

Förord

Långtidsutredningen 2003 utarbetas inom Finansdepartementets strukturenhet. I samband med utredningen genomförs ett antal specialstudier, vilka publiceras som bilagor till utredningens huvudbetänkande.

I denna bilaga behandlas ett antal frågeställningar kring inkomstfördelning och den offentliga sektorns omfördelningseffekt. Författarna belyser särskilt hur resultaten förändras om tidsperspektivet utvidgas till att omfatta hela livscykeln i stället för endast ett enskilt år. Förutom tidsperspektivet nyanseras också det traditionella inkomstbegreppet i och med att effekten av offentligt finansierade tjänster läggs till den disponibla inkomsten i analysen.

I appendix till bilagan finns en dokumentation av den dynamiska mikrosimuleringsmodell – SESIM – som har använts i beräkningarna.

Bilagan har utarbetats av Thomas Pettersson vid Finansdepartementets fördelningsanalysenhet och Tomas Pettersson vid Statisticon AB. Ansvar för långtidsutredningens bilagor och de bedömningar dessa innehåller vilar på respektive författare. Av det kommande huvudbetänkandet framgår hur bilagorna använts i utredningens arbete.

Stockholm i november 2003

Yvonne Fredriksson
Departementsråd

Innehåll

1	Inledning.....	9
2	Metod & Data	13
2.1	Alternativa metoder	13
2.2	Tidigare studier	14
2.3	Vad är en mikrosimuleringsmodell?	17
2.4	SESIM	19
2.5	Simuleringsmiljö och scenarier	20
2.6	Datakällor	22
	2.6.1 LINDA.....	22
	2.6.2 HEK/HINK	23
2.7	Subventioner av offentliga tjänster	24
2.8	Indirekt beskattning	25
3	Inkomstfördelningsanalys – års- och livsinkomster	27
3.1	Definitioner och begrepp	28
3.2	Resultat och analys.....	34
4	Inkomstkomponenters omfördelande egenskaper i tvärsnittet och över livet.....	39
4.1	Metod.....	39
4.2	Resultat	43

4.2.1	Progressiva inkomstslag	47
4.2.2	Regressiva inkomstslag	50
4.2.3	Inkomstslag som byter tecken	51
4.2.4	Ginibidrag.....	52
5	Omfördelning mellan individer och över livscykeln	55
5.1	Bakgrund.....	55
5.2	Metod.....	57
5.3	Resultat	62
5.3.1	Beräknade omfördelningskomponenter	62
5.3.2	Skatter, transfereringar och subventioner över livscykeln	66
5.3.3	Omfördelning mellan olika grupper	69
5.3.4	Omfördelning genom ålderspensioner och offentliga subventioner	74
6	Sammanfattning och diskussion.....	81
6.1	Inledning.....	81
6.2	Inkomstfördelningen – effekter av offentlig konsumtion och undersökningsperiodens längd.....	82
6.3	Omfördelningseffekter	83
6.4	Den offentliga sektorns omfördelande och periodiserande roll.....	84
6.5	Slutsatser/diskussion.....	87
	Appendix A: SESIM – en dynamisk mikrosimuleringsmodell ...	89
A.1	Inledning.....	89
A.2	Vad är en mikrosimuleringsmodell?.....	90
A.2.1	Statisk mikrosimulering.....	90
A.2.2	Dynamisk mikrosimulering.....	92
A.3	SESIM:s modellstruktur	94
A.4	Datakällor	99
A.4.1	LINDA – en longitudinell databas	99

A.4.2	Korrigerig av hushållens sammansättning	101
A.4.3	Tillägg av utlandsboende med svensk pensionsrätt.....	105
A.4.4	Imputering av subventioner för offentlig konsumtion	106
A.5	SESIM – en stokastisk simuleringsmodell.....	113
A.6	Simulering av inkomster	118
A.7	Teknisk plattform	122
A.8	Förteckning över stokastiska modeller i SESIM.....	124
Appendix B: Validering av simulerade inkomster		141
B.1	Status, stockar och flöden	141
B.2	Inkomst	143
B.2.1	Tvärsnittsfördelning	144
B.2.2	Livsinkomstfördelning	150
B.2.3	Inkomströrlighet	152
Begrepp och definitioner		161
Referenser		165

1 Inledning

Den offentliga sektorn omfördelar årligen stora summor mellan olika individer och hushåll. Omfördelningen sker via ett stort antal skatter, transfereringar och bidrag samt genom offentligt finansierad privat konsumtion. De olika systemen har vitt skilda motiv varav många i första hand inte alls syftar till omfördelning av ekonomiska resurser. För t.ex. socialbidrag, bostadsbidrag och förmögenhetsskatt är det uppenbart att den direkta inkomstomfördelningseffekten är ett huvudsyfte, andra system såsom sjukpenning och arbetsmarknadsstöd omfördelar mer än något annat *risk*. Socialförsäkringarna omfattar de flesta som arbetar men endast de som drabbas (blir sjuka, arbetslösa etc.) erhåller någon ersättning. Kostnaden att drabbas är stor om försäkring saknas, på samma sätt som i en vanlig försäkring delar alla försäkrade på risken. Vidare har stora delar av den offentliga konsumtionen helt andra syften än att direkt utjämna fördelningen av ekonomiska resurser. Grundskolan utjämnar t.ex. i första hand framtida *inkomstmöjligheter* hos eleverna. Pensioner, barnbidrag och studiemedel är exempel på system vars främsta syfte är att omfördela *över livscykeln*. Samtliga system medför dock i större eller mindre utsträckning omfördelning i ett tvärsnittsperspektiv, det vill säga inom ramen för ett enstaka år, och analyseras därför ofta i ett sådant perspektiv. För många av systemen är det dock lika relevant att analysera dem över längre tidsperioder.

Denna rapports syfte är att, utifrån beräkningar gjorda med den dynamiska mikrosimuleringsmodellen SESIM, belysa följande frågeställningar.

- Hur påverkas inkomstfördelningen om inkomstbegreppet utvidgas till att omfatta offentligt finansierad privat konsumtion och av att mätperioden förlängs?

- Vilka direkta omfördelningseffekter har olika skatter, transfereringar och offentlig konsumtion, och påverkas dessa omfördelningseffekter av mätperiodens längd?
- Hur stor del av den omfördelning som sker via den offentliga sektorn medför en faktisk omfördelning av resurser mellan individer och hur stor del kan betraktas som rundgång eller periodisering?

Inkomstfördelningen

Hur jämnt eller ojämnt levnadsstandarden är fördelad är en fråga som ofta diskuteras såväl inom forskningen som i den allmänna debatten. I en överväldigande majoritet av de studier av denna fördelning som produceras används inkomsten som indikator på levnadsstandard. Med inkomst avses då oftast disponibel årsinkomst justerad för försörjningsbörda. Att likställa årsinkomst och levnadsstandard är inte oproblematiskt.

För det första går det att på goda grunder argumentera för att inkomsten under ett år inte säger så mycket om levnadsstandarden och att längre mätperioder rimligen bättre fångar den materiella levnadsnivå som man egentligen vill studera. Tidigare studier har visat att fördelningen för period- eller livsinkomster är jämnare än den för ett givet år. Hur stor skillnaden är påverkas framför allt av mätperiodens längd och vilka inkomstkomponenter som betraktas.

För det andra täcks inte alla nyttigheter som hushållen kan tillgodogöra sig in av den disponibla inkomsten. Hushållen erhåller offentligt finansierade varor och tjänster i stor omfattning, hemproduktion och svartarbete genererar värden som inte ingår i offentlig statistik.

I rapporten kvantifieras den utjämning som uppkommer dels av att mätperioden förlängs, dels av att värdet av den offentliga konsumtionen adderas till inkomsten. På detta sätt görs ett försök att komma något närmare individernas och hushållens egentliga levnadsstandard. Värdet av hemproduktion och svartarbete hanteras inte i denna rapport, värdet av fritid beaktas inte heller.¹

¹ I bilaga 13 till denna långtidsutredning behandlas bland annat värdet av fritid och hemproduktion.

Omfördelningseffekter av skatter, transfereringar och offentlig konsumtion

En skatt eller en transferering kan antingen bidra till ojämnheten i inkomstfördelningen eller motverka densamma. Genom att, via *dekomponering*, dela upp den totala inkomstojämnheten med avseende på inkomstslag studeras de omfördelande egenskaperna hos olika subventioner, skatter och transfereringar. Ett inkomstslag som verkar starkt utjämnande på inkomstfördelningen ett givet år kan ha andra (t.o.m. motsatta) egenskaper då motsvarande dekomponering i stället görs på livsinkomsten. Här jämförs inkomstslagets fördelningsegenskaper i ett livsinkomstperspektiv med motsvarade tvärsnittsegenskaper.

Omfördelning, periodisering och rundgång

Den offentliga sektorn kan tänkas omfördela ekonomiska resurser antingen från en individ till en annan eller från en individ till samma individ. Den senare typen av omfördelning kan ske antingen under ett och samma år, t.ex. genom skattepliktiga transfereringar, eller genom en periodisering av individens inkomster över livet.

För varje individ kan ett årligt netto mot den offentliga sektorn beräknas. Nettot kan vara positivt vissa år och negativt andra år. På så sätt kan en del av de transfereringar och de offentligt finansierade varor och tjänster som hushållen mottar sägas vara *självfinansierade*. Med självfinansierad menas att de har finansierats av individens egna skatter och avgifter vid någon tidpunkt i livet. I rapporten kvantifieras hur stor denna del är i förhållande till den totala summan av omfördelade medel över en kohorts² livscykel.

I kapitel 2 presenteras den metod som använts för att generera data över livsinkomster och mikrosimuleringsmodellen SESIM samt de datamaterial som ligger till grund för beräkningarna och utifrån vilka SESIM konstruerats. I kapitel 3, 4 och 5 presenteras och analyseras resultaten. Rapporten sammanfattas i kapitel 6, och där diskuteras även vilka slutsatser som kan dras och möjliga tolkningar av resultaten.

² En kohort utgörs av alla som har samma födelseår.

2 Metod & Data

I detta kapitel beskrivs de metoder och datamaterial som används i rapporten. Beskrivningen är allmänt hållen, för den som är intresserad finns en mer detaljerad och teknisk beskrivning i appendix A.

2.1 Alternativa metoder

För att studera inkomstfördelningen över livsrykeln behövs ett datamaterial som beskriver inkomstförloppet hos ett antal individer över hela livsrykeln, eller åtminstone stora delar av den. De metoder som kan användas för att skapa sådana inkomstprofiler kan grovt delas upp i *historiska* och *hypotetiska* metoder.³

Den historiska metoden innebär att en sekvens av faktiska inkomster för ett antal individer utgör grunden för analysen. Den största fördelen med detta angreppssätt är att det handlar om faktiska historiska inkomster och alla observerade samband är sanna. En nackdel är att det krävs långa tidsserier eller *paneler* av hög kvalitet där inkomsterna definitionsmässigt är någorlunda jämförbara över tiden. Ett annat problem med den historiska metoden är att endast fördelningen hos avlidna personer, alternativt ofullständiga livsinkomster, kan studeras. Med ofullständig livsinkomst menas här att endast en begränsad del av de ingående individernas hela livsinkomster kan studeras, livsförloppen kapas vid en viss ålder eller åldrar. Den historiska metoden innebär också att en analys av regelsystemens egenskaper kommer att beskriva historiska, varierande, regelsystem. Om analysen huvudsakligen berör års- kontra livsinkomstfördelning kan historiska metoder vara ändamålsenliga men regel- och orsaksanalyser blir desto svårare.

³ Blomquist, N.S. [1976].

Den hypotetiska metoden utgår vanligen från ett tvärsnitt av individer och använder någon typ av funktion eller modell för att simulera framtida inkomster. Med denna metod undviks problemen med ofullständiga livsinkomster och i en regelanlys kan valfritt regelverk analyseras på ett konsistent sätt. Nackdelen med simulerade data är att livsförloppen blir konstruerade och samband över tiden och mellan olika variabler överensstämmer inte nödvändigtvis med de sanna sambanden. Validering av modellens egenskaper är därför av avgörande betydelse för analysens trovärdighet.

I denna studie används den hypotetiska metoden, mer specifikt baseras analyserna på resultat simulerade med den dynamiska mikrosimuleringsmodellen SESIM. Dynamisk mikrosimulering, och SESIM-modellen, beskrivs konceptuellt senare i detta kapitel.

2.2 Tidigare studier

Ett flertal studier av livsinkomstfördelning och omfördelnings-effekter av skatter och transfereringar över livscykeln har publicerats. Som en introduktion till denna litteratur presenteras här kortfattat några av dessa och den använda metodiken kommenteras kort. Översikten är på intet sätt heltäckande utan speglar endast ett urval av de artiklar som ligger till grund för vårt val av metod. De första studierna är internationella och kan, åtminstone delvis, sägas vara de som utvecklat de metoder för datagenerering och analys som vi senare kommer att använda. Därefter följer några tidigare svenska studier som vi kommer att referera löpande i våra analyser och med vilka vi kommer att jämföra våra resultat.

I flera tidigare studier har enstaka system, såsom pensions-system, analyserats. Bland de första som på ett systematiskt sätt försökte att konstruera hela livsinkomster och analysera fler skatter och transfereringar i ett livscykelerspektiv kan nämnas Falkingham och Lessof (1992) samt Harding (1993). I Falkingham och Harding (1996) kombinerades resultaten från dessa modeller. Utgående från två *dynamiska kohortmodeller* med delvis gemensam modellstruktur jämfördes omfördelningsegenskaperna i Australien och Storbritannien 1986 respektive 1985. En kohort bestående av 2 000 män och 2 000 kvinnor fick leva hela sina liv i den ekonomiska miljö och under de regler som gällde de aktuella åren. Ett antal transfereringssystemens omfördelningsegenskaper analyserades

i både års- och livsinkomstperspektiv. I artikeln presenteras även en metod för att kvantifiera hur stor del av omfördelningen som faktiskt flyttar resurser mellan individer (interpersonell omfördelning) och hur stor del som är periodisering eller rundgång från och till samma individ (intrapersonell omfördelning). Denna metod presenteras närmare och används på svenska data i kapitel 5.

Ett annat exempel är Nelissen (1998) där en dynamisk mikrosimuleringsmodell (NEDYMAS) används för att analysera omfördelningsegenskaperna hos ett stort antal transfereringssystem. Nelissen utgår från ett tvärsnitt från 1947 som sedan får åldras. Livsinkomsterna för kohorterna 1930 och 1950 jämförs sedan med tvärsnittsfördelningen 1991. Till skillnad från Falkingham och Harding som utförde sina analyser i en "steady-state" miljö är ett av Nelissens uttalade syften att se hur den demografiska och ekonomiska miljön påverkar utfallet. Fördelen med Falkingham och Hardings ansats är att effekterna av regelsystemen isoleras, nackdelen är att den inte säger så mycket om situationen för den nuvarande befolkningen. Nelissens ansats tar däremot hänsyn till såväl den demografiska som de ekonomiska utvecklingen och effekten av t.ex. en åldrande befolkning modelleras på ett trovärdigt sätt. Nackdelen är att regelsystemen förändrats mycket under den analyserade perioden och det blir därför svårt att uttala sig om de nuvarande systemens faktiska egenskaper.

Ett senare exempel på en dynamisk modell som omfattar de flesta transfereringssystemen är en avhandling av O'Donoghue (2001b) där bland annat självfinansieringsgraden i de irländska transfereringssystemen undersöks. Avhandlingen innehåller en omfattande dokumentation av den använda modellen och ingående analyser av allehanda egenskaper hos de irländska skatte- och transfereringssystemen, både simulerade livsinkomstdata och faktiska tvärsnittsdata används.

Ett av de tidigaste försöken att simulera livsinkomster i Sverige återfinns i Blomqvist (1976). Efter en inledande diskussion av för och nackdelar med olika metodologiska ansatser genereras livstidsinkomster med hjälp av en "funktion". Funktionen, som i praktiken är en enkel mikrosimuleringsmodell, skapar en livsinkomst som består av löner och kapitalinkomster (förräntning av förmögenhet) efter skatt samt arv och gåvor. Studien är begränsad till anställda män som var 25 år 1970. Kvinnor, företagare, arbetslösa, studerande och pensionärer modelleras ej.

Till långtidsutredningen 1992 studerade Björklund sambandet mellan års- och livsinkomster. I stället för att, som Blomqvist, simulera inkomstbanor utgick Björklund från faktiska data, levnadsnivåundersökningen (LNU) för åren 1968, 1974, 1981 och 1991 kompletterade med registerbaserade inkomstuppgifter fr.o.m. 1951. I undersökningen ingår kohorterna 1936 till 1942 och undersökningsperioden sträcker sig från 1951 till 1989. Med *livsinkomster* avses här alltså 38 årsinkomster. Analysvariabel var *sammanräknad nettoinkomst* för *individer*. Skillnaderna mellan fördelningen av livs- och årsinkomster studeras genom att kohorternas livsinkomstfördelning jämförs med tvärsnittsfördelningen för de som i tvärsnittet har samma ålder som kohorten hade under perioden. Kohort 1936 var t.ex. 16–53 år under perioden, deras livsinkomstfördelning jämförs alltså med tvärsnittsfördelningen hos denna åldersgrupp. Kohort 1973 jämförs med 17–54 åringar osv. Björklund studerade även korrelationen mellan års- och livsinkomst och fann att korrelationen var låg eller obefintlig i åldrarna 25–30, för att därefter bli förhållandevis hög. En slutsats var att årsinkomsten är en god indikator på livsinkomsten för de som etablerat sig på arbetsmarknaden men en dålig indikator för yngre.

Den enda tidigare svenska studie som specifikt studerar intra-kontra interpersonell omfördelning över livsrytmen på ett systematiskt sätt är Hussénius och Selén (1994).⁴ Denna rapport kan sägas ha varit utgångspunkten för arbetet med föreliggande rapport. För att generera livsinkomster används successiva två-årspaneler som genom statistisk matchning sammanfogas till hela livsbanor. En person som var 18 år gammal år ett i tvåårspanelen kommer år två att vara 19 år, i grunddata är alltså personens inkomst vid 18 och 19 års ålder kända. För att generera en inkomst vid 20 års ålder lokaliserar en person som har liknande karaktäristika och som år ett i panelen var 19 år. Dennes inkomst år två, vid 20 års ålder, länkas ihop med vår ursprungsindivid varefter man fortsätter med en individ som var 20 år under panelens första år osv. På detta sätt kan hela livsbanor skapas. I studien kvantifierades hur stor del av all omfördelning inom en kohort som är genuin omfördelning mellan individer och hur stor del som är rundgång, dvs. hur stor del av de transfereringar som betalas ut är självfinansierade. I rapporten uppdateras dessa resultat med en mer

⁴ Ds 1994:135.

avancerad inkomstgenereringsmetodik (dynamisk mikrosimulering) och ett utvidgat inkomstbegrepp som omfattar i stort sett samtliga skatter, transfereringar samt viss offentlig konsumtion.

2.3 Vad är en mikrosimuleringsmodell?

Termen mikrosimulering används för att beskriva en typ av beräkningsmodeller i vilka beräkningarna utgår från den minsta beståndsdel i en population. Den typ av modeller som berörs här utgår ifrån den enskilda individen och är uteslutande konstruerade för att användas i analyser av samhällsekonomisk karaktär, till exempel att studera olika aspekter av skatte- och transfereringssystem. Man brukar vanligen göra skillnad mellan de två huvudtyperna statiska och dynamiska mikrosimuleringsmodeller.

Statisk mikrosimulering

Den statiska mikrosimuleringsmodellen utgår vanligen från en databas innehållande detaljerad information om individernas olika egenskaper såsom till exempel hushållets sammansättning, utbildningsnivåer och olika typer av inkomstuppgifter.

I modellen finns detaljerade beskrivningar av de aktuella skatte- och transfereringssystemen vilket möjliggör att individernas skatter, transfereringar och disponibla inkomster kan beräknas.

Det primära syftet med den statiska mikrosimuleringsmodellen är att utvärdera fördelnings- eller budgeteffekter av olika reformförslag. Det kan röra sig om ändringar av de befintliga skatte- och transfereringssystemen eller införandet av nya komponenter i dessa (eller helt nya system).

På grund av eftersläpningen i insamlingen av deklara-tionsuppgifter är de mest aktuella databaserna minst två år gamla. För att kunna göra beräkningar som återspeglar en aktuell omvärld med avseende på till exempel befolkningens sammansättning, inkomster och sysselsättningsnivåer behöver prognoser två eller fler år in i framtiden göras. Dessa prognoser görs genom att tillämpa så kallat *statiskt åldrande* på modellbefolkningen. Med statistiskt åldrande av befolkningen avses tekniker som förändrar modellbefolkningens sammansättning på ett önskat sätt utan att förändra egenskaperna hos de ingående individerna. Detta åstadkoms genom att i stället

förändra individernas urvalsvikter. På så sätt kan förändringar i demografisk sammansättning, arbetskraftsdeltagande och andra nyckeltal kontrolleras.

En konsekvens av att individernas egenskaper inte förändras i simuleringen är att endast de statiska, eller direkta, effekterna av en systemförändring kan beräknas. Eventuella beteendeförändringar som uppstår till följd av en reform fångas inte upp. Till exempel kommer en nivåförändring av inkomstbeskattningen inte ha någon effekt på arbetsutbudet.

Ett exempel på en statisk mikrosimuleringsmodell är statistiska centralbyråns (SCB) fördelningsmodell FASIT⁵ som i över ett decennium har använts inom regeringskansliet och under de senaste åren även fått en spridning utanför statsförvaltningen.

Dynamisk mikrosimulering

Den primära skillnaden mellan den statiska och dynamiska modellen är det sätt på vilken modellpopulationen framskrivs. Till skillnad från den statiska modellen så förändrar den dynamiska modellen hela tiden egenskaperna hos den ingående modellpopulationen, ett *dynamiskt åldrande*. På så sätt söker modellen avbilda det sätt på vilket en faktisk population åldras - individer blir äldre, föder barn, flyttar hemifrån, börjar arbeta och så vidare.

Sådana händelser antas inträffa med sannolikheter som i sin tur är beroende av individernas egenskaper. Sambandet mellan individernas egenskaper och deras sannolikheter för olika händelser skattas med hjälp av statistiska modeller. De skattade sannolikheterna kan sedan användas för att slumpa ut olika händelser bland individerna i modellpopulationen. Statistiska modeller används även för att simulera värden för andra typer av variabler, till exempel inkomster.

I de flesta dynamiska simuleringsmodeller uppdateras modellpopulationens egenskaper stegvis i jämna tidsintervall, oftast ett år. Inom tidsintervallet antas sedan händelser kunna inträffa i en viss bestämd sekvens. Det innebär att en händelse som placeras mitt i sekvensen kan vara beroende av händelser som placeras tidigare i sekvensen men inte av händelser som placeras senare – en så kallad rekursiv struktur.

⁵ FördelningsAnalytisk Statistiksystem för Inkomster och Transfereringar.

I vissa modeller används en teknik där en individs hela liv simuleras innan simuleringen för nästa individ påbörjas, det vill säga en *individvis* simulering. Det vanligaste är dock att använda en *periodvis* simulering, där hela modellpopulationen genomlöps en gång per tidsintervall.

Det primära syftet med de dynamiska modellerna är ofta att göra beräkningar på mycket lång sikt för att analysera system där tidsaspekten är avgörande. Ett viktigt exempel är analyser av ålderspensionssystem för vilka det krävs att beräkningar av individens årliga avsättning till systemet kan göras likväl som beräkningar av hur det totala pensionskapitalet växer över tiden fram till pensioneringen. För sådana systemanalyser har de dynamiska modellerna en fördel gentemot de statiska eftersom åldrandeprocessen mer efterliknar den i verkliga livet.

För en mer detaljerad diskussion kring dynamisk mikrosimulering samt genomgång av ett antal befintliga modeller se Rake och Zaidi (2001) samt O'Donoghue (2001a).

2.4 SESIM

Samtliga beräkningsresultat som presenteras i rapporten baseras på den dynamiska mikrosimuleringsmodellen SESIM. Modellens basbefolkning utgörs av ett stickprov draget 1999 ur den svenska befolkningen med tillägg av ett motsvarande stickprov av utlandsboende individer med svensk pensionsrätt. Datakällan är en longitudinell databas vilket har gjort det möjligt att komplettera informationen i stickprovet med historiska uppgifter för urvalsindividerna.

Modellpopulationens egenskaper uppdateras årligen i simuleringarna, det vill säga en *periodvis simulering*. Detta innebär att samtliga modellindividers egenskaper uppdateras under en beräkningsperiod innan nästa beräkningsperiod påbörjas. Uppdateringen sker genom simulering av ett antal statistiska modeller samt genom tillämpning av olika algoritmer. De senare används till exempel för att beskriva regelverk för skatte- och transfereringssystemen.

Modellen är stokastisk och data från de ingående statistiska modellerna genereras genom Monte Carlo-simulering.

SESIM har en rekursiv struktur med ett antal moduler som genomlöps i tur och ordning:

- Demografi
- Utbildning
- Arbetsmarknad
- Förmögenhet & bostadsmarknad
- Skatter & transfereringar
- Offentligt subventionerade tjänster

Dessa moduler innehåller i sin tur ett antal beräkningsprocesser. Modellens minsta beståndsdel är individen men på samma sätt som i den verkliga befolkningen är individerna sammanlänkade i hushåll. Hushållet är en viktig komponent i SESIM då många processer refererar till hushållets egenskaper likväl som individens. Framför allt gäller detta regler för beräkning av olika skatter och bidrag men även olika statistiska modeller.

För en mer detaljerad beskrivning av SESIM hänvisas till appendix A. I appendix B valideras de inkomster som genereras av SESIM. De huvudsakliga slutsatserna är att SESIM genererar en inkomstfördelning som i stort överensstämmer med de som observeras i verkligheten. Detta gäller såväl för enstaka år som över längre perioder. Modellen ger dock en något för låg andel med låg ekonomisk standard. Inkomströrligheten i SESIM är på kort sikt (6–7 år) något för hög och på längre sikt något för låg. Den sammantagna bedömningen är att de inkomster som genereras, i för analyserna i denna rapport relevanta avseenden, rimligt väl speglar hur inkomster genereras i verkligheten.

2.5 Simuleringsmiljö och scenarier

Alla simuleringar som sker i en mikrosimuleringsmodell måste ske i en ekonomisk miljö. Genom att kalibrera⁶ olika förlopp kan variabler göras *exogena*, alternativt får modellutfallet bestämma utvecklingen och den aktuella variabeln blir då *endogen*. Syftet med denna rapport är i första hand att kartlägga och kvantifiera de omfördelningseffekter som offentliga ingrepp i hushållssektorns ekonomi ger upphov till. För att isolera de nuvarande regelsystemens effekter hålls därför den övriga miljön så konstant som möjligt. Observera att detta är ett avsteg från övriga delar av långtidsutredningen.

⁶ Modellutfallet tvingas att sammanfalla med ett exogent givet antagande.

I beräkningarna uppdateras regler för skatter och transfereringar fram t.o.m. 2003 och hålls därefter konstanta. Fr.o.m. 2003 antas ingen tillväxt eller inflation men den demografiska utvecklingen tillåts få fullt genomslag.⁷ På detta sätt kan vi analysera de nuvarande regelsystemen i 2003 års ekonomiska miljö men med en befolkningsutveckling som med stor sannolikhet ligger framför oss. Motivet till att avvika från Steady-state-miljön i början av perioden är att det medför valideringsmöjligheter, genom att försöka efterbilda den faktiska utvecklingen 1999–2003 kan modellens prediktionsförmåga utvärderas.

I analyser av livsinkomster, eller inkomster över en längre period, är det vanligt att inkomstströmmarna diskonteras så att inkomster genererade senare i livet viktas ned. Motiven för en sådan diskontering kan vara att ta hänsyn till inflation, tillväxt och riskaversion. Eftersom vi varken har inflation eller tillväxt (i meningen ökande reallöner) i våra simuleringar faller dessa två motiv. Det tredje motivet, riskaversion, innebär att en säker inkomst i dag antas vara mer värd än en osäker inkomst i framtiden. Den aktuella inkomsten värderas alltså vid en given tidpunkt högre än en framtida lika stor inkomst. Skillnaden i värdering beror på individens riskaversion. I de simuleringar som genomförs i denna rapport utvärderas livsinkomsten ”i efterhand” när alla i populationen ingående individer redan har dött. Även riskaversionsmotivet faller alltså här och de genererade inkomsterna diskonteras därför inte.

Antaganden om utebliven tillväxt och inflation inverkar även på den antagna utvecklingen av olika offentliga system, till exempel skatte- och transfereringssystemen. Belopp knutna till basbeloppet förblir oförändrade efter 2003 på grund av inflationsantagandet. Övriga belopp, vilka förändras genom politiska beslut (till exempel barnbidrag och bostadsbidrag), antas i beräkningarna även de vara oförändrade.

Oförändrade nivåer antas även för de offentliga subventioner som beaktas i analyserna. Befolkningens totala subventioner kan dock förändras över tiden beroende på att dess sammansättning med avseende på de faktorer som bestämmer konsumtionsbenägenhet och konsumtionsvolymmer eventuellt förändras (till exempel genom en generellt ökande medelålder).

⁷ Befolkningen utvecklas enligt SCB:s befolkningsprognos BE SM 0201.

Premiepensionssystemet antas ge en viss avkastning under simuleringsperioden (2 procent årligen för samtliga individer). Vidare har normen för följsamhetsindexering av utgående ålderspensioner satts till noll. Detta sänker pensionerna något i början av utbetalningsperioden⁸ genom att delningstalet blir högre men förhindrar i stället att utbetalningarna minskar över tiden, något som annars blir följderna av en simulering utan tillväxt. Pensionsnivåerna konstanthålls alltså på samma sätt som övriga transfereringar. De genomsnittliga utbetalningarna över livet förändras inte av dessa antaganden.

2.6 Datakällor

2.6.1 LINDA

LINDA-databasen utgör den primära datakällan för SESIM, beträffande såväl estimering av statistiska modeller som konstruktion av modellens baspopulation.

Databasen består av ett urval om ca 3,5 procent av den svenska befolkningen vilket för år 1999 motsvarar ca 308 000 individer. Till dessa urvalsindivider har eventuella hushållsmedlemmar adderats. Totalt omfattar urvalet därför omkring 786 000 individer år 1999. De individer som ingår i urvalet följs upp på en årlig basis varvid data samlas in och infogas i databasen. Även information för förfluten tid samlas in. Databasen har alltså en så kallad longitudinell struktur där de ingående individerna kan observeras vid upprepade tillfällen. Urvalsindivider som av någon anledning försvinner från databasen ett år, t.ex. genom dödsfall eller emigration, ersätts med nya individer på ett sätt som säkerställer databasens representativitet gentemot den svenska befolkningen det givna året.

