

Går det att göra rättvisande utvärderingar av den omfattande arbetsmarknadspolitik som har förts i Sverige under 1990-talet? Kenneth Carling och Laura Larsson hävdade i föregående nummer av *Arbetsmarknad & Arbetsliv* att sådana utvärderingar i princip är omöjliga att göra. I detta nummer kommer två inlägg i debatten av Lars Behrenz och Anders Harkman, som reser invändningar mot denna negativa syn på utvärderingens möjligheter inom arbetsmarknadsområdet. Kenneth Carling och Laura Larsson ger avslutningsvis en kommentar.

Lars Behrenz

Låt inte det bästa bli det godas fiende

– om att utvärdera svensk arbetsmarknadspolitik

Att utvärdera arbetsmarknadspolitik är inte enkelt.¹ Olika typer av projekt eller program ska behandlas som ett nytt "piller" utan de möjligheter som t ex medicinare har att genomföra kontrollerade experiment.² Utvärderingsmetoden utgår ofta från jämförelser mellan försöks- och kontrollgrupper över tiden. I Kenneth Carlings och Laura Larssons artikel i *Arbetsmarknad & Arbetsliv* nr 3 år 2000 tas utvärderingsproblematiken vad gäller jämförbarheten mellan olika studerade grupper upp. Författarna menar att utformningen av den svenska arbetsmarknadspolitikens medför problem när det gäller att besvara de grundläggande utvärderingsfrågorna. De hävdar att man i Sverige

inte kan besvara de klassiska utvärderingsfrågorna:

- (1) "vilket är deltagarnas utfall i en värld med eller utan program",
- (2a) "vilket är utfallet för deltagarna i program P i en värld med alla program jämfört med utfallet i en värld helt utan program?" och
- (2b) "vilket är utfallet för deltagarna i program P i en värld med alla program jämfört med utfallet i en värld med alla program förutom P ?"

1. Se t ex Delander & Niklasson (1987) för en genomgång av metoderna för utvärderingar av arbetsmarknadspolitik. Se Heckman & Smith (1996) och Heckman m fl (1999) för en genomgång av de teoretiska och empiriska erfarenheterna vid effektanalyser.

2. Det är otroligt att det endast vid ett enda tillfälle har genomförts ett riktigt experiment, på det arbetsmarknadspolitiska området, i Sverige, Delander (1978).

Lars Behrenz är fil dr i nationalekonomi och föreståndare för Centrum för Arbetsmarknadspolitisk forskning (CAFO) vid Växjö universitet.

Carling och Larsson påstår att den svenska arbetsmarknadspolitikens utformning står i strid med den vedertagna utvärderingsteorin, vilket innebär att de klassiska utvärderingsfrågorna inte kan besvaras. Författarnas huvudinvändning mot uppläggningsen att jämföra olika grupper som tar del av olika åtgärder med varandra är att det krävs att jämförelsepersonerna endast har en period i livet då de har möjlighet att delta i det aktuella programmet. De menar att nästan samtliga arbetslösa, genom den svenska arbetsmarknadspolitikens utformning, någon gång blir erbjudna att delta i ett program och då skulle inte det kontrafaktiska tillståndet existera.

Carlings och Larssons poäng är att det är svårt att göra analyser av program kontra inget program alls. Man kan dock hävda att: "rätt svar är viktigt men Carling och Larssons fråga är inte mest intressant". Jag tycker att vi ska ha en aktiv arbetsmarknadspolitik och då blir den intressanta frågan: "Vilket program fungerar bäst för aktuell målgrupp?" Det "kontrafaktiska skeendet" blir olika alternativa insatser och då blir problemen inte lika svårlösta som när vi jämför med inga program alls. Carlings och Larssons huvudinvändning mot den vanliga utvärderingsmetoden med jämförelsegrupper vad gäller fråga 2b (jämförelser mellan olika program) är väldigt dåligt underbyggd. De hävdar att man inte kan jämföra olika grupper med varandra pga att jämförelsegruppen inte endast har en period i livet där de kan delta i aktuellt program. Det är en mycket märklig uppfattning. Jag kan inte förstå att en jämförelse mellan deltagarna i ett visst program och deltagare i andra program inte skulle vara möjlig. Om man ser till så att jämförelsegruppen inte består av personer som får ta del av "försöksåtgärden" och om man kan kontrollera för

