed skulle tre aktörer beröras: departementet som ansvarar för politikutformningen, en myndighet (till exempel Nutek eller Vinnova) som ansvarar för att genomföra politiken och en myndighet med ansvar att utvärdera den (t ex ITPS). Genom att tilldela olika ansvar (genomförande och utvärdering) till de två myndigheterna kan departementet öka oberoendet och tillförlitligheten i utvärderingsresultaten och minska eventuell partiskhet vilket påverkar trovärdighet hos utvärderaren, men få högre kostnader eftersom det finns två aktörer.
- 2 Det andra alternativet kan vara en enda myndighet som ansvarar för att genomföra, analysera och utvärdera politiken. Berörda aktörer skulle därmed endast vara två: departementet och myndigheten och allt skulle göras internt enligt principerna i Nuteks

system Nytt. Detta alternativ har fördelen att de administrativa kostnaderna minskas med en myndighet, men samtidigt finns det större risk för att myndigheten är partisk, i synnerhet när den utför ex ante-utvärderingar, för att öka sin budget.

Oavsett vilken väg som väljs är det uppenbart att fokus mot en ny utvärderingskultur behöver göras som riktar sig mot angivna syften och kvantitativa mål är förståeliga; att man ser bortom enkla metoder för uppföljning och förbättrar instrumenten så hänsyn tas till urvalsskevhet och leder till en inlärningsprocess.

Den andra delen av arbetet är inriktat mot utvärderingsmetoder och innehåller ett förslag på en modell som ITPS skulle kunna utveckla i framtiden. Den viktigaste slutsatsen i den andra delen är att den kontrafaktiska situationen, som besvarar dilemman ”Vad skulle ha hänt om åtgärden inte genomförts?”, kan analyseras på både mikro- och makronivå.

För att uppnå regleringsbrevsuppdragets syfte föreslås att ITPS utvecklar en modell för att undersöka effekterna av statligt stöd på tre olika nivåer: företagsnivå, regional nivå och nationell nivå. Ett lämpligt sätt för att utveckla modellen är att inleda en företagsspecifik analys av varje företag i ekonomin, gruppera dem på regional nivå för att undersöka regionala effekter och slutligen sammanställa uppgifterna från de olika regionerna för att studera effekterna på hela ekonomin. Dessvärre är ett sådant förfarande både tidsödande och tekniskt omöjligt. Vi föreslår därför att ett flexibelt system, i likhet med Riksbankens RAMSES/BVAR, utvecklas på regionala nivåer, där effekterna av olika åtgärder på mindre grupper av företag kan undersökas.

I rapporten diskuteras vilka resurser som krävs för att utveckla modellen. Det krävs 6-8 analytiker/forskare under 3 till 4 år. Inom ITPS finns redan kompetens, men nyrekrytering är nödvändig. För att utveckla modellen krävs ekonometriker, nationalekonomer och statistiker.

Den totala kostnaden för att utveckla modellen uppskattas till ca 5 miljoner kr. per år eller 20 miljoner kr. under en fyraårsperiod.

Om regeringen beslutar att ge ITPS i uppdrag att utveckla modellen kommer ITPS att arbeta på följande sätt:

- 1 att genomföra ett test inom sex månader för att kontrollera om modellen ger avsedda resultat, dvs. om modellen är lämplig eller ej
- 2 att genomföra en testutvärdering av en policyåtgärd inom 18 månader
- 3 inom 36–48 månaders beräknas modellen vara färdig för att utvärdera olika policyåtgärder.

Referenser

- Adolfson M., Laséen S., Lindé J. and Villani M. (2007). RAMSES-A new general equilibrium model for monetary policy analysis in *Sveriges Riksbank Economic Review* 2007:2
- Ahn, S.C. and Schmidt, P. (1995). Efficient Estimation of Models for Dynamic Panel Data in *Journal of Econometrics*, vol.68
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Kluwer Academic Publishers.
- Armington, P.S. (1969). A Theory of Demand for Products Distinguished by Place of Production in *International Monetary Fund Staff Papers*, vol. 16
- Armstrong, H.W. (1995). An appraisal of the evidence from cross-sectional analysis of the regional growth process within the European Union. I Armstrong. H.W. och R.W. Vickerman (eds) *Convergence and Divergence Among European Regions*. London: Pion Ltd.
- Aronsson, T., Lundberg, J. and Wikström, M. (2001). The impact of regional public expenditures on the local decision to spend in *Regional Science and Urban Economics*, vol. 33
- Arrelano, M. and Bond S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations in *Review of Economic Studies*, vol. 58
- Aten, B. (1996). Evidence of spatial autocorrelation in international prices in *Review of Income and Wealth*, vol. 42
- Baltagi, B.H. (2001). *Econometric analysis of panel data*.
- Bergvall, A., Forsfält, T., Hjelm, G., Nilsson, J. and Vartiainen, J. (2007). KIMOD 1.0 Documentation of NIER's Dynamic Macroeconomic General Equilibrium Model of the Swedish Economy, Working paper No. 100, Konjunkturinstitutet.
- Bernt, E.R. (1991). *The Practice of Econometrics: Classic and Contemporary*, Addison-Wesley.
- Besley, T. and Case, A. (1995). Incumbent behaviour: vote-seeking, tax-Setting, and yardstick competition in *American Economic Review*, vol. 85
- Bivand, R. and Szymanski S. (1997). Spatial dependence through local yardstick competition: Theory and testing in *Economics Letters*, vol. 55
- Boyle, R. (2005). Evaluation capacity development in the republic of Ireland. ECD working paper series no. 14, June 2005, Washington: *World Bank*
- Brett, C. and Pinkse J. (1997). Those taxes are all over the map: A test for spatial dependence of municipal tax rates in British Columbia in *International Regional Science Review*, vol. 20
- Brockwell, P.J. and Davis, R.A. (2002). *Introduction to Time Series and Forecasting*. Springer.