SESIM:s baspopulation utgörs av ett stickprov om ca 104 000 individer draget ur 1999 års LINDA-databas. Till detta har ett urval om ca 8 000 individer draget från Riksförsäkringsverkets pensionspoängsregister adderats. Tilläggsurvalet avser individer som är bosatta utomlands men har svensk pensionsrätt (se appendix A).

De olika statistiska modeller som utgör grunden för SESIM har i de flesta fall estimerats med hjälp av LINDA-data. De flesta modeller beskriver dynamiska processer som bygger på information från en sammanhängande period av två eller fler år. För att

⁸ I beräkningarna antas samtliga individer gå i pension det år de fyller 65 år.

kunna estimeras sådana modeller krävs följaktligen tillgång till longitudinella data.

I appendix A ges en utförligare beskrivning av databasen och dess viktigaste egenskaper.

Korrigeringar av LINDA-data

I konstruktionen av SESIM:s baspopulation ingår två huvudsakliga bearbetningar, eller justeringar, av LINDA-databasen. Syftet med dessa bearbetningar är att förbättra hushållsbegreppet så att det bättre stämmer överens med verkliga förhållanden. Korrigeringarna avser individer i åldrarna 18–29 år. Dels korrigeras andelen i dessa åldrar som bor kvar hos sina föräldrar, dels andelen som är gifta/samboende. I appendix A beskrivs justeringarna i detalj.

2.6.2 HEK/HINK

SCB genomför årligen en undersökning av inkomstfördelningen, undersökningen om hushållens ekonomi, HEK.⁹ Tidigare, t.o.m. inkomståret 1998, kallades undersökningen för inkomstfördelningsundersökningen, HINK

Undersökningen som omfattar ca 30 000 individer består av såväl register- som enkätinformation. Tack vare enkätinformationen är bl.a. hushållsbegreppet av högre kvalitet än i t.ex. LINDA.

HEK-data har använts för estimering av några samband i SESIM och resultaten i rapporten kommer i några fall att jämföras med studier baserade på HEK/HINK. Framst berör detta situationer där information rörande hushållens sammansättning är viktig. I HEK kan det så kallade kosthushållsbegreppet¹⁰, vilket ligger nära SESIM:s hushållsbegrepp, användas.

⁹ För en utförlig beskrivning se SCB [2003].

¹⁰ Ett kosthushåll omfattar individer som är bosatta i samma bostad samt har en gemensam ekonomi och kosthållning.

2.7 Subventioner av offentliga tjänster

I rapportens analyser kompletteras inkomstbegreppet genom att hushållens inkomster ökas med värdet av viss konsumtion av offentliga tjänster.

Från tidigare studier av den offentliga konsumtionens effekter på fördelningen av hushållens ekonomiska standard finns information om utnyttjande av offentligt finansierade tjänster samt nettosubventionen för dessa i datamaterialet HEK för åren 1997 och 1999¹¹. Nettosubventionen beräknas genom att produktionskostnaden för respektive tjänst minskas med eventuella avgifter.

De tjänster som inkluderas i analysen är sådana tjänster som kan relateras till en specifik individ. Kostnaden för de personliga tjänsterna uppgick 1999 till ca 330 miljarder kronor att jämföra med en total produktionskostnad för offentliga tjänster på omkring 536 miljarder kronor.

Nedan anges de offentliga tjänster som inkluderas i analysen. För aggregerade poster anges de ingående komponenterna inom parentes.

- skola (förskoleklass, grundskola)
- gymnasium
- komvux
- vuxenskola (särvtux, statens skola för vuxna, svenska för invandrare, folkbildning/folkhögskola, kvalificerad yrkesutbildning)
- universitet / högskola (inklusive forskarutbildning)
- barnomsorg
- äldreomsorg (hemtjänst, särskilt boende)
- arbetsmarknadsåtgärder
- vård (primärvård, slutenvård, tandvård)
- läkemedel

I LINDA finns ingen information om utnyttjande av offentliga tjänster eller relaterade kostnader. I stället imputeras dessa data från 1999 års HEK. Imputeringen baseras på observerade samband mellan å ena sidan de variabler som anger bruk av offentliga tjänster och storleken på subventionerna för dessa tjänster och å andra sidan ett urval av övriga variabler som finns i båda databaserna.

¹¹ Regeringens proposition [1999/00:1], bilaga 4 samt regeringens proposition [2001/02:100], bilaga 3.

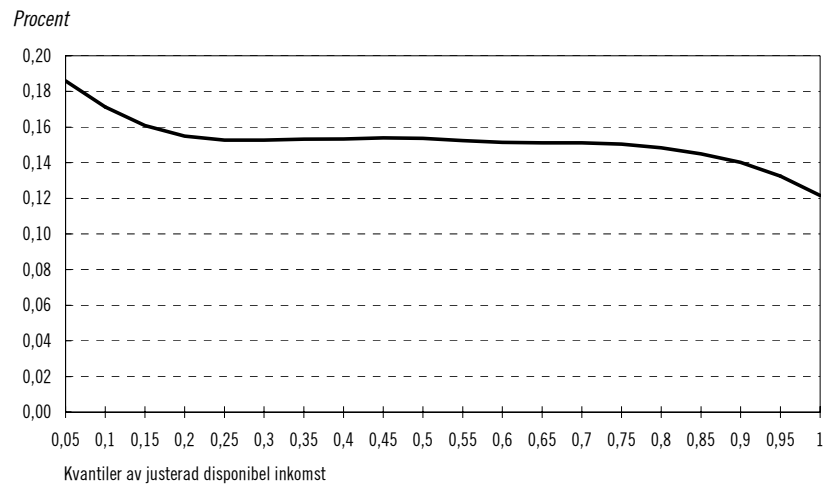
Dessa samband specificeras med hjälp av statistiska modeller. Dessa modeller och dess egenskaper beskrivs i appendix A.

I vissa fall, t.ex. vid sjukvård eller äldreomsorg, är det inte rimligt att anta att hushållens nytta av en subventionerad tjänst är lika stor som subventionens belopp. På grund av detta har i stället en ”försäkringsansats” använts i analyserna för dessa komponenter. Subventionsvärdet uppskattas då som kostnaden av premien för en hypotetisk försäkring, som mottagaren hade behövt teckna om inte tjänsten varit subventionerad, i stället för värdet av den faktiskt utgående subventionen. Beloppen beräknas som den genomsnittligt imputerade nettosubventionen per år, kön och ålder¹².

2.8 Indirekt beskattning

De indirekta skatter hushållen betalar kan inte observeras direkt i LINDA utan beräknas som en viss andel av hushållets disponibla inkomst. Andelen är i sig beroende av hushållets ekonomiska ställning. För de hushåll med starkast (svagast) ekonomisk ställning utgör de indirekta skatterna relativt sett en mindre (större) del av den disponibla inkomsten. Andelarna har skattats genom att använda data från SCB:s undersökning av hushållens konsumtionsvanor (HBS) vilka har imputerats till SCB:s undersökning av hushållens ekonomi (HEK). Diagram 2.1 visar hur stor del av hushållets disponibla inkomst de indirekta skatterna utgör per kvantil av hushållets (justerade) disponibla inkomst.

¹² Denna ansats har använts i tidigare studier, se regeringens proposition [1999/00:1], bilaga 4, regeringens proposition [2001/02:100], bilaga 3 och Smeeding, T., m. fl. [1993].

Diagram 2.1 Indirekta skatter som andel av hushållens disponibla inkomst

Källa: HEK samt egna beräkningar.

3 Inkomstfördelningsanalys – års- och livsinkomster

Analys av hur de disponibla inkomsterna är fördelade bland hushållen under ett år är vanliga såväl i den akademiska litteraturen som i mer policyorienterade sammanhang. Inte sällan används då den disponibla årsinkomsten som en approximation för nytta. Genom att förlänga mätperioden och utvidga inkomstbegreppet kan en fördelning som ligger närmare den nyttofördelning som i dessa sammanhang ofta är den mest intressanta åstadkommas och analyseras. I detta kapitel analyseras hur inkomstfördelningen påverkas av att mätperioden förlängs från ett kalenderår till att omfatta hela livscykelns samt av att inkomstbegreppet kompletteras med värdet av subventionerad offentlig konsumtion. Års- och livsinkomstfördelningen av olika inkomstbegrepp analyseras och jämförs.

När fördelningen av inkomster eller ekonomisk standard skall analyseras måste ett antal val avseende definitioner och metod göras. Det är på intet sätt invändningsfritt hur inkomster skall mätas eller vilka mått och definitioner som bör användas för att bäst spegla hushållens faktiska nytta av inkomsten. Normalt försöker man skapa ett begrepp som på ett så realistiskt sätt som möjligt skall spegla individens eller hushållets konsumtionsutrymme. Kapitlet inleds därför (avsnitt 3.1) med en översiktlig beskrivning av de metoder och definitioner som används i kapitel 3, 4 och, i viss mån, 5.

3.1 Definitioner och begrepp

Inkomstbegrepp

För att komma så nära hushållets faktiska konsumtionsmöjligheter som möjligt används regelmässigt *disponibel inkomst*. Denna definieras som hushållets totala inkomst från arbete, kapital och transfereringar minus betald (direkt) skatt och återbetalt studiemedel.¹³ Erhållet studiemedel betraktas som transferering och ingår i disponibel inkomst. Analyserna i denna rapport utgår också från den disponibla inkomsten men denna har kompletterats med värdet av individuellt konsumerade offentligt finansierade varor och tjänster, i fortsättningen kallad *offentlig konsumtion*. I kapitel 2 redovisas vilka varor och tjänster som ingår. Från den disponibla inkomsten dras här även betald indirekt skatt i form av moms och punktskatter.

Inkomstenhet

Målet är att mäta individernas nytta eller inkomst och hur stora skillnader det är mellan olika individer i detta avseende men det inses lätt att individen då inte kan betraktas som oberoende av övriga hushållsmedlemmar. En persons ekonomiska standard är beroende även av inkomsterna hos andra i familjen eller hushållet. I de flesta hushåll torde en viss gemensam ekonomi förekomma, åtminstone delar nog de flesta på de hushållsgemensamma utgifterna. Kännedomen om hur inkomstfördelningen ser ut inom hushållen är liten, det standardmässiga antagandet är här att alla hushållsmedlemmar lever på samma materiella standard. Detta är ett starkt antagande som förmodligen i många fall är felaktigt men enligt de flesta bedömare är detta förhållningsätt det minst missvisande. Inkomsterna summeras därför till hushållsnivå.

Stordriftsfördelar

För att kunna jämföra konsumtionsmöjligheter i familjer med olika sammansättning måste hushållets inkomst justeras det antal personer som skall leva på dessa inkomster. Förekomsten av

¹³ Normalt dras även betald underhållsbidrag och återbetalt socialbidrag från den disponibla inkomsten. Dessa inkomstslag modelleras inte i SESIM.

stordriftsfördelar gör att det inte ger realistiska och jämförbara inkomstnivåer om hushållets inkomst endast divideras med antalet hushållsmedlemmar. En bostad för en familj om fem personer kostar normalt inte fem gånger så mycket som en bostad av motsvarande standard för en person. I stora hushåll delar man ofta på vissa varor såsom TV, telefon, dagstidning etc.

För att beräkna jämförbara inkomster används en så kallad *ekvivalensskala* som ger varje hushåll en viss "konsumtionsvikt" beroende på hur många vuxna och barn som lever tillsammans. För att erhålla jämförbara inkomster divideras hushållets inkomst med dess konsumtionsvikt. I litteraturen förekommer en mängd olika skalor, en del baserade på vad hushållen faktiskt konsumerar, andra på expertbedömningar av vad olika hushåll behöver äta m.m. Den ekvivalensskala som används i Sveriges officiella statistik ger, jämfört med andra länder, stor tyngd till barn och tar liten hänsyn till stordriftsfördelar. Skalan tar inte heller hänsyn till hur boendekostnader varierar med hushållsstorlek. Vid internationella jämförelser och i den akademiska världen används ofta den s.k. kvadratrotsskalan, konsumtionsvikten sätts helt enkelt till kvadratroten ur antalet hushållsmedlemmar. Den vanligaste kritiken mot denna skala är att den överdriver stordriftsfördelarna i stora familjer samt att den inte skiljer på konsumtionsvikt mellan barn och vuxna. Ingen av dessa skalor framstår som optimal för denna rapportens analyser, i stället används en skala kallad PEL-skalan (se faktaruta 3.1).

Faktaruta 3.1 PEL-skalan

Vad som kan anses vara en skälig lägsta standard för familjer med olika sammansättning beslutades av riksdagen 1997 genom införandet av riksnormer för socialbidragen. Den nya normen innebär bl.a. att för familjen gemensamma belopp införs för olika kostnadsposter. Därmed tar normen hänsyn till stordriftsfördelar. Det kan vara rimligt att jämförelser av konsumtionsmöjligheterna i familjer med olika sammansättning utgår från riksnormen. Genom sin komplicerade konstruktion med många olika belopp beroende på ålder, kostnadspost m.m. är det emellertid svårt att direkt översätta denna riksnorm till en överskådlig ekvivalensskala.

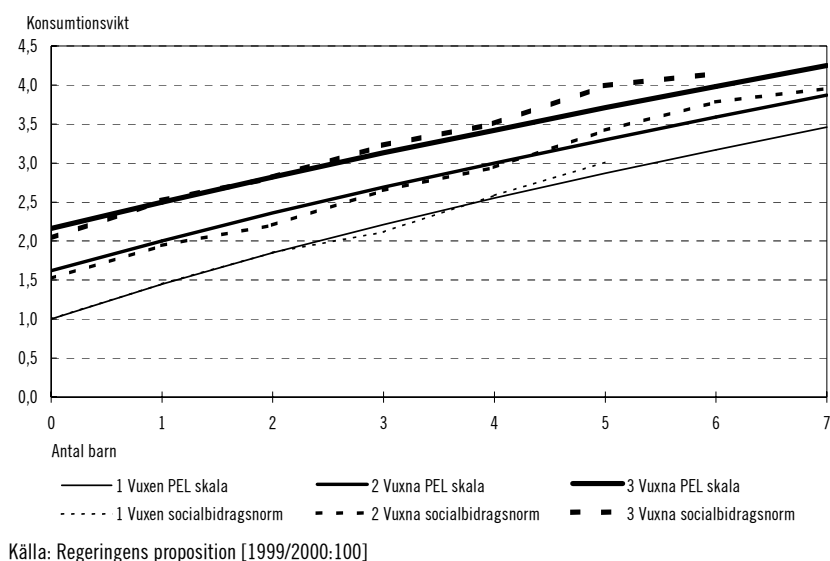
Forts. Faktaruta 3.1

Till vårpropositionen 2000¹⁴ skapades en parametriserad skala utgående från riksnormen. För varje hushåll i HINK97 beräknades ett normbelopp inklusive en schablonmässig boendekostnad och schablonmässiga men normala utgifter utöver normen för kostnader för vård och möbler. Utifrån denna "empiriska" skala anpassades en modell med antal vuxna och antal barn som förklarande variabler. Följande ekvivalensskala avspeglar väl riksnormen.

Hushållets konsumtionsvikt = $(N+0,7 \cdot n)^{0,7}$
 där N = antal vuxna, n = antal barn

Jämförelser mellan den "empiriska" skalan och normbelopp beräknade enligt denna ekvivalensskala, i fortsättningen kallad PEL-skalan, visar överlag en god överensstämmelse för familjer med upp till fyra barn. Familjer med fler barn erhåller en något för låg relativ vikt (diagram 3.1).

Diagram 3.1 Empirisk konsumtionsvikt enligt socialbidragsnormen jämfört med parametriserade värden, beräknade i HINK97



¹⁴ Regeringens proposition [1999/2000:100].

Justerad disponibel inkomst och total inkomst

PEL-skalan används för att justerat den disponibla inkomsten och betald indirekt skatt. Den offentliga konsumtionen räknas däremot per capita. Motivet för detta tillvägagångssätt är att ”värdet” av den offentliga konsumtionen inte påverkas av familjens storlek eller sammansättning och att det därför inte uppkommer några stor-driftsfördelar.¹⁵ Disponibel inkomst per konsumtionsenhet benämns i fortsättningen *justerad disponibel inkomst*. Justerad disponibel inkomst minus indirekt skatt per konsumtionsenhet plus offentlig konsumtion per capita benämns *total inkomst*.

Redovisningsenhet

Inkomsterna räknas samman till hushållsnivå men redovisningen avser, om inget annat anges, individer. Den justerade disponibla inkomsten eller den totala inkomsten påförs samtliga familjemedlemmar.

Livsinkomst

Ovan har beskrivits hur justerad disponibel inkomst respektive total inkomst beräknas i ett tvärsnitt. Livsinkomster, eller inkomster över en längre period, beräknas som individens medelvärde av de årsvisa nivåerna. Medelvärden, i stället för summor, används för att korrigera för variation i livslängd mellan olika individer. Vid livsinkomstanalyser ingår endast individer som levt större delen av sina liv i Sverige, individer med fler än nio år utomlands exkluderas. De beräknade medelvärdena avser endast de år individen bott i Sverige.

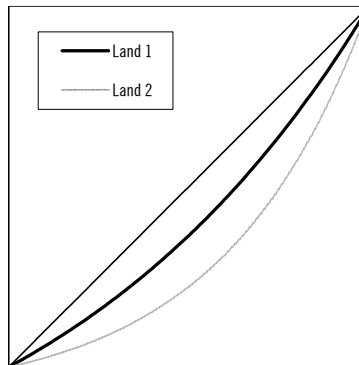
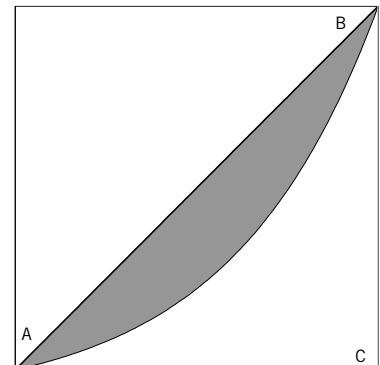
Gini-koefficienten

Gini-koefficienten är det vanligaste måttet på ojämnheten i inkomstfördelningen och antar värdet 0 när inkomsten är lika för alla och värdet 1 vid maximal ojämnhet, dvs. när en person får alla inkomster. Se vidare faktaruta 3.2.

¹⁵ För ett mer ingående resonemang kring detta hänvisas till Smeeding, T. m.fl. [1993].

Faktaruta 3.2 Gini-koefficienten och Lorentzkurvan

Gini-koefficienten definieras enklast utgående från den s.k. Lorenz-kurvan. För att definiera denna så rangordnas först alla individer efter inkomstens storlek så att den med lägst inkomst kommer först och den med högst inkomst kommer sist. Sedan beräknas den kumulativa inkomsten, dvs. i den första positionen placeras den fattigaste individens inkomst, på position 2 summan av de två fattigaste individernas inkomster, på position 3 summan av de tre fattigaste individernas inkomster osv. I den sista positionen placeras den totala summan av alla inkomster. I diagram 4.3 nedan har två serier med kumulativa inkomster plottats för två påhittade länder. De streckade och prickade linjerna är Lorenz-kurvorna för de två påhittade länderna. Om perfekt jämlikhet råder så har alla exakt lika stora inkomster, Lorenz-kurvan sammanfaller då med den räta diagonallinjen. Ju mer kurvan avviker från diagonallinjen desto ojämnare är fördelningen, i exemplet nedan är alltså inkomsterna jämnare fördelade i land 1 än i land 2.

Diagram 3.2 Lorenz-kurvor för två hypotetiska länder**Diagram 3.3 Lorenz-kurvan och Gini-koefficienten**

Forts Faktaruta 3.2

I diagram 3.3 ovan har en Lorenz-kurva ritats. Gini-koefficienten definieras som den skuggade arean i förhållande till arean av triangeln ABC.¹⁶ Vid perfekt jämlikhet är den skuggade arean noll och Gini-koefficienten blir 0. Vid maximal ojämnhet, dvs. när en individ har alla inkomster täcker den skuggade arean hela triangeln och Gini-koefficienten blir 1. För att översätta denna något abstrakta definition till mer konkreta siffror kan Gini-koefficienten sägas visa hur stor förväntad inkomstskillnad det är mellan två slumpmässigt valda individer/hushåll. Om Gini-koefficienten är 0,250 och medelinkomsten för alla är 90 000 kronor skall man vänta sig att den genomsnittliga skillnaden är $2 \cdot 0,250 \cdot 90\,000$, dvs. 45 000 kronor.

Decilkvot

I några analyser kompletteras Gini-koefficienten med ett enklare spridningsmått. Befolkningen delas först in i tio lika stora grupper, decilgrupper, där de 10 procent med lägst inkomst hamnar i decilgrupp 1, de med näst lägst standard i decilgrupp 2 osv. ända upp till decilgrupp 10 som består av de 10 procent av befolkningen som har högst inkomst. Decilkvoten definieras som inkomsten som avgränsar dem med högre ekonomisk standard dividerad med inkomsten som avgränsar dem med lägre standard. Olika decilkvoter kan på detta sätt konstrueras genom att välja olika decilgränser. I denna rapport används endast den kvot som erhålls om den åttonde decilgränsen divideras med den andra. Detta mått anger avståndet mellan den inkomstgräns som avgränsar de tjugo procenten med lägst inkomst och den gräns som avgränsar de tjugo procent som har högst inkomst.

¹⁶ Den formella definitionen är: $Gini = \left(\frac{1}{2n^2 \bar{y}} \right) \sum_{i=1}^n \sum_{j=1}^n |y_i - y_j|$ där

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ och y_i betecknar den i :te individens inkomst.

3.2 Resultat och analys

I appendix B jämförs de inkomster som genereras av SESIM med motsvarande inkomster i LINDA och HEK. När inkomstfördelningen i sin helhet, uttryckt som Gini-koefficienten för justerad disponibel inkomst, studeras är resultatet tillfredsställande. SESIM genererar Gini-koefficienter för årsinkomster som väl överensstämmer med de som kan observeras i verkligheten. Under detta så döljer sig dock två problem som motverkar varandra.

För det första underskattas andelen med mycket låga inkomster i SESIM, detta leder till att inkomstojämnheten underskattas. Det andra problemet är att även andelen med något högre, men fortfarande låga, inkomster överskattas. Detta leder till en över-skattning av ojämnheten, sammantaget tar dessa problem i stort ut varandra och ett sammanfattande fördelningsmått hamnar på ”rätt” nivå.

Den offentliga konsumtionens effekt på inkomstfördelningen

För analyser av årsinkomster läggs data för perioden 2000 till 2109 samman till en datamängd.¹⁷ Gini-koefficienten för justerad disponibel inkomst är 0,217 i tvärsnittet. Motsvarande för total inkomst är 0,189. Att gå från justerad disponibel inkomst till total inkomst minskar alltså inkomstojämnheten med ca 10 procent. Detta är betydligt lägre än vad som tidigare rapporterats, i flera studier har effekten av att addera offentlig konsumtion till justerad disponibel inkomst kvantifierats till ca 20 procent.¹⁸ En förklaring är att analysen här inkluderar indirekt beskattning, denna verkar regressivt och tar ut en del av utjämnings effekten av offentlig konsumtion. Om analysen delas upp i två steg, där den indirekta skatten dras först, ger den offentliga konsumtionen en utjämning på 15 procent.

¹⁷ Eftersom alla inkomster är uttryckta i samma valuta kan tvärsnittet poolas på detta sätt.

¹⁸ Regeringens proposition [2001/02:100], bilaga 3, regeringens proposition [1999/2000:1], bilaga 4 samt [Ds 1994:86].

Livsinkomstfördelningen

Nästa steg i analysen är att gå från års- till livsinkomster. Gini-koefficienten för genomsnittet över livet för total inkomst är 0,086.¹⁹ Ojämnheten minskar alltså med knappt 60 procent när livs- i stället för årsinkomsten analyseras. Detta resultat ligger väl i linje med Hussénius, och Selén (1994)²⁰ som erhöll en drygt 60 procentig minskning av Gini-koefficienten för justerad disponibel inkomst när perspektivet förflyttades från års- till livsinkomster.

Björklund (1992) beräknade, utgående från faktiska data, utjämningseffekten av att gå från års- till livsinkomster till ca 40 procent. En förklaring till den lägre effekten kan vara att Björklund undersökte sammanräknad nettoinkomst på individnivå, ett inkomstbegrepp som ligger ganska långt från både justerad disponibel inkomst och total inkomst, undersökningen var även begränsad tidsmässigt och inga hela livsinkomster kunde studeras.

I tabell 3.1 nedan sammanfattas hur inkomstfördelningen påverkas när mätperioden förlängs.

Tabell 3.1 Gini-koefficienten (och decilkvot) för års- och livsinkomster för olika inkomstbegrepp

Inkomstslag	Årsinkomst	Livsinkomst	Förändring
Faktorinkomst	0,490 (15,6)	0,196(1,73)	- 60%
Justerad disponibel inkomst	0,217 (1,90)	0,102 (1,35)	-53%
- indirekt skatt	0,224 (1,95)	0,104 (1,36)	-54%
+ offentlig konsumtion (total inkomst)	0,189 (1,75)	0,086 (1,28)	-55%

Anm.: Decilkvoterna (inom parentes) anger hur mycket större inkomsten är vid den åttonde decilgränsen jämfört med vid den andra decilgränsen för respektive inkomstbegrepp.

Källa: SESIM

Att fördelningen av livsinkomster är jämnare än motsvarande fördelning av årsinkomsterna beror på att individernas position i inkomstfördelningen varierar mellan åren. Hur mycket jämnare fördelningen blir av att mätperioden förlängs beror på graden av inkomströrlighet. Ju högre rörlighet desto större blir utjämnningen. Om livsinkomsterna analyseras med ett sammanfattande fördelningsmått (Gini-koefficienten) så är dess ojämnheter mindre än

¹⁹ Samtliga undersökta kohorter (födda år 2000–2010) läggs samman till en datamängd.

²⁰ Ds 1994:135.

hälften av den ojämnhets som återfinns när tvärsnittsinkomsterna analyseras. Ett sätt att tolka detta är att den förväntade inkomstskillnaden mellan två slumpmässigt utvalda individer halveras.

Av appendix B framgår att den långsiktiga inkomströrligheten i SESIM förmodligen är mindre än den verkliga inkomströrligheten vilket här skulle betyda att utjämnings-effekten på drygt 50 procent sannolikt är en underskattning.

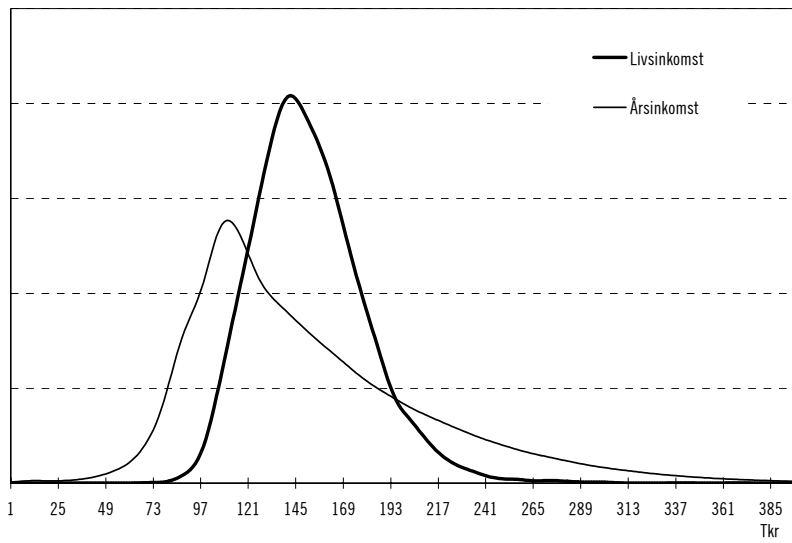
Sammantagen effekt

Den årsvisa fördelningen av faktorinkomster, som här utgörs av arbets- och kapitalinkomster, är av naturliga skäl mest ojämnt fördelad. Faktorinkomsterna, som är de inkomster som genereras på marknaden, är vid den fjärde kvartilgränsen (eller åttonde decilgränsen) mer än 15 gånger så hög som vid den första kvartilgränsen. När hänsyn tas till den utjämnings-effekt som sker via den offentliga sektorn och som uppkommer av att mätperioden förlängs minskar avståndet till knappt 30 procent, *mer än 90 procent av den ursprungliga skillnaden försvinner*. Gini-koefficienten minskar med drygt 80 procent.

Var i fördelningen sker utjämnings-effekten?

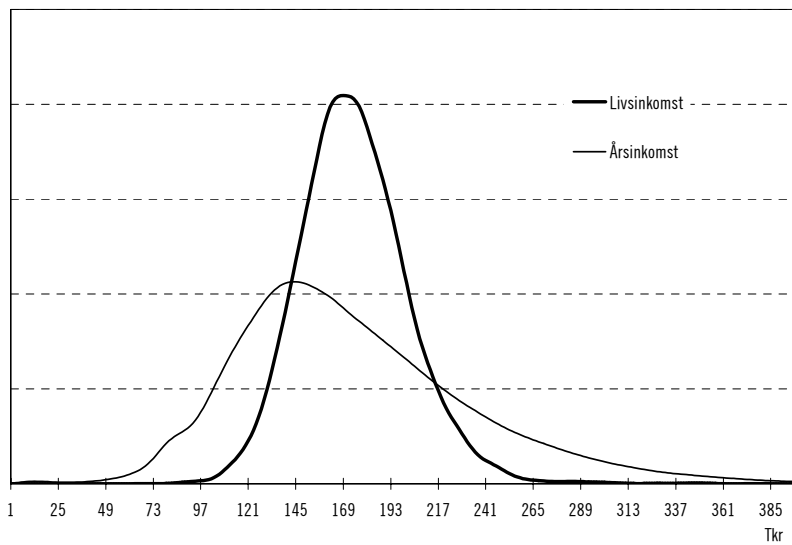
En utjämnings-effekt kan uppkomma på flera sätt, fördelningen kan pressas samman underifrån genom att de som har låga inkomster ett år kompenserar detta med högre inkomster under andra år. På samma sätt kan sammanpressningen komma uppifrån om de som har högst inkomst i tvärsnittet har mer måttliga inkomster under andra år. I diagram 3.4 och 3.5 illustreras hur fördelningen förändras när mätperioden förlängs från ett kalenderår till hela liv. Fördelningen av justerad disponibel inkomst i tvärsnittet och kohorterna visas i det första diagrammet, i det andra visas motsvarande fördelningar av total inkomst.