skillnader mellan grupperna t ex ifråga om selektion och heterogenitet så får man väl robusta resultat? Tolkningen av jämförelser mellan olika grupper får göras under förutsättning av vi inte kommer att kunna garantera att jämförelsegruppen aldrig får ta del av det aktuella programmet. Det finns också andra aspekter som bör vägas in. Samtliga program ändras mer eller mindre mycket över tiden vilket innebär att ett program idag inte är detsamma om t ex fem år. Regler, lagar och behörigheter kan ändras i framtiden. Carling och Larsson hävdar att de speciella förhållandena i Sverige gör det omöjligt att jämföra olika program (eller program mot inga program alls) med varandra. På vilket sätt är Sverige unikt? I samtliga länder med någon form av utvecklad arbetsmarknadspolitik kommer man väl, som långtidsarbetslös, att få ta del av någon åtgärd?

För mig framstår problemet med att jämförelsegruppen eventuellt kan få ta del av åtgärden ifråga i framtiden som ett olösligt problem för samtliga typer av experiment och försök inom samtliga vetenskaper om man inte kan modellera dessa förutsättningar. Tar man Carling och Larsson exakt på orden så är väl den första åtgärden att propagera för att samtliga experiment/försök inom samtliga vetenskaper bör avvecklas så länge man inte kan "blockera bort" jämförelsegruppen från det aktuella "pillret" för all framtid. Carling och Larsson anger "blockering" och jämförelser mellan "specialfall" (t ex 24 åringar mot 25 åringar) där den ena gruppen, per definition, inte kan få tillgång till aktuell åtgärd. Som jag ser det finns det anledning att försöka utveckla dessa idéer. Jag anser dock att utvärderingsproblemen kan vara minst lika stora som vid jämförelser mellan individer som fått ta del av olika program (t ex kan "block-

ering” endast ske av nya program och hur kan vi vara säkra på att jämförelsegruppen inte, genom regeländringar, får ta del av aktuell åtgärd i framtiden) som vid Carling och Larssons förslag. För det tredje hävdar jag att utvärderingar måste göras, punkt slut. Vi måste ha beslutsunderlag. Genom att påtala de brister som finns i t ex databaser och allmänna projektuppläggningar kan utvärderingsförutsättningarna bli bättre i framtiden men tills världen är helt perfekt kan man inte isolera sig från den, man får göra sitt bästa.

Varför utvärdera?

Utvärderingar genomförs, normalt sett, av dels olika typer av försöksverksamheter, dels av pågående program. Huvudsyftet med att genomföra en utvärdering är, enligt min uppfattning, dels att ta fram ett bra beslutsunderlag, dels att ta fram ny kunskap för att utgöra en liten ”byggkloss” i den empiriska kunskapen om arbetsmarknaden. Ofta behöver det inte råda någon konflikt mellan de båda huvudsyftena med utvärderingar.

Utvärderingar av olika typer av arbetsmarknadspolitiska program kan genomföras av flera orsaker. Dimensioneringen mellan olika program kan bli grundas på olika utvärderingar. För mig blir huvudutgångspunkten, vid utvärderingar av program, att svara på frågan: ”är det aktuella programmet sämre eller bättre jämfört med andra program som fyller samma syfte?”

I de fall där en konflikt kan uppstå mellan å ena sidan politikens behov av bra beslutsunderlag och forskningens krav på stringent genomförda utvärderingar måste man enligt min åsikt, trots alla problem, hitta någon lösning. Det kan finnas situationer där man som utvärderare har stora svårigheter att genomföra en helt invändningsfri utvärde-

ring.³ Man kan komma för sent in i utvärderingen, det går inte att hitta något helt perfekt jämförelseobjekt, tillgängliga data saknar viktig information osv. Då ställs man som utvärderingsforskare i en valsituation: ska man genomföra utvärderingen trots att uppläggningsen har brister? Ur mitt perspektiv måste vi som utvärderare svara ja på frågan, om det finns förutsättningar för att utvärderingen kommer att förbättra underlaget för framtida beslut och ställningstagande i den aktuella frågan. Beslut måste fattas. Sedan är det givetvis så att som forskare har man ett ansvar att endast uttala sig om det som det går att säga något om. Vad det handlar om är att göra klart för tjänstemän och politiker att ju större problem som finns desto osäkrare blir slutsatserna.