-
- Brueckner, J.K. (1998). Testing for strategic interaction among local governments: The case of growth controls in *Journal of Urban Economics*, vol.44
- Brännlund, R. and Lundgren, T. (2004). Kyoto och Basindustri – modellering och simulering, *Arbetsrapport 347*, Institutionen för Skogsekonomi, SLU Umeå.
- Brännlund, R. and Lundgren T. (2005). Swedish Industry and Kyoto – An Assessment of the Effects of the European CO2 Emission Permit Trading System, *UES 668*, Department of Economics, Umeå University.
- Brännlund, R. (1997). Industrins efterfrågan på energi. Bilaga 4 till *SOU 1997:11*.
- Brännlund, R. (2006). Grön skatteväxling- Framgångsväg eller återvändnings-gränd? *SNS förlag*.
- Case, A. (1991). Spatial patterns in household demand in *Econometrica*, vol. 59
- Case, A., Rosen, H.S. and Hines J. R. (1993). Budget Spillovers and fiscal policy interdependence: Evidence from the States in *Journal of Public Economics*, vol. 52
- Cornet M., Vroomen B. and van der Steeg M. (2006). Do innovation vouchers help SMEs to cross the bridge towards science? *CPB Discussion Paper No. 58*.
- Cox, C. J., Ingersoll, Jr. J.E. and Ross, S A. Ross (1985). An Intertemporal General Equilibrium Model of Asset Prices in *Econometrica*, vol. 53
- Cozzens S. (2003). Frameworks for Evaluating Science and Technology Policies in the United States. In *Learning from Science and Technology Policy Evaluation - Experiences from the United States and Europe*, edited by Philip Shapira and Stefan Kuhlmann. Edward Elgar, Cheltenham UK and Northampton, USA.
- Dargay, J. (1983). The Demand for Energy in Swedish Manufacturing in *Energy in Swedish Manufacturing*. IUI Stockholm.
- Department of Enterprise, Trade and Employment (2004), Expenditure review of IDA Ireland's property programme.
- Duignan, P. (2007), Seven possible outcome designs, <http://www.parkerduignan.com/oiiwa/model/>
- ESV (2001). Verksamhetslogik. *ESV 2001:16*
- ESV (2006). Måluppfyllelseanalys. *ESV 2006:7*
- European Community Treaty, http://www.europa.eu.int/comm/competition/state_aid/legislation/aid3.html
- Fabiani, S., Locarno, A., Oneto, G. and Sestito, A.P. (2000).The Sources of Unemployment Fluctuations: An Empirical Application to the Italian Case. European Central Bank: Working Paper No.29.
- Fingleton, B. (2001). Equilibrium and economic growth: Spatial econometric models and simulations. *Journal of Regional Science* 41, s 117-147.
- Forfås (1995). Evaluation of the Applied Research Programme 1988-1994.
- Forfås (1996). Evaluation of AMT Ireland.
- Forfås (2003). Review of the IDA's research and development capability grants scheme.

-
- Forfås (2004). Evaluation of agency supports for R&D in the business sector.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Heckman, J., Ichimura H. and Todd, P. (1997). Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme in *The Review of Economic Studies*, vol. 64 (4).
- Holtz-Eakin, D. (1994). Public-sector capital and the productivity puzzle in *Review of Economics and Statistics*, vol. 76
- Holtz-Eakin, D., Newey, D. and Rosen H.S. (1988). Estimating vector autoregressions with panel data in *Econometrica*, vol. 56(6)
- ITPS (2006). Statligt stöd till näringslivet 2005. S2006:007
- ITPS (2007). Sveriges Konkurrenskraft – att förstå och mäta nationell konkurrenskraft, A2007:002.
- ITPS (2007). Statligt stöd till näringslivet 2006. S2007:007
- Jansson, P., J. Hansson and M. Löf (2005). Business survey data: Do they help in forecasting GDP growth? in *International Journal of Forecasting*, vol.21.
- Kehoe, P.J. and Kehoe T.J. (1995). A Primer on Static Applied General Equilibrium Models, in P.J. Kehoe and T.J. Kehoe (eds.), *Modeling North American Economic Integration*, Boston: Kluwer Academic Publishers, 1-31.
- Kydland, F. and Prescott E. (1982). Time to build and aggregate fluctuations in *Econometrica*, vol. 50.
- Lundberg, J. (2003). On the determinants of average income growth and net migration at the municipal level in Sweden in *The Review of Regional Studies*, vol. 33(2)
- Lundberg, J. (2006a). Using spatial econometrics to analyze local growth in Sweden in *Regional Studies*, vol. 40(3)
- Lundberg, J. (2006b). A spatial interaction model of benefit spillovers from locally provided public services in *Regional Studies*, vol. 40(6)
- Mas-Colell, A, Whiston, M. D. and Green, J.R. (1995). *Microeconomic Theory*. Oxford University Press, Inc.
- Miles, I. and Cunningham, P. (2006). SMART Innovation: A Practical Guide to Evaluating Innovation Programmes, DG Enterprise and Industry, European Commission
- Ministry of Economic Affairs (2005). Monitoring Innovatievouchers 2004.
- Murdoch, J.C., Rahmatian, M. and Thayer, M.A. (1993). A spatially autoregressive median voter model of recreation expenditures in. *Public Finance Quarterly*, vol. 21
- Nutek (2006). Effektvärdering av produktutvecklingsprogrammet, Dnr 70-2006-435
- Nutek (2006). Nyttä – Nuteks system för uppföljning och utvärdering
- Poot T., den Hertog P., Grosfeld T. and Brouwer E. (2005). Evaluation of a major Dutch Tax Credit Scheme (WBSO)
www.visionerianet.org/files/55/evaluationofDutchR_Dtaxcreditscheme.pdf