Diagram 3.4 Fördelning av justerad disponibel inkomst



Källa: SESIM

Diagram 3.5 Fördelning av total inkomst



Källa: SESIM

När mätperioden ändras från års- till livsinkomst förändras fördelningens utseende på ungefär samma sett oavsett om justerad disponibel inkomst eller total inkomst analyseras. Fördelningen pressas samman både underifrån och från toppen. Bilderna speglar helt enkelt den rörlighet som förekommer i fördelningen, de som har låg inkomst under ett år kompenserar ofta det med högre inkomster under andra år och omvänt har de med höga inkomster under ett år ofta lägre inkomster andra år.

I de lägre inkomstnivåerna trycks fördelningen framåt (åt höger i diagrammen), detta är tydligast när disponibel inkomst studeras. Fördelningens utseende förändras inte så mycket utan de lägsta inkomsterna blir helt enkelt lite högre när medelvärden över livet används i stället för utfallet ett enstaka år. I toppen av fördelningen är förändringen något annorlunda, kurvan pressas nedåt (mot x-axeln) och den högra delen av kurvan blir brantare. Sett över livet är förekomsten av mycket höga inkomster alltså betydligt lägre än under ett år. I stället samlas höginkomsttagarna ihop på något lägre nivåer.

4 Inkomstkomponenters omfördelningsegenskaper i tvärsnittet och över livet

För att förstå hur inkomstfördelningen är uppbyggd och vilka möjligheter det finns att med politiska medel påverka densamma krävs insikt om hur fördelningen påverkas av olika inkomstslag. I detta syfte ”dekomponeras” i detta kapitel såväl tvärsnitts- som livsinkomstfördelningen med avseende på inkomstslag. Syftet med detta är att beskriva hur olika skatter och bidrag omfördelar ekonomiska resurser mellan hushåll, samt att avgöra hur denna omfördelning påverkas om mätperioden varieras. Även effekterna av viss offentligt finansierad privat konsumtion (t.ex. sjukvård, utbildning och olika typer av omsorg) samt indirekt beskattning redovisas.

Skatter, bidrag och offentlig konsumtion påverkar inkomstfördelningen både direkt och indirekt. Fördelningen av faktorinkomster (löner, företagariinkomster och kapitalinkomster) påverkas indirekt dels via individernas val av arbetstid, dels genom pris- och lönepåverkan. Om skatte- och bidragssystemen inte fanns eller var annorlunda utformade skulle sannolikt individernas val av arbetstid, lönestrukturen och avkastningskraven på kapital m.m. ha sett annorlunda ut. På lång sikt påverkar t.ex. subventionerad utbildning fördelningen av inkomstmöjligheter och således även inkomsterna. Ambitionsnivån här sträcker sig inte längre än till att analysera de direkta effekterna.

4.1 Metod

Den gängse metoden för att beskriva skatternas och bidragens direkta fördelningseffekter är att dekomponera inkomstspridningen. Den totala inkomstspridningen delas då upp med avseende på de inkomstslag som bygger upp den totala inkomsten. Ett alternativ till dekomponering är att beräkna spridningen stegvis. Först beräknas ojämnheten hos ett inkomstslag, t.ex. marknads-

inkomst, därefter läggs inkomstslagen till ett i taget. Efter varje ny komponent beräknas ojämnheten på nytt och jämförs med det tidigare resultatet. På detta sätt kan man avgöra om ojämnheten ökar eller minskar när inkomstslaget adderas till inkomsten. Effekterna vid en sådan analys kommer dock att bero på i vilken ordning komponenterna adderas vilket undviks med dekomponeringsmetoden. Den stegvisa metodiken kan sägas beskriva ett inkomstslags fördelning i förhållande till fördelningen av närmast föregående aggregat medan dekomponeringsmetoden beskriver samma inkomstslags fördelning i förhållande till fördelningen av total inkomst.

I dekomponeringsanalysen tas ingen hänsyn till vad de inbetalda skatterna går till eller hur transfereringarna och subventionerna finansierats. Varje komponents effekt analyseras isolerat. En alternativ ansats är att kombinera effekten av en transferering med dess finansiering, effekten av transfereringen blir då i regel mer progressiv. Detta beror på att finansieringen i regel är progressiv. I kapitel 5 analyseras ålderspensionerna och den offentliga konsumtionen med hänsyn tagen till finansieringen.

Dekomponering

Det finns flera tänkbara alternativa dekomponeringsmetoder att tillgå, den som används här presenterades först av Kakwani (1977). I analysen delas den totala inkomstojämnheten upp i ett antal klart avgränsande delkomponenter (inkomstslag). Varje delkomponent är i sin tur en produkt av två faktorer, dels en vikt som anger komponentens relativa storlek (eller betydelse) och dels ett koncentrationsindex.²¹

Koncentrationsindexet, vilket kan anta värden mellan -1 och 1, är ett mått på hur ett inkomstslag är fördelat i förhållande till den totala inkomsten. Ett negativt koncentrationsindex betyder att det aktuella inkomstslaget motverkar ojämnheten i inkomstfördelningen, dvs. inkomsterna från inkomstslaget är koncentrerade till hushåll med relativt låga totala inkomster. Ett positivt värde betyder att inkomstslaget bidrar till ojämnheten i fördelningen av total inkomst. Vikten är inkomstens genomsnittliga andel av den

²¹ Gini-koefficienten kan skrivas som: $Gini = \sum_{i=1}^n v_i * k_i$ där v_i och k_i står för vikt respektive koncentrationsindex för inkomstkomponent i .

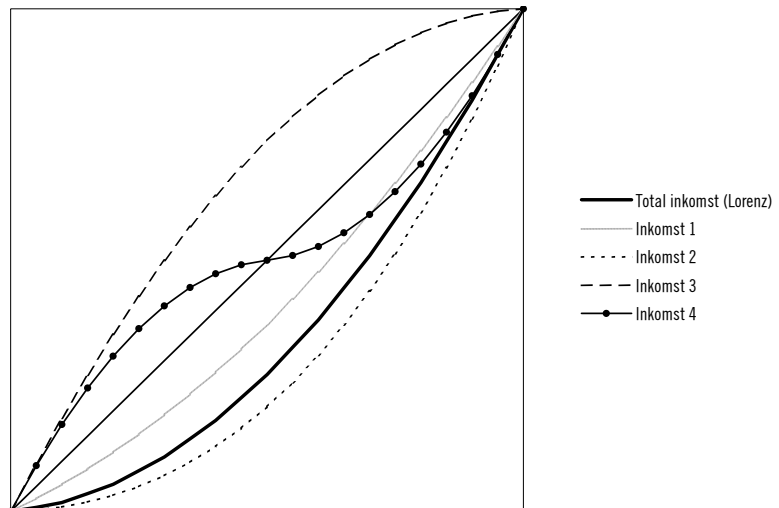
totala inkomsten. Skatternas vikt antar negativa värden eftersom de utgörs av transfereringar från hushållen. Vid tolkning av skatternas bidrag till inkomstspridningen utifrån koncentrationsindexet blir därför värdet omvänt.

Koncentrationsindexet skall hellre tolkas som inkomstslaget fördelning givet omfattning och fördelningen av övriga inkomster. Om ett inkomstslag vid låg omfattning är utjämnande kan det trots detta bidra till ökad ojämnheter om omfattningen ökar. Om man t.ex. tänker sig ett bidrag som endast går till den fattigaste medborgaren (när detta bidrag inte räknas in i inkomsten) så skulle detta verka utjämnande vid låga nivåer. Om detta bidrag sedan höjs successivt kommer den totala inkomsten (inkl. bidraget) till slut att bli så hög att mottagaren av bidraget flyttas upp till den övre delen av inkomstfördelningen. Bidraget kan då övergå till att i stället bidra till ojämnheten i inkomstfördelningen. Ett positivt koncentrationsindex behöver alltså inte betyda att fördelningen skulle vara jämnare om bidraget inte fanns. Resonemanget kan utvidgas till de flesta inkomstslag. Vid någon nivå kommer en transferering som till en början har minskat ojämnheten i inkomstfördelningen att få omvänd effekt om omfattningen ökar.

Ett positivt index behöver inte heller med nödvändighet innebära att en marginell ökning av bidraget leder till ökad ojämnheter. Om omfattningen av ett inkomstslag som har en jämnare fördelning än den totala inkomsten (koncentrationsindex som är mindre än Gini) ökas, kommer den totala ojämnheten att minska trots att inkomstslaget koncentreringsindex är positivt. Dess bidrag till den totala ojämnheten kommer dessutom, genom en högre vikt, att öka. När koncentrationsindex skall tolkas är dess relation till Gini-koefficienten därför mer intressant än dess tecken.

Faktaruta 4.1 Koncentrationsindex och koncentrationskurvor

Den formella definitionen av koncentrationsindexet utgår från respektive inkomstslags *koncentrationskurvor*. Dessa är nära besläktade med Lorenzkurvan (se faktaruta 3.1). Koncentrationskurvan är den kumulativa summan av ett inkomstslag när individerna sorterats efter storleken på total inkomst. Koncentrationsindexet definieras, i likhet med Gini-koefficienten, som arean mellan koncentrationskurvan och diagonallinjen i förhållande till triangeln under diagonallinjen. Eftersom individerna sorterats efter total inkomst och inte efter respektive inkomstslag behöver en koncentrationskurva inte nödvändigtvis ligga under diagonallinjen, ett inkomstslag vars koncentrationskurva ligger ovanför diagonallinjen har ett negativt koncentrationsindex. I diagram 4.1 nedan visas koncentrationskurvor för ett antal påhitade inkomstslag.

Diagram 4.1 Koncentrationskurvor för ett antal hypotetiska inkomstslag

Forts Faktaruta 4.1

För total inkomst är Lorenz- och koncentrationskurvan per definition samma sak.

I diagrammet representerar inkomst 1 ett inkomstslag som är mer jämnt fördelad än den totala inkomsten, koncentrationskurvan är mindre konvex än Lorenz-kurvan och dess koncentrationsindex är mindre än Gini-koefficienten.

Inkomst 1 har ett positivt koncentrationsindex och bidrar därför till den totala ojämnheten men fördelningen är jämnare än fördelningen av total inkomst. Om inkomst 1 ökas proportionellt kommer dess koncentrationskurva och Lorenzkurvan att närma sig varandra, detta får till följd att koncentrationsindex för inkomst 1 ökar och Gini-koefficienten minskar.

Koncentrationskurvan för inkomst 2 är mer konvex än Lorenz-kurvan och dess koncentrationsindex således större än Gini-koefficienten. Ett inkomstslag med en sådan fördelning är lättare att tolka, det bidrar till ojämnheten och en ökning skulle medföra en mer ojämn fördelning av total inkomst.

Koncentrationskurvan för inkomst 3 ligger helt ovanför diagonallinjen, inkomstslaget är alltså utjämnande och dess koncentrationsindex är negativt.

Koncentrationskurvan för inkomst 4 korsar diagonallinjen, koncentrationsindexet definieras då som arean under linjen minus arean ovanför i förhållande till triangeln nedanför diagonalen.

4.2 Resultat

I tabellerna 4.1 och 4.2 nedan har total inkomst i tvärsnitt respektive över livet dekomponerats med avseende på inkomstslag. Koncentrationsindex multiplicerat med vikt ger inkomstslagets respektive bidrag till Gini-koefficienten. Vikterna för samtliga inkomstslag summerar till 1 och ginibidragen summerar till Gini-koefficienten.²²

²² I dekomponeringsanalysen kan inte värdet på ett inkomstslag ha varierande tecken mellan olika individer. Om en individ har ett negativt värde på ett inkomstslag som normalt är positivt sätts värdet till noll. Detta får till följd att Gini-koefficienterna i tabell 4.1 och 4.2 inte är identiska med de som redovisades i tabell 3.1. Förfarandet kallas "bottenkodning" och tillämpas här endast på faktorinkomst. Förekomsten av negativa faktorinkomster beror på negativa kapitalinkomster.

I tvärsnittsanalysen har åren 2000 till 2109 lagts ihop till en datamängd. I livsinkomstanalysen har kohorterna 2000 till 2010 lagts samman. Tabellerna analyseras inte var för sig utan i ett sammanhang.

Tabell 4.1 Inkomstslagets bidrag till ojämnheten i total årsinkomst

Inkomstslag	Koncentrationsindex	Vikt	Ginibidrag
Faktorinkomst	0.2924	0.8750	0.2559
Avtalspension	0.2343	0.0398	0.0093
Ålderspension	0.0601	0.1622	0.0097
Förtidspension	-0.2309	0.0387	-0.0089
Efterlevandepension	0.3307	0.0013	0.0004
Arbetsmarknadsstöd	-0.0337	0.0283	-0.0010
Sjukpenning	0.0711	0.0325	0.0023
Föräldrapenning	-0.4806	0.0121	-0.0058
Barnbidrag	-0.0960	0.0202	-0.0019
Bostadsbidrag	-0.5072	0.0032	-0.0016
Bostadstillägg för pensionärer	-0.2025	0.0080	-0.0016
Underhåll (mottaget)	-0.1622	0.0060	-0.0010
Studiemedel	-0.1217	0.0105	-0.0013
Socialbidrag	-0.6378	0.0048	-0.0031
Sjukvård	0.1076	0.0640	0.0069
Barnomsorg	-0.1282	0.0212	-0.0027
Äldreomsorg	0.5763	0.0526	0.0303
Gymnasium	0.1364	0.0110	0.0015
Arbetsmarknadsåtgärder	0.0261	0.0073	0.0002
Komvux	-0.3954	0.0003	-0.0001
Läkemedel	0.0652	0.0109	0.0007
Grundskola	0.0263	0.0345	0.0009
Vuxenutbildning	-0.0156	0.0004	0.0000
Universitet	-0.1397	0.0051	-0.0007
Inkomstskatt	0.2575	-0.3316	-0.0854
Övriga direkta skatter	0.3461	-0.0283	-0.0098
Återbetalt studielån	0.1830	-0.0097	-0.0018
Indirekta skatter	0.1178	-0.0804	-0.0095
Total inkomst		1.0000	0.1821

Källa: SESIM

Tabell 4.2 Inkomstslagens bidrag till ojämnheten i total livsinkomst

Inkomstslag	Koncentrationsindex	Vikt	Ginibidrag
Faktorinkomst	0.1306	0.8517	0.1112
Avtalspension	0.3271	0.0442	0.0145
Ålderspension	0.1060	0.1767	0.0187
Förtidspension	-0.3190	0.0415	-0.0132
Efterlevandepension	0.0044	0.0006	0.0000
Arbetsmarknadsstöd	-0.0234	0.0275	-0.0006
Sjukpenning	0.0526	0.0344	0.0018
Föräldrapenning	0.0239	0.0113	0.0003
Barnbidrag	-0.0407	0.0188	-0.0008
Bostadsbidrag	-0.2264	0.0032	-0.0007
Bostadstillägg för pensionärer	-0.3166	0.0097	-0.0031
Underhåll (mottaget)	-0.1226	0.0058	-0.0007
Studiemedel	0.0317	0.0104	0.0003
Socialbidrag	-0.2907	0.0047	-0.0014
Sjukvård	0.0459	0.0683	0.0031
Barnomsorg	-0.0315	0.0200	-0.0006
Äldreomsorg	0.1985	0.0564	0.0112
Gymnasium	-0.0025	0.0104	0.0000
Arbetsmarknadsåtgärder	-0.0503	0.0071	-0.0004
Komvux	-0.1341	0.0003	0.0000
Läkemedel	0.0279	0.0117	0.0003
Grundskola	-0.0323	0.0327	-0.0011
Vuxenutbildning	-0.0038	0.0005	0.0000
Universitet	0.0669	0.0051	0.0003
Inkomstskatt	0.1368	-0.3361	-0.0460
Övriga direkta skatter	0.1454	-0.0230	-0.0034
Återbetalt studielån	0.0727	-0.0104	-0.0008
Indirekta skatter	0.0596	-0.0834	-0.0050
Total inkomst		1.0000	0.0842

Källa: SESIM

Omfördelningsegenskaperna beror på inkomstslagets fördelning i förhållande till fördelningen av total inkomst samt dess vikt. Det första steget i analysen utgår från koncentrationsindex, utifrån vilket det går att avgöra vilka inkomstslag som verkar regressivt respektive progressivt. De inkomstslag vars index är större än Gini-koefficienten betraktas som *regressiva*, det är alltså dessa som drar isär inkomstfördelningen.²³ Inkomstslag med index som är mindre

²³ För skatter, som har negativa vikter, gäller det omvända. En skatt med koncentrationsindex som är *mindre* än Gini är regressiv.

än Gini verkar åt motsatt håll och trycker ihop fördelningen, dessa inkomstslag betraktas som *progressiva*.

I tvärsnittet är det äldreomsorg, efterlevandepension, faktorinkomster, indirekta skatter och avtalspensioner som drar isär fördelningen. När motsvarande analys gjordes på ett faktiskt datamaterial (HINK) var det äldreomsorg, betalt underhåll, faktorinkomst och moms som var regressiva.²⁴ I analysen av HINK-data ingick alla pensioner i en sammanslagen post varför resultaten inte är fullt jämförbara. Betalt underhåll modelleras för närvarande inte i SESIM. I ett livsinkomstperspektiv är avtalspensioner, äldreomsorg, faktorinkomster, indirekt beskattning ålderspension och återbetalt studiemedel regressiva inkomstslag. Regressiviteten i återbetalt studielån är dock svag och inkomstslaget är mer eller mindre neutralt.

De inkomstslag som inte verkar regressivt är i olika utsträckning progressiva och verkar alltså utjämnande. I ett tvärsnittsperspektiv är socialbidrag, bostadsbidrag, föräldrapenning, subventionerna för komvuxstudier, förtidspension som är mest progressiva. Motsvarande i HINK-analysen var socialbidrag, bostadsbidrag, statlig skatt, föräldrapenning och arbetsmarknadsstöd.

De inkomstslag som är mest progressiva i ett livsinkomstperspektiv är förtidspension, bostadstillägg för pensionärer, socialbidrag, bostadsbidrag och subventioner vid komvuxstudier.

För vissa inkomstslag är det lätt att förstå varför de är progressiva eller regressiva, för andra är det mindre uppenbart. I vissa fall kan sättet på vilket inkomstslaget hanteras i SESIM ha betydelse. Nedan grupperas inkomstslagen efter fördelningseffekt. Först kommenteras de progressiva inkomstslagen, därefter de regressiva och sist de som har olika egenskaper i års- respektive livsperspektiv.

²⁴ Regeringens proposition [2000/20001:1], bilaga 4.

4.2.1 Progressiva inkomstslag

Socialförsäkringar

De flesta analyserade inkomstslagen är progressiva såväl under ett år som över livet. De försäkringsmässiga transfereringarna har generellt sådana egenskaper.

Förtidspensionernas progressivitet tilltar när mätperioden förlängs. Under ett år har förtidspensionärer låga totalinkomster och förtidspensionen i sig är ofta förhållandevis låg. Det är därför naturligt att effekten i tvärsnittet är progressiv. Benägenheten att bli förtidspensionerad är högre för lågutbildade/låginkomsttagare och förtidspensionering ger permanent låga inkomster varför den progressiva effekten blir kraftigare över livet.

Arbetslöshetsersättningen är lägre än tidigare lön och låginkomsttagare har dessutom större sannolikhet att bli arbetslösa. Sannolikheten att bli arbetslös ökar dessutom om man varit det tidigare. Progressionen var mycket kraftigare i HINK, detta förklaras av att de arbetslösa i SESIM erhåller subventioner för arbetsmarknadsåtgärder i betydligt större omfattning än vad de gör i HINK-analysen.²⁵ Detta placerar de arbetslösa högre upp i inkomstfördelningen.

Högre sjuklighet för lågutbildade/låginkomsttagare medför att sjukpenningen är progressiv i tvärsnittet. Utjämningseffekten är dock ganska svag, sannolikt motverkas effekten av att ersättningen är knuten till den försäkrade inkomsten. Över livet är utjämnings-effekten något svagare.

Endast kvinnor får föräldrapenning i SESIM och har under det år de är föräldralediga inga andra inkomster (men eventuella subventioner). Detta förklarar den kraftiga progressiviteten i tvärsnittet. Kvinnor har visserligen lägre inkomster över livet än män vilket skulle kunna motivera en progressivitet även över livet. Detta tycks dock i stor utsträckning uppvägas av de subventioner som utges till barnfamiljer över livet.

²⁵ HINK-analysen avser data från 1997 där arbetsmarknadsutbildning för 3,4 mdkr ingår medan imputeringen av offentlig konsumtion i SESIM baseras på data från 1999 där arbetsmarknadsåtgärder för 19,1 mdkr ingår. Se regeringens proposition [1999/2000:1], bilaga 4 respektive proposition [2001/02:100], bilaga 3.

Bidrag

Barnbidraget är progressivt både i tvärsnittet och över livet. Barnfamiljernas försörjningsbörda gör att de pressas nedåt i inkomstfördelningen, detta medför en lägre livs- och tvärsnittsinkomst jämfört med barnlösa. Kvinnans inkomstbortfall under föräldraledigheten (endast föräldrapenning) bidrar också i viss mån till resultatet.

De inkomstprövade bidragen, bostadsbidrag, BTP och socialbidrag är till sin konstruktion inkomstomfördelning och är således tydligt progressiva i ett tvärsnittsperspektiv. Hushåll med mycket låga inkomster tenderar att ha låga inkomster under långa perioder vilket medför att bidragen är progressiva även över livet.

För BTP förstärks effekten över livet på grund av att de låginkomstpensionärer som får BTP även har haft låga inkomster under den aktiva delen av sina liv. De egenskaper som ökar sannolikheten att få socialbidrag är även förknippade med låg inkomst i allmänhet (låg utbildning, utrikes född etc.), effekten håller därför i sig över livet.

Underhåll betalas till ensamstående kvinnor med barn, dessa har ofta låg inkomst i förhållande till sin försörjningsbörda. Dessa hushåll tenderar att ha låga inkomster under långa perioder varför progressiviteten tilltar över livet.

Studiemedel (lån och bidrag) verkar progressivt i tvärsnittet och något progressivt över livet. Unga studenter har oftast inga eller låga andra inkomster och försörjs helt av studiemedel. Nivån på studiemedlen är betydligt lägre än genomsnittliga inkomster varför studenterna hamnar i botten av inkomstfördelningen. Den subventionerade utbildningen minskar den progressiva effekten. De fattiga studenterna blir senare välbetalda inkomstagare med hög livsinkomst och det förväntade är därför att studiestödet är regressivt sett över livet. Gruppen studenter är dock heterogen, den består inte bara av universitetsstudenter med höga förväntade framtida inkomster utan också av vuxenstudenter med låga historiska och framtida inkomster. De som hoppar av studierna och inte tar någon examen får (i SESIM) ingen inkomstökande effekt av sina studier, detta motverkar också den förväntade regressiviteten.

Skatter

Inkomstskatten är progressiv i tvärsnittet såväl som över livet. Utjämningseffekten över livet skiljer sig inte nämnvärt från effekten i ett tvärsnitt. Björklund (1992) kom fram till samma resultat med en stegvis metod²⁶. Övriga direkta skatter består av skatt på kapitalinkomster, förmögenhets- och fastighetsskatt. Dessa inkomst- och tillgångsslag är mer ojämnt fördelade än arbetsinkomster varför skatterna på dem, i ett tvärsnitt, blir mer progressiva än inkomstskatten. Över livet är fördelningen dock mer utjämnad, få föds med förmögenhet och fastighetsförsäljningar genererar högst tillfälliga inkomster. I ett livsinkomstperspektiv är dessa skatter ungefär lika progressiva som inkomstskatten.

Offentlig konsumtion

Alla erhåller samma subvention för grundskoleutbildning som barn, subventionen blir över livet därmed relativt sett värd mer för de som har lägre livsinkomster. I tvärsnittet förklaras progressionen av att barnfamiljerna har större försörjningsbörda (se även barnbidrag). Samma resonemang kan tillämpas på gymnasieutbildningen som är mer eller mindre neutralt i tvärsnittet och lätt progressivt över livet.

Även subventioner till universitets- och högskoleutbildning verkar progressivt i tvärsnittet. Detta är förväntat då mottagarna är studenter som lever på (låga) studiemedel. Även över livet är subventionerna lätt progressiva, detta är något förvånande då universitetsstuderande senare under livet har högre inkomster än övriga. Förklaringen är att de som hoppar av studierna (i SESIM) inte förbättrar sina framtida inkomstmöjligheter (jmf studiemedel).²⁷

Subventioner till Komvux verkar kraftigt progressivt både i tvärsnitt och över livet. Under studieperioden erhålls en låg ersättning som placerar mottagaren långt ner i fördelningen, detta förklarar effekten i tvärsnittet. Effekten över livet beror på att de som är aktuella för komvux (i SESIM) har låg utbildning och därmed förhållandevis låga inkomster, åtminstone före utbild-

²⁶ Se avsnitt 4.1 för en kort beskrivning av stegvis metodik.

²⁷ Av de som någon gång bedriver universitetsstudier tar, i SESIM, ca en fjärdedel inte examen. I verkligheten torde dock även avbrutna universitetsstudier kunna leda till högre framtida inkomster, det gör de ej i modellen.

ningsinsatsen. Övrig subventionerad vuxenutbildning är också progressiv men effekterna är små.

Arbetsmarknadsåtgärder verkar progressivt, något tydligare över livet än i tvärsnittet. De som är aktuella för arbetsmarknadsåtgärder är arbetslösa som lever på arbetsmarknadsstöd (hela året i SESIM) och saknar lön, de återfinns därför i fördelningens nedre del. Egenskaper som ökar sannolikheten för arbetslöshet ger dessutom lägre arbetsinkomster under perioder i arbete, detta gör att progressionen blir starkare över livet (jämför med arbetsmarknadsstöd).

Sjukvårdssubventionerna är, i likhet med äldreomsorgen, ”försäkringsmässigt” utlagd så att alla med samma ålder och kön erhåller en lika stor subvention. Motivet för detta förfarande är att det inte går att likställa ett högt utnyttjande av sjukvård med nytta samt att även de som inte behöver utnyttja sjukvården i någon större utsträckning ändå kan sägas erhålla den nytta det innebär att omfattas av fri sjukvård. En nackdel med denna försäkringsansats är att de inkomstmönster som faktiskt finns i utnyttjandet av sjukvård suddas ut.²⁸ Den svaga progression som ändå uppkommer beror på att kvinnor erhåller större sjukvårdssubventioner än män. Läkemedel läggs ut som genomsnitt per ålder och kön, även här beror progressionen på att kvinnor erhåller högre subventioner än män.

4.2.2 Regressiva inkomstslag

Fyra inkomstslag är regressiva i både års- och livsperspektiv: faktorinkomst, avtalspensioner, äldreomsorg och indirekt beskattning.

Faktorinkomster är de inkomster som genereras på marknaden, i SESIM utgörs de av löner och kapitalinkomster. Att dessa inkomster är ojämnt fördelade är föga förvånande, om marknadens fördelning var jämn finns inget behov av offentliga ingrepp av inkomstutjämningskäl. Regressiviteten är något kraftigare i tvärsnittet. Detta förklaras av mobilitet, en hög inkomst ett visst år behöver inte betyda hög livsinkomst.

Pensionärer med hög avtalspension har också relativt höga offentliga pensioner och därmed totala inkomster. En hög offentlig pension tyder även på en hög livsinkomst vilket ger en ännu högre regressivitet över livet.

²⁸ Se t.ex. regeringens proposition [2001/02:100], bilaga 3.

Äldreomsorgen har den mest regressiva fördelningen av alla inkomstslag i tvärsnittet, detta resultat erhöles också i den tidigare refererade studien baserad på HINK-data. Även över livscykeln verkar dessa subventioner regressivt. Att den subventionerade äldreomsorgen är så kraftigt regressiv beror på att de värden som påförs mottagarna i vissa fall är mycket stora. Äldreomsorgen är här "försäkringsmässigt" utlagd. Detta innebär att alla, oavsett egen konsumtion, erhåller genomsnittliga värden för det egna könet och den egna åldersklassen. Trots denna "utjämning" blir subventionerna höga i de aktuella åldersklasserna. Som exempel kan nämnas att män som är 90 år eller äldre påförs ett nettovärde (efter betalda avgifter) motsvarande ca 180 000 kronor per år. I kombination med (även en låg) pension lyfter detta mottagaren högt upp i fördelningen. Det finns ett positivt samband mellan livsinkomst och livslängd, ju högre inkomst desto längre förväntad livslängd. Även detta bidrar till regressiviteten.

Indirekta skatter tas ut på konsumtion. Den andel av disponibel inkomst som konsumeras minskar med ökande inkomst. Att indirekt beskattning verkar regressivt är därför förväntat.

4.2.3 Inkomstslag som byter tecken

De flesta inkomstslag har samma omfördelningseffekter över livet som under ett år. Effekternas styrka påverkas ibland av tidsperspektivet men ytterst få inkomstslag byter tecken.

Ett undantag är efterlevandepension som enligt tabellerna är regressivt i tvärsnittet men progressivt över livet. Detta är dock i huvudsak en effekt av den modellerade perioden. I tvärsnittsanalysen ingår alla individer som återfinns i modellpopulationen under de aktuella åren (2000–2109). Bland dessa finns änkor med rätt till änkepension och denna verkar regressivt. När livsinkomsterna analyseras ingår endast de som föddes mellan år 2000 och 2010, dessa får aldrig rätt till änkepensioner och efterlevandepensionerna består då endast av barnpensioner som verkar progressivt.

Ålderspensionerna verkar progressivt i tvärsnittet men regressivt över livet. Ålderspensionärer har i snitt relativt låg totalinkomst eftersom de inte har några övriga inkomster²⁹ vilket ger

²⁹ Detta är i viss mån en effekt av modellspecifika omständigheter, SESIM tillåter för närvarande inte deltidspension och pensionärer kan heller inte erhålla arbetsinkomster.

progressivitet i tvärsnittet. En hög pension ger i allmänhet inte en inkomst i toppen av fördelningen utan någonstans mellan mitten och toppen. De som har höga pensioner har dock haft höga inkomster när de var aktiva varför de återfinns i toppen av livs-inkomstfördelningen. Ålderspensionerna verkar därför regressivt över livet. Observera att analysen här endast avser *utbetalad* pension och inte tar hänsyn till inbetalda avgifter. En stegvis analys av ålderspensionssystemet i sin helhet visar att effekten sett i ett livscykelperspektiv är progressiv. Gini-koefficienten för total livs-inkomst uppgår, enligt tabell 4.1, till 0,086. Om hela ålderspensionssystemet tas bort³⁰ ökar ojämnheten till 0,110, en ökning med 28 procent. Ålderspensionssystemets progressivitet undersöks med en annan metodik i kapitel 5. Även med denna metod blir resultatet att pensionssystemet i sin helhet verkar progressivt i ett livscykelperspektiv.