Slutkommentarer

Carling och Larsson (2000) har startat en diskussion kring möjligheten att utvärdera arbetsmarknadspolitik. De invändningar mot utvärderingar som Carling och Larsson har tagit fram till diskussion gäller möjligheten att jämföra olika grupper av individer under den situationen att även jämförelsegruppen kan få tillgång till aktuellt program i framtiden. Denna invändning mot utvärderingar av arbetsmarknadspolitik är helt relevant när det gäller jämförelsen mellan program och inget program alls. Carlings och Larssons invändningar mot utvärderingar som bygger på jämförelser mellan olika program förefaller däremot konstig.

3. Jag har själv erfarenhet av att man, under pågående projektperiod, flyttade personal som utgjorde en personalförstärkning så att det som skulle utvärderas försvann och ett antal försökskontor fick strykas ur utvärderingen, se Behrenz (1998).

Varför kan man inte jämföra personer som får ta del av olika "försökspiller" givet att man kontrollerar för skillnader mellan grupperna i olika avseenden? Program ändras också över tiden, beslut fattas på basis av situationen idag och vi måste fatta beslut idag. Stora resurser satsas på arbetsmarknadspolitik och för att kunna göra bedömningar om "det är värt pengarna" måste vi ha uppgifter om effekterna som underlag i t ex olika samhällsekonomiska kalkyler.⁴

Vi som bedriver utvärderingsforskning har ett ansvar att utföra utvärderingar, givetvis har vi också ett ansvar för att dessa utvärderingar skall vara pålitliga. Oftast när olika förutsättningar för "den perfekta utvärderingen" inte är uppfyllda är det ändå min uppfattning att den metod som används vid utvärderingar av arbetsmarknadspolitik ger ett bättre beslutsunderlag än att säga till politikerna: "det går inte, jämförelsegruppen kan eventuellt få tillgång till aktuell åtgärd i framtiden!" Det handlar om att vara mycket tydlig vad gäller de reservationer som måste göras på grund av de, eventuellt, bristfälliga utvärderingsförutsättningarna.

Referenser

- Behrenz L (1998) *Essays on the Employment Service and Employers' Recruitment Behaviour*. Lund Economic Studies 79, Doktorsavhandling, Lunds universitet.
- Carling K & Larsson L (2000) "Att utvärdera arbetsmarknadsprogram i Sverige: Rätt svar är viktigt, men vilken var nu frågan?" *Arbetsmarknad & Arbetsliv*, årg 6, nr 3, s 185–192.
- Delander L (1978) *Studier kring den arbetsförmelande verksamheten*. SOU 1978:60, Fritzes, Stockholm.
- Delander L & Niklasson H (1987) *Metoder för evalueringar av arbetsmarknadspolitik*. NAUT-rapport 1987:4, Nordiska arbetsmarknadsutskottet, Köpenhamn.
- Delander L & Niklasson H (1996) "Cost-Benefit Analysis" i Schmid G m fl (red) *International Handbook of Labour Market Policy and Evaluation*, Edward Elgar, Cheltenham.
- Heckman J & Smith J A (1996) "Experimental and Nonexperimental Evaluation" i Schmid G m fl (red) *International Handbook of Labour Market Policy and Evaluation*, Edward Elgar, Cheltenham.
- Heckman J, Lalonde R J & Smith J A (1999) "The Economics and Econometrics of Active Labor Market Programmes" i Ashenfelter O & Card D (red) *Handbook of Labor Economics*. Volume IIIA, Elsevier Science BV, Amsterdam.

4. Se Delander & Niklasson (1996) för en beskrivning av "Cost-Benefit analys".

Går det att utvärdera svensk arbetsmarknadspolitik?

I en artikel i förra numret av *Arbetsmarknad & Arbetsliv* (år 6, nr 3) hävdar Kenneth Carling och Laura Larsson att det är omöjligt att med svenska data från nittio-talet besvara frågan om effekten av att delta i ett arbetsmarknadspolitiskt program jämfört med att inte delta. Orsaken är den omfattning som de aktiva åtgärderna haft och som innebär att det är omöjligt att hitta en kontrollgrupp som representerar det som skulle ha hänt om man inte deltagit ("det kontrafaktiska tillståndet"). Som exempel på studier som trots detta försökt besvara den aktuella frågan nämns två där jag har medverkat (AMS 1997 och AMS 1999). En färsk utvärdering av Larsson (2000) räddas från kritiken genom att den påstås besvara en annan fråga än AMS-rapporterna.