-
- Restad, T (2006). LUMOD - En Långsiktig UtvecklingsModell för Svensk Ekonomi, Stockholms Universitet.
- Revelli, F. (2002). Local taxes, national politics and spatial interactions in English district election results in *European Journal of Political Economy*, vol. 18
- Rey, S.J. and Montouri, B.D. (1999). US regional income convergence: A spatial econometric perspective in *Regional Studies*, vol. 33(2)
- Rosenbaum, P and Rubin, D.B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects in *Biometrika*, vol.70.
- Rosenbaum, P and Rubin, D. (1985). Construction a Control Group Using Multivariate Matched Sampling Methods that Incorporate the Propensity Score in *American Statistician*, vol.39.
- Rubin, D. (1978). Bayesian Inference for Causal Effects: the Role of Randomization in *Annals of Statistics*, vol.7.
- Rubin, D. and Thomas N. (1992). Characterizing the Effects of Matching using Linear propensity Score Methods with Normal Distribution in *Biometrika*, vol.79.
- Shoven, J.B. and Whalley, J.L. (1984). Applied General Equilibrium Models of Taxation and International Trade: An introduction and Survey in *Journal of Economic Literature*, vol.22
- SIKA (2004). SAMPERS och SAMGODS – Nationella Modeller för Prognoser och Analyser inom Transportsektorn, SIKA 2004.
- SIKA (2004). The Swedish National Freight Model, Samplan 2004:1.
- Smets, F.and Wouters, R. (2003). An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area in *Journal of the European Economic Association*, vol.1
- Storey, D.J. (2000). Six Steps to Heaven: evaluating the impact of public policies to support small businesses in developed economies, in Landstrom H and Sexton D.L. (eds.) *Handbook of Entrepreneurship*, Blackwells, Oxford, pp. 176-194.
- Topa, G. (1996). *Social interactions, local spillovers and unemployment*. Ph.D. Dissertation, Department of Economics, University of Chicago.
- Varian, H.R. (1992). *Microeconomic Analysis*, 3rd Ed. W.W. Norton & Company, Inc.
- Vedung E (1997). *Utvärdering i politik och förvaltning*, Studentlitteratur
- Vinnova (2007). Vinnovas samlade strategi för uppföljning, utvärdering och effektanalys” DNR 2006-04079
- Östblom, G. (1991). *An Environmental Medium Term Economic Model-EMEC*. Working paper No. 69. Konjunkturinstitutet.

Bilaga 1. Diskussionsämnen för intervjuer.

A. Utvärderingskontext (det allmänna sammanhang då politiska program arbetas fram och övervägs)

1. Hur regleras utvärderingen av statligt stöd i landet? Normer, föreskrifter, lagar ...
2. Vilken slags utvärderingar utförs mest? Ex ante (förhandsbedömning) – Mid term (efter halva tiden) – Ex post (utvärdering i efterhand).

B. Relationen mellan utformare och verkställare (hur nya politiska program utformas och hur utvärderingsresultat beaktas i denna process)

1. Beskriv processen för att utarbeta, utforma och specificera ett nytt statligt stödprogram Berörda aktörer. Maktförhållanden och hierarki. Kommunikationsnivå och samordning mellan aktörer.
2. Graden av detaljer i program och riktlinjer. Använder ni programlogik eller modeller för händelsekedjor?
3. Graden av frihet på genomförandenivå.
4. Kan du ge mig ett konkret exempel på beslutsprocessen i ett statligt stödprogram?

C. Programmets målsättning (förhållandet mellan syften och mål och processen att sätta upp mål)

1. Beskriv processen att ange målsättningar och lämpliga indikatorer för dem, vilka resurser som är tillgängliga för att uppnå dem och angivna förväntningar om måluppfyllelse. Denna process kopplas ofta samman med processen att sätta upp milstolpar med indikatorer som bekräftar att man är på rätt väg.
2. Arbetar ni efter riktlinjer som t.ex. SMART (specifika, mätbara, adekvata, relevanta och tidsatta)?
3. Anser ni att de resurser som tilldelats för programmet i allmänhet är väl anpassade för de syften som angetts/valts?