Återbetalning av studiemedel verkar i tvärsnittet mer eller mindre neutralt. Detta beror möjligen på att många av de som betalar är föräldrar med en försörjningsbörda som, trots en hög inkomst, pressar ner dem till mitten av fördelningen. Vid en jämförelse med övriga subventioner och transfereringar som utgår till universitets- studerande är återbetalningarna det enda inkomstslag som över livet är (lätt) regressiva. Även här är förklaringen avbrutna studier, de som studerat och tagit studielån men ej examen betalar tillbaka sina lån men får ingen avkastning på sin utbildning.

4.2.4 Ginibidrag

Hittills har beskrivningen helt fokuserats på riktningen av inkomstslagets omfördelningsegenskaper, nästa steg är att även ta hänsyn till hur omfattande inkomstslagen är. Det är då kolumnen *Gini-bidrag* som ska studeras. Då kolumnerna summerar till olika nivåer blir jämförelsen mellan livs- och årsinkomster svår, i tabell 4.4 nedan har ginibidragen normerats med Gini-koefficienten så att kolumnerna summerar till 100 procent.

Detta, i kombination med att den version av SESIM som används här inte heller modellerar privata pensioner, leder till att ålderspensionärernas relativa position underskattas.

³⁰ Total inkomst minskas med utbetalda pensioner och ökas med allmänna egenavgifter samt arbetsgivaravgifter avseende ålderspensioner.

Tabell 4.3 Inkomstslagens relativa bidrag till ojämnheten i total års- och livsinkomst

Inkomstslag	Årsinkomst	Livsinkomst
Faktorinkomst	141%	132%
Avtalspension	5%	17%
Ålderspension	5%	22%
Förtidspension	-5%	-16%
Efterlevandepension	0%	0%
Arbetsmarknadsstöd	-1%	-1%
Sjukpenning	1%	2%
Föräldrapenning	-3%	0%
Barnbidrag	-1%	-1%
Bostadsbidrag	-1%	-1%
Bostadstillägg för pensionärer	-1%	-4%
Underhåll (mottaget)	-1%	-1%
Studiemedel	-1%	0%
Socialbidrag	-2%	-2%
Sjukvård	4%	4%
Barnomsorg	-1%	-1%
Äldreomsorg	17%	13%
Gymnasium	1%	0%
Arbetsmarknadsåtgärder	0%	0%
Komvux	0%	0%
Läkemedel	0%	0%
Grundskola	0%	-1%
Vuxenutbildning	0%	0%
Universitet	0%	0%
Inkomstskatt	-47%	-55%
Övriga direkta skatter	-5%	-4%
Återbetalt studielån	-1%	-1%
Indirekta skatter	-5%	-6%
Total inkomst	100%	100%

Källa: SESIM

I tvärsnittet drivs inkomstojämnheten nästan helt av faktorinkomsterna och i viss mån av äldreomsorgen. Faktorinkomsterna är det överlägset största inkomstslaget, det motsvarar volymmässigt 88 procent av total inkomst. Faktorinkomsternas bidrag till ojämnheten överstiger 100 procent, vilket betyder att övriga inkomstslag sammantaget verkar utjämnande. Äldreomsorgens fördelningseffekt har redan diskuterats. Övriga inkomstslag som lämnar ett märkbart positivt bidrag är ålders- och avtalspensioner

samt sjukvårdssubventionerna. Kraftigast utjämnande effekt har inkomstskatten. Över livet är mönstret detsamma men pensionernas effekter accentueras, ålders- och avtalspensionernas bidrag ökar på bekostnad av faktorinkomsten.

5 Omfördelning mellan individer och över livscykeln

5.1 Bakgrund

Ekonomiska transaktioner mellan de enskilda hushållen och den offentliga sektorn leder till en omfördelning av ekonomiska resurser i samhället. Främst sker detta genom skatte- och transfereringssystemen men även genom olika former av subventionerad verksamhet som helt eller delvis finansieras av offentliga medel. I vissa delsystem finns en uttalad målsättning att utjämna fördelningen av ekonomiska resurser genom att omfördela dessa från hushåll med en stark ekonomisk ställning till hushåll med en svagare ekonomisk ställning. Andra har en mer försäkringsmässig karaktär och syftar i första hand till att hjälpa hushållen att periodisera sina ekonomiska resurser över livscykeln.

I den första kategorin återfinns till exempel systemen för social- och bostadsbidrag samt den progressiva inkomstbeskattningen. Det kanske tydligaste exemplet på ett system av det andra slaget är ålderspensionssystemet i vilket individer sparar pengar under de förvärvsaktiva åren vilka sedan förbrukas under åren som pensionär.

I det följande benämns den första typen av omfördelning, mellan olika individer i populationen, *interpersonell* omfördelning. Den omfördelning som görs från en individ till samma individ benämns *intrapersonell* omfördelning. Den senare omfattar två delkomponenter, omfördelning inom ett och samma år för en och samma individ samt omfördelning över individens livscykel.

För vissa komponenter i skatte- och transfereringssystemen kan en klassificering, med avseende på vilken form av omfördelning komponenten huvudsakligen bidrar till, göras på intuitiva grunder. För andra är uppgiften något svårare.

Med hjälp av resultat från SESIM görs nedan en uppskattning av storleken hos de nämnda omfördelningskomponenterna för den offentliga sektorn som helhet. Dessutom görs för några av de

viktigaste komponenterna i de analyserade offentliga systemen en uppskattning av omfördelningsegenskaperna hos just dessa komponenter.

I analysen utökas individernas inkomster med subventionsvärdet för vissa offentliga tjänster. Detta ger en mer komplett bild av den offentliga sektorns totala resursomfördelning jämfört med att endast betrakta skatte- och transfereringssystemen. En explicit uppskattning av värdet av de offentliga tjänsterna för de enskilda hushållen kan förstås göras på många olika sätt. I dessa analyser likställs värdet med den beräknade nettosubventionen, definierad som tjänstens produktionskostnad minus eventuellt betalda avgifter.³¹

Tabell 5.1 innehåller en förteckning över de inkomstslag som inkluderas i analysen.

Tabell 5.1 Ingående transfereringar, skatter och subventioner

Transfereringar	Subventioner	Skatter/avgifter
Studiemedel (lån och bidrag)	Barnomsorg	Inkomstskatt
Underhållsbidrag	Grundskola	Indirekta skatter
Barnbidrag	Gymnasium	Socialavgifter (arbetsgivaravgifter, egenavgifter)
Bostadsbidrag	Universitetsstudier	Fastighetsskatt
Bostadstillägg för pensionärer	Komvuxstudier	Förmögenhetsskatt
Arbetslöshetsersättning	Vuxenstudier	Skatt på kapitalinkomster
Föräldrapenning	Arbetsmarknadsåtgärder	Återbetalt studielån
Sjukpenning	Äldreomsorg	
Socialbidrag	Vård	
Ålderspension (offentlig)	Läkemedel	
Förtidspension		
Efterlevandepension		

I analyserna betraktas i första hand individen. Individens ekonomiska situation är dock beroende av den ekonomiska situationen för övriga hushållsmedlemmar. Ett antagande som görs i dessa analyser, på samma sätt som i föregående kapitel, är därför att hushållets samtliga resurser fördelas mellan dess medlemmar. Eftersom

³¹ Se avsnitt 2.7.

fokus här ligger på storleken av olika omfördelningar och inte hushållens ekonomiska standard i sig görs ingen justering för olikheter i försörjningsbördan. Resurserna delas i stället per person inom hushållen. Detta gäller samtliga komponenter som presenteras i tabell 5.1. Exempel på konsekvenser av antagandet är att föräldrar kommer att påföras en viss del av sina barns subventioner för utbildning och att barnen påföras en viss del av sina föräldrars betalda skatter.

5.2 Metod

Beräkningen av de inter- respektive intrapersonella omfördelningskomponenterna görs enligt en metod som tidigare använts av bland andra Hussénius och Selén (1994)³², Falkingham och Harding (1996) samt O'Donoghue (2001b).

Enligt metoden antas varje kohort³³ balansera sig själv finansiellt över livscykeln, det vill säga att summan av inbetalda skatter och avgifter exakt motsvarar summan av utbetalda transfereringar och subventioner för offentliga tjänster. Genom att låta varje kohort uppnå finansiell balans bortser metoden från eventuella statsfinansiella under- eller överskott (för en given kohort över dess livstid) som uppstår under analysperioden. Sådana kan dock tänkas förekomma inom ramen för de delar av offentlig sektor som inte inkluderas i analysen.

Kohorternas totala inbetalade skatter kommer i simuleringen att överstiga summan av transfereringar och subventioner beroende på att skatterna även ska finansiera vissa offentliga kostnadsposter vilka inte inkluderas i analysen. För att uppnå finansiell balans över livscykeln för varje kohort måste alltså en viss anpassning av det totala skatteuttaget göras. Eftersom en sådan anpassning kan åstadkommas på många olika sätt och resultaten i viss mån kan påverkas av valet av metod kommer två olika metoder att användas i analyserna.

Enligt den första metoden finansieras den offentliga sektorns utgifter av samtliga socialavgifter³⁴ och en så stor andel av direkta och indirekta skatter som krävs för att kohorten ska balanseras.

³² Ds 1994:135.

³³ Med kohort avses här de individer som är födda under samma kalenderår.

³⁴ Arbetsgivaravgiften betraktas här som en skatt även om det formellt inte är hushållet som erlägger avgiften.

Enligt den andra metoden används samtliga indirekta skatter och en så stor andel av socialavgifter och direkta skatter som krävs (se Tabell 5.2).

Metod 1 lägger en större finansieringsbörda på individer som har arbetsinkomster då socialavgifterna där väger tyngst. Här kommer alltså hushåll som har låga, eller helt saknar, förvärvsinkomster att ge låga bidrag. Metod 2, som i stället lägger störst vikt vid de indirekta skatterna, fördelar finansieringsbördan något jämnare i populationen. De indirekta skatterna beräknas, som tidigare nämnts, som en andel av hushållens disponibla inkomst vilket innebär att i stort sett alla hushåll bidrar i någon utsträckning.

Tabell 5.2 Metoder för tilldelning av skatter och avgifter

	Belopp	Metod 1	Metod 2
Socialavgifter	A	100 %	β %
Direkta skatter	D	α %	β %
Indirekta skatter	I	α %	100 %
Tilldelade skatter/avgifter ³⁵		A + α (D+I)	I + β (A+D)

Anm.: A, D och I representerar här summor över samtliga levnadsår för alla individer i respektive kohort.

Givet denna metod för tilldelning av skatter och avgifter kan de inter- respektive intrapersonella omfördelningskomponenterna beräknas.

Individer kan varje år få transfereringar (och subventioner för offentliga tjänster) samtidigt som de betalar skatter (och avgifter). Om en individ får större transfereringar än hon betalar skatter är individens netto gentemot den offentliga sektorn positivt det året. Omvänt uppkommer negativa netton genom att individen betalar högre skatter än hon får igen i form av transfereringar. Summan av individens årliga netton över livet är lika med individens *livsnetto*. Eftersom varje kohort antas betala lika mycket skatter som de får i transfereringar, över livscykeln, är summan av de positiva livsnettona lika med (minus) summan av de negativa livsnettona. Summan

³⁵ Låt T beteckna en kohorts totala transfereringar och subventioner över livscykeln. Uttagsandelarna, enligt respektive metod, kan då beräknas som

$$\alpha = \frac{T - A}{D + I} \text{ och } \beta = \frac{T - I}{A + D} .$$

motsvarar de medel som omfördelas *mellan individer* (interpersonell omfördelning) inom kohorten.

Om en individ under ett visst år får transfereringar samtidigt som hon betalar skatt kan det ses som om transfereringarna finansieras av de betalade skatterna, här benämnt *årlig rundgång*. Om skatterna överstiger transfereringarna finansieras dessa helt av de betalade skatterna, det vill säga den årliga rundgången är lika stor som transfereringarna. Om transfereringarna överstiger skatterna finansieras bara den del som inte överstiger skatterna, den årliga rundgången motsvarar då de betalade skatterna.

En individ som har ett positivt livsnetto gentemot den offentliga sektorn kan ändå för vissa år ha negativa netton, det vill säga då skatterna överstiger transfereringarna. Dessa skatter kan då antas finansiera erhållna transfereringar under andra år i individens liv, här benämnt *livscykelrundgång*. Omvänt kan en individ med ett negativt livsnetto ha enstaka år med positiva netton. Dessa års transfereringar kan då ses som finansierade av andra års betalade skatter.

Sammantaget ger den årliga rundgången och livscykelrundgången ett mått på hur mycket av de erhållna transfereringarna som är *självfinansierade*, det vill säga har finansierats av individen själv vid något tillfälle i livet.

För en mer formell definition av beräkningarna se faktaruta 5.1.

Faktaruta 5.1 Beräkning av inter- respektive intrapersonella omfördelningskomponenter

Beräkningar av inter- och intrapersonella fördelningskomponenter utförs för varje kohort. För att förenkla notationen används dock ingen indexering med avseende på de olika kohorterna.

Låt S_{it} beteckna uttaget av skatter och avgifter från individ i vid ålder t . Låt vidare T_{it} beteckna summan av transfereringar och subventioner som betalas till individ i vid ålder t . Individens livsnetto, betecknat N_i , gentemot den offentliga sektorn kan nu beräknas som summan av erhållna transfereringar och subventioner minskade med summan av betalade skatter och avgifter:

$$N_i = \sum_{t=0}^{K_i} T_{it} - \sum_{t=0}^{K_i} S_{it},$$

forts. Faktaruta 5.1

där K_i betecknar åldern vid vilken individ i avlider. Den interpersonella omfördelningen i kohorten beräknas som summan av alla positiva livsnetton i kohorten:

$$INTER = \sum_{i=1}^N (N_i | N_i > 0)$$

Observera att detta exakt motsvarar (minus) summan av alla negativa livsnetton i kohorten eftersom kohortens totala netto är lika med noll.

Den intrapersonella omfördelningen består av två komponenter. Dels den omfördelning som går från en individ tillbaka till samma individ *under samma år*, här kallad årlig rundgång. Dels den omfördelning som går från en individ ett visst år till samma individ *under ett annat år*. Den årliga rundgången uppstår genom till exempel skattepliktiga transfereringar där en viss del av transfereringarna kan antas vara finansierade av de skatter och avgifter individen betalar samma år.

Den årliga rundgången för individ i vid ålder t betecknas RA_{it} och beräknas som minimum av erlagda skatter/avgifter och erhållna transfereringar och subventioner:

$$RA_{it} = \min(T_{it}, S_{it})$$

Den totala årliga rundgången för kohorten beräknas genom att summera den årliga rundgången över kohortens sammanlagda levnadsår:

$$INTRA_{AR} = \sum_{i=1}^N \sum_{t=1}^{K_i} RA_{it}$$

Den del av den intrapersonella omfördelningen, för individ i , som görs mellan olika åldrar i individens liv beräknas på olika sätt beroende på om individen har ett positivt eller negativt livsnetto.

forts. Faktaruta 5.1

$$RL_i = \begin{cases} \sum_{t=i}^{K_i} (S_{it} - RA_{it}) & , \text{ om } N_i > 0 \\ \sum_{t=i}^{K_i} (T_{it} - RA_{it}) & \text{ om } N_i \leq 0 \end{cases}$$

Summerat över samtliga individer i kohorten ges den totala livscykelrundgången som

$$INTRA_{LIV} = \sum_{i=1}^N RL_i$$

Sammanfattningsvis gäller nu, för respektive kohort, att summan av utbetalda transfereringar och subventioner är lika stor som summan av inbetalda skatter och avgifter. Vidare är dessa summor lika stora som summan av de intra- respektive interpersonella omfördelningarna i kohorten:

$$\sum_{i=1}^N \sum_{t=0}^{K_i} T_{it} = \sum_{i=1}^N \sum_{t=0}^{K_i} S_{it} = INTER + INTRA_{AR} + INTRA_{LIV}$$

Ett exempel

För att åskådliggöra tekniken presenteras i tabell 5.3 ett räkneexempel för en fiktiv population bestående av två individer vilka observeras under fem perioder.

Den första individen har totalt tilldelats 10 enheter i transfereringar och subventioner (T) och betalat 20 enheter i skatt (S) vilket ger ett livsnetto om minus 10 enheter. För den andra individen är dessa siffror omvända. Totalt sett har i populationen alltså omfördelats $10+20=30$ enheter. Dessa kan dekomponeras i en interpersonell omfördelning om 10 enheter (summan av positiva livsnetton), en årlig rundgång (RÅ) om $8+8=16$ enheter samt en intrapersonell livsomfördelning (RL) om $2+2=4$ enheter. I exempelpopulationen utgör alltså den interpersonella omfördelningen en tredjedel av den totala omfördelningen. Frånräknat den årliga rundgången uppgår den i stället till ca 71 procent ($10/14$).

Tabell 5.3 Exempelberäkning av omfördelningskomponenter

Individ	Period	T	S	RÅ	RL
1	1	2	2	2	0
	2	3	2	2	1
	3	4	3	3	1
	4	1	5	1	0
	5	0	8	0	0
	Summa	10	20	8	2
2	1	5	2	2	0
	2	2	4	2	2
	3	2	2	2	0
	4	5	1	1	0
	5	6	1	1	0
	Summa	20	10	8	2

Anm.: T står för transfereringar, S för skatter, RÅ för årlig rundgång och RL för rundgång över livet.

5.3 Resultat

Följande resultat baseras på beräkningar avseende individer födda mellan 2000 och 2010 vilka har simulerats till och med år 2119. Resultaten för de olika födelsekohorterna är mycket lika sinsemellan och presenteras här av den anledningen i stället som ett genomsnitt av de ingående kohorterna. Samtliga belopp anges i 2003 års fasta priser.

5.3.1 Beräknade omfördelningskomponenter

I nedanstående tabeller presenteras de beräknade beloppen för respektive omfördelningskomponent. I beräkningarna regleras skatteuttaget med hjälp av metod 1.

Av tabell 5.4 framgår att det är den årliga rundgången som är den enskilt största omfördelningskomponenten och utgör 45 procent av samtliga omfördelade medel. Den årliga rundgången uppstår framför allt genom de skattepliktiga transfereringarna och genom att skattebetalande hushåll erhåller offentliga subventioner. Den påverkas därför i hög grad av de antaganden som görs angående resursfördelningen inom hushållen, till exempel genom att föräldrar antas få del av barnens utbildningssubventioner och genom att barn antas betala en del av föräldrarnas skatt. På grund

av detta redovisas omfördelningskomponenterna i det följande även frånräknat den årliga rundgången³⁶.

Den interpersonella omfördelningen beräknas till 18 procent, vilket gör att andelen självfinansierade transfereringar och subventioner uppgår till 82 procent. Drygt åtta av tio kronor som individen får i transfereringar och subventioner över livet har, för en genomsnittlig individ, alltså finansierats av individen själv vid någon tidpunkt.

Frånräknat den årliga rundgången uppgår de inter- respektive intrapersonella omfördelningskomponenterna i stället till 32 procent och 68 procent. Sammantaget visar dessa resultat att de offentliga systemen huvudsakligen åstadkommer en periodisering av resurser över livscykeln.

Tabell 5.4 Beräknade omfördelningskomponenter.

Genomsnitt per individ

Komponent	Belopp (tkr)	Andel (%)	Andel, exklusive årlig rundgång (%)
INTER	1194	18	32
INTRA (år)	3024	45	
INTRA (liv)	2540	38	68
Totalt	6758	100	100

Källa: SESIM

Hussénus och Selén (1994) visar i liknande beräkningar på en något högre interpersonell omfördelning, ca 24 procent. Frånräknat den årliga rundgången redovisas även där en interpersonell omfördelning på 32 procent vilket helt överensstämmer med resultaten i tabell 5.4.³⁷

I liknande internationella studier visar Falkingham och Harding (1996) att den interpersonella omfördelningen ligger mellan 48 och 62 procent för Australien samt mellan 29 och 38 procent för Storbritannien beroende på vilken metod som används för att

³⁶ Detta ger en analys där hushållens årliga netton betraktas och inte de faktiska bruttoströmmarna av skatter och transfereringar.

³⁷ Betydande skillnader i beräkningsförutsättningar föreligger dock mellan dessa båda analyser. Hussénus, J. & Selén, J. [Ds 1994:135] baserar sina beräkningar på 1994 års skatte- och transfereringssystem och använder en annan typ av modell för simulering av befolkningen. Vidare beaktas inte värdet av offentliga subventioner. Vissa skillnader finns även i antaganden om delning av medel inom hushållen.

bestämma skatteuttaget. Enligt O'Donoghue (2001b) är motsvarande siffror för Irland och Italien 45 procent respektive 24 procent.

Den förhållandevis låga interpersonella omfördelningen i Sverige (åtminstone i förhållande till Storbritannien, Australien och Irland) kan förklaras av att en stor del av transfereringarna här består av socialförsäkringar som är knutna till den försäkrades inkomster. Endast ålderspensionssystemet utgör till exempel i beräkningarna ca 56 procent av samtliga transfereringar. En relativt stor andel av transfereringarna är i Sverige skattepliktiga vilket bidrar till den intrapersonella omfördelningen (i form av årlig rundgång).

Tabell 5.5 visar genomsnittliga flöden per individ mellan individer och offentlig sektor uppdelat på kvintiler för genomsnittlig livsinkomst. Livsinkomstnivån beräknas som den genomsnittliga justerade disponibla inkomsten per levnadsår. Den genomsnittliga livsinkomsten används här i stället för den summerade för att undvika att kortlivade individer ska hamna långt ned i den relativa livsinkomstfördelningen oberoende av deras faktiska årliga inkomstnivå.

Tabellen visar att skatteuttaget ökar med stigande livsinkomst. Jämfört med det genomsnittliga skatteuttaget för samtliga kvintiler är det i den första kvintilen ca 37 procent lägre och i den femte ca 40 procent högre.

Transfereringar och subventioner för offentlig konsumtion är betydligt mer jämnt fördelade över livsinkomstnivåerna. Transfereringar till den första livsinkomstkvintilen är ca 3 procent högre än det totala genomsnittet medan det för den femte kvintilen är ca 7 procent lägre. Offentliga subventioner är störst för den andra och tredje livsinkomstkvintilen, 6 procent högre än genomsnittet, och lägst för den femte, 9 procent lägre än genomsnittet. Notera även att den lägsta kvintilen har lägre subventioner än det totala genomsnittet.

De genomsnittliga livsnettona är negativa i de högre livsinkomstkvintilerna och positiva i de lägre livsinkomstkvintilerna vilket visar att den offentliga sektorn har en utjämnande inverkan även på livsinkomster. För populationen i stort är det genomsnittliga livsnettot noll eftersom kohorterna balanseras finansiellt.

Andelen självfinansierade transfereringar och subventioner³⁸ stiger med ökande livskomst. I den första kvintilen uppgår dessa till ca 61 procent medan de i den femte kvintilen uppgår till ca 98 procent.³⁹ Även bland individer i den lägsta livsinkomstkvintilen finansieras alltså i genomsnitt drygt sex av tio kronor som betalas ut i form av transfereringar och subventioner över livscykeln genom egna skatter och avgifter.

Tabell 5.5 Omfördelningskomponenter samt livsnetto per livsinkomst (tkr)

<i>Genomsnitt per individ</i>						
Kvintil av livsinkomst	1	2	3	4	5	alla
Skatter	4264	5741	6597	7691	9483	6758
%	-37	-15	-2	+14	+40	
Transfereringar	3924	3983	3915	3770	3543	3827
%	+3	+4	+2	-2	-7	
Subventioner	2840	3109	3109	2922	2679	2932
%	-3	+6	+6	0	-9	
INTRA (år)	2603	3057	3143	3153	3162	3024
INTRA (liv)	1536	2374	2784	3093	2907	2540
Livsnetto	2500	1351	428	-999	-3262	0

Anm.: Rader markerade med % avser förhållande till totalt populationsgenomsnitt (procentuell skillnad).

Källa: SESIM

Som tidigare nämnts skiljer sig de båda metoderna för tilldelning av skatter genom att metod 1 belastar hushåll/individer med arbetsinkomster hårdare medan metod 2 fördelar finansieringsbördan jämnare bland hushållen⁴⁰. Tabell 5.6 visar resultat från beräkningar där metod 2 använts.

Den huvudsakliga skillnaden mot de ovan redovisade beräkningarna är att den årliga rundgången ökar med drygt tre procentenheter. En förklaring till detta är att resurssvaga hushåll i högre utsträckning bidrar till finansieringen enligt metod 2 jämfört med

³⁸ Andelen självfinansierade transfereringar och subventioner beräknas som summan av intrapersonella omfördelningskomponenter (år och liv) dividerat med summan av transfereringar och subventioner.

³⁹ Ett positivt samband mellan andelen självfinansierade transfereringar och livsinkomster visas även av Hussénus, J. & Selén, J. [Ds1994:135], O'Donoghue, C. [2001b] samt Falkingham, J. & Harding, A. [1996] i analyser av de svenska, irländska, australiensiska och brittiska skatte- och transfereringssystemen.

⁴⁰ Se tabell 5.2.

metod 1. Eftersom dessa hushåll ofta erhåller transfereringar och subventioner skapas en ökad intrapersonell omfördelning genom det ökade skatteuttaget.

Frånräknat den årliga rundgången ger metoden endast marginella skillnader i interpersonell respektive intrapersonell livsomfördelning jämfört med metod 1.

Tabell 5.6 Beräknade omfördelningskomponenter. Skattetilldelning enligt metod 2

<i>Genomsnitt per individ</i>			
Komponent	Belopp (tkr)	Andel (%)	Andel, exklusive årlig rundgång (%)
INTER	1146	17	33
INTRA (år)	3233	48	
INTRA (liv)	2379	35	68
Totalt	6758	100	100

Källa: SESIM

Skatteuttaget är i dessa beräkningar något mer jämnt fördelat mellan livsinkomstkvintilerna. Den första kvintilens genomsnittliga skatter ökar med ca 2 procent medan den femte kvintilens minskar med ca 1 procent. Detta gör även att det genomsnittliga livsnettot här är något mer jämnt fördelade jämfört med metod 1 och följaktligen att den utjämnande effekten av den offentliga sektorn inte är lika stor. Skillnaderna är dock relativt små. I det följande används metod 1 genomgående.

5.3.2 Skatter, transfereringar och subventioner över livscykeln

Beräkningarna ovan tyder på att de ekonomiska transaktionerna mellan individerna/hushållen och den offentliga sektorn skapar en betydande livscykelomfördelning. För att förstå hur denna omfördelning uppkommer behövs information om när, under livscykeln, dessa transaktioner inträffar.

Diagram 5.1 visar genomsnittliga skatter, transfereringar, subventioner och individnetto per ålder⁴¹. Beräkningarna baseras som tidigare på kohorterna 2000 till 2010 vilka simuleras över hela livscykeln.

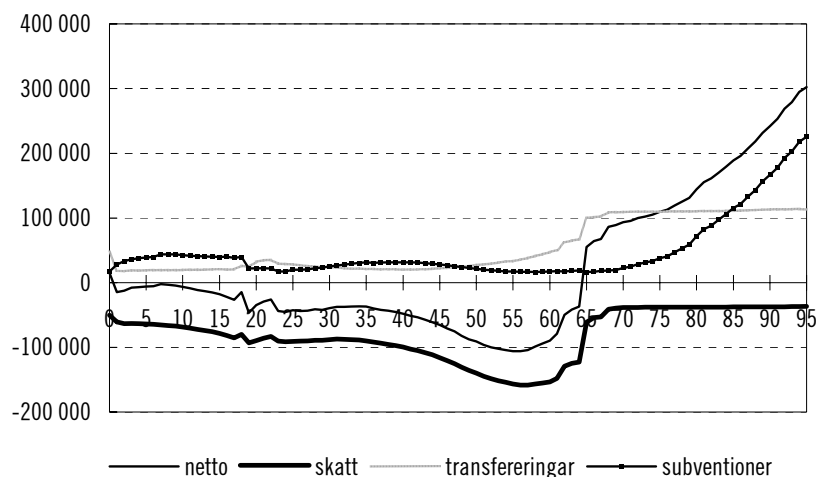
Transfereringarna är, främst genom ålders- och förtidspensions-systemen, starkt koncentrerade till höga åldrar. Beräkningar visar att ca 65 procent av samtliga transfereringar erhålls av individer som är 60 år eller äldre.

Skatterna utvecklas under de förvärvsaktiva åldrarna ungefär i samma takt som den allmänna inkomstutvecklingen i dessa åldrar. Efter pensioneringen är skatteuttaget betydligt lägre då inkomsterna minskar samtidigt som inga socialavgifter längre betalas.

Det genomsnittliga årliga individnettot är negativt fram till pensioneringen. Därefter gör den sammantagna effekten av de minskade skatterna och de ökande transfereringarna att nettot genast blir positivt. Eftersom subventionerna ökar relativt snabbt i de högsta åldrarna ges även en motsvarande ökning av individnettot.

Diagram 5.1 Skatter, transfereringar, subventioner och netto per ålder

Genomsnitt per ålder



Källa: SESIM

⁴¹ I de följande diagrammen redovisas inte resultat för åldrar över 95 år. Detta på grund av att beräkningarna baseras på få observationer och därför är behäftade med en hög osäkerhet.

Av diagrammet framgår även effekterna av att hushållsmedlemmar antas ha en delad ekonomi. Till exempel verkar småbarn betala skatt, något som förstås visar barnens andel av föräldrarnas skatter snarare än barnens "egna" skatter. Om medel inte delades i hushållet skulle alltså de genomsnittliga skatterna i föräldraåldrarna ökas med det belopp som i denna analys i stället påförs under barnåldrarna.

Hushållens delning av resurser påverkar även profilen för de offentliga subventionerna vilken i diagrammet är mer utjämnad över åldrarna än den som presenteras i appendix A. Barnens utbildningssubventioner blir genom delningen något lägre och föräldrarna får en motsvarande höjning av sina subventioner.

Diagram 5.1 visar att huvuddelen av skatterna för en genomsnittlig individ tas ut vid en lägre ålder än den där huvuddelen av transfereringarna och subventionerna utbetalas⁴². Individen har ett negativt netto under den första delen av sitt liv för att under den senare delen få ett positivt netto.

Detta förhållande leder till frågan om vid vilken ålder den genomsnittliga individen uppnår ett permanent positivt netto gentemot den offentliga sektorn. Det intressanta måttet är då inte individens årliga netto utan snarare det kumulativa nettot vid en viss ålder, dvs. summan av årliga individnetton från det första levnadsåret till den aktuella åldern.