Carlings och Larssons huvudpoäng är att det normalt inte går att med en jämförelse mellan deltagare och icke-deltagare med svenska data besvara frågan vad som skulle hända om programmet inte fanns (undantag är det specialfall som författarna föreslår att man kan utnyttja). Det går nämligen inte – enligt författarna – att hitta jämförelsegrupper som i praktiken lever i en värld utan program (eller det aktuella programmet).

Carling och Larsson menar alltså att det krävs att kontrollgruppen lever i en värld

där de inte har möjlighet att delta i det program som skall utvärderas för att man skall kunna besvara frågan om effekten av att delta i programmet jämfört med att inte delta. Tyvärr redovisar författarna inte några egentliga skäl till varför detta skulle vara ett nödvändigt villkor utan hänvisar endast till att den bakomliggande teorin är sådan.¹ Själv kan jag tänka mig två skäl till att villkoret skulle vara nödvändigt. Det första hänger samman med hur man preciserar själva frågeställningen och det andra hänger samman med mer "tekniska" mätproblem. Jag menar dock att inget av dessa skäl motiverar Carlings och Larssons ståndpunkt att det är näst intill omöjligt att besvara frågan om effekten av att delta respektive inte delta i de arbetsmarknadspolitiska programmen i Sverige.

Låt mig börja med frågeställningen. Vad menas med "effekten av att delta jämfört med att inte delta"? Frågan kan tolkas på (åtminstone) två olika sätt. En tolkning, som framskymtar i artikeln, är att man vill undersöka *vad som skulle ha hänt med deltagarna om de inte haft möjlighet att delta men där de har möjlighet att välja bland alla*

Anders Harkman är nationalekonom och utredare på Arbetsmarknadsstyrelsen.

1. Presentationen av den modell som Carling och Larsson hänvisar till i Heckman m fl (1999) inleds med "To simplify the discussion, [...] assume that persons have only one period in their lives – period k – where they have the chance to take job training" (s 1 914, min kursivering). Detta indikerar ganska tydligt att man där inte menar att det är ett nödvändigt villkor.

andra alternativ som finns i samhället. För att se vilka val man gör är det då av central betydelse att man faktiskt är utestängd från att även senare delta i det aktuella programmet. En möjlighet att senare delta kan starkt påverka de val man gör idag. Anta att en person kan skaffa sig en yrkesutbildning antingen via ett arbetsmarknadspolitiskt program som är gratis och där deltagarna får ersättning motsvarande arbetslöshetsersättningen eller via en reguljär utbildning där deltagarna själva måste betala för uppehållet under kurs tiden. Man kan då anta att många som annars skulle kunna tänka sig att delta i den reguljära utbildningen väntar i det längsta med detta om de har chansen att få delta i det arbetsmarknadspolitiska programmet.

En annan tolkning av frågan är istället att man vill mäta programmets *behandlings effekt*, dvs i vilken utsträckning deltagaren blir mer attraktiv på arbetsmarknaden genom att delta i programmet. I exemplet med yrkesutbildningen vill man veta om de kunskaper och färdigheter som deltagaren får genom programmet är värda de kostnader som det förorsakar. I det här fallet finns, såvitt jag kan förstå, inte något principiellt skäl mot att personer i en kontrollgrupp senare kan delta i det aktuella programmet. Det som är viktigt i det här fallet är att jämförelsegruppen inte deltar i attraktivitetshöjande aktiviteter i annan utsträckning än deltagargruppen under uppföljningsperioden och att den enda skillnaden mellan grupperna är att den ena deltagit i det aktuella programmet och den andra inte. Personerna i jämförelsegruppen kan fungera som jämförelse ända tills de går in i någon annan aktivitet som syftar till att höja attraktiviteten.

Skillnaden mellan de båda frågeställningarna är att den första tillåter deltagarna i jämförelsegruppen att kompensera det ute-

blivna deltagandet i det arbetsmarknadspolitiska programmet med andra aktiviteter (eller andra program) medan den andra förutsätter att deltagarna avhåller sig från detta. Frågorna ger samma svar bara i det fall då det inte finns några som helst alternativ till programmet. Den senare varianten är dessutom den som är mest intressant t ex för dem som ska fatta beslut om arbetsmarknadspolitikkens utformning och inriktning och också den som utvärderare i allmänhet försöker besvara. Svaret på den frågan räcker för att fatta beslut om ett program ska finnas kvar eller dras in. Ett program som inte har någon effekt som motsvarar kostnaden är inte värt att behålla (om det inte finns andra icke-ekonomiska värden som ändå motiverar programmet).