D. Uppföljning och utvärdering (Utvärderingsmetoder: vilka är mest effektiva enligt era erfarenheter?)

1. Vilken slags utvärdering utför ni? Exempel (måluppfyllelse, sidoeffekter, påverkansbedömning, produktivitet, effektivitet, ...)
2. Uppföljning (mottagarnas synpunkter och åsikter om vilken skillnad programmet medfört)? Har ni problem med att uppföljningsdata saknas eller är olämpliga? Konflikt mellan att samla information och administrera programmet (fördela stöd)? Hur hanterar ni detta?
3. Utvärdering? På vilken nivå? Jämförelse av resultat hos dem som mottagit stöd med typiska företag? Jämförelse med ”matchade” företag? Redogörelse för urvalskevhet?

-
4. Vem utför utvärderingen? Privata konsultföretag, universitet, statliga myndigheter?
 5. Gemensamma förhållningssätt till utvärdering. Standardmetod för alla program?
 6. Vem definierar de olika stegen i utvärderingsgången? Uppdragsgivaren eller agenten? Hur mycket kan de personer som ansvarar för genomförandet styra programmet? Hur skickliga är de på det? Vet de vad de försöker åstadkomma för varje steg?

E. Inlärningsprocess (besluts som fattats på grund av information som kommit fram i utvärderingar)

1. Hur kommer tidigare erfarenheter från utvärderingar in i bilden? Skillnader mellan ex ante, mid term, och ex post?
2. Är utvärderingen användbar vid beslutsfattande? Används den effektivt och tar man hänsyn till den?
3. Leder utvärderingen till ändrat handlingsätt?
4. Vilka är de vanligaste problem ni möter vid utvärdering? Vilka av dessa uppträder vid ex ante utvärderingar? Hur löser ni dem?

Ge exempel från verkliga program då ovanstående frågor besvaras.

Bilaga 2. Några program och metoder från andra länder

A2.1. Nederländerna

Innovationskuponger

Det nederländska programmet med innovationskuponger från 2004 syftar till att öka samverkan mellan små och medelstora företag och offentliga kunskapsinstitutioner som universitet och tekniköverföringsinstitutioner. Kupongen är en kreditnota värd 7 500 euro som ska användas på sådana kunskapsinstitutioner och är avsedd att inriktas på sådant som uppfattas som bristande samverkan mellan kommersiella företag och forskningsinstitut.

Programmet är framför allt intressant på grund av sitt speciella upplägg, som ger en bra utgångspunkt för att bedöma det politiska instrumentets kausala inverkan. I september 2004 delades 100 kuponger ut slumpmässigt bland 1 044 sökande små och medelstora företag. På grund av den slumpmässiga fördelningen beror eventuella skillnader i innovativt beteende mellan företag som fått och inte fått en kupong helt och hållet på kupongens kausala effekt och har inte något samband som kan förklaras av andra faktorer.

Den centrala frågan är därför om små och medelstora företag som fått en kupong ger forskningsinstitutioner fler uppdrag än små och medelstora företag utan kuponger. För att besvara sådana frågor beräknar Marten Cornet, Björn Vroomen och Marc van der Steeg⁵¹ från CPB "interaktion" som det antal uppdrag som små och medelstora företag ger till offentliga kunskapsinstitutioner och definierar effektivitet som skillnaden i antalet uppdrag som lämnats av företag med och utan kuponger. I totalresultatet särskiljs uppdragens antal, storlek och val av tidpunkt.

För att kunna fastställa innovationskupongernas effektivitet samlade CPB in information om forskningsuppdragen från de företag som deltog i kupongprojektet 2004. En del information hämtades från ansökningsformuläret och en del från en speciellt framtagen enkät som genomfördes per telefon bland ett urval av företag, både sådana som fått en kupong (vinnare) och företag som inte fått någon kupong (förlorare).

Med denna informationsinsamlingsmetod var det möjligt att skilja mellan den totala gruppen av sökande och undergruppen av dem som deltog i kartläggningen. Ansökningsformuläret informerar om företagets omsättning, storlek (i antal anställda), bransch och region. Däremot finns information om forskningsuppdrag endast för de företag som deltog i undersökningen. Information som samlats in från den sistnämnda gruppen användes för att undersöka innovationskupongens effektivitet.

Frågeformuläret utarbetades i samarbete med SenterNovem för att få information om forskningsuppdrag som små och medelstora företag gett till forskningsinstitutioner. Företagen ombads lämna information om tidpunkten för uppdraget, dess värde och vilken institution som hanterade forskningsfrågan. De fick också flera andra frågor som syftade till att mäta hur nöjda de var med olika aspekter av uppdraget. Dessutom ombads företagen reagera på ett antal uttalanden om innovationskupongen och fick frågan vad de skulle ha gjort om de hade fått, eller i tillämpliga fall, inte hade fått kupongen. Intervjuerna var

⁵¹ Cornet M., Vroomen B. och van der Steeg M. (2006), "Do innovation vouchers help SMEs to cross the bridge towards science?". CPB diskussionspapper nr 58.

kvalitativa och halvstrukturerade, vilket ger möjlighet att diskutera mer komplicerade ämnen och i högre grad använda det som kallas ”öppna” frågor.