I diagram 5.2 visas det beräknade genomsnittliga kumulativa individnettot för varje ålder mellan 0 och 95 år. Diagrammet visar på ett tydligt sätt hur den genomsnittliga individen under den första delen av livet bygger upp ett negativt kumulativt netto gentemot den offentliga sektorn. Nettot "försämras" snabbast i 50-årsåldern då individen betalar mycket skatt jämfört med de transfereringar och subventioner hon erhåller. Först under pensionsåldern vänder utvecklingen och det kumulativa nettot avtar. I dessa beräkningar uppnår den genomsnittliga individen ett permanent nollnetto först vid 87 års ålder.⁴³

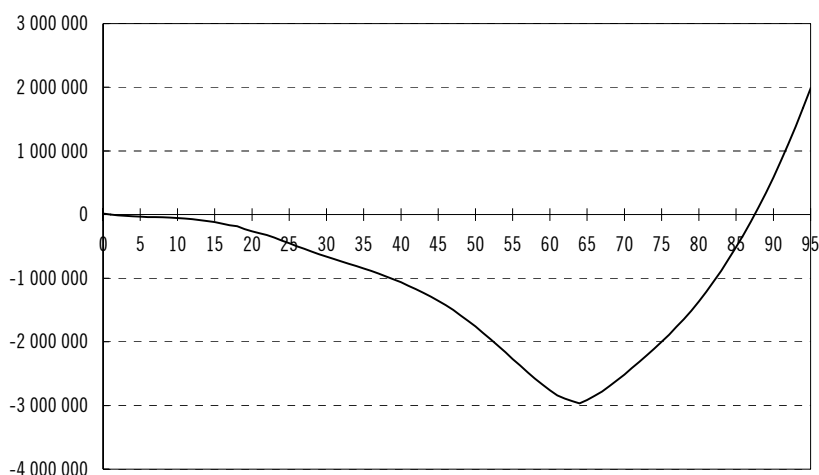
En konsekvens av åldersmönstret hos de kumulativa nettona är att det är mer sannolikt att en individ som lever till hög ålder är

⁴² Beräkningar visar på en "medelålder" för skatteuttag på 40 år och en "medelålder" för transfereringar och subventioner på 48 år vid en livslängd av 80 år. För en livslängd på 90 år beräknas motsvarande till 43 respektive 59 år.

⁴³ Den genomsnittliga livslängden för de studerade kohorterna uppgår till ca 84 år. På grund av att fördelningen för de kumulativa livsnettona vid varje given ålder är något skev, med lägre medelvärde än median, förskjuts tidpunkten för uppnått nollnetto något till högre åldrar.

nettobidragstagare, dvs. har ett positivt netto gentemot den offentliga sektorn, jämfört med en individ som dör tidigt. På så sätt åstadkoms en omfördelning från kortlivade individer till långlivade individer⁴⁴. Effekten av denna omfördelning minskas dock av det positiva sambandet mellan livsinkomst och förväntad medellivslängd.

Diagram 5.2 Genomsnittligt kumulativt livsnetto i åldrarna 0–95 år



Källa: SESIM

5.3.3 Omfördelning mellan olika grupper

Tabell 5.5 visade att det över livscykeln omfördelas medel från individer i de högre livsinkomstkvintilerna till individer i de lägre livsinkomstkvintilerna. Ett alternativt sätt att beskriva denna omfördelning är genom att studera beräknade genomsnittliga kumulativa livsnetton för respektive livsinkomstkvartil. På så sätt framgår det hur individer i olika inkomstlägen genomsnittligt förhåller sig ekonomiskt till den offentliga sektorn, totalt sett sedan födseln, vid olika tidpunkter i livet.

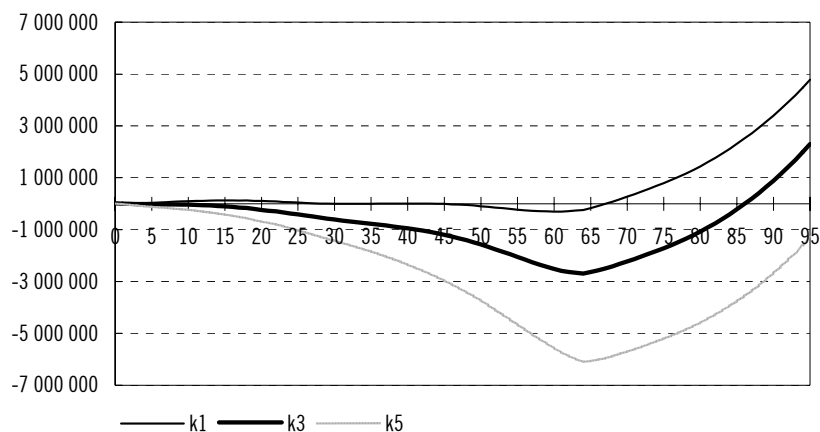
⁴⁴ Även Harding, A. m.fl. [2000] noterar denna typ av omfördelning, med avseende på livslängd, i en australiensisk studie av hälso- och sjukvårdssubventioners livsomfördelningseffekter. De kumulativa nettoprofilerna i studien uppvisar stora likheter med de i diagram 5.3.

Diagram 5.3 visar genomsnittliga kumulativa livsnetton vid olika åldrar för den första, tredje och femte livsinkomstkvintilen.

Generellt ligger det kumulativa livsnettot vid en given ålder på en lägre nivå ju högre livsinkomsten är. Den första kvintilen har ett svagt positivt kumulativt netto under största delen av livet bortsett från en period innan pensionsåldern. Efter 67 års ålder är det kumulativa nettot permanent positivt. För den tredje och femte kvintilen är de kumulativa nettona negativa och minskar dessutom fram till pensionsfallet. För den tredje kvintilen blir det kumulativa nettot positivt vid 87 års ålder och för den femte kvintilen först vid 99 års ålder.

Det tar alltså i genomsnitt 32 år längre för en individ från den högsta livsinkomstkvintilen att uppnå ett positivt netto mot den offentliga sektorn än det gör för en individ från den lägsta livsinkomstkvintilen.

Diagram 5.3 Genomsnittligt kumulativt livsnetto per ålder och livsinkomstkvartil



Källa: SESIM

Omfördelning mellan män och kvinnor

Då de genomsnittliga arbetsinkomsterna är högre för män än för kvinnor kan det även vara motiverat att undersöka om det även

finns en omfördelning mellan könen och, om så är fallet, beräkna dess storlek.

Tabell 5.7 visar de beräknade genomsnittliga omfördelningskomponenterna per individ och kön. Jämfört med det genomsnittliga skatteuttaget i populationen är männens i genomsnitt 6 procent högre och kvinnornas motsvarande 6 procent lägre. För transfereringar är skillnaderna ungefär lika stora men här ligger kvinnorna över populationsgenomsnittet och männen under. Den största skillnaden mellan könen kan ses i fördelningen av offentliga subventioner där kvinnornas genomsnitt ligger 11 procent över populationsgenomsnittet och männens 11 procent under.

Andelen transfereringar och subventioner som är självfinansierade, dvs. betalas av individens egna skatter och avgifter någon gång under livet, är ca 89 procent för män och 77 procent för kvinnor.

Livsnettot visar att det under en livstid, genom den offentliga sektorn, överförs något under en miljon kronor från en genomsnittlig man till en genomsnittlig kvinna. Till detta kommer den omfördelning som sker inom hushållen.

Tabell 5.7 Omfördelningskomponenter samt livsnetto för män och kvinnor (tkr)

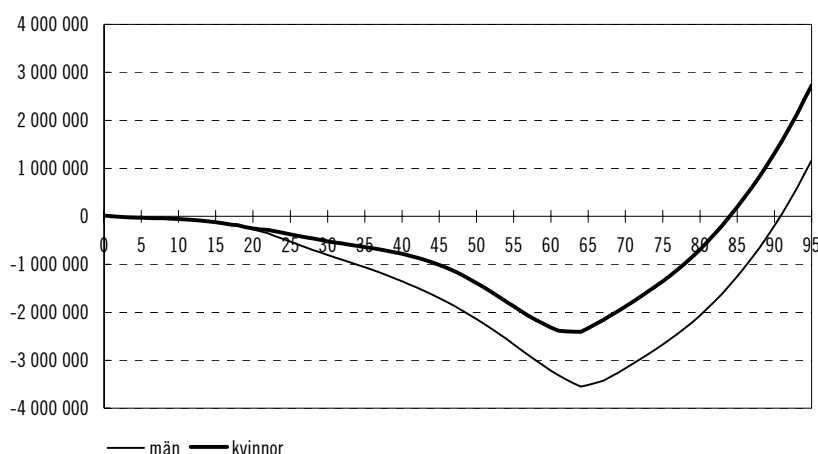
<i>Genomsnitt per individ</i>			
	Män	Kvinnor	Alla
Skatter	7136	6389	6758
%	+6	-6	
Transfereringar	3562	4085	3827
%	-7	+7	
Subventioner	2611	3245	2932
%	-11	+11	
INTRA (år)	2920	3125	3024
INTRA (liv)	2547	2534	2540
Livsnetto	-963	941	0

Anm.: Rader markerade med % avser förhållande till totalt populationsgenomsnitt (procentuell skillnad).

Källa: SESIM

Diagram 5.4 visar att det genomsnittliga kumulativa livsnettot för en kvinna är permanent positivt efter 84 års ålder medan det för en man är permanent positivt först efter 90 års ålder.

Diagram 5.4 Genomsnittligt kumulativt livsnetto per ålder och kön



Källa: SESIM

Omfördelning mellan olika utbildningsnivåer⁴⁵

Eftersom arbetsinkomster i SESIM uppvisar ett positivt samband med individens högsta utbildningsnivå kan det, utifrån ovanstående resultat, förväntas ske en omfördelning från högutbildade individer till individer med lägre utbildning. Tabell 5.8 visar att så är fallet. Det finns ett tydligt negativt samband mellan utbildningsnivå och genomsnittligt livsnetto – ju högre utbildning desto lägre netto.

Det genomsnittliga skatteuttaget ökar med ökande utbildningsnivå. Genomsnittet för de lägst utbildade är 15 procent lägre och för de högst utbildade 9 procent högre än populationsgenomsnittet. Mönstret gäller även för subventionerna där genomsnittet för de lägst utbildade är 10 procent lägre och för de högst utbildade 6 procent högre än populationsgenomsnittet. De högre subventionerna för högutbildade kan till stor del förklaras med utbildningssubventionerna. En ytterligare bidragande orsak är det positiva sambandet mellan utbildningsnivå och medellivslängd som gör att de höga subventionerna sent i livet kommer högutbildade till del i högre utsträckning än övriga. Minst skillnad mellan utbild-

⁴⁵ Utbildningsnivåerna låg, mellan och hög avser här den under livet högst uppnådda utbildningsnivån och indelas efter grundskola (låg), gymnasieexamen (mellan) samt universitetsexamen inklusive forskarutbildning (hög).

ningsnivåerna ses i fördelningen av transfereringar där de lägst och högst utbildade ligger något över populationsgenomsnittet.

Andelen självfinansierade transfereringar och subventioner stiger med utbildningsnivån och är ca 76 procent, 82 procent och 85 procent för låg-, medel- och högutbildade.

Tabell 5.8 Omfördelningskomponenter samt livsnetto för olika högsta utbildningsnivåer (tkr)

Genomsnitt per individ

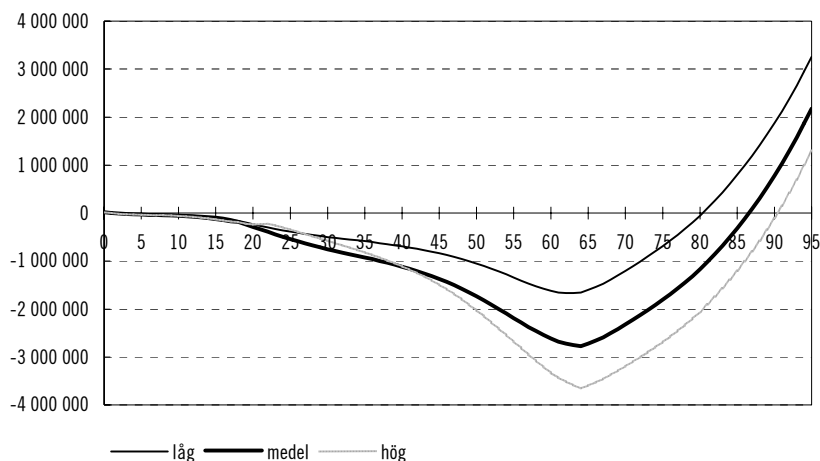
	Låg	Mellan	Hög	Alla
Skatter	5733	6547	7394	6758
%	-15	-3	+9	
Transfereringar	3917	3792	3845	3827
%	+2	-1	+1	
Subventioner	2640	2884	3095	2932
%	-10	-2	+6	
INTRA (år)	2908	3012	3079	3024
INTRA (liv)	2084	2433	2840	2540
Livsnetto	824	129	-453	0

Anm.: Rader markerade med % avser förhållande till totalt populationsgenomsnitt (procentuell skillnad).

Källa: SESIM

Diagram 5.5 visar genomsnittliga kumulativa livsnetton i olika åldrar för de olika utbildningsnivåerna. Generellt sett innebär högre utbildning lägre kumulativt netto i varje given ålder. Mönstret bryts dock av de högst utbildade som, genom bland annat sina utbildningssubventioner från studietiden och fram till ca 40 års ålder har lika högt eller högre netto som individer med lägre utbildningsnivå.

För låg-, mellan- respektive högutbildade blir det kumulativa livsnettot positivt vid 80, 87 och 91 års ålder.

Diagram 5.5 Genomsnittligt kumulativt livsnetto per ålder och utbildningsnivå

Källa: SESIM

Observera att resursfördelningen inom hushållen leder till att den uppskattade omfördelningen mellan de ovan studerade grupperna blir mindre omfattande än vad den skulle ha varit i en helt individ-baserad analys. Detta eftersom hushållets skatter, transfereringar och subventioner delas per capita och registreras på hushållets samtliga medlemmar oavsett grupptillhörighet.

5.3.4 Omfördelning genom ålderspensioner och offentliga subventioner

De beräknade omfördelningskomponenterna har hittills avsett den sammantagna effekten av transfereringssystem och offentliga subventioner (samt deras respektive finansiering). Två delsystem vars livsomfördelningssegenskaper är intressanta, inte minst för deras vikt i totalanalyserna, är ålderspensionssystemet och de offentliga subventionerna⁴⁶.

Ett sätt att modifiera den använda beräkningsmetodikerna i syfte att undersöka fördelningssegenskaperna hos olika delsystem är att

⁴⁶ Ålderspensionerna utgör ca 56 procent av de totalt utbetalade transfereringarna över de studerade kohorternas livstider. Ungefär 32 procent av alla medel som omfördelas går via pensionssystemet. För de offentliga subventionerna är motsvarande siffra 43 procent.

utföra en partiell analys. I den partiella analysen elimineras ett delsystem (inklusive dess finansiering) varvid omfördelningskomponenterna åter beräknas. Genom att studera förändringen i de beräknade omfördelningskomponenterna mellan de två analyserna erhålls en indirekt uppskattning av omfördelningskomponenterna för det delsystem som undersöks.

I de partiella beräkningarna antas implicit att det delsystem som studeras är i finansiell balans över livscykeln. Detta följer av att motsvarande antagande gäller för hela den offentliga sektorn. Vidare är analysen i viss mån beroende av den totala omfattningen av de studerade (del-)systemen. En större förändring av det totala skatteuttaget, som följer då ett delsystem med stor omfattning elimineras, kan även ge förändringar av skatternas omfördelningsegenskaper i och med att fördelningen mellan socialavgifter, direkta och indirekta skatter förändras.

I följande partiella beräkningar har ålderspensionssystemet och de offentliga subventionerna exkluderas, en åt gången. I beräkningarna utan ålderspensionssystemet exkluderats även samtliga ålderspensionsavgifter samt de inkomstskatter som betalas på ålderspensioner. Observera att analysen är statisk i den meningen att bortfallet av ålderspensioner inte kompenseras av andra transfereringar (t.ex. BTP). I beräkningarna exklusive de offentliga subventionerna görs en allmän nedjustering av skatteuttaget eftersom subventionerna antas finansieras genom allmänna skattemedel. I båda delanalyserna har skatteuttaget fördelats enligt metod 1.

Omfördelning genom ålderspensioner

Tabell 5.9 visar de (indirekt) beräknade genomsnittliga omfördelningskomponenterna per livsinkomstkvintil för ålderspensionssystemet. Totalt sett har ålderspensionssystemet en något högre självfinansieringsgrad jämfört med samtliga system, 86 procent mot 82 procent. För pensionssystemet består dessutom en mindre andel av den intrapersonella omfördelningen av den årliga rundgången, endast 44 procent jämfört med 54 procent för alla system. Detta innebär att den rena livscykelomfördelningen (exklusive den årliga rundgången) inom pensionssystemet är betydligt högre än den är i samtliga system sammantaget, 77 procent jämfört med 68 procent.

Ålderspensioner bidrar alltså i betydligt högre grad till att periodisera medel över livet än till att omfördela medel mellan

individer⁴⁷. Detta är förstås inte förvånande med tanke på systemets utformning och målsättning.

Tabell 5.9 Omfördelningskomponenter samt livsnetto per livsinkomst för ålderspensionssystemet (tkr)

Genomsnitt per individ

Kvintil av livsinkomst	1	2	3	4	5	Alla
Skatter	1399	1856	2112	2425	2933	2145
%	-35	-14	-2	+13	+37	
Transfereringar	1848	2146	2248	2260	2225	2146
%	-14	0	+5	+5	+4	
INTRA (år)	727	816	826	821	878	813
INTRA (liv)	604	923	1073	1251	1261	1023
Livsnetto	449	290	138	-165	-708	0

Anm.: Rader markerade med % avser förhållande till totalt populationsgenomsnitt (procentuell skillnad).

Källa: SESIM

Det relativa skatteuttaget per livsinkomstkvintil i förhållande till populationsgenomsnittet är något mer jämnt fördelat jämfört med i analysen av samtliga system (se tabell 5.5). Individer med låg livsinkomst bidrar i genomsnitt något mer och individer med hög livsinkomst något mindre.

Ålderspensionerna är, i förhållande till den genomsnittliga ålderspensionen, relativt sett mindre i de lägre inkomstkvintilerna än i de högre. Detta beror förstås på att huvuddelen av de utbetalda ålderspensionerna är inkomstgrundade.

Sammantaget ger detta livsnetton som skiljer sig mindre åt mellan de högre och lägre inkomstkvintilerna inom pensionsystemet jämfört med i samtliga system. Detta är förstås delvis en effekt av att omfattningen av samtliga system nödvändigtvis är större än den för ålderspensionssystemet.

Ålderspensionssystemet bidrar dock till en utjämning av livsinkomster eftersom de lägre inkomstkvintilerna har positiva genomsnittliga livsnetton mot pensionssystemet medan de högre har negativa genomsnittliga livsnetton. I avsnitt 4.2.3 visar en

⁴⁷ Hussénus, J. & Selén, J. (Ds 1994:135) når samma slutsats i liknande beräkningar och lyckas även påvisa kraftigare effekter då ålderspensionssystemet exkluderas. Detta torde dock delvis bero på att beräkningarna inte inkluderar offentliga subventioner, något som ger ålderspensionssystemet större relativ vikt. Vidare antas inte medel delas inom hushållet vilket bidrar till att förstärka ålderspensionssystemets intrapersonella omfördelning.

stegvis dekomponering av den disponibla livsinkomsten på samma effekt – inkomstspridningen ökar då ålderspensioner (inklusive finansiering) tas bort.

Det bör noteras att den så kallade garantipensionen i dessa beräkningar räknas in i ålderspensionen. Garantipensionen utgör en grundnivå för individer med låg inkomstgrundad pension. Dessa finansieras genom allmänna skattemedel vilka huvudsakligen är progressiva. Vidare betalas arbetsgivaravgifter på inkomster som överstiger det så kallade intjänandetaket och därmed inte ger motsvarande pensionsförsäkring. Dessa faktorer bidrar till att skapa en interpersonell omfördelning inom ålderspensionssystemet vilket annars bör vara fördelningsmässigt neutralt över livscykeln då inbetalda avgifter i genomsnitt motsvaras av utbetald pension.

Omfördelning genom offentliga subventioner

Tabell 5.10 visar de (indirekt) beräknade genomsnittliga omfördelningskomponenterna per livsinkomstkvintil för de offentliga subventionerna. Självfinansieringsgraden för de offentliga subventionerna är 85 procent vilket är något högre än självfinansieringsgraden för samtliga system. Huvuddelen av den intrapersonella omfördelningen, 80 procent, består här av en årlig rundgång. Denna beror i hög utsträckning på att individer antas ha en delad ekonomi inom hushållen (barn och ungdomars subventioner och föräldrars skatter). Frånräknat den årliga rundgången ges en interpersonell omfördelning på 45 procent (och således en livscykelomfördelning om 55 procent) vilket ska jämföras med en nivå på 32 procent för samtliga system.

Detta stämmer väl överens med bilden av de offentliga subventionerna som ett heterogent system bestående av både delar som uppfattas som typiskt livsomfördelande (t.ex. barnomsorg, grundskola och äldreomsorg) och typiskt individomfördelande (t.ex. arbetsmarknadsåtgärder).

Tabell 5.10 Omfördelningskomponenter samt livsnetto per livsinkomst för offentliga subventioner (tkr)

Genomsnitt per individ

Kvintil av livsinkomst	1	2	3	4	5	alla
Skatter	1939	2523	2861	3286	4043	2931
%	-34	-14	-2	+12	+38	
Subventioner	2840	3109	3109	2922	2679	2932
%	-3	+6	+6	0	-9	
INTRA (år)	1709	2026	2089	2115	2131	2014
INTRA (liv)	228	442	595	666	516	490
Livsnetto	901	586	249	-364	-1365	0

Anm.: Rader markerade med % avser förhållande till totalt populationsgenomsnitt (procentuell skillnad).
Källa: SESIM

Det genomsnittliga skatteuttaget per livsinkomstkvintil liknar det för ålderspensionssystemet, det vill säga något mer jämnt fördelat mellan kvintilerna jämfört med i analysen av samtliga system. För de genomsnittliga subventionsbeloppen gäller att kvintilerna ett och (i synnerhet) fem har lägre nivåer än populationsgenomsnittet medan kvintilerna två och tre ligger något över.

Fördelningen av genomsnittligt livsnetto per inkomstkvintil visar på större skillnader mellan den högsta och lägsta kvintilen jämfört med ålderspensionssystemet vilket tyder på att de offentliga subventionerna åstadkommer en större omfördelning av livsinkomster än ålderspensionssystemet. Även här beror detta delvis på att omfattningen av offentliga subventioner är större än omfattningen av pensionssystemet.

Omfördelning mellan olika grupper inom olika system

Tabell 5.11, tabell 5.12 och tabell 5.13 visar genomsnittliga (per individ) livsnetton uppdelat på kön, utbildningsnivå och livsinkomstkvintil. De totala genomsnittliga livsnettona för den offentliga sektorn har dekomponerats i ålderspensioner, offentliga subventioner samt övriga transfereringar. Posten övriga transfereringar har beräknats som differensen mellan netton för samtliga system och summan av netton för ålderspensioner och offentliga subventioner från de partiella analyserna. Beräkning av posten genom en analys där ålderspensioner och offentliga subventioner

exkluderas simultant ger inte exakt samma resultat. Anledningen är att omfattningen av de system som exkluderas påverkar skatteuttagets omfördelningsegenskaper i de kvarvarande delarna av den offentliga sektorn. Detta eftersom fördelningen av olika skatteslag, med sinsemellan olika fördelningsegenskaper, förändras om den totala omfattningen förändras enligt den metod som används för bestämning av skatteuttaget (se tabell 5.2). Skillnaderna mellan beräkningarna i detta fall är dock små och påverkar inte de generella slutsatserna.

Tabell 5.11 Livsnetton för olika delar av den offentliga sektorn för män och kvinnor

Genomsnitt per individ

	Män	Kvinnor
Ålderspensioner	-292	286
Subventioner	-427	418
Övrigt	-244	237
Samtliga system	-963	941

Källa: SESIM

Tabell 5.11 visar att nära hälften av de medel som över en livstid i genomsnitt som överförs från män till kvinnor gör det via de offentliga subventionerna. Av den kvarvarande överföringen går en något större del via ålderspensionssystemet jämfört med övriga transfereringar.

Tabell 5.12 Livsnetton för olika delar av den offentliga sektorn för olika utbildningsnivåer

Genomsnitt per individ

	Låg	Mellan	Hög
Ålderspensioner	-4	9	-10
Subventioner	141	63	-134
Övrigt	687	57	-312
Samtliga system	824	129	-453

Källa: SESIM

Tabell 5.12 visar att mycket lite av den överföring som görs mellan utbildningsnivåerna görs via ålderspensionssystemet. Faktum är att de med lägst utbildning har ett negativt genomsnittligt netto mot ålderspensionssystemet. Denna effekt torde huvudsakligen uppstå genom att den förväntade medellivslängden är lägre för individer med lägre utbildningsnivå.

Subventionerna bidrar till en viss omfördelning från högutbildade till lågutbildade men det faktum att subventionerna tenderar att stiga med utbildningsnivån (se tabell 5.8) minskar storleken på denna omfördelning.

Störst omfördelning mellan utbildningsnivåerna görs via övriga transfereringssystem.

Tabell 5.13 Livsnetton för olika delar av den offentliga sektorn för olika livsinkomstkvintiler

<i>Genomsnitt per individ</i>					
Kvintil av livsinkomst	1	2	3	4	5
Ålderspensioner	449	290	138	-165	-708
Subventioner	901	586	249	-364	-1365
Övrigt	1150	475	41	-470	-1189
Samtliga system	2500	1351	428	-999	-3262

Källa: SESIM

Tabell 5.13 visar att den omfördelning som görs från individer med hög livsinkomst till individer med låg livsinkomst inte i första hand görs via ålderspensionssystemet utan genom offentliga subventioner och övriga transfereringar, vilket är konsistent med tidigare resultat. Omfördelningen som görs via subventioner och övriga transfereringar är av ungefär samma storleksordning.

6 Sammanfattning och diskussion

6.1 Inledning

I denna bilaga till långtidsutredningen behandlas ett antal frågeställningar kring inkomstfördelning och inkomstomfördelning. Tidsperspektivet är centralt i analyserna och livscykelperspektivet lyfts fram och kontrasteras mot det mer traditionella ettårs-perspektivet som oftast används vid fördelningsanalyser. Förutom tidsperspektivet analyseras också inkomstbegreppet betydelse, och då i synnerhet vilken effekt offentligt finansierade tjänster har på inkomstfördelningen.

Analyserna delas in i tre separata kapitel där indelningen grundas dels på frågeställning och dels på analysmetod. I kapitel 3 analyseras hur inkomstfördelningen påverkas om inkomstbegreppet utvidgas till att omfatta offentligt finansierad privat konsumtion och av att mätperioden förlängs från ett år till hela livscykler.

I kapitel 4 kartläggs vilka direkta omfördelningseffekter olika skatter, transfereringar och offentlig konsumtion har och hur dessa effekter påverkas av mätperiodens längd.

I kapitel 5, slutligen, undersöks hur stor del av den omfördelning som sker via den offentliga sektorn som medför en faktisk omfördelning av resurser mellan individer och hur stor del som kan betraktas som rundgång eller periodisering.

Samtliga analyser baseras på inkomster genererade av den dynamiska mikrosimuleringsmodellen SESIM. I modellen genereras hela livsförlopp för ett stort antal individer. Livsförloppen beskrivs i termer av ett stort antal variabler rörande individernas (och hushållens) egenskaper, varav inkomsterna utgör en del.

Dynamisk mikrosimulering i allmänhet och SESIM i synnerhet beskrivs översiktligt i kapitel 2 och en mer ingående beskrivning återfinns i appendix A.

6.2 Inkomstfördelningen – effekter av offentlig konsumtion och undersökningsperiodens längd

Den justerade disponibla årsinkomsten används ofta som en approximation för nytta. Genom att förlänga mätperioden och utvidga inkomstbegreppet görs i denna rapport ett försök att komma lite närmre, om inte individernas nytta, så åtminstone deras faktiska ekonomiska standard.

När värdet av subventionerad offentlig konsumtion läggs till den disponibla inkomsten minskar ojämnheten i inkomstfördelningen, mätt som Gini-koefficienten⁴⁸, från 0,217 till 0,189 i ett tvärsnittsperspektiv. Inkomstojämnheten minskar alltså med cirka 10 procent. Denna ganska måttliga effekt⁴⁹ förklaras till viss del av att indirekt beskattning här inkluderas. Denna verkar regressivt och tar ut en del av utjämningseffekten av offentlig konsumtion.

Att förändra tidsperspektivet och analysera livsinkomster i stället för årsinkomster har desto större effekt. Gini-koefficienten minskar från 0,189 för årsinkomster till 0,086 för livsinkomster, en *minskning på nästan 60 procent*. Att fördelningen av livsinkomster är jämnare än motsvarande fördelning av årsinkomsterna beror på att individernas position i inkomstfördelningen varierar mellan åren, t.ex. kompenseras ofta låga inkomster ett givet år av högre inkomster under andra år.

Fördelning av de inkomster som genereras på marknaden är mycket ojämn. Faktorinkomsten (summan av löner, företagarinkomster och kapitalinkomster) är vid den fjärde kvintilgränsen⁵⁰ mer än 15 gånger så hög som vid den första kvintilgränsen. När hänsyn tas till den utjämnning som sker via den offentliga sektorn och som uppkommer av att mätperioden förlängs minskar avståndet till knappt 30 procent, *mer än 90 procent av den ursprungliga skillnaden försvinner*. Gini-koefficienten minskar med drygt 80 procent.

⁴⁸ Gini-koefficienten är ett vanligt förekommande mått på ojämnheten i inkomstfördelningen. Måttet kan anta värden mellan 0 och 1 där 1 motsvarar perfekt jämlikhet, alla har lika stora inkomster, och 0 total ojämlikhet, en person uppbär samtliga inkomster. Koefficientens definition och egenskaper beskrivs ingående i kapitel 3.

⁴⁹ Motsvarande effekt har i tidigare studier kvantifierats till cirka 20 procent.

⁵⁰ Den fjärde kvintilgränsen avser den inkomstnivå som krävs för att tillhöra de 20 procent av befolkningen som har högst inkomster. De 20 procent som har lägst inkomst har, på motsvarande sätt, inkomster som understiger den första kvintilgränsen.

6.3 Omfördelningseffekter

Skatter, bidrag och offentlig konsumtion påverkar inkomstfördelningen både direkt och indirekt. Fördelningen av faktorinkomster påverkas indirekt dels via individernas val av arbetstid, dels genom pris- och lönepåverkan. Om skatte- och bidragssystemen inte fanns eller var annorlunda utformade skulle sannolikt individernas val av arbetstid, lönestrukturen och avkastningskraven på kapital m.m. sett annorlunda ut. På lång sikt påverkar t.ex. subventionerad utbildning fördelningen av inkomstmöjligheter och således även inkomsterna. I denna rapport analyseras enbart de direkta effekterna genom att såväl tvärsnitts- som livsinkomstfördelningen delas upp med avseende på inkomstslag. Syftet med detta är att beskriva hur olika skatter och bidrag omfördelar ekonomiska resurser mellan hushåll, samt att avgöra hur denna omfördelning påverkas om mätperiodens längd varierar.