Det andra tänkbara skälet till att man behöver en "värld utan program" – ett skäl som mer har att göra med mättekniska problem – är att en mätning av effekten kan göras först lång tid efter att man deltagit i programmet. Det kräver då i praktiken att jämförelsegruppen lever i "en värld utan program". Här ställer den omfattande arbetsmarknadspolitiken onekligen till problem. Så gott som alla som misslyckas med att hitta jobb eller en annan lösning utanför arbetsmarknaden kommer, som Carling och Larsson visar, förr eller senare att gå in i ett program. Det innebär att man snabbt får svårt att hitta icke-deltagare som inte gjort olika kompenserande val (gått in i något program). Det är alltså praktiskt omöjligt att direkt läsa av programmets effekter genom att jämföra utfallet för deltagarna med motsvarande utfall för icke-deltagarna en lång tid efter programmet.²

2. Vid den tidpunkten har säkert också politiker och andra beslutsfattare tappat intresset för frågan.

Det finns dock inget principiellt problem med att man har möjlighet att vid ett senare tillfälle delta i ett visst program. En jämförelsegrupp av personer med samma utgångsläge som deltagarna är en relevant jämförelsegrupp ända tills de har deltagit i andra attraktivitetshöjande aktiviteter. Man kan försöka att på ett intelligent sätt utnyttja en kortare tidsperiod efter programmet för jämförelsen. I vilken utsträckning detta ger rättvisande resultat beror på hur nära i tiden effekterna kan läsas av. Min personliga uppfattning är att man kan komma ganska långt den vägen. Det rör sig ju inte om någon livslång behandling utan en ganska kort, väl avgränsad, period som dessutom förväntas ge ganska snabbt resultat i form av att man snabbare lämnar arbetslösheten.

Även om författarna menar att det är näst intill omöjligt att besvara frågan om effekten av att delta respektive inte delta i ett program framkommer att de menar både att man kan jämföra effektiviteten av olika program och jämföra programdeltagande med att "skjuta upp" deltagandet. Eftersom de här möjligheterna inte närmare diskuteras i artikeln så är det inte heller helt klart vad som motiverar den ståndpunkten. Även vid jämförelser mellan olika program råkar man ut för det problem som skapas av att nästan alla som misslyckas med att hitta jobb förr eller senare går in i ett nytt program vilket gör *långsiktiga* jämförelser omöjliga. Ändå är den svårigheten överkomlig och jämförelser av den typen politiskt relevanta och av stort intresse.

Vilken fråga är det man besvarar genom att jämföra personer som i anslutning till arbetslöshet går in i ett program med sådana som fortsätter att söka som öppet arbetslösa? Rent formellt kan man förstås säga att det är effekten av att skjuta upp

deltagandet. Jag har dock svårt att inse varför inte detta också är en relevant utgångspunkt för att mäta behandlingseffekten av ett program. Åtminstone hittar jag inte i artikeln några övertygande argument mot att en jämförelse av den typen faktiskt mäter behandlingseffekten av programmet.

Sammantaget anser jag att författarna i artikeln ger en alltför pessimistisk bild av möjligheterna att utvärdera den svenska politiken. Tvärtemot Carlings och Larssons uppfattning menar jag att det i princip inte finns något hinder för att besvara frågan om effekten av att delta jämfört med att inte delta i ett program om man med detta menar den behandlingseffekt som programmet har. Det är också den frågan som är mest intressant för beslutsfattarna. Den omfattande användningen av arbetsmarknadspolitiska program innebär problem med att mäta effekten men utgör inte något principiellt hinder mot att besvara frågan.

Referenser

- AMS (1997) "Den yrkesinriktade arbetsmarknadsutbildningen 1994 – effekter för individen" Ura 1997:10, Arbetsmarknadsstyrelsen, Solna.
- AMS (1999) "Åtgärdsundersökningen 1998 – åtgärdernas effekter på deltagarnas sysselsättning och löner" Ura 1999:1, Arbetsmarknadsstyrelsen, Solna.
- Carling K & Larsson L (2000) "Att utvärdera arbetsmarknadsprogram i Sverige: Rätt svar är viktigt, men vilken var nu frågan?" *Arbetsmarknad & Arbetsliv*, årg 6, nr 3, s 185–192.
- Heckman J, Lalonde R L & Smith J A (1999) "The Economics and Econometrics of Active Labor Market Programs" i Ashenfelter O & Card D (red) *Handbook of Labor Economics*, Volume III, Elsevier Science BV, Amsterdam.
- Larsson L (2000) "Evaluation of Swedish Youth Labour Market Programmes" Working Paper 2000:1, IFAU, Uppsala.