Forskning som utförts av Ministry of Economic Affairs⁵² (ekonomidepartementet) visar att gruppen på 1 044 som ansökte om innovationskuponger representerar ett tvärsnitt av små och medelstora företag i Nederländerna. 313 företag av de 1 044 som ansökte om kuponger kontaktades. Av de 313 deltagarna i undersökningen hade 142 gett forskningsinstitutioner ett eller fler uppdrag, vilket gav totalt 158 rapporterade uppdrag. Av de 71 kupongvinnarna som deltog i undersökningen hade 66 lämnat totalt 76 uppdrag. De flesta av de företag som inte lämnat något forskningsuppdrag sa att kostnaden hade varit ett allvarligt hinder. Bland de företag som tilldelades en kupong var en tredjedel inte nöjda med förhållandet mellan pris och kvalitet.

Ur frågeformulären har CPB emellertid kunnat ta fram två typer av information om beteendet hos företag när det gäller forskningsuppdrag: faktisk tilldelning av uppdrag (faktiskt beteende) och rapporterade åsikter (hypotetiskt beteende). De två informationskällorna används komplementärt för att besvara de tre forskningsfrågorna, nämligen effekten på antalet uppdrag, effekten på värdet av uppdragen och effekten på valet av tidpunkt för uppdragen.

Båda källorna ger samma resultat när det gäller kupongens additionalitet, som bedöms till åtta av tio. En av tio kuponger används inte och en av tio används för uppdrag som skulle ha tilldelats i alla fall. Att döma av svaren på förslagen finns det inga indikationer på att kupongerna har någon effekt på uppdragens värde. Men dessa svar ger viss indikation på att det finns en svag tidsrelaterad effekt, i den meningen att ett begränsat antal uppdrag lades fram.

WBSO- Promotion of Research and Development Act

1994 trädde en lag i kraft i Nederländerna för att främja forskning och utveckling, Promotion of Research and Development Act (WBSO). WBSO erbjuder ett skatteincitament som minskar lönekostnaderna för FoU-anställda genom att minska löneskatt och sociala avgifter och är den viktigaste åtgärden för att utveckla FoU-verksamhet i företagen, i fråga om omfattning och budget. Med hänsyn till syftet med WBSO är en central fråga i utvärderingen⁵³ huruvida och i vilken omfattning WBSO leder till att företagen genomför fler FoU-aktiviteter (primäreffekt) och bli mer innovativa (sekundäreffekt). Dessutom tittade utvärderingen på effekterna av WBSO när det gällde företagets innovationer (sekundäreffekt), effekter på företagets resultat (tertiäreffekt), i vilken utsträckning de omständigheter som ledde till att WBSO introducerades fortfarande är relevanta, penetreringen av WBSO:s målgrupp, uppfattningar om genomförandet och slutligen WBSO-användares uppfattningar om potentiella (budgetneutrala) förändringar av WBSO-planens utformning.

För att besvara de frågor som nämns ovan användes en kombination av metoder:

- Ekonometrisk analys. 1998 gjorde Bartels Bureau en analys av WBSO tillsammans med Nederländernas statistiska centralbyrå, Statistics Netherlands (CBS). Ett av de metodologiska resultaten var att databegränsningarna gjorde det svårt att kvantifiera effekterna av WBSO. I utvärderingarna av Poot et al (2002) gjordes en ekonomisk

⁵² Ministry of Economic Affairs (2005), "Monitoring Innovatievouchers 2004".

⁵³ Poot T., den Hertog P., Grosfeld T. och Brouwer E., "Evaluation of a major Dutch Tax Credit Scheme (WBSO)" www.visionerianet.org/files/55/evaluationofDutchR_Dtaxcreditscheme.pdf

analys som byggde på erfarenheterna från nämnda undersökning och utifrån bättre data om WBSO:s användarprofiler, som byggts upp under ett antal år. Detta gjorde det möjligt att utvärdera primära och sekundära effekter av WBSO.

- Telefonundersökning. I en detaljerad fältstudie frågade man 500 företag (nettosvar) som använde eller hade använt WBSO-incident om deras beslutsfattande när det gällde FoU, effekterna av att använda WBSO-systemet, erfarenheter av genomförandet av WBSO-planen och potentiella förbättringar av dess utformning. Då resultaten bearbetades skilde författarna mellan flera olika dimensioner som storlekskategori, sektor, FoU-intensitet, WBSO-intensitet, typ av WBSO-användare (t.ex. strukturell, sporadisk, nykomling), typ av projekt, användning av mellanhänder/stödrådgivare och om företaget är ett högteknologiskt startupföretag eller inte.
- Skrivbordsundersökning. Utöver den detaljerade analys av WBSO-utvärderingar som genomfördes 1998 har de senaste vetenskapliga insikterna och politiska studierna inom området kvantitativ utvärderingsforskning samt användningen av skatteavdragssystem noterats och inkluderats i utformningen av utvärderingen och analysen av resultaten.
- Intervjuer. Halvstrukturerade intervjuer gjordes med representanter för ett begränsat antal företag och forskningsinstitut. Dessa intervjuer var användbara framför allt för att belysa de första insikterna och för att få mer kunskaper om hur WBSO används, i synnerhet av företag som inte ingick i fältstudien eller den ekonomiska analysen.