Ett inkomstslags effekt på inkomstfördelning beror dels på dess fördelning i förhållande till fördelningen av övriga inkomster, dels på dess omfattning. Ett inkomstslag som har en väldigt tydlig hög- eller låginkomstprofil kan ha svaga eller obefintliga omfördelningsegenskaper totalt sett om dess omfattning är ringa.

Inkomstslagen kan delas in i *regressiva*, dessa inkomstslag medverkar till ojämnheten och drar alltså isär inkomstfördelningen, och *progressiva* som verkar åt motsatt håll och trycker ihop fördelningen.

I tvärsnittet, dvs. under ett år, är det äldreomsorg, efterlevandepension, faktorinkomster, indirekta skatter och avtalspensioner som verkar regressivt. I ett livsinkomstperspektiv är avtalspensioner, äldreomsorg, faktorinkomster, indirekt beskattning och ålderspension regressiva inkomstslag.

De inkomstslag som inte verkar regressivt är i olika utsträckning progressiva och verkar alltså utjämnande. I ett tvärsnittsperspektiv är det socialbidrag, bostadsbidrag, föräldrapenning, subventionerna för komvuxstudier och förtidspension som är mest progressiva. De inkomstslag som är mest progressiva i ett livsinkomstperspektiv är förtidspension, bostadstillägg för pensionärer, socialbidrag, bostadsbidrag och subventioner vid komvuxstudier.

De flesta inkomstslag påverkar inkomstfördelningen åt samma håll över livet som under ett år, ytterst få inkomstslag byter tecken. Effekternas styrka påverkas dock av tidsperspektivet. Generellt har inkomstslag som på något sätt relaterar till livsinkomsten, såsom

t.ex. ålders- och avtalspensioner, starkare effekter över livet än under ett enstaka år.

Ett inkomstslag som uppvisar motsatt effekt över livet jämfört med under ett år är utbetald ålderspension som verkar progressivt under ett år men regressivt över livet. Ålderspensionärer har i snitt relativt låg totalinkomst eftersom de inte har några övriga inkomster vilket ger progressivitet i tvärsnittet. En hög pension ger i allmänhet inte en inkomst i toppen av fördelningen utan någonstans mellan mitten och toppen. De som har höga pensioner har dock haft höga inkomster när de var aktiva varför de återfinns i toppen av livsinkomstfördelningen. Ålderspensionerna verkar därför regressivt över livet. En stegvis analys av ålderspensionssystemet i sin helhet, där hänsyn även tas till finansieringen, visar att effekten sett i ett livscykelperspektiv är progressiv. Detta beror på att finansieringen är progressiv.

6.4 Den offentliga sektorns omfördelning och periodiserande roll

Ekonomiska transaktioner mellan de enskilda hushållen och den offentliga sektorn leder till en omfördelning av ekonomiska resurser i samhället. Främst sker detta genom skatte- och transfereringssystemen men även genom olika former av subventionerad verksamhet som helt eller delvis finansieras av offentliga medel.

I rapporten analyseras den omfördelning som i genomsnitt görs under en livstid. Omfördelning mellan olika individer i populationen kallas här *interpersonell* omfördelning. Den omfördelning som görs från en individ till samma individ benämns *intrapersonell* omfördelning. Den senare omfattar två delkomponenter, omfördelning inom ett och samma år för en och samma individ samt omfördelning över individens livscykel.

Interpersonell omfördelning

Individer kan få transfereringar och subventioner för offentliga tjänster samtidigt som de betalar skatter och avgifter. Om en individ får större transfereringar än hon betalar skatter är individens netto gentemot den offentliga sektorn positivt det året. Omvänt uppkommer negativa netton genom att individen betalar

högre skatter än hon får igen i form av transfereringar. Summan av individens årliga netton över livet är lika med individens *livsnetto*. Summan av de positiva livsnettona motsvarar de medel som omfördelas *mellan individer* (interpersonell omfördelning).

Intrapersonell omfördelning

Om en individ under ett visst år får transfereringar samtidigt som hon betalar skatt kan det ses som om transfereringarna finansieras av de betalade skatterna, här benämnt *årlig rundgång*. En individ som har ett positivt livsnetto gentemot den offentliga sektorn kan ändå för vissa år ha negativa netton, det vill säga då skatterna överstiger transfereringarna. Dessa skatter kan då antas finansiera erhållna transfereringar under andra år i individens liv, här benämnt *livscykelrundgång*. Sammantaget ger den årliga rundgången och livscykelrundgången ett mått på hur mycket av de erhållna transfereringarna som är *självfinansierade*. Det vill säga har, via skatter, finansierats av individen själv vid något tillfälle i livet.

Hur stor är den reella omfördelningen?

Beräkningarna visar att drygt 80 procent av alla transfereringar och subventioner är självfinansierade, dvs. de betalas av individens egna skatter vid något tillfälle under livet. Endast 18 procent av den omfördelning som sker via skatter, transfereringar och offentlig konsumtion är genuin omfördelning mellan individer.

Livsnettot mot den offentliga sektorn, dvs. summan av alla subventioner och transfereringar som erhålls under livet minus motsvarande skatter, är i genomsnitt positivt för individer med låg livsinkomst och negativt för individer med hög livsinkomst. Detta innebär att det sker en utjämning även av livsinkomster via den offentliga sektorn. En individ ur den högsta livsinkomstkvintilen gör över livet i genomsnitt en nettoförlust på ca 3,3 mkr mot den offentliga sektorn och en individ ur den lägsta kvintilen i genomsnitt en nettovinst på ca 2,5 mkr.

Omfördelningen sker i första hand genom offentliga subventioner och övriga transfereringar och inte genom ålderspensionssystemet.

Graden av självfinansiering av transfereringar och subventioner ökar med individernas livsinkomst men även individerna med de lägsta 20 procent livsinkomsterna finansierar över 60 procent av sina transfereringar och subventioner själva vid någon tidpunkt i livet.

Utveckling över livsrymden

Den genomsnittlige skattebetalaren är yngre än den genomsnittlige uppbäraren av transfereringar och subventioner. Detta medför att det kumulativa livsnettot mot den offentliga sektorn i genomsnitt är negativt under större delen av livet och positivt först i höga åldrar. Den genomsnittlige individen uppnår ett permanent positivt netto mot den offentliga sektorn vid 87 års ålder. Detta medför att det sker en omfördelning av resurser från individer med korta liv till individer med längre liv. Detta motverkas i viss mån av det positiva samband mellan utbildning (och därmed livsinkomst) och förväntad livslängd som kan observeras i verkligheten och som efterbildas i SESIM.

Omfördelning mellan olika grupper

Förutom den omfördelning mellan individer med olika livsinkomster som görs via den offentliga sektorn förekommer även omfördelningar mellan grupper. Omfördelning från män till kvinnor medför att *män under sin livstid i genomsnitt gör en nettoförlust på nästan 1 miljon kronor mot den offentliga sektorn medan kvinnor gör en lika stor nettovinst*. De offentliga subventionerna genererar nära hälften av denna omfördelning.

Omfördelning från individer med hög utbildning till individer med lägre utbildning medför att en individ med eftergymnasial utbildning i genomsnitt gör en nettoförlust på ca 0,45 miljoner kronor mot den offentliga sektorn och en individ med grundskoleutbildning gör en nettovinst på 0,8 miljoner kronor. Denna omfördelning uppkommer inte i första hand via ålderspensioner eller offentliga subventioner utan via övriga transfereringar.

Ålderspensionssystemet bidrar naturligtvis i huvudsak till att periodisera individens inkomster över livsrymden. Självfinansieringsgraden är för ålderspensioner högre än för hela den offentliga

sektorn. En viss omfördelning från individer med hög livsinkomst till individer med låg livsinkomst görs dock. Främst beror detta på garantipensioner och på de arbetsgivaravgifter som betalas över intjänandetaket.

De offentliga subventionerna består av olika system som är relativt heterogena med avseende på dess omfördelningsegenskaper. Totalt sett genererar subventionerna en hög årlig rundgång. Frånräknat den årliga rundgången är den interpersonella omfördelningen genom subventioner betydligt högre än för den offentliga sektorn som helhet.

6.5 Slutsatser/diskussion

Denna rapport beskriver omfördelningsprocessens egenskaper i ett antal dimensioner. Ambitionen har inte varit att dra några slutsatser om vilka implikationer våra analyser har för de olika systemens effektivitet eller måluppfyllelse. Några förslag till regeländringar eller reformer lämnas inte heller. Förhoppningen är att analyserna kan fungera som underlag för en förutsättningslös diskussion om den offentliga sektorns roll för inkomstfördelningen. Utan att föregripa en sådan diskussion kan det vara intressant att lyfta fram några slutsatser och ställa några frågor som resultaten ger upphov till.

Fördelningen av livsinkomster är betydligt jämnare än fördelningen av årsinkomster och den offentliga sektorns omfördelningseffekt är större under ett givet år än över livet. Vid analyser av olika omfördelningssystem och vid reformer av desamma är det därför av avgörande betydelse att livscykelperspektivet beaktas. En ensidig fokusering på effekterna under ett kalenderår leder lätt till felaktiga slutsatser och reformer riskerar att, på lite längre sikt, ha andra effekter än de avsedda.

Självfinansieringsgraden hos offentliga subventioner och transfereringar är omfattande och endast en liten del av all den omfördelning som sker via den offentliga sektorn medför en faktisk resursomfördelning. Mot bakgrund av den förväntade framtida demografiska utvecklingen och de problem med finansiering av den offentliga sektorn som redan i dag kan förutses är det därför viktigt att diskutera den offentliga sektorns åtagande. Prioriteringar kom att behöva göras, en del av periodiseringen kan rimligen överlämnas till individerna själva.

Appendix A: SESIM – en dynamisk mikrosimuleringsmodell

A.1 Inledning

Utvecklingen av SESIM påbörjades under 1997. Syftet var då att skapa ett beräkningsverktyg att använda i arbetet med att reformera studiemedelsystemet. Arbetet resulterade i, förutom beräkningsunderlag till reformarbetet, rapporten ”Studiebidraget i det långa loppet”, Ericson och Hussenius (2000)⁵¹.

Sedan 2000 har fokus för modellarbetet ändrats från studiemedel till ålderspensioner. Detta har föranlett en vidareutveckling av SESIM för att dels förbättra de befintliga modellkomponenter som bedöms viktiga för den nya analysinriktningen, dels för att introducera komponenter som tidigare saknats.

En långsiktig målsättning är att SESIM ska utgöra en generell plattform där ett relativt brett spektrum av analyser kan utföras antingen direkt eller efter mindre modifieringar av modellen.

Detta appendix syftar till att komplettera den korta beskrivning av SESIM som ges i rapportens huvudtext. I avsnitt A.2 ges en allmän introduktion till området mikrosimulering där både statiska och dynamiska modeller diskuteras kort. I avsnitt A.3 presenteras SESIM:s grundläggande modellstruktur. Därefter diskuteras de datakällor som används i SESIM i avsnitt A.4, samt de modifieringar av dessa som görs. Avsnitt A.5 och A.6 diskuterar stokastisk simulering och dess egenskaper, särskilt för den modell som används för simulering av inkomster.

Avsnitt A.7 beskriver SESIM:s tekniska plattform och i avsnitt A.8 presenteras en förteckning över de stokastiska modeller som ingår i SESIM.

⁵¹ Ds 2000:19.

A.2 Vad är en mikrosimuleringsmodell?

Termen mikrosimulering används för att beskriva en typ av beräkningsmodeller i vilka beräkningarna utgår från den minsta beståndsdelen i en population. Vanligen utgörs detta av den enskilda individen även om det i vissa fall kan vara andra ekonomiska enheter som till exempel företag eller något helt annat som t.ex. fiskar, atomer eller bilar.

Den typ av modeller som berörs här utgår ifrån den enskilda individen och är uteslutande konstruerade för att användas i analyser av samhällsekonomisk karaktär, till exempel att studera olika aspekter av skatte- och transfereringssystem.

I begreppet mikrosimulering ryms flera olika typer av modeller. Man brukar vanligen göra skillnad mellan de två huvudtyperna statiska och dynamiska mikrosimuleringsmodeller. Denna kategorisering är dock något flytande och ibland kan särdrag från båda dessa modelltyper finnas representerade i en och samma modell.

A.2.1 Statisk mikrosimulering

Den statiska mikrosimuleringsmodellen utgår vanligen från en databas innehållande detaljerad information om individernas olika egenskaper såsom till exempel hushållets sammansättning, utbildningsnivåer och olika typer av inkomstutgifter. Databasen utgör ofta ett stickprov draget från den population vilken analysen syftar till.

I modellen finns detaljerade beskrivningar av de aktuella skatte- och transfereringssystemen vilket möjliggör att individernas skatter, transfereringar och disponibla inkomster kan beräknas.

Det primära syftet med den statiska mikrosimuleringsmodellen är att utvärdera effekter av olika reformförslag. Det kan röra sig om ändringar av de befintliga skatte- och transfereringssystemen eller införandet av nya komponenter i dessa (eller helt nya system). Med effekter avses vanligen såväl totala budgeteffekter för den offentliga sektorn som effekter för den enskilde individen - vem vinner/förlorar på reformen? Som en följd av regelförändringarna kan förslagens fördelningsprofil bestämmas, det vill säga hur systemförändringen bidrar till en omfördelning av ekonomiska resurser mellan individer/hushåll i olika inkomstlagen.

På grund av eftersläpningen i insamlingen av deklarationsuppgifter är de mest aktuella databaserna minst två år gamla. För att kunna göra beräkningar som återspeglar en aktuell omvärld med avseende på till exempel befolkningens sammansättning, inkomster och sysselsättningsnivåer finns därför behov av att kunna göra prognoser två eller fler år in i framtiden. Dessa prognoser görs genom att tillämpa så kallat *statiskt åldrande* på modellbefolkningen. Med statistiskt åldrande av befolkningen avses tekniker som förändrar modellbefolkningens sammansättning på ett önskat sätt utan att förändra egenskaperna hos de ingående individerna. Detta åstadkoms genom att individerna tilldelas nya urvalsvikter. På så sätt kan förändringar i demografisk sammansättning, arbetskraftsdeltagande och andra nyckeltal kontrolleras. En individ som i det ursprungliga datamaterialet var 17 år och bedrev gymnasiestudier kommer således att fortsätta vara 17 år gammal och bedriva gymnasiestudier oavsett hur långt fram i tiden prognosen görs. Den enda skillnaden är hur många individer i totalpopulationen individen antas representera.

Ett exempel på en statistisk mikrosimuleringsmodell är Statistiska centralbyråns (SCB) fördelningsmodell FASIT som i över ett decennium har använts inom regeringskansliet och under de senaste åren även fått en spridning utanför statsförvaltningen.

Liknande modeller används i ett flertal länder, däribland Norge (LOTTE), Danmark (LOV), Australien (STINMOD) samt Storbritannien (POLIMOD, TAXBEN).

En konsekvens av att individernas egenskaper inte förändras i simuleringen är att endast de statiska, eller direkta, effekterna av en systemförändring kan beräknas. Eventuella beteendeförändringar som uppstår till följd av en reform fångas inte upp. Till exempel kommer en nivåförändring av inkomstbeskattningen inte ha någon effekt på arbetsutbudet.

Vissa statistiska modeller har därför infogat så kallade beteendeeckvationer vilka gör det möjligt att återge en viss dynamik, eller beteendeförändringar, i modellpopulationen. I FASIT-modellen finns till exempel möjligheter att göra beräkningar av individernas/hushållens arbetsutbud i stället för att använda det faktiskt observerade. Detta antas då bero på hur de offentliga systemen påverkar hushållets disponibla inkomst vid en viss sysselsättningsnivå. Modeller med så kallade beteendeeckvationer kan sägas vara hybrider mellan statistiska och dynamiska mikrosimuleringsmodeller.

A.2.2 Dynamisk mikrosimulering

Liksom den statiska modellen utgår även den dynamiska modellen från en databas över den aktuella analyspopulationen, vanligen ett stickprov. Här används två olika strategier – antingen används ett så nydraget stickprov som möjligt eller ett som daterar sig ett antal år tillbaka i tiden. Fördelen med det nya stickprovet är att analyserna utgår från en så färsk datakälla som möjligt. Det gamla stickprovet kan emellertid ge värdefulla möjligheter till utvärdering av modellens funktionalitet genom att en framskrivning av datamaterialet först utförs med hjälp av modellen varvid en jämförelse mellan framskrivna data och faktiskt observerade data kan göras.

Den primära skillnaden gentemot den statiska modellen är det sätt på vilken modellpopulationen framskrivs. Till skillnad från den statiska modellen så förändrar den dynamiska modellen hela tiden egenskaperna hos den ingående modellpopulationen. På så sätt söker modellen avbilda det sätt på vilket en faktisk population åldras - individer blir äldre, föder barn, flyttar hemifrån, börjar arbeta och så vidare. Åldrandeprocessen utgår ifrån ett förutbestämt antal olika händelser en individ kan utsättas för. Sannolikheten att en viss individ ska utsättas för en viss händelse estimeras med hjälp av olika statistiska modeller. De statistiska modellerna kan sedan användas för att återge de i datakällorna observerade händelseförloppen i syfte att skriva fram modellpopulationen. Detta åstadkoms vanligen genom att så kallad *Monte Carlo-simulering* används.

De flesta dynamiska simuleringsmodeller låter modellpopulationens egenskaper uppdateras stegvis i jämna tidsintervall, oftast ett år. Valet av uppdateringsfrekvens är dock starkt beroende av vilka datakällor som finns tillgängliga för estimering av de statistiska modeller som driver simuleringen. I vissa modeller används kortare tidsintervall, till exempel månader, och i andra (sällsynta) fall antas händelser kunna inträffa när som helst. Många av de processer som modelleras pågår sannolikt i kontinuerlig tid eller i små diskreta tidsintervall. Här får modellering i ettårssteg ses som en (grov) approximation. För en diskussion av val av uppdateringsfrekvens i dynamisk mikrosimulering se Galler (1997).

Givet att ett visst tidsintervall används i modellen ansätts en sekvens i vilken de olika händelserna kan inträffa. Detta ger modellen en rekursiv struktur där händelser som antas inträffa senare under året endast kan använda information från händelser som inträffat

tidigare. Ett exempel på detta är antagandet att endast kvinnor som överlever året kan komma ifråga för barnafödande. Simuleringen bör då hantera mortalitet före fertilitet och i den senare endast utsätta överlevande kvinnor för risk. Den ordning som väljs för simuleringen är godtycklig men ofta förläggs de grundläggande händelserna, till exempel de som har en demografisk karaktär, först i sekvensen. Det är viktigt att de statistiska modellerna estimeras under beaktande av den sekvens som ska användas i simuleringen. Detta för att riskgruppen vid estimering ska vara konsistent med den vid simulering.

I vissa modeller används en teknik där en individs hela liv simuleras innan simuleringen för nästa individ påbörjas, det vill säga en *individvis simulering*. Det vanligaste är dock att använda en *periodvis simulering*, det vill säga en struktur där hela modellpopulationen genomlöps en gång per tidsintervall. På så sätt underlättas hanteringen av situationer där händelser för en individ är beroende av händelser för en annan individ. Ett sådant exempel är så kallad *kalibrering*, där modellen tvingas att producera ett i förväg bestämt resultat, avseende en viss analysvariabel. I ett första steg simuleras då hela modellpopulationen varvid ett visst antal händelser genereras. För att uppnå det förutbestämda resultatet manipuleras sedan de estimerade modellerna, eller de sannolikheter modellerna producerar, varvid modellpopulationen simuleras igen.

Liksom hos de statiska modellerna finns i de dynamiska vanligen de offentliga systemen avbildade, dock i något mindre detaljerat utförande. Det primära syftet med de dynamiska modellerna är ofta att göra beräkningar över mycket lång sikt för att analysera system där tidsaspekten är avgörande. Ett viktigt exempel är analyser av ålderspensionssystem för vilka det krävs att beräkningar av individens årliga avsättning till systemet kan göras likväl som beräkningar av hur det totala pensionskapitalet växer över tiden fram till pensionsfallet. Ett annat exempel på sådana långsiktiga system är studiemedelssystemet. För analys av system där kumulativa processer är viktiga har de dynamiska modellerna en fördel gentemot de statiska eftersom åldrandeprocessen mer efterliknar den i verkliga livet.

Beteendekvationer används även i vissa dynamiska modeller även om huvuddelen av de estimerade relationerna i någon mening kan sägas vara rent statistiska. Skillnaden består i att beteendemodellen låter individen, givet hennes egenskaper, reagera på förändringar i omvärlden, till exempel genom ändringar i regelverken för

olika skatter och transfereringar. Den statistiska modellen utgår endast från individens egenskaper.

I dynamisk mikrosimulering används ofta externa prognoser av den makroekonomiska utvecklingen som indata till simuleringen. Detta beror huvudsakligen på att mikrosimuleringsmodellen i sig inte utgör en komplett beskrivning av ekonomin vilket begränsar dess förmåga att generera makroekonomiska prognoser. Till exempel modelleras individernas/hushållens arbetsutbud ofta utan att efterfrågan på arbetskraft modelleras. Ett annat exempel kan vara att efterfrågan på vårdplatser modelleras men inte tillgången på sådana. Implicit antas att den modellerade komponenten är styrande; det efterfrågas så mycket arbetskraft som finns till hands och det finns så många vårdplatser som efterfrågas. Styrning av simuleringsmodellen till externa prognoser åstadkoms vanligtvis med hjälp olika tekniker för kalibrering. Kalibrering av modellen kan alltså ses som ett sätt att implementera olika scenarier i beräkningarna.

Ytterligare en fråga att ta ställning till i dynamisk mikrosimulering är om modellen ska simuleras i ett stabilt tillstånd, på engelska *steady state*, eller inte. Att simulera i ett stabilt tillstånd innebär att de processer som simuleras antas fortgå på samma sätt som under den senaste perioden (snarare: under den tidsperiod på vilka de ingående modellerna har estimerats). De samband som gäller under estimeringsperioden kommer således antas gälla för all framtid. För modeller som inte simuleras i ett stabilt tillstånd används i regel olika externa informationskällor som beskriver omvärldens utveckling under prognosperioden, till exempel en demografisk prognos eller en prognos över arbetskraftens storlek och sammansättning. Då dessa resultat påtvingas simuleringsmodellen kan det ses som en (implicit) ändring av de estimerade processerna.

För en mer detaljerad diskussion kring dynamisk mikrosimulering samt genomgång av ett antal befintliga modeller se Rake och Zaidi (2001) samt O'Donoghue (2001a).

A.3 SESIM:s modellstruktur

SESIM kan beskrivas som en uppsättning deterministiska regler, algoritmer och stokastiska modeller som tillsammans används för att stegvis uppdatera egenskaperna hos den initiala modellpopu-

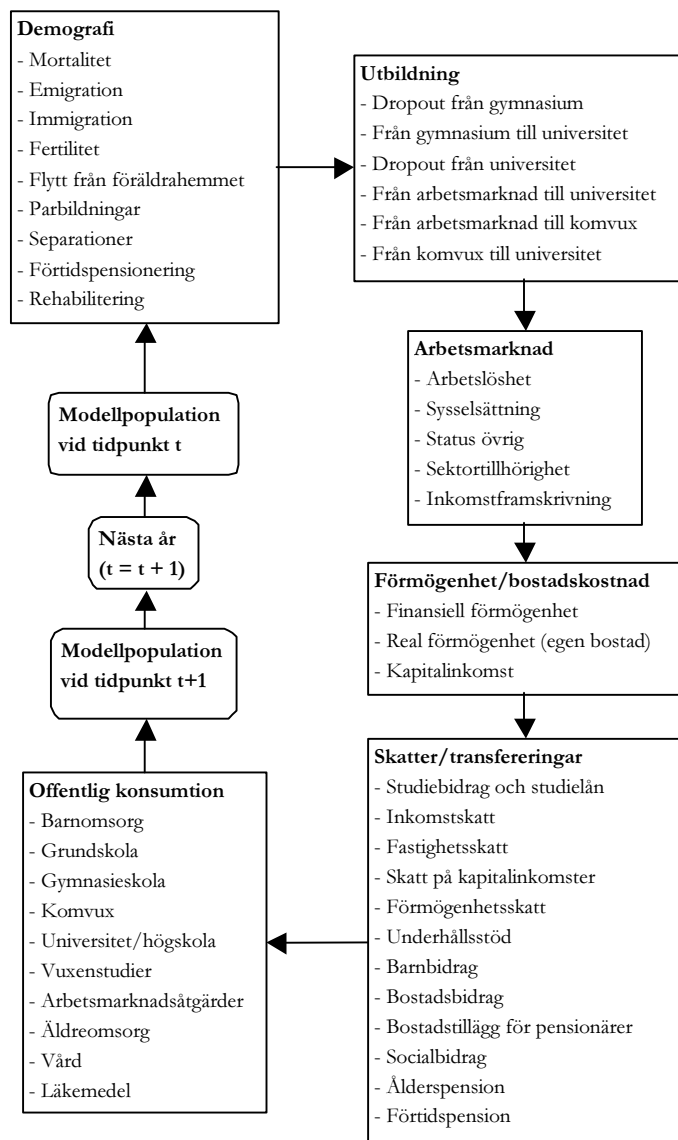
lationen och därigenom skriva fram den över tiden. Uppdateringen tar hänsyn till de processer från den verkliga världen som bedöms vara viktiga att avbilda eftersom de antas påverka de variabler som ska analyseras.

Modellen simulerar de olika processerna en efter en i sekvens, så kallad *rekursiv simulering*. Simuleringen utförs stegvis och varje steg avser ett kalenderår. Inom varje simuleringsår uppdateras varje individ en gång, en *periodvis simulering*. Detta innebär att de händelser som simuleras maximalt kan inträffa en gång under ett år. Somliga händelser leder till absorberande tillstånd och kan således inträffa maximalt en gång. Exempel på detta är dödsfall och utflyttning från föräldrahemmet.

Den rekursiva simuleringssekvensen innebär att ett antagande om den kausala riktningen (och därmed ordningen) mellan de olika händelser som modelleras måste göras. De händelser som antas vara de mest basala placeras därför först i sekvensen. Detta gäller framför allt de händelser som har en demografisk karaktär, till exempel mortalitet och fertilitet.

En direkt effekt av valet av simuleringssekvens är att samtliga modeller måste estimeras under beaktande av den valda sekvensen. Om händelse A simuleras innan händelse B och riskpopulationen för händelse B påverkas av utfallet av händelse A bör estimering av händelse B vara betingad på utfallet av händelse A. Figur A.1 beskriver översiktligt den sekvens som används vid simulering i SESIM. Här framgår i vilken ordning de olika modulerna exekveras men också ordningsföljden för de respektive modellerna inom varje modul.

Figur A.1 Översiktlig beskrivning av beräkningsordning i SESIM



Simuleringarna resulterar i att varje individ i modellpopulationen årligen klassificeras efter individens *huvudsakliga sysselsättning* under året. Användning av en årlig indelning av individernas sysselsättning ger en förenklad bild av verkligheten eftersom många individer har flera sysselsättningar under ett år. Till exempel är det

vanligt att studenter arbetar under delar av året, i synnerhet de år då studierna påbörjas och avslutas. Dessutom förekommer det att olika sysselsättningar pågår simultant, till exempel deltidsstudier och förvärvsarbete. Förenklingen begränsar modellens komplexitet betydligt och gör den mer överblickbar. Nedanstående lista beskriver de olika klassificeringar av individer som används i SESIM.

1. **barn:** avser individer som är yngre än 16 år.
2. **ålderspensionär:** avser individer som påbörjat uttag av ålderspension.
3. **student:** avser individer som bedriver studier vid gymnasium, komvux eller universitet/högskola.
4. **förtidspensionär:** avser individer som uppbär förtidspension eller sjukbidrag.
5. **föräldraledig:** avser kvinnor som föder barn under året.
6. **arbetslös:** avser individer som uppbär någon form av arbetslöshetsersättning. Här inkluderas även individer i arbetsmarknadsåtgärder.
7. **övrig:** avser individer som kan anses falla utanför ramen för den övriga klassificeringen eller som inte har någon *huvudsaklig* sysselsättning under året.
8. **sysselsatt:** avser förvärvsarbetande individer.
9. **emigrerad:** avser individer bosatta utomlands som har svensk pensionsrätt. Observera att klassificeringen i detta fall inte är unik då dessa individer även kan erhålla förtidspension eller ålderspension.

Eftersom simuleringarna sker på individnivå utgörs den minsta analysenheten av just den enskilda individen. Analyser av enskilda individer är dock sällan intressanta varför det vanliga är att individer aggregeras till lämpliga analysgrupper.

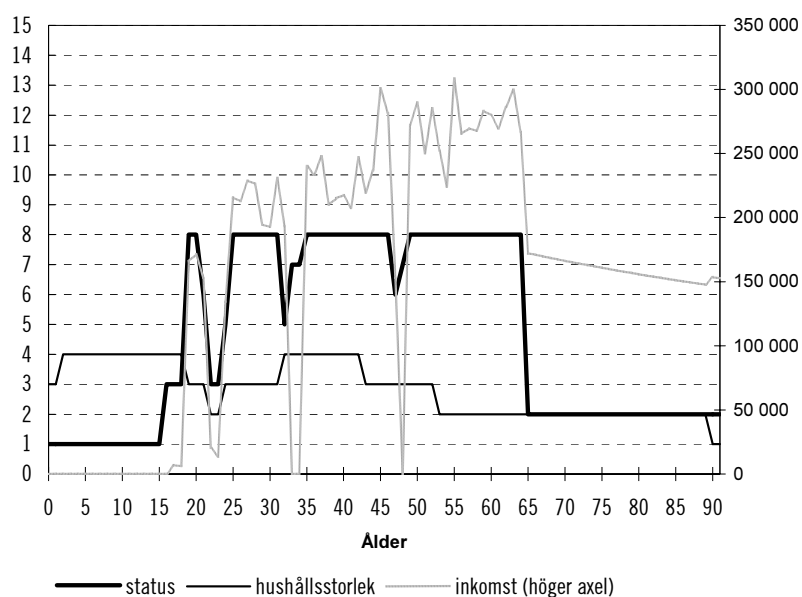
En viktig faktor i de flesta analyser är *hushållen* och dess sammansättning. I SESIM är modellpopulationens individer sammankopplade till hushåll vilka sedan åldras tillsammans med individerna på ungefär samma sätt som i verkligheten; hushåll bildas då individer flyttar ihop och splittras då individer avlider eller separerar. I baspopulationen utgörs modellhushållen av faktiskt observerade hushåll (med vissa modifieringar, se avsnitt A.4.2). Information om hushållstillhörighet och hushållens sammansättning är viktig i SESIM då flera av de stokastiska modellerna använder hushållens egenskaper vid simulering. En annan orsak till

hushållens betydelse är att vissa transfereringar beräknas utifrån information om hushållet snarare än den enskilda individen. Exempel på detta är bostadsbidrag, socialbidrag och barnbidrag.