Replik till Lars Behrenz och Anders Harkman

Vid Institutet för arbetsmarknadspolitisk utvärdering (IFAU) – där vi bägge två är verksamma – bedrivs utvärderingar av den svenska arbetsmarknadspolitiken. Det problem vi tog upp i debattartikeln i föregående nummer av *Arbetsmarknad & Arbetsliv* (årg 6, nr 3) har därför största relevans i vår verksamhet. Faktum är att problemet tydliggjordes när en av författarna skulle tolka den egna studiens uppmätta effekter (Larsson 2000). Avsikten med att skriva debattartikeln var att beskriva problemet och att pröva våra argument mot andras. Vi är därför tacksamma för att Lars Behrenz och Anders Harkman har försökt finna argument mot vår slutsats i debattartikeln. Tyvärr kan vi inte se att det finns skäl att ändra slutsatsen, utan endast att (möjligen) modifiera den något.

Behrenz förefaller i sak dela vår slutsats, men framför några kompletterande synpunkter. I vårt svar bemöter vi punkt för punkt sex synpunkter som han framför. Harkman delar tydligen inte vår uppfattning, utan han menar att "Personerna i jämförelsegruppen kan fungera som jäm-

förelse ända tills de går in i någon annan aktivitet som syftar till att höja attraktiviteten". I vår replik menar vi oss visa att det påståendet är sant endast i en mycket speciell och orealistisk situation.

Lars Behrenz

(i) Behrenz menar att vi skall ha en aktiv arbetsmarknadspolitik och därför är den klassiska utvärderingsfrågan mindre intressant än andra utvärderingsfrågor, såsom frågor rörande programmens målgrupp, omfattning och utformning. Resonemanget är bakvänt. De andra frågorna blir aktuella efter det att man visat att arbetsmarknadsprogrammen har en positiv effekt. Förvisso kan man postulera att programmen har en positiv effekt, men det är ett antagande som inte vilar på empirisk grund.

(ii) Behrenz menar att vi inte visat empiriskt att fråga 2b inte låter sig besvaras. Vi instämmer och tar därmed tillbaka det påståendet.

(iii) Behrenz undrar, med syftning på problemet, på vilket sätt Sverige är unikt. Sverige är inte unikt, och förmodligen är inte heller 1990-talet det. J Smith beskrev ett likartat problem (personlig kommunikation 2000), med hänvisning till flera olika länder, i samband med en utvärderingskonferens i London i juli 2000. Unikt i detta sammanhang är däremot Schweiz: Efter två år som arbetssökande (och behörig) avfördes den sökande från arbetskraften och blir

Kenneth Carling är docent i statistik vid Institutet för arbetsmarknadspolitisk utvärdering, **Laura Larsson** är doktorand i nationalekonomi vid Uppsala universitet och Institutet för arbetsmarknadspolitisk utvärdering.

därmed obehörig till programmen (personlig kommunikation Lechner M 2000).

(iv) Behrenz uppfattar problemet som att det skulle vara allmängiltigt för alla vetenskaper och typer av experiment, eftersom jämförelsegruppen aldrig skall få ta del av åtgärden. Det är felaktigt uppfattat (se repliken till Harkmans kommentarer).

(v) Behrenz menar att ansatsen med "blockering" och "specialfall" kan innebära lika stora problem som nuvarande observationsstudier. Blockering (läs socialt experiment) är i praktiken svårt och i vissa fall omöjligt. Heckman m fl (2000) presenterar en uttömmande lista över fallgroparna. Men fallgropar kan utvärderaren undvika genom en noggrant genomtänkt design, problemet med observationsstudierna kan han däremot inte undvika.

(vi) Behrenz menar att beslut måste fattas, och att därför ska utvärderingar genomföras närhelst underlaget för framtida beslut och ställningstagande kan förbättras. Självfallet instämmer vi. Vi vill dock (magsurt) påpeka att det omvända också bör gälla, nämligen att utvärderingar inte ska genomföras när underlaget för framtida beslut och ställningstagande inte kan förmodas bli bättre. Vi avslutar med en fråga till Behrenz: Vilken eller vilka svenska utvärderingsstudier av programmens individeffekter har inneburit att vi idag bättre än för femtio år sedan kan avgöra om arbetsmarknadsprogram är att föredra framför öppen arbetslöshet?