Trots metodernas fullständighet och utvärderingarnas objektivitet måste författarna göra flera antaganden om substitution och additionalitet⁵⁴ och står inför en del praktiska metodproblem:

- Trots att man använt en bättre uppsättning data än den som användes i utvärderingen 1998 har uppgifterna fortfarande begränsningar som påverkar den ekonomiska analysen.
- En av de viktigaste frågor som diskuteras är kausaliteten mellan WBSO och en ökning av FoU-aktiviteter i företagen. Denna punkt fick exceptionellt stor uppmärksamhet i utvärderingen men för en mer definitiv slutsats om kausaliteten skulle det behövas paneldata för att göra det möjligt att följa upp företagen under en längre tidsperiod.
- De data som är tillgängliga medger inte alltid studier av effekterna av WBSO på företagsresultatet med ekonomiska metoder.

⁵⁴ 1) Först och främst att det finns en direkt koppling mellan arbetskostnader och totala utgifter för FoU-verksamhet. När de beaktar möjliga effekter utgår de ifrån att detta förhållande är konstant och anser att arbetskostnader för FoU representerar de totala FoU-utgifterna och tar inte hänsyn till vilka metoder som använts för att bevilja WBSO-anslag. 2) För det andra utgår de ifrån att företag med egna FoU-utgifter (inklusive arbetskostnader för FoU) också ansöker om WBSO. Med andra ord hade företag som inte ansökt om WBSO-medel under ett visst år inga, eller en ytterst låg nivå av FoU-utgifter. 3) Det tredje antagandet är att de planerade arbetskostnaderna för FoU under ett visst år är den avgörande faktorn för det totala beloppet arbetskostnader för FoU som man ansökt om inom ramen för WBSO. 4) För det fjärde att den reduktion man mottagit är samma som den man ansökt om.

-
- Ekonometriska utvärderingar av WBSO och liknande system skulle gynnas av att ha tillgång till paneldata under en tillräckligt lång tidsperiod.
 - Effekterna på företagets resultat (tertiära effekter) kan endast fastställas om man har tillgång till en längre tidsserie för fler variabler på individuell nivå, till exempel när det gäller marknadsförväntningar.

A2.2. Irland

På myndighetsnivå operationaliseras den fastlagda politiken i olika typer av program. Enterprise Ireland driver ett antal olika stödprogram bland annat:

- Ett program för utveckling e-handel i företag. Består av konsultcheckar till företag om max 10 000 kr
- Ett program som syftar till att förbättra små och medelstora företags konkurrenskraft genom att förbättra exportpotentialen. Ett företag kan få stöd med maximalt 4 650 000 kr varav maximalt 1 860 000 kr får avse kapitaltillskott.
- Ett program som stöder produktutveckling i tillverkande företag. Totalt stöd per företag är ca 6 000 000 kr varav ca 2 000 000 kr ska betalas tillbaka.
- Stöd till forskningsprojekt inom företag som beräknas till över 27 000 000 kr. Ersättning lämnas till maximalt 45 procent av de stödbara kostnaderna.
- Konsultcheckar för småföretag i syfte att stimulera innovativa lösningar på affärsproblem. Totalt stöd ca 45 000 kr.
- Enterprise Ireland har intressen i ett antal riskkapital fonder där stödet varierar. Om stöd ges för första gången varierar kapitaltillskottet mellan 700 000 kr till 20 000 000 kr.
- Ett program som syftar till att stödja framtagande av nya affärsidéer eller bearbetning av nya marknader. Maximalt stöd ca 600 000 under en tvåårsperiod.

IDA Ireland som motsvarar Invest in Sweden agency erbjuder företag som vill direktinvestera på Irland fyra olika typer av stöd. Anställningsstöd, kompetensutvecklingsstöd, FoU-stöd och kapitalstöd. Anställningsstödet varierar mellan 1250 € och 12500 € per anställd. Kompetensutvecklingsstödet täcker lönekostnader för en anställd under överenskommen utbildningstid. FoU-stödet täcker upp till 35 procent av stödbara kostnader för ett FoU-projekt eller maximalt ca 4 000 000 kr. Kapitalstödet ska stödja inköp av mark eller byggnation av en ny produktionsanläggning och stödet varierar mellan 7 procent och 40 procent av stödbara kostnader.

Utöver detta erbjuder Irland en av Europas lägsta företagsskatter på 12,5 procent. När det gäller statliga stöd har Näringsdepartementet de senaste åren arbetat med att förändra regelverket i syfte att säkerställa att de företag som behöver stöd får det.

En genomgång av ett antal utvärderingar som utförda av Forfås visar att utvärderingarna fokuseras till analyser av input och output från programmet. Någon koppling till regionalekonomiska eller makroekonomiska sammanhang görs inte. Genomgången nedan visar att kvalitativa metoder dominerar utvärderingarna.