Ett syntetiskt livsöde

I diagram A.1 visas ett exempel på ett livsöde som genererats av SESIM. Figuren visar hushållsstorlek, inkomst och status (det vill säga årlig klassificering enligt föregående avsnitt) för en kvinna i modellpopulationen under varje år i hennes liv.

Diagram A.1 Ett livsöde genererat av SESIM



Källa: SESIM

Efter att ha varit klassificerad som barn upp till 15 års ålder deltar kvinnan i studier mellan 16 och 24 års ålder. Studierna avbryts av en kort period av förvärvsarbete. Vid 22 års ålder lämnar kvinnan föräldrahemmet och flyttar samman med en man. Vid 24 års ålder får paret sitt första barn. Kvinnan är föräldraledig under året och återgår sedan till sitt arbete. Det andra barnet kommer då kvinnan

är 32 år gammal. De två år som följer denna föräldraledighet är kvinna klassificerad som övrig vilket i detta fall skulle kunna tolkas som att hon stannar hemma med barnen. Efter dessa år återgår kvinnan till arbetet igen och arbetar, med undantag från två år då hon är arbetslös respektive övrigklassificerad, tills hon går i pension vid 65 års ålder.

Kvinnans inkomst varierar något mellan åren, till exempel beroende på att antalet arbetade timmar varierar, men ökar successivt i takt med att hennes arbetslivserfarenhet ökar. De år då kvinnan studerar eller är övrigklassificerad är inkomsterna låga. Under pensionsåldrarna bestäms inkomsten av reglerna för ålderspension och uppvisar därför inte den variation som är typisk för de förvärvsaktiva åldrarna.

När kvinnan är 43 år lämnar det första barnet föräldrahemmet följt av det andra barnet tio år senare. När kvinnan är 90 år gammal avlider hennes man, två år innan kvinnan själv avlider.

På detta sätt alstras under simuleringen en stor mängd information kring modellpopulationens individer och deras hushåll. Den i diagram A.1 redovisade informationen baseras på tre av ca trehundra variabler som berör individen, hushållet eller populationen i SESIM.

A.4 Datakällor

I avsnittet beskrivs de datakällor som huvudsakligen används i SESIM, för estimering och konstruktion av modellpopulation. Vidare beskrivs de korrigeringar av datakällorna som i vissa fall görs samt de metoder för imputering av subventioner för offentlig konsumtion som används.

A.4.1 LINDA – en longitudinell databas⁵²

LINDA-databasen utgör den primära datakällan för SESIM, beträffande såväl estimering av statistiska modeller som konstruktion av modellens baspopulation.

Databasen består av ett urval om ca 3,5 procent av den svenska befolkningen vilket för år 1999 motsvarar ca 308 000 individer. Till

⁵² Longitudinal Individual Data for Sweden. För detaljerad dokumentation se Edin, P.A. & Fredriksson, P. [2000].

dessa urvalsindivider har eventuella hushållsmedlemmar adderats. Totalt omfattar urvalet därför omkring 786 000 individer år 1999. De individer som ingår i urvalet följs upp på en årlig basis varvid data samlas in och infogas i databasen. Även information för förfluten tid samlas in. Viss information, till exempel pensionspoängshistorik, kan följas så långt tillbaka i tiden som till 1960. Databasen har alltså en så kallad longitudinell struktur där de ingående individerna kan observeras vid upprepade tillfällen. Urvalsindivider som av någon anledning försvinner från databasen ett år, t.ex. genom dödsfall eller emigration, ersätts med nya individer på ett sätt som säkerställer databasens representativitet gentemot den svenska befolkningen det givna året.

Databasen innehåller ett stort antal variabler som hämtas från olika register; inkomst- och förmögenhetsstatistiken, löneregistret, folk- och bostadsräkningen, pensionsregistret, pensionspoängsregistret, sjukfallsregistret, föräldrapenningsregistret, arbetslöshetsregistret, arbetslöshetsersättningsregistret, högskole-registret och komvuxregistret.

SESIM:s baspopulation utgörs av ett stickprov om ca 104 000 individer draget från 1999 års LINDA-databas. Till detta har ett urval om ca 8 000 individer draget från Riksförsäkringsverkets pensionspoängsregister adderats. Tilläggsurvalet avser individer som är bosatta utomlands men har svensk pensionsrätt (se avsnitt A.4.3).

De olika statistiska modeller som utgör grunden för SESIM har i de flesta fall estimerats utifrån LINDA. De flesta modeller beskriver dynamiska processer som bygger på information från en sammanhängande period av två eller fler år. För att kunna estimeras sådana modeller krävs följaktligen tillgång till longitudinella data där individer kan observeras upprepade gånger under olika undersökningsår.

I några fall har SCB:s inkomstundersökning HEK⁵³ använts för estimering. Framst berör detta situationer där information rörande hushållens sammansättning är viktig. Eftersom HEK är baserad på en enkätundersökning, vilken sedan kompletteras med registerbaserad information, kan det så kallade kosthushållsbegreppet⁵⁴,

⁵³ SCB:s inkomstfördelningsundersökning, är en årlig enkätundersökning omfattande ca 30 000 individer vilken kompletteras med en mängd information från olika register. Se även avsnitt 2.6.2.

⁵⁴ Ett kosthushåll omfattar individer som är bosatta i samma bostad samt har en gemensam ekonomi och kosthållning.

vilket ligger nära SESIM:s hushållsbegrepp, användas i dessa analyser.

I konstruktionen av SESIM:s baspopulation ingår två huvudsakliga bearbetningar, eller justeringar, av LINDA-databasen. Syftet med dessa bearbetningar är att göra datamaterialet konsistent med vissa definitioner som används i SESIM. Bearbetningarna beskrivs nedan.

A.4.2 Korrigering av hushållens sammansättning

Hushållet i LINDA definieras utgående från befolkningsregistret⁵⁵ och "flergenerationsregistret". Dessa hushåll stämmer i vissa fall inte överens med verkliga förhållanden. Till att börja med förutsetts alla bo där de är folkbokförda, för det andra så behandlas vuxna som bor ihop utan att vara gifta eller ha gemensamma barn som separata hushåll. En jämförelse med andra datakällor där hushållsdefinitionen bedöms vara av högre kvalitet ger att bristerna i hushållsdefinitionen medför två strukturella fel som helst bör åtgärdas.⁵⁶ För det första överskattas andelen ungdomar mellan 18 och 29 som bor kvar hemma. Många som i verkligheten flyttat hemifrån är fortfarande folkbokförda hos sina föräldrar. För det andra underskattas andelen samboende utan barn, framförallt bland de yngre.

Kvarboende

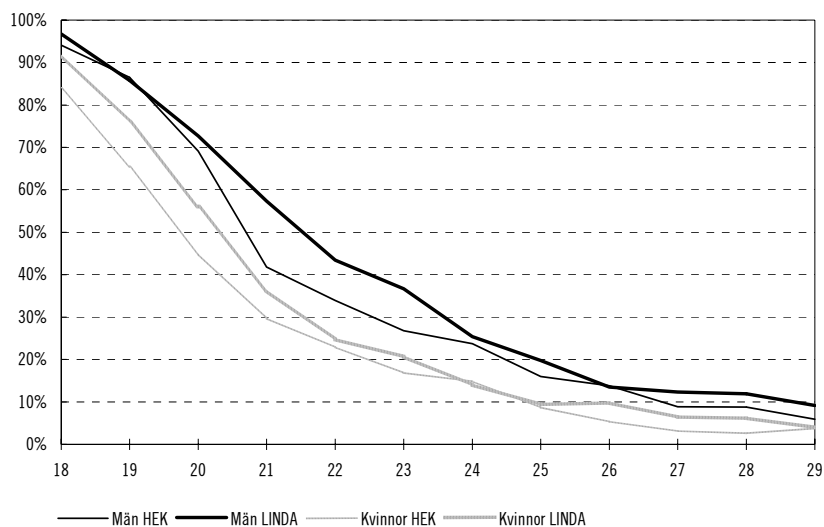
I diagram A.2 nedan jämförs andelen kvarboende, dvs. 18–29 åringar som bor kvar hos sina föräldrar, i LINDA respektive HEK 1999.

⁵⁵ Med befolkningsregistret avses registret över Rikets TotalBefolkning (RTB).

⁵⁶ LINDA jämförs här med inkomstfördelningsundersökningen, HEK. Hushållsbegreppet baseras där på intervjuer och kan därför i större utsträckning antas spegla faktiska förhållanden.

Diagram A.2 Andelen kvarboende i LINDA respektive HEK

Procentuell andel



Källa: SCB, HEK, LINDA, Finansdepartementets beräkningar.

Andelen kvarboende underskattas som synes i de flesta åldrar. Störst är avvikelsen för 21-åriga män där 57 procent redovisas som kvarboende i LINDA jämfört med 42 procent i HEK. För att korrigera detta estimeras först sannolikheten att inte bo hemma för 18–29 åringar i HEK.⁵⁷ Utgående från dessa estimat beräknas sannolikheten att lämna hemmet för de 18–29 åringar som i LINDA är registrerade som kvarboende. Därefter rankas de kvarboende individerna efter sannolikheten att lämna hemmet. Slutligen avskiljs det antal individer som krävs för att erhålla samma frekvenser av kvarboende per ålder och kön som i HEK från föräldrarnas hushåll. De individer som har högst sannolikhet att lämna hemmet är de som väljs.

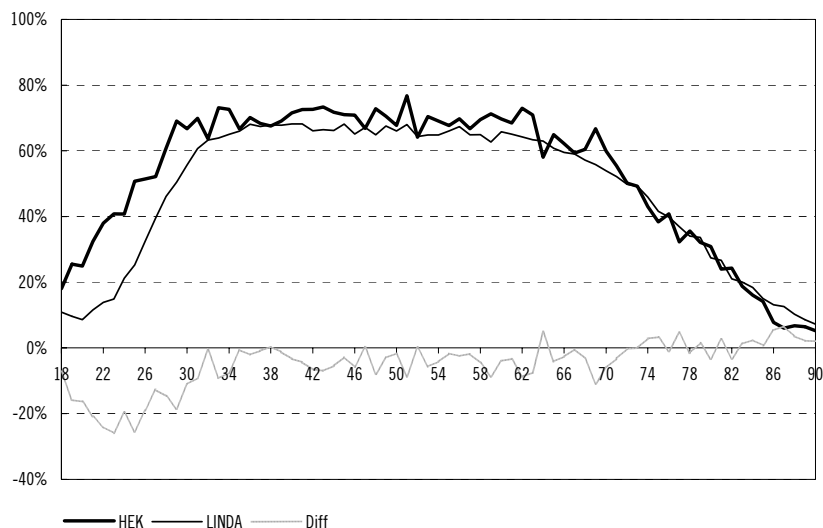
⁵⁷ Sannolikheterna skattas med hjälp av logistisk regression. De förklarande variablerna är följande: kön, ålder, ålder i kvadrat, född i Sverige (dummy), studiemedel/10000, inkomst av tjänst/10000, högsta utbildning (dummys för grundskola och gymnasium) samt interaktionstermer mellan ålder och utbildning.

Samboende

I diagram A.3 nedan jämförs gifta/samboende kvinnor i LINDA respektive HEK 1999.

Diagram A.3 Andelen samboende kvinnor i LINDA respektive HEK

Procentuell andel

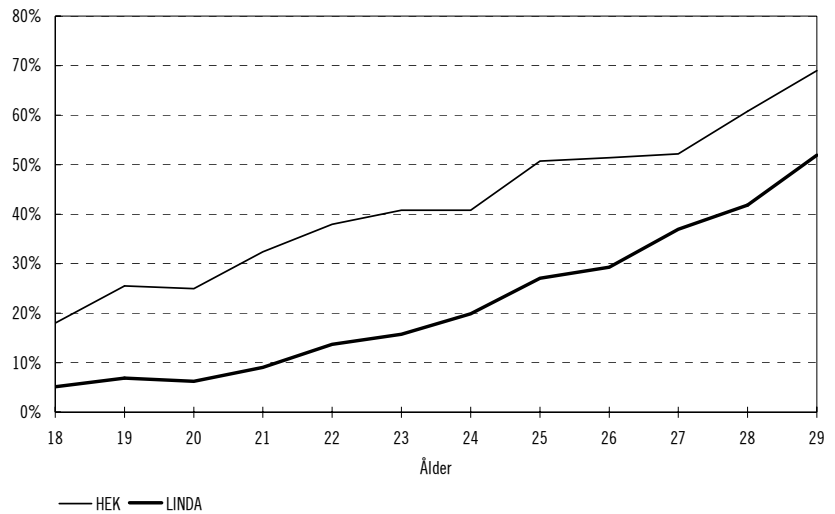


Källa: SCB, HEK, LINDA, Finansdepartementets beräkningar.

Avvikelserna är störst i yngre åldrar, detta är förväntat då samboende utan äktenskap eller barn är vanligare i dessa åldrar. Därför korrigeras samboende endast i åldrarna 18–29 år. Mönstret är ungefär detsamma för män.

När kvarboende flyttas ut från sina föräldrar så blir de automatiskt klassade som ensamstående. Underskattningen av samboende blir därför ännu större och det är därför viktigt att dessa två steg görs i rätt ordning. I diagram A.4 nedan visas andelen gifta/samboende kvinnor i LINDA, efter att andelen kvarboende justerats, respektive HEK.

Diagram A. 4 Andelen gifta/samboende kvinnor i LINDA (efterjusterade kvarboendefrekvenser) respektive HEK



Källa: SCB, HEK, LINDA, Finansdepartementets beräkningar.

För att korrigera de nu ganska stora skillnaderna estimeras först sannolikheten att vara gift/samboende bland 18–29-åriga kvinnor utan barn i HEK.⁵⁸ Med skattningarna från HEK beräknas sannolikheten att bilda par bland 18–29-åriga ensamstående kvinnor utan barn i LINDA. Dessa individer sorteras efter sannolikheten att bilda par. De X kvinnor med högst sannolikhet att bilda par matchas med en ensamstående man utan barn. Mannen antas vara två år äldre och ha samma utbildningsnivå som kvinnan. X bestäms så att frekvenserna överensstämmer med HEK per ålder.

⁵⁸ Sannolikheterna skattas med hjälp av logistisk regression. De förklarande variablerna är följande: ålder, ålder i kvadrat, född i Sverige (dummy), studiemedel/10000, inkomst av tjänst/10000, högst utbildning (dummys för grundskola och gymnasium) samt interaktionstermer mellan ålder och utbildning.

A.4.3 Tillägg av utlandsboende med svensk pensionsrätt

Individer som vid något tillfälle arbetar in pensionsrätt i Sverige behåller den resten av livet även om individen senare väljer att emigrera. Efter pensionering betalas ålderspensionen ut till individen oavsett i vilket land denne då är bosatt. En konsekvens av detta är att SESIM måste ta hänsyn till utrikes migration samt att individer som utvandrar under simuleringsperioden måste kvarstå i modellpopulationen. Detta för att inte generera en genomgående underskattning av de årliga kostnaderna för utbetalda ålderspensioner.

Eftersom LINDA-databasen inte innehåller individer som är bosatta i utlandet under undersökningsåret har i stället ett urval ur Riksförsäkringsverkets pensionspoängsdata använts. Databasen omfattar samtliga individer med registrerad svensk pensionsrätt oavsett vistelseort.

Databasen har dock egenskaper som skiljer sig från SESIM:s övriga baspopulation. Nedan följer en beskrivning av de åtgärder som vidtas för att, så långt det är möjligt, eliminera dessa olikheter.

För det första är urvalet daterat från 1995 medan resten av SESIM:s baspopulation baseras på 1999 års LINDA. För att korrigera för detta skjuts emigranternas pensionspoängshistorik fram fyra år så att 1994 års pensionspoäng i stället kommer att avse år 1998 och så vidare.

För det andra är emigranturvalet individbaserat. Två gifta/sammanboende individer kommer i databasen att representeras av två separata observationer utan information om kopplingen mellan individerna. För att urvalet ska bli konsistent med den övriga modellpopulationen imputeras en hushållssammansättning baserat på den som observeras i 1997 års HINK-data. Eftersom den utlandsboende populationen till stor del består av individer som är utlandsfödda utgår imputeringen ifrån observerade hushåll där någon medlem är född utomlands.

Skattningarna av den observerade hushållssammansättningen utgår från den kvinnliga delen av populationen. Kvinnor boende i hushåll där någon medlem är född i utlandet delas in i femåriga åldersgrupper varvid andelen gifta/sammanboende i respektive grupp beräknas. För de gifta/sammanboende kvinnorna beräknas sedan medelvärde och standardavvikelse för differensen mellan mannens och kvinnans ålder. Vidare beräknas, per civilstånd och (kvinnans) åldersgrupp, andelen hushåll med noll, ett, två eller tre

barn. Andelen hushåll med fyra eller fler barn är mycket liten och ignoreras därför. För kvinnor boende i hushåll innehållande minst ett barn beräknas därefter genomsnitt och standardavvikelse för barnens ålder.

Avsikten är nu att skapa hushåll bland emigranterna på ett sådant sätt att den i HINK observerade hushållssammansättningen replikeras så nära som möjligt. För att åstadkomma detta delas emigrantkvinnorna in i femåriga åldersgrupper varvid kvinnor som ska bilda hushåll slumpas ut efter den skattade andelen gifta/sammanboende i respektive grupp. För de utvalda kvinnorna slumpas den tilltänkte mannens ålder ifrån en normalfördelning med medelvärde och standardavvikelse motsvarande de skattade värdena. Män i korrekta åldrar lokaliserar (slumpmässigt) i emigranturvalet och matchas därefter ihop med kvinnorna.

För att tilldela hushållen barn i rätt antal och ålder slumpas förekomst av noll, ett, två eller tre barn utifrån de skattade andelarna per civilstånd och kvinnans ålder. För varje utslumpat barn tilldelas en ålder utifrån en (avrundad) normalfördelning med ovan skattat medelvärde och standardavvikelse.

Efter dessa korrigeringar läggs de utlandsboende till den övriga baspopulationen. Det bör dock noteras att simuleringen behandlar emigranterna på ett något förenklat sätt jämfört med den i Sverige boende delen av modellpopulationen. Till exempel simuleras inte bildandet av nya hushåll eller upphörandet av befintliga hushåll för dessa individer så länge de bor i utlandet.

A.4.4 Imputering av subventioner för offentlig konsumtion

I huvudtextens analyser kompletteras inkomstbegreppet genom att hushållens inkomster ökas med värdet av viss konsumtion av offentliga tjänster. Uppgifter om sådan offentlig konsumtion finns dock inte i LINDA.

Genom tidigare studier av subventionernas effekter på fördelningen av hushållens ekonomiska standard finns däremot information om utnyttjande av offentligt finansierade tjänster samt nettosubventionen⁵⁹ för dessa i HEK för åren 1997 och 1999⁶⁰.

⁵⁹ Nettosubventioner beräknas genom att produktionskostnaden för respektive tjänst minskas med eventuella avgifter.

⁶⁰ Regeringens proposition [1999/00:1], Bilaga 4 samt regeringens proposition [2001/02:100], Bilaga 3.

För att få tillgång till denna information i SESIM görs en imputering av data från HEK (1999). Nedan listas de (netto-) subventioner som imputeras till LINDA. För aggregerade poster anges de ingående komponenterna inom parentes.

- skola (förskoleklass, grundskola)
- gymnasium
- komvux
- vuxenskola (särvtux, statens skola för vuxna, svenska för invandrare, folkbildning/folkhögskola, kvalificerad yrkesutbildning)
- universitet/högskola (inklusive forskarutbildning)
- barnomsorg
- äldreomsorg (hemtjänst, särskilt boende)
- arbetsmarknadsåtgärder
- vård (primärvård, slutenvård, tandvård)
- läkemedel

Imputeringen baseras på observerade samband mellan å ena sidan de variabler som anger bruk av offentliga tjänster och storleken på subventionerna för dessa tjänster och å andra sidan ett urval av övriga variabler som finns i båda databaserna. Dessa samband specificeras med hjälp av statistiska modeller.

För några av de offentliga tjänsterna genererar SESIM information om vilka individer som använder vilka tjänster; framför allt gäller detta för olika typer av utbildningar. För dessa tjänster behöver därför endast information om subventionens storlek imputeras. För andra tjänster finns varken information om utnyttjande eller subventionsbelopp i SESIM varvid båda måste imputeras.

Imputeringen utförs initialt 1999 vilket är simuleringarnas basår. Informationen måste dock uppdateras då modellen skriver fram de ingående individerna över tiden. De estimerade modellerna används därför för att göra en årlig imputering.

Modeller och simulering

Nedan följer en översiktlig beskrivning av de samband som använts för att utföra imputeringen. Generellt kan sägas att två typer av statistiska modeller används. Benägenheten att utnyttja respektive

tjänst modelleras med logistisk regression. Modellen beskriver hur sannolikheten att ha en subvention förändras med värdet på ett antal förklarande variabler. Under simulering slumpas sedan förekomst av subventioner för individerna utifrån de skattade sannolikheterna. Som en följd av detta kan det hända att utfallet för två, vad förklarande variabler beträffar, identiska individer blir olika.

För att modellera de belopp som ska tilldelas individer som utnyttjar respektive tjänst används skattade väntevärden, givet individernas egenskaper, från regressionsmodeller. Metoden tar på detta sätt inte hänsyn till att subventionerna för två individer med samma värde på de förklarande variablerna kan vara olika stora. Anledningen är att den observerade fördelningen för många av tjänsterna har visat sig mycket svår att modellera på ett meningsfullt sätt.

Subventionernas storlek och fördelning

I det följande redovisas hur förekomsten av, och storleken hos, imputerade nettosubventioner för olika offentliga tjänster fördelar sig bland individer i SESIM. Samtliga angivna subventionsbelopp anges i 1999 års priser och har beräknats med utgångspunkt ifrån den sammansättning av olika karakteristika som gäller för SESIM:s basbefolkning under 1999. Under simuleringarna antas subventionerna förändras i takt med den allmänna prisutvecklingen.

Alla barn mellan 6–15 års ålder antas delta i skolundervisning (förskoleklass och grundskola). Subventionsbeloppet varierar med barnets ålder men ligger i åldrarna 8–15 år relativt stabilt kring drygt 55 000 kronor.

Samtliga individer antas påbörja gymnasiestudier det år de fyller 16 år men kan sedan välja att avsluta dessa innan examen uppnåtts. Subventionen för gymnasiestudier är något högre för pojkar än för flickor samt något högre för svenskfödda jämfört med utlandsfödda. Den genomsnittliga subventionen per student är ca 65 000 kronor.

I simuleringen bestäms vilka individer som under året deltar i komvuxstudier. Subventionen för studierna är högre för kvinnor än för män och även högre för svenskfödda än för utlandsfödda. Beloppen avtar något med stigande ålder. Den genomsnittliga subventionen är ca 13 000 kronor per student.

Deltagande i vuxenstudier slumpas ut bland individer i åldrarna 20–50 år som inte deltar i andra studieformer. Benägenheten att delta i vuxenstudier avtar med stigande ålder. Kvinnor är något oftare vuxenstudenter än män, speciellt i yngre åldrar. Studiebenägenheten är högre för individer med endast grundskoleutbildning jämfört med mer högutbildade. Den tenderar dessutom att vara betydligt högre för utlandsfödda än för svenskfödda. Subventionsbeloppet för de slumpade brukarna är lägre för män än för kvinnor och avtar med stigande ålder. Högutbildade och svenskfödda har något högre subventionsbelopp än lågutbildade och utlandsfödda. Den genomsnittliga subventionen är ca 15 000 kronor per student.

Av simuleringen följer vilka individer som under året bedriver universitetsstudier. Subventionen för dessa studenter är något högre för män än för kvinnor. Beloppen avtar med stigande ålder; något snabbare för män än för kvinnor. Den genomsnittliga subventionen är drygt 45 000 kronor per student i åldern 20–29 år.

Barnomsorg utnyttjas av ca 60 procent av alla barn i åldrarna 1–12. I SESIM slumpas utnyttjandet ut i den aktuella åldersgruppen. Bruksbenägenheten varierar med barnets ålder och är högre för svenskfödda barn än för utlandsfödda. Barn till högutbildade föräldrar och barn med förvärvsarbetande föräldrar utnyttjar oftare barnomsorg än andra. Subventionens belopp varierar med barnets ålder. Utlandsfödda barn och barn med lågutbildade föräldrar tenderar att ha högre subventioner än svenskfödda barn och barn till högutbildade föräldrar. Den genomsnittliga subventionen för ett barn i åldern 1–5 är något under 60 000 kronor.

Individer som utnyttjar äldreomsorg slumpas ut bland de individer som är 65 år och äldre. Bruksbenägenheten ökar med stigande ålder och ca 23 procent av 80-åringarna utnyttjar äldreomsorgen. Kvinnor erhåller oftare subventioner än män. Detsamma gäller svenskfödda individer vilka utnyttjar äldreomsorg i högre grad än utlandsfödda. Faktorer som verkar reducerande på nyttjandegraden är hög pension, att vara gift/sammanboende samt att kvarstå i arbete. För individer i åldern 80–84 år är den genomsnittliga subventionen ca 245 000 kronor.

I SESIM ingår individer som deltar i någon arbetsmarknadsåtgärd i gruppen arbetslösa och subventioner slumpas därför ut inom gruppen. Deltagandet är högre för män än för kvinnor, högre för lågutbildade än för högutbildade samt högre för utlandsfödda än för svenskfödda. Subventionsbeloppen avtar med stigande ålder,

minskar med stigande utbildning och är högre för svenskfödda jämfört med utlandsfödda. Den genomsnittliga subventionen är ca 53 000 kronor.

Samtliga individer i åldrarna 0–19 år antas utnyttja offentliga vårdtjänster. Subventionsbeloppen är högst under småbarnsåren och i slutet av tonåren. Flickor får högre subvention än pojkar och utlandsfödda högre subvention än svenskfödda. Vidare är subventionsbeloppen högre för barn och ungdomar med lågutbildade föräldrar än för de med högutbildade föräldrar. Den genomsnittliga subventionen i åldersgruppen är ca 4 000 kronor.

I åldrar över 19 år slumpas subventioner ut för utnyttjandet av vårdtjänster till vissa individer. I åldersgruppen 20–64 år är nyttjandegraden relativt låg men ökar med stigande ålder och är högre för kvinnor än för män. Faktorer som verkar förhöjande på nyttjandegraden är att vara förtidspensionär, att vara svenskfödd samt att vara gift/sammanboende. Bruksbenägenheten ökar även med stigande inkomst. För subventionerna gäller att beloppen ökar med stigande ålder. Kvinnor får högre subvention än män, speciellt i de mest fertila åldrarna. Subventionerna är förhöjda för förtidspensionärer samt för utlandsfödda. Subventionerna minskar med stigande utbildning och med stigande inkomster. Åldersgruppens genomsnittliga subvention är omkring 8 000 kronor.

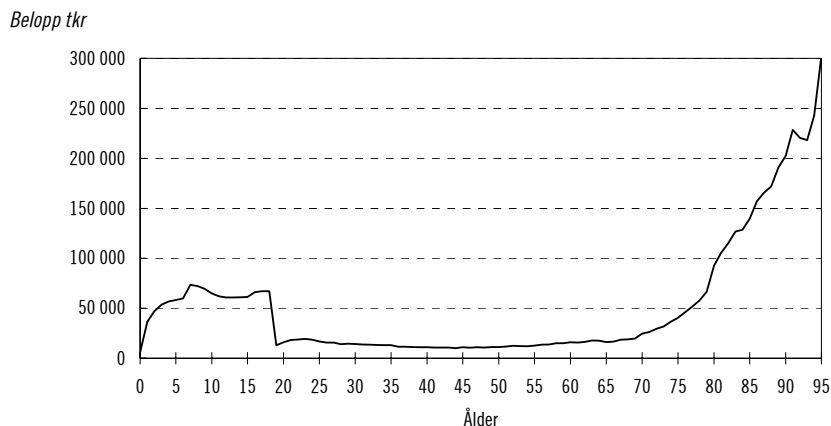
Även i gruppen 65 år och äldre slumpas subventionsbelopp ut till vissa individer. Också här ökar nyttjandegraden med stigande ålder och är högre för kvinnor än för män. För svenskfödda, ensamstående och individer med låg pension är bruksbenägenheten förhöjd. Subventionsbeloppen ökar med stigande ålder och är även högre för män än för kvinnor. Förhöjda subventioner tilldelas även utlandsfödda och ensamstående individer. Den genomsnittliga subventionen i åldersgruppen är ca 23 000 kronor men för de äldsta är beloppen närmare det dubbla.

Subventioner för bruk av läkemedel påförs samtliga individer och varierar med åldersgrupp och kön. Beloppen är högre för män än för kvinnor i åldrarna upp till 18 år samt över 65 år. I övriga åldrar är beloppen högre för kvinnor. För båda könen stiger subventionsbeloppen med åldern. Den genomsnittliga subventionen i åldern 75–79 år uppgår till ca 4 000 kronor.

Som framgår av diagram A.5 erhålls de största genomsnittliga subventionerna för konsumtion av offentliga tjänster under barn- och skolåldern samt i de högsta åldrarna. Observera att dessa

belopp är beräknade över hela populationen och därigenom även inkluderar individer som inte erhåller några subventioner.

Diagram A.5 Genomsnittlig subvention av offentliga tjänster i SESIM (1999)



Källa: SCB, HEK samt Finansdepartementets beräkningar.

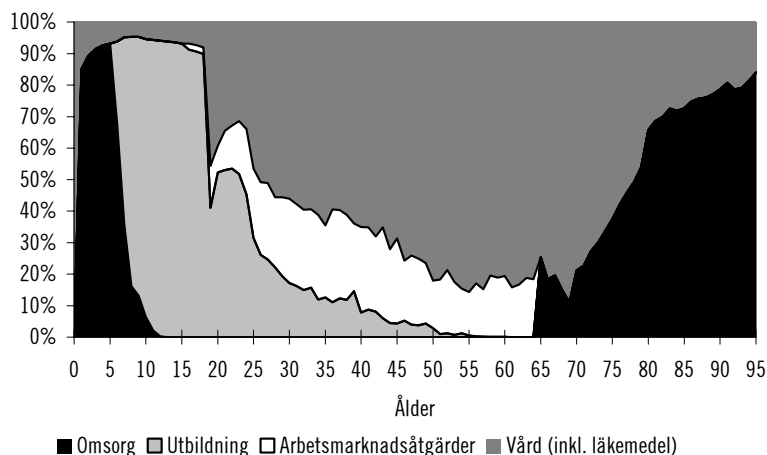
För en genomsnittlig individ, med en livslängd på 80 år, är ca 74 procent av den totala subventionen över livet förlagd till åldrarna under 20 år samt över 64 år. För en individ som lever till och med 95 år förändras siffran till ca 88 procent.

Fördelningen av olika typer av offentliga tjänster för en viss ålder kan studeras i diagram A.6 där tjänsterna delas in i aggregaten vård (inklusive läkemedel), arbetsmarknadsåtgärder, utbildning och omsorg.

Före skolåldern, samt i åldrarna 79 och uppåt dominerar subventioner för omsorg, något mer uttalat för barn än för åldringar. För de äldsta utgör även subventioner av vård en stor del av det totala beloppet. Vård är även den dominerande posten i åldern 25–78, speciellt mellan 50–70 års ålder. Utbildning tar, naturligt nog, den största andelen av subventionsbeloppet under skolåldern, ända upp till 24 års ålder varefter en successiv avtrappning sker.

Diagram A.6 Åldersvis fördelning av genomsnittlig subvention för olika typer av tjänster

Procentuell andel



Källa: SCB, HEK samt Finansdepartementets beräkningar.

Tidigare analyser⁶¹ av det använda datamaterialet visar på en viss underskattning av subventioner för slutenvård och särskilt boende i äldreomsorgen beroende på låga svarsfrekvenser bland brukare av dessa tjänster. De ovan angivna subventionsbeloppen för de äldsta utgör således sannolikt en försiktig uppskattning av de faktiska beloppen.

Nettosubventionen för flertalet tjänster uppvisar en stark regional variation och justeringar för detta görs genomgående i estimeringen av imputeringsmodellerna. Eftersom regionala omflyttningar inte simuleras explicit i SESIM används inte sambanden vid imputering. Nivåer för andel brukare och genomsnittligt belopp per brukare inom varje tjänst har anpassats till en genomsnittlig nationell nivå genom justering av intercepttermen i respektive modell.

⁶¹ Regeringens proposition [2001/02:100], Bilaga 3.

A.5 SESIM – en stokastisk simuleringsmodell

De processer som används för att åldra SESIM:s modellpopulation beskrivs med hjälp av olika statistiska modeller. Modellerna bestämmer egenskaperna hos en specifik process och anger hur den varierar med individernas egenskaper. Ett exempel är sannolikheten att vara arbetslös ett givet år som är högre för individer som var arbetslösa året innan jämfört med andra individer, allt annat lika. Effekterna av individernas egenskaper på den beskrivna processen bestäms genom att modellernas parametrar estimeras utifrån ett datamaterial. Detta ger modellerna, och därmed simuleringarna, en empirisk bas.

SESIM är en *stokastisk simuleringsmodell* vilket innebär att processerna även inkluderar en slumpmässig komponent. Simulering av processerna åstadkoms genom så kallad *Monte Carlo-simulering* vilket är en teknik som används för att generera data från stokastiska processer.

Den vanligast förekommande stokastiska processen i SESIM, och allmänt i dynamisk mikrosimulering, är den där utfallsvariabeln är binär. Variabeln kan då sägas följa bernoullifördelningen, eller mer formellt $Y_i \sim \text{bernoulli}(\pi_i)$, där $\Pr[Y_i = 1] = \pi_i$ och $\Pr[Y_i = 0] = 1 - \pi_i$.

Y_i kan här exempelvis beteckna tillstånd med avseende på arbetslöshet för den i :te individen i modellpopulationen under den aktuella perioden. För att förenkla notationen används här inte något index för att markera tidpunkten. Låt $Y_i = 1$ innebära arbetslöshet och $Y_i = 0$ sysselsättning. π_i betecknar då sannolikheten att individen är arbetslös under året. Händelsen simuleras genom att ett rektangulärfördelat slumpstal $u_i \in [0,1]$ genereras och jämförs med π_i . Om $u_i < \pi_i$ genereras en händelse, det vill säga den i :te individen blir arbetslös, annars inte.

Benägenheten att bli arbetslös kopplas till individernas egenskaper genom att låta π_i vara en funktion av dessa, vanligen genom *logistisk regression* eller *probitregression*. Den logistiska regressionsmodellen formuleras $\pi_i = [1 + \exp(-X_i\beta)]^{-1}$, där X_i är en (rad-)vektor som innehåller den i :te individens värden på de oberoende variablerna och β är en (kolumn-)vektor bestående av modellparametrar.

På grund av Monte Carlo-simuleringens stokastiska egenskaper kommer, vid upprepade simulering, antalet genererade händelser inte alltid bli detsamma⁶².

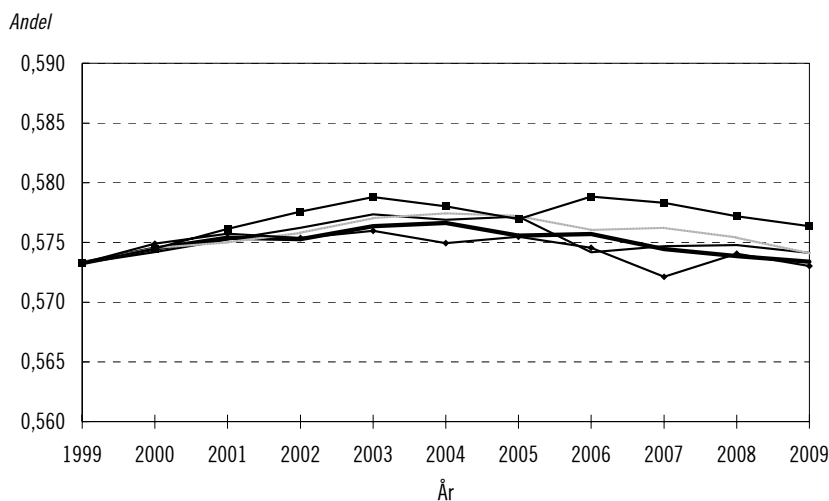
Låt T beteckna det totala antalet individer i en population av storlek N som upplever den simulerade händelsen, dvs. $T = \sum_{i=1}^N Y_i$. Om individerna simuleras oberoende av varandra kommer det förväntade antalet händelser vara $E(T) = \sum_{i=1}^N \pi_i$ och variansen $Var(T) = \sum_{i=1}^N \pi_i(1 - \pi_i)$. Om N är stort nog, och π_i inte är mycket små eller stora, kan T antas vara approximativt normalfördelad. Antag att en händelse inträffar med 10 procents sannolikhet (och för enkelhetens skull att samtliga individer har samma risk). Vid ett stort antal upprepade simuleringar av en population om 10 000 individer kommer då antalet individer som upplever händelsen i ca 95 procent av fallen⁶³ ligga mellan 941 och 1 059.

Som ett exempel på variationen i en simulerad process i SESIM visas i diagram A.7 hur andelen sammanboende bland vuxna individer realiserar vid fem oberoende simuleringar. Beräkningarna baseras på knappt 30 000 individer. Under simuleringarnas startår 1999 är samtliga resultat identiska eftersom hushållen förändras första gången år 2000. Från och med år 2000 tilltar successivt andelen hushåll som bildats av SESIM vilket resulterar i en motsvarande årlig ökande spridning mellan de olika simuleringarna. Efter några år har variationen stabiliserats.

⁶² Detta gäller under förutsättning att det så kallade slumpfelsfröet ändras mellan simuleringarna.

⁶³ Om antalet händelser approximativt kan beskrivas med en normalfördelning beräknas ett 95 procent konfidensintervall enligt: $10\,000 * 0,1 \pm 1,96 \sqrt{0,1 * 0,9 * 10\,000}$

Diagram A.7 Fem simuleringar av andelen sammanboende mellan 1999–2009



Källa: SESIM

Effekten av Monte Carlo-variationen kan vara besvärlig att hantera i vissa analysituationer. Efter en ändring av någon komponent i modellen kan det till exempel vara svårt att avgöra hur stor del av förändringen i resultaten som är rena effekter av modelländringen och vad som beror på Monte Carlo-simuleringens variansbidrag. Ett sätt att hantera problemet är att upprepa simuleringen ett antal gånger och analysera det *genomsnittliga* resultatet. Härigenom reduceras effekten av den stokastiska simuleringen.

Andra sätt att hantera problemet är att använda någon av de metoder för *variansreduktion* som används för att minska variansen vid Monte Carlo-simulering. I SESIM används en metod för variansreduktion som är intimt förknippad med begreppet *kalibrering*. Kalibrering används för att få en modell att vid simulering generera ett visst, på förhand specificerat, utfall. För en binär responsvariabel innebär det att generera ett visst antal positiva utfall i populationen eller i en delpopulation. En kalibrering åstadkoms genom att sannolikheten för ett positivt utfall för de ingående individerna förändras på ett sådant sätt att det förväntade antalet positiva utfall blir det önskade.

Ett enkelt sätt att åstadkomma detta, och vanligt förekommande, är en proportionell justering av sannolikheterna, $\pi^*_i = \alpha\pi_i$, där π^*_i är den justerade sannolikheten och α är justeringsfaktorn. Ett problem med ansatsen är att den inte garanterar att π^*_i är ett tal mellan 0 och 1. Om då i stället $\pi^*_i = \min(1, \alpha\pi_i)$ används innebär detta att de individer där $\pi^*_i = 1$ helt säkert kommer att få positiva utfall, något som kan ge oönskade effekter i olika situationer (till exempel att samtliga individer med vissa egenskaper avlider ett visst år).

En alternativ ansats är att göra justeringen på en annan skala. I SESIM används en additiv justering på logit-skalan vilket även kan ses som en justering av intercepttermen i den logitmodell från vilka sannolikheterna beräknades⁶⁴. Här har vi alltså i stället $\text{logit}(\pi^*_i) = \alpha + X_i\beta$, där $\text{logit}(x) = \log[x/(1-x)]$, eller $\pi^*_i = [1 + \exp(-\alpha - X_i\beta)]^{-1}$ vilket leder till att $\pi^*_i \in [0,1]$.

I båda fallen behöver justeringsfaktorn α beräknas. Detta låter sig enkelt göras i det första fallet (om inte den genomsnittliga sannolikheten är hög vilket vid ett stort α leder till en hög andel censurerade sannolikheter) och med något större besvär i det andra.

De ovan beskrivna metoderna gör visserligen att det förväntade antalet positiva utfall motsvarar det önskade men fortfarande kan Monte Carlo-variationen ge avvikelser till det faktiskt erhållna resultatet. Den metod för variansreduktion som används i SESIM eliminerar dessa avvikelser genom att använda det α som, givet de slumpade värdena på u_i genererar det önskade antalet positiva utfall n .

Problemet med att explicit beräkna α undviks genom att observera att ovanstående är ekvivalent med att sortera variabeln $v_i = \text{logit}(u_i) - \text{logit}(\pi_i)$ i stigande ordning och låta de individer med de lägsta värdena få positiva utfall.

En viss kritik finns mot användning av kalibrering av dynamiska mikrosimuleringsmodeller, t.ex. Klevmarken (1998). Ett argument är att avvikelser från förväntade resultat kan bero på misspecifierade modeller och snarare bör hanteras genom att dessa specificeras, och estimeras om, snarare än att simuleringarna kalibreras.

⁶⁴ Här kan förstås andra funktioner $f(x)$ som är monotont växande i x och antar värden på intervallet $[0,1]$ användas. Lämpliga funktioner kan vara andra så kallade länkfunktioner från klassen av generaliserade linjära modeller från vilken logit- och probitmodellerna är specialfall.

Trots kritiken är det ovanligt att modeller som används i faktiska analyser inte använder sig av denna funktionalitet. En primär funktion hos kalibreringen är att säkerställa att modellen replikerar en antagen demografisk utveckling eftersom analyserna ofta betingas på en vedertagen demografisk prognos.

Ett alternativt synsätt är att se kalibrering som en möjlighet att implementera olika scenarion i simuleringarna. Detta kan annars vara svårt att åstadkomma på ett kontrollerat sätt genom att till exempel manipulera parametrar i olika modeller beroende på simuleringsmodellens totala komplexitet.

Vilka modeller/processer i SESIM som använder kalibrering och/eller variansreduktion framgår av den specifikation av de stokastiska modellerna som ges i avsnitt A.8

Monte Carlo-variansen är dock inte den enda varianskomponenten i en mikrosimuleringsmodell se Klevmarken (1998). Eftersom simuleringarna baseras på ett stickprov draget från någon population introduceras även här en felkälla. Ytterligare en källa till osäkerhet är de estimerade parametrar som ingår i de statistiska modellerna – även dessa är behäftade med en viss osäkerhet⁶⁵. För att kunna göra inferens i någon mer formell mening bör alltså osäkerheten i alla dessa komponenter beaktas. Beroende på komplexiteten i en dynamisk mikrosimuleringsmodell är det tyvärr svårt att härleda analytiska uttryck för variansen hos de statistikor som utgör resultatvariabler. I stället kan man tänka sig ansatser baserade på datorintensiva inferensmetoder (till exempel *bootstrap*, se Davison och Hinkley (1997)).

⁶⁵ På grund av de ofta mycket stora datamaterial som används vid estimering av de statistiska modellerna i SESIM är denna osäkerhetskälla sannolikt relativt begränsad.

A.6 Simulering av inkomster

Eftersom inkomster är en av de viktigaste variablerna i de flesta analyser ges i detta avsnitt en mer detaljerad beskrivning av den modell som används för att simulera inkomster i SESIM.

En ansats till att beskriva hur inkomster utvecklas över tiden är att observera inkomster under ett visst år och ansätta en regressionsmodell för att förklara arbetsinkomsternas fördelning med olika variabler såsom till exempel ålder, kön, utbildningsnivå och/eller andra socioekonomiska faktorer. Den ansatta modellen kan sedan användas för att göra prediktioner för inkomster, eller väntevärdet för dessa, vid olika åldrar (givet övriga förklarande variabler).

Ett problem med ansatsen är att endast information om inkomster vid en viss given tidpunkt (ett *tvärsnitt*) används – inget kan sägas om eventuella samband mellan inkomster från olika tidpunkter för en och samma individ. Vidare kan dessa skattningar av inkomsternas genomsnittliga förändring mellan olika åldrar under det observerade året vara dåliga approximationer av den genomsnittliga inkomstförändringen över tiden som föreligger för en given kohort.

Ett alternativt angreppssätt är att basera analysen på *longitudinella data*. Här observeras inkomster i stället under ett antal sammanhängande år, det vill säga en och samma individ förekommer flera gånger i datamaterialet. Modellens formulering blir då en något modifierad version av den klassiska regressionsmodellen för tvärsnittsdata:

$$Y_{ij} = \mathbf{X}_{ij}\boldsymbol{\beta} + \gamma_i + \varepsilon_{ij}, \text{ där } \gamma_i \sim N(0, \tau^2) \text{ och } \varepsilon_{ij} \sim N(0, \sigma^2).$$

Komponenterna γ_i och ε_{ij} antas vara inbördes och sinsemellan oberoende. Y_{ij} betecknar här inkomsten för den i :te individen vid den j :te tidpunkten. Termen γ_i ingår inte i formuleringen av den typiska regressionsmodellen för tvärsnittsdata och kan ses som en *stokastisk interceptterm*. Avsikten med formuleringen är att fånga upp den variation som genereras av att det, givet samma förklarande variabler, finns individer vars inkomster tenderar att genomgående ligga över eller under det genomsnittliga värdet.

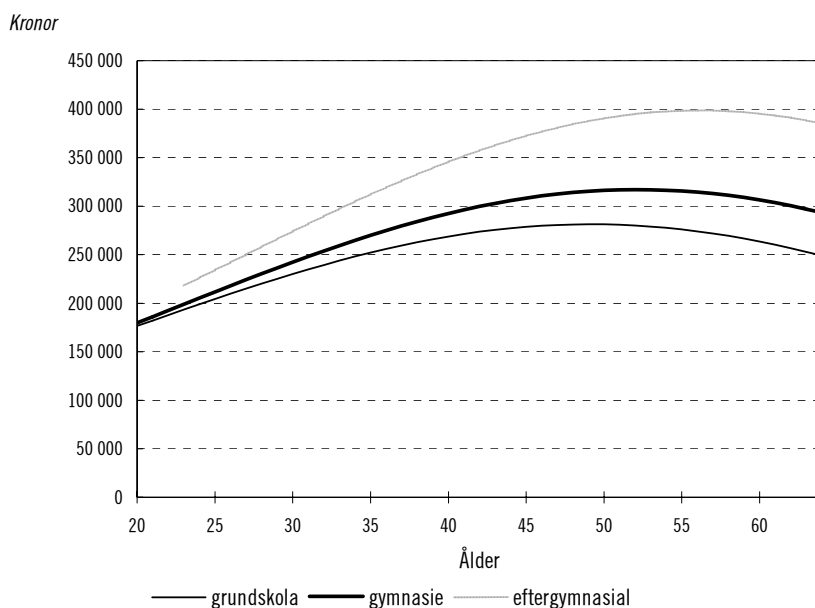
Modellen implicerar att en individs inkomster inte är oberoende över tiden. Däremot är olika individers inkomster oberoende av varandra, på samma sätt som i modellen för tvärsnittsdata. Detta ger att kovariansmatrisen för $\mathbf{Y} = [Y_{11}, Y_{12}, \dots, Y_{1J}, Y_{21}, \dots, Y_{IJ}]^T$ blir blockdiagonal. Varje block, representerande data från en individ, har diagonalelementen $\tau^2 + \sigma^2$ och övriga element τ^2 . Detta innebär i sin tur att korrelationen mellan inkomster från olika år, för samma individ, är lika med $\rho = \tau^2 / (\tau^2 + \sigma^2)$. Korrelationen kommer alltså att vara hög om variansen för den stokastiska intercepttermen är stor i förhållande till den övriga residualvariansen. Notera även att korrelationen inte förändras av längden på den mellanliggande perioden. För vidare diskussion av statistiska modeller för longitudinella data, se t.ex. Diggle, Liang och Zeger (1994).

Den i SESIM använda specifikationen innehåller de förklarande variablerna arbetslivserfarenhet, högsta utbildningsnivå, arbetsmarknadssektor, civilstånd samt indikator för nationalitet. Separata modellskattningar görs för respektive kön. Vidare görs separata skattningar av varianskomponenterna τ^2 och σ^2 för respektive arbetsmarknadssektor.

Samtliga skattningar görs på logaritmtransformerade arbetsinkomster. Transformationen garanterar positiva prediktioner av inkomster.

Diagram A.8 beskriver den förväntade inkomsten för svenskfödda, sammanboende män som arbetar som privatanställda tjänstemän vid olika utbildningsnivåer och åldrar. Eftersom modellen innehåller arbetslivserfarenhet men inte explicit ålder antas här att arbetslivet påbörjas vid 16, 19 och 23 års ålder för de olika utbildningsnivåerna och därefter fortgår kontinuerligt tills 64 års ålder.

Diagram A.8 Förväntad inkomst – 1999 års priser



Källa: SESIM

Diagrammet visar ett tydligt positivt samband mellan utbildning och inkomst. Vidare är inkomstskillnaden mellan eftergymnasial och gymnasial utbildning betydligt större än den mellan gymnasial och förgymnasial utbildning. Den förskjutning av kurvans maximum som tydligt ses för individer med eftergymnasial utbildning är en effekt av att dessa antas påbörja sin karriär senare beroende på den längre studietiden.

Tabell A.1 visar effekter av samtliga förklarande variabler (bortsett från arbetslivserfarenhet) på den förväntade inkomsten⁶⁶. Enligt modellen är utbildningspremien något högre för kvinnor än för män, allt annat lika. Spannet mellan lägst och högst utbildning är för kvinnor 34 procent mot 29 procent för män. Modellen ger även en rangordning av arbetsmarknadssektorer där egna företagare ligger lägst och där privatanställda tjänstemän ligger högst. Härkomst och civilstånd har mycket små effekter för män men ger ett

⁶⁶ Tabellvärdena är beräknade som inkomstens väntevärde för den angivna nivån av den förklarande variabeln dividerat med väntevärdet för referensgruppen (1,00 i tabellen) givet identiska värden på övriga förklarande variabler.

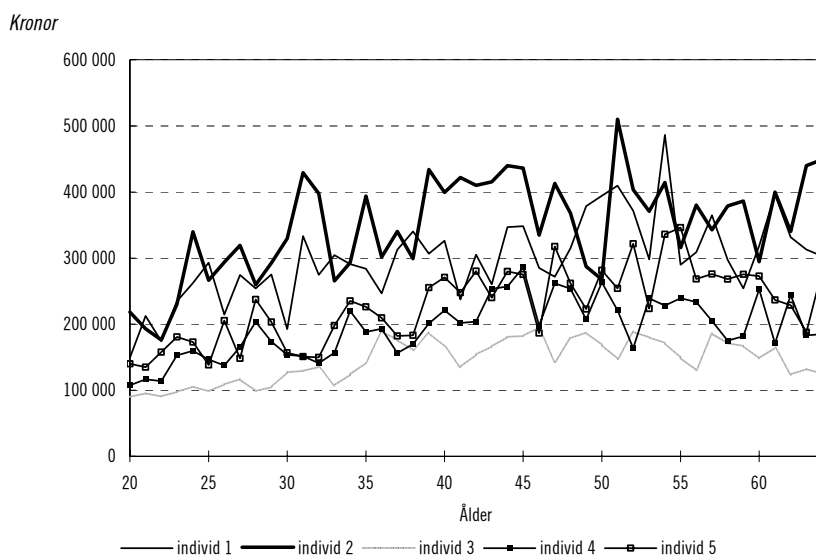
visst utslag för kvinnor genom att utlandsfödda och ensamstående har något högre inkomster än svenskfödda och sammanboende.

Tabell A.1 Effekter av förklarande variabler i inkomstmodeller

Variabel	Kommentar	Män	Kvinnor
Utbildning	Förgymnasial	0,71	0,66
	Gymnasial	0,80	0,75
	Eftergymnasial	1,00	1,00
Sektor	Privatanställd arbetare	1,60	1,43
	Privatanställd tjänsteman	2,02	1,80
	Statligt anställd	1,66	1,49
	Kommunalt anställd	1,51	1,26
Härkomst	Egen företagare	1,00	1,00
	Sverige	1,00	0,97
Civilstånd	Utlandet	1,00	1,00
	Ensamstående	1,01	1,06
	Sammanboende	1,00	1,00

Källa: SESIM, egna beräkningar

Simuleringar av inkomstmodellen utgår från individens egenskaper \mathbf{X}_{ij} , de skattade parametrarna $\hat{\boldsymbol{\beta}}$ samt slumpalen $\tilde{\gamma}_i$ och $\tilde{\epsilon}_{ij}$. De sistnämnda dras från normalfördelningar med respektive skattade varianser $\hat{\sigma}^2$ och $\hat{\sigma}^2$. Den simulerade inkomsten beräknas som $\tilde{Y}_{ij} = \mathbf{X}_{ij}\hat{\boldsymbol{\beta}} + \tilde{\gamma}_i + \tilde{\epsilon}_{ij}$. Eftersom $\tilde{\gamma}_i$ är specifik för en enskild individ och inte förändras över tiden görs endast en dragning av detta slumpstal medan $\tilde{\epsilon}_{ij}$ dras på nytt varje simuleringsperiod (och för varje ny individ). Diagram A.9 visar fem oberoende realisationer av inkomster för svenskfödda, sammanboende män som arbetar som privatanställda tjänstemän och har gymnasial utbildning. Observera att den mittersta kurvan i diagram A.8 alltså motsvarar seriernas väntevärde. Som framgår av bilden finns alltså för individerna en variation mellan inkomster från olika år. Det framgår även relativt tydligt att serierna har olika generella genomsnittsnivåer. Detta är en direkt effekt av modellens stokastiska interceptterm.

Diagram A.9 Fem simulerade inkomstprofiler - 1999 års priser

Källa: SESIM

Notera att de inkomster som genereras av den ovan presenterade framskrivningsmodellen är *arbetsinkomster*. Dessa genereras endast för individer som klassificeras som sysselsatta i SESIM. För andra individer beräknas inkomster på andra sätt, till exempel som olika transfereringar (vid arbetslöshet, föräldradighet osv). Inkomstgenereringen i SESIM är på så sätt inte uteslutande en produkt av modellen för inkomstframskrivning utan påverkas även av de modeller som bestämmer sannolikheten att en individ ska vara sysselsatt eller inte ett visst år.

I appendix B presenteras en genomgång av egenskaperna hos olika simulerade inkomstslag.

A.7 Teknisk plattform

SESIM är skriven i programmeringsspråket Visual Basic. Språkets relativa enkelhet gör att det mesta av kodningsarbetet kan skötas utan hjälp av programmeringsexperten. Den kompilerade koden är

dessutom tillräckligt effektiv för att de relativt beräkningsintensiva analyserna ska kunna utföras utan orimligt långa väntetider⁶⁷.

Windowsapplikationen SESIM har ett användarvänligt grafiskt gränssnitt. Det har designats för att möjliggöra datauttag från SESIM på ett enkelt sätt. Ett flertal statistikfunktioner finns tillgängliga, t.ex. univariata statistikor, frekvenstabeller, populationshistogram, kernelestimatorer av fördelningar samt formulär för att direkt undersöka data för enskilda individer/hushåll.

Ytterligare flexibilitet i analysarbetet uppnås genom att SESIM kan exportera de genererade datamängderna i flera olika format vilket gör att befintliga programvaror för statistiska analyser lätt kan användas för att analysera data från SESIM.

Utöver windowsapplikationen finns även en särskild rapportgenerator, implementerad i Excel, som tillåter enkel och snabb presentation av SESIM-genererad data med hjälp av Excels inneboende funktionalitet. Ett huvudsakligt användningsområde för rapportgeneratoren är modellvalideringsarbete. I förväg iordninggjorda kalkylark där simuleringsresultat ställs mot andra datakällor kan lätt uppdateras varje gång något förändras i modellen. På så sätt underlättas valideringsarbetet väsentligt.

Excel används även för hantering av indata till SESIM. Detta görs genom att indata sammanställs på ett särskilt format i kalkylark vilka sedan läses in av SESIM under simuleringen. Främst kontrolleras härigenom olika aspekter av den demografiska utvecklingen men även andra variabler som t.ex. årliga pris- och inkomstförändringar, antagen avkastning för premiepensionsfonder, generella arbetslöshets- och sysselsättningsnivåer. Indata för olika beräkningsscenarion kan lagras i olika kalkylark. På så vis byts indata enkelt genom ett byte av kalkylark. Kalkylarken utgör även en god dokumentation av de beräkningsscenarion som används.

⁶⁷ För de mer krävande beräkningarna används effektiva FORTRAN-rutiner vilka har länkats till Visual Basic via ett så kallat dynamiskt länkat bibliotek (DLL).

A.8 Förteckning över stokastiska modeller i SESIM

I det följande ges en översiktlig beskrivning av samtliga stokastiska modeller som ingår i SESIM. Av utrymmesskäl begränsas beskrivningen till följande egenskaper:

- **Händelse/utfall:** i de fall detta inte framgår av rubrik eller sammanhang specificeras den aktuella händelsen eller utfallsvariabeln här.
- **Riskgrupp:** de individer (eller hushåll) som är ”under risk”. Med detta avses att deras egenskaper kan påverkas av modellens utfall.
- **Modelltyp:** avser en grov klassificering av olika typer av statistiska modeller som används. De i särklass vanligaste modelltyperna i SESIM är linjära regressionsmodeller för kontinuerliga responsvariabler och logistisk regression alternativt probitregression för binära responsvariabler.
- **Kovariater:** de förklarande variabler som används för att, genom modellen, länka individers och hushålls egenskaper till responsvariabeln. Här beskrivs endast vilka förklarande variabler som förekommer, inte vilken funktionell form de har eller om de bildar interaktionstermer med andra förklarande variabler.
- **Kommentar:** generella kommentarer. Till exempel beträffande olika alternativ för kalibrering av modellen.
- **Algoritm:** i vissa fall kombineras stokastiska modeller med olika regler eller algoritmer för att beskriva ett händelseförlopp.

Mortalitet

Riskgrupp: individer i åldern 0–29 år.

Modelltyp: ettåriga dödsrisker enligt SCB:s befolkningsprognosmodell.

Kommentar: på grund av de mycket låga dödsfallsfrekvenserna i dessa åldrar är det mycket svårt att hitta en lämplig modell på empiriska grunder. Av detta skäl ansätts i stället de av SCB skattade genomsnittliga dödsriskerna per kön och ålder. Dessa bygger på observerad dödlighet för samtliga individer i respektive ålder/kön.

Riskgrupp: individer i åldrarna 30–64 år.

Modelltyp: logistisk regression.

Kovariater: kön, ålder, indikator för förtidspension, pensionsgrundande inkomst (kvintil), civilstånd.

Riskgrupp: individer i åldrarna över 64 år.

Modelltyp: logistisk regression.

Kovariater: kön, ålder, indikator för förtidspension vid 64 års ålder, civilstånd, högsta utbildningsnivå.

Kommentar: för mortalitet görs en kalibrering till de dödsrisker per kön och åldersgrupp som gäller enligt SCB:s långsiktiga befolkningsprognos⁶⁸.

Adoption (av föräldralösa)

Händelse/utfall:

Adopterande hushåll slumpas från mängden av hushåll där ålderssammansättningen är lämplig.

Riskgrupp:

1) barn (yngre än 18) som blivit föräldralösa under året.

2) adopterande hushåll med en, för de vuxna, lägsta ålder överstigande 24 år och en högsta ålder understigande 51 år.

Modelltyp: regression.

Kovariater: indikator för hushållstyp (ensamstående kvinna, ensamstående man, sammanboende), kvinnans ålder (mannens ålder för ensamstående män).

Algoritm: De adopterande hushållens simulerade antal barn (utifrån den estimerade modellen) jämförs med hushållets faktiska antal barn. De föräldralösa barnen adopteras av hushåll med stora skillnader mellan simulerat och faktiskt antal barn.

⁶⁸ Se SCB:s [2002] demografiska prognos (BE18SM0201) avseende en framskrivning av Sveriges befolkning till år 2050.

Emigration

Riskgrupp: tidigare invandrade hushåll.

Modelltyp: logistisk regression.

Kovariater: hushållets högsta ålder, antal barn, antal vuxna, hushållets högsta utbildningsnivå, tid sedan invandring.

Riskgrupp: svenskfödda hushåll.

Modelltyp: logistisk regression.

Kovariater: hushållets högsta ålder, antal barn, antal vuxna, hushållets högsta utbildningsnivå