Anders Harkman

Harkman påstår att möjligheten att vid ett senare tillfälle delta i programmet inte medför något principiellt problem till att utvärdera programmets behandlingseffekt. Dessvärre är han något otydlig i sin argumentation. Vårt intryck är att hans försvar

till att vi faktiskt kan identifiera program-effekten bygger på en rad implicita antaganden om bland annat när effekten går att avläsa, hur länge den varar och hur ofta individerna erbjuds möjlighet att delta. Vi har därför valt att bemöta hans kommentarer genom att systematiskt gå igenom de antaganden som enligt vår tolkning ligger till grund för hans argumentation.

Låt oss börja med att tydliggöra med hjälp av *figur 1* vad vi menar med program-effekten såsom definierad av den klassiska frågan. Antag att utfallet mäts i termer av individens produktivitet vilket i sin tur speglas i inkomster. För enkelhetens skull antas att programmet börjar och slutar vid tidpunkt t så att effekten inträder omedelbart. Den heldragna linjen visar då inkomstutvecklingen över tiden om han deltar och den streckade motsvarande utveckling om han inte deltar. Effekten av programmet är ytan mellan de två linjerna. Empiriskt bestäms effekten genom att jämföra en deltagares inkomstutveckling med en *jämförbar* icke-deltagares.

Vidare antas:

- *endast ett tillfälle att delta*

Icke-deltagaren kommer aldrig att delta i programmet (eller påverkas av det), dennes utfall efter programmet sammanfaller med det före. Därmed representerar icke-deltagare efter tidpunkt t en värld utan program.

- *en konstant eller långvarig programeffekt och en begränsad uppföljningsperiod*

Ytan mellan den hela och den streckade linjen anger individens intäkter av att delta i programmet. Detta är av betydelse för en samhällsekonomisk nyttokalkyl där kostnaden för programmet vägs mot den totala intäkten. I fallet som beskrivs i *figur 1* skulle det vara nödvändigt att följa de två perso-

nera för all framtid, vilket i praktiken är omöjligt. Säg att personerna följs fram till T . Antingen får man anta att effekten upphör därefter, eller extrapolera den observerade effekten som uppmäts mellan t och T . Vid ett experiment väljer man blockerings-tiden $T-t$ med hänsyn taget till hur länge man tror effekten varar. I studier med retrospektiva data är det ofta tillgången på data som styr uppföljningsperioden.

• *ingen annonseringseffekt*

Fram till programmet beter sig deltagaren och icke-deltagaren som om det inte fanns något program, de har inga förväntningar som förändrar deras beteende före programstart. Detta illustreras i figuren med att utfallet före program för de två sammanfaller med icke-deltagarens utfall efter program som i sin tur representerar en värld där programmet inte finns.

Är det möjligt att fortfarande identifiera programeffekten om vi lättar på något av dessa antaganden?

(i) *Senare deltagande*

Figur 2 illustrerar fallet då icke-deltagaren har möjlighet att delta vid en senare tidpunkt t_2 . Vi kan identifiera den del av programeffekten som observeras mellan tidpunkterna t_1 och t_2 – givet att uppfölj-

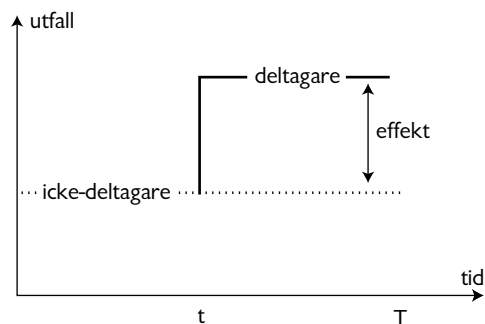
ningsperioden är längre än t_2-t_1 – men vi kan inte vara säkra på att identifiera hela effekten. Detta är i analogi med diskussion under konstant effekt. Notera dock att utvärderaren inte kan påverka t_2-t_1 och i praktiken är intervallet mycket kort (kanske någon månad som mest).

(ii) *Senare deltagande och en kortvarig programeffekt*

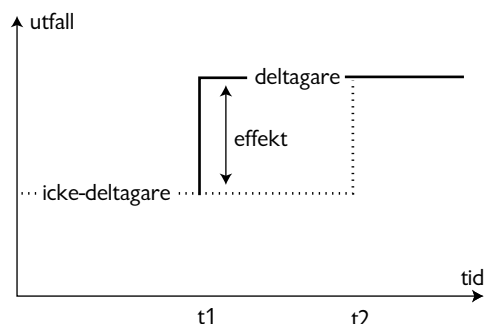
Om programeffekten är kortvarig – programmet ger nya idéer och krafter till jobbsökandet för en viss tid framöver, och effekten mäts i termer av sannolikhet att hitta ett jobb – kan den bestämmas förutsatt att vi hinner observera hela effekten före icke-deltagaren börjar sitt deltagande. Figur 3 illustrera ett sådant fall. Harkman antyder att det faktum att målet med de flesta program är sysselsättning skulle innebära att verkligheten mest liknar fallet i figur 3. Antingen erhåller man ett jobb direkt efter programmet eller så gör man det inte, vilket förstås kan observeras under en mycket kort period.

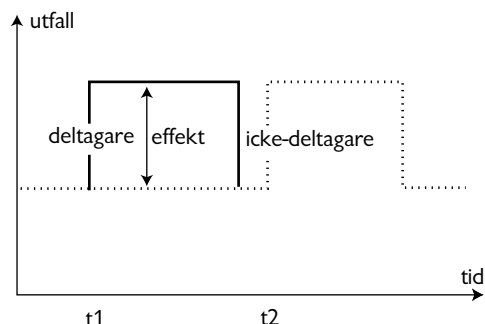
Vi påstår att det är en felaktig förståelse av effekten, som inte heller leder till någon rättvis samhällsekonomisk kalkyl. Den omedelbara effekten kan knappast uppväga programmets kostnad. I stället ska man mäta *hur mycket snabbare* individen erhåller

Figur 1. Den klassiska programeffekten.



Figur 2. Deltagande vid ett senare tillfälle.



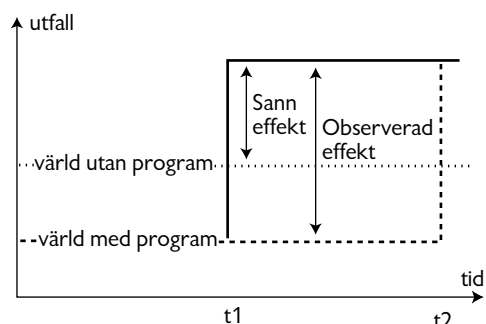
Figur 3. Kortvarig programeffekt.


ett jobb (som förhoppningsvis är varaktigt) tack vare programmet, vilket kräver en betydligt längre uppföljningsperiod.

(iii) Annonseringseffekt

Antar vi att de två personerna är framåtblickande¹ och anpassar sitt beteende till förväntningarna om framtiden kan vi inte längre identifiera programeffekten. *Figur 4* illustrerar fallet där de två har förväntningar om att kunna delta i programmet någon gång i framtiden, och de förmodar att programmet har en positiv effekt. Därför är deltagarens utfall före t_1 , och icke-deltagarens utfall före t_2 lägre än de skulle ha varit i en värld utan program. Den observerade effekten blir därmed större än den sanna effekten oavsett om effekten är bestående eller kortvarig.

Annonseringskostnader ensam räcker inte till att omöjliggöra identifieringen av programeffekten. I ett fall med annonseringskostnader men endast ett tillfälle att delta skulle icke-deltagaren "hoppa upp" på den övre streckade linjen efter t_1 och vi skulle kunna identifiera den sanna effekten.

Figur 4. Annonseringseffekt.


Hur ser verkligheten ut i 1990-talets Sverige? Som i figur 3 verkar Harkman påstå. Vi håller inte med honom, utan tror att figur 4 ger den mest rättvisande bilden av förhållandena. Människorna lever inte i ett vacuum där programmet kommer som en överraskning, utan de blickar framåt och känner till chansen att delta vid ett senare tillfälle. Därmed förblir vår slutsats densamma som tidigare: För att kunna identifiera den grundläggande programeffekten krävs en ny teori.

Referenser

- Heckman J, Lalonde R J & Smith J A (1999) "The Economics and Econometrics of Active Labor Market Programmes" i Ashenfelter O & Card D (red) *Handbook of Labor Economics*. Volume III, Elsevier Science BV, Amsterdam.
- Larsson L (2000) "Evaluation of Swedish Youth Labour Market Programmes" Working Paper 2000:1, IFAU, Uppsala.

1. I den ekonomiska teorin baseras beslutet om deltagande på framtida utfall och därmed tas det som självklart att personerna är framåtblickande.