-
- Utvärdering av det tillämpade forskningsprogrammet 1988-1994⁵⁵. Utvärderingen omfattar en undersökning av uppnådda prestationer genom att använda data från programmet, en enkät till deltagande företag, intervjuer med nyckelpersoner och platsbesök vid berörda tekniska högskolor.
 - Utvärdering av AMT Ireland. Irlands äldsta program för att stödja och utveckla avancerad teknologi. Metodologiskt ligger tyngdpunkten på ”peer review” av projekt, intervjuer med programavnamare, genomgång av programmets processtyrning. Utvärderarna kommenterar den valda ansatsen som även använts tidigare.⁵⁶
 - *In 1993, we returned to monitor the implementation of our recommendations and the subsequent progress of the programme. We believe that this process of evaluation and follow-up has been among the major factors leading to the significant improvements in programme performance. It reinforces our faith in so-called 'formative' evaluation, where the job of the evaluator is not merely to award marks for performance, like an Olympic judge, but to intervene to help improve performance, like a sports coach.*
 - Utvärdering av IDAs stödprogram för FoU-möjligheter. Programmet syftar till att stödja företag som redan har etablerat sig på Irland med att utveckla FoU-verksamheten. Utvärderingen använder sig av intervjuer av olika grupper, genomgång av all tillgänglig information om programmet.⁵⁷
 - Utvärdering av myndighetsstöd till FoU utfört av näringslivet. Utvärderingen bygger på både kvalitativa och kvantitativa metoder. Intervjuer, fokusgrupper och platsbesök med intervjuer samt litteraturstudier tillhör den kvalitativa ansatsen. Projektdata har använts för kvantitativa analyser. Dödviktsdiskussion förs utifrån intervjudata.⁵⁸

⁵⁵ *Evaluation of the Applied Research Programme 1988-1994. 1 dec 1995 Forfás.*

⁵⁶ *Evaluation of AMT Ireland, Forfás 1996*

⁵⁷ *Review of the IDAs research and development capability grants scheme, 9 may 2003, Forfás*

⁵⁸ *Evaluation of agency supports for R&D in the business sector, 25 october 2004, Forfás*

Bilaga 3. Olika ekonometriska program för rumslig analys

Det finns en rad olika ekonometriska programpaket, alla mer eller mindre specialiserade på olika saker. Generellt gäller att ju mer användarvänliga programmen är desto mindre specialiserade tenderar de att vara och även mindre flexibla vad gäller programmering av test och estimatorer som inte redan finns inbyggda i programmet. För att kunna nyttja de allra senaste estimationsmetoderna inom ett område krävs ofta ett flexibelt program med en hög grad av egen programmering. Detta ställer dock högre krav på användarens kompetens. Dels ska användaren kunna ta del av de senaste forskningsrönen inom området, dels kunna använda detta i praktiken genom egen programmering.

Allmänt gäller att alla program som tillåter egen programmering, matrisalgebra samt icke-linjär optimering går att använda för att skatta rumsliga ekonometriska modeller och test. Den stora fördelen med specialprogrammen är möjligheterna att på ett enkelt sätt skapa olika typer av viktmatriser. Exempelvis är det inte alltid helt enkelt att programmera procedurer för att radstandardisera eller skapa ”cut off” värden i en viktmatris. Dessutom går det av estimeringstekniska skäl inte att använda den vanliga minsta kvadratmetoden (OLS) för att estimerade rumsliga modeller. Istället används ofta Maximum Likelihood (ML) eller instrumentvariable estimatorer (exempelvis 2SLS eller GMM) vilka är betydligt mer komplicerade att på egen hand programmera upp. Egen programmering av både deskriptiva test för rumsligt beroende och regressionsmodeller med rumsligt beroende kräver mycket goda kunskaper i ekonometri samt vana att programmera i olika typer av statistikprogram.

Allmänna ekonometriprogram

Ett allmänt ekonometriprogram är användbart för att skatta andra typer av modeller än rumsliga. Det kan naturligtvis vara så att det inte föreligger något rumsligt beroende och då är ett allmänt ekonometriprogram att föredra då specialprogrammen ofta har ett begränsat utbud av andra typer av estimatorer. Nackdelen är att de allmänna programmen ofta endast har få inbyggda applikationer för att skatta rumsliga modeller och därmed ställer högre krav på egen programmering av användaren.

Stata

Stata är ett mycket kompetent allmänt ekonometriprogram som både ger stora möjligheter till egen programmering samtidigt som det finns en stor mängd färdiga estimatorer och olika typer av ekonometriska test. I grundutförandet finns dock inga speciella applikationer för rumslig ekonometri, men det går att i efterhand ladda hem Moran's *I*, Geary's *C*, de olika *G*-statistikerna samt Maximum Likelihood estimatorer för både lag- och errormodellerna från Statas officiella hemsida. I princip går det att med hjälp av de mycket flexibla möjligheterna till egen programmering och matrisalgebra att på egen hand programmera upp och ändra viktmatriserna samt estimerade olika typer av rumsliga regressionsmodeller. Detta ställer dock höga krav på användarens kompetens.

GAUSS

GAUSS är i grunden ett matrisalgebraprogram och utgjorde till en början stommen för SpaceStat (se mer detaljerad information om SpaceStat nedan). GAUSS är extremt flexibelt samtidigt som det är tämligen användarvänligt i den bemärkelsen att det mesta måste användaren programmera upp själv. Fördelen är att programmet i princip inte sätter

några restriktioner på vad som kan göras eller vilka estimatorer som kan användas. Nya test och estimatorer kan för en van användare enkelt programmeras upp. Nackdelen med programmet är att det ställer höga krav på användaren vad gäller ekonometrisk kompetens och programmeringsvana.

SAS

SAS är även det ett mycket kraftfullt ekonometri- och statistikprogram som ger användaren stora möjligheter till egen programmering. Programmets styrka ligger i tidsserieanalys och prognoser. Det finns dock möjlighet till matris algebra och optimering, men det finns inga färdiga applikationer för rumslig ekonometri.

Limdep

Limdep är ett tämligen flexibelt program samtidigt som det innehåller en stor mängd förprogrammerade estimatorer och procedurer. Ursprungligen var Limdep utvecklat för att beräkna så kallade tobit estimatorer, därav namnet *Limited Dependent variable models*. Programmet innehåller även en mängd analysverktyg och estimatorer för tvärsnitt och paneldata. Limdep går att använda för att estimerarumsliga modeller, men i princip allt måste programmeras upp då det inte finns några förprogrammerade applikationer.

SPSS

SPSS är kanske det mest användarvänliga programmet av dem som behandlas här, och därmed samtidigt det mest begränsade. Programmet har sin styrka i andra typer av modeller, inte rumslig ekonometri.

Matlab

Matlab är ett mycket flexibelt matematikprogram med en rad inbyggda optimeringsrutiner. Matlab klarar i princip alla typer av matematiska beräkningar, både tillämpad, symbolisk och teoretisk matematik. Matlab har även väl utbyggda simuleringsrutiner. Trots att det i Matlabs basutförande inte finns några specifika funktioner för rumslig ekonometri går nästan allt att programmera. Dessutom finns ett stort utbud av applikationer som går att ladda hem från Internet och allt fler som håller på med rumslig ekonometri på avancerad nivå går idag över från andra program till Matlab.

TSP och E-views

E-views var ursprungligen en variant av TSP anpassad för vad som i början av 1980-talet kallades för micro-datorer. Eftersom den tidens TSP i original utförande var för stort för att kunna installeras på micro-datorer utvecklades vad som då kallades för microTSP, vilket sedermera blev E-views. Styrkan hos båda dessa program är dynamiska modeller och panel data. En viss mån av flexibilitet finns i programmen genom att det är möjligt att utföra vissa matrisalgebraiska operationer. Däremot är programmen inte lämpliga för att omvandla viktmatriser eller mer avancerade algebraiska funktioner. Ingen av dessa program har inbyggda moduler för rumslig ekonometri och får egentligen anses som tämligen olämpliga för den här typen av modeller.

S-PLUS

S-PLUS är ett bra ekonometriprogram där det går att lägga till en specifik modul för rumsliga modeller. Programmet går att länka till kartritningsprogrammen ArcView och Grassland för visuell framställning av resultaten i kartform. S-PLUS innehåller band annat globala och lokal test för rumslig korrelation (Moran's I , Geary's C , LISA) samt en mängd olika regressionsmodeller. I programmet finns färdiga applikationer för att både definiera en viktmatris och sedan göra om den utifrån olika kriterier. S-PLUS i kombination med exempelvis ArcView måste anses vara en god kombination för rumsliga analyser.

Speciella ekonometriprogram för rumsliga modeller

SpaceStat

SpaceStat är ett specialprogram för rumslig ekonometri med ett flertal inbyggda estimatorer och test. En av programmets starkaste sidor är att när en viktmatris väl laddats in i programmet är det mycket enkelt att transferera om matrisen betingat på olika kriterium. Exempelvis går det enkelt att radstandardisera, införa ”cut off” värden och på annat sätt manipulera viktmatrisen. Dessutom finns Moran’s I , Geary’s C , de olika G_i statistikerna, olika typer av ML och IV estimatorer (inklusive GMM) redan förprogrammerade i programmet. Det finns även möjlighet att utföra enklare matrisalgebra och införa strukturell instabilitet i regressionsmodellerna. SpaceStat har även fördelen att det går att koppla till ArcView, vilket är ett GIS-baserat kartritningsprogram. Därmed kan resultaten från exempelvis de lokala testen för rumsligt beroende enkelt visualiseras i kartor. I SpaceStat klarar också av att skatta många ”vanliga” modeller utan rumsliga effekter. SpaceStat är ett mycket kompetent program för rumslig analys.

GeoDa

GeoDa finns att tillgå utan kostnad via Internet (<https://www.geoda.uiuc.edu/>). Programmet, vilket är tämligen användarvänligt, är främst avsett för deskriptiv analys och grafisk framställning av data i form av kartor. Test för rumslig korrelation och indikatorer för rumsliga outliers (både lokala och globala test för rumslig korrelation) finns inbyggt i programmet. Det finns även väl utvecklade rutiner för att transformera viktmatrisen efter olika kriterier. Däremot finns inga möjligheter till regressionsanalys vilket måste anses vara en stark begränsning för mer avancerad analys. Trots detta kan programmet med fördel användas som ett komplement till andra program, speciellt som programmet är fritt tillgängligt via Internet.

STARS

STARS, Space-Time Analysis of Regional Systems, är ett program för deskriptiv analys av datamängder där det både finns ett beroende mellan regioner och över tiden. Programmet är tämligen användarvänligt och innehåller applikationer för grafisk framställning av olika förlopp. Programmet går att ladda ner gratis från Internet (<http://regionalanalysislab.org/?n=STARS>). STARS har sina fördelar i vad gäller grafisk framställning av regionala utvecklingstendenser över tiden samt deskriptiva test, både lokala och globala test för rumslig korrelation. Programmets svaghet ligger i avsaknaden av regressionsanalytiska verktyg. I likhet med GeoDa måste därför STARS ses mer som ett komplement till andra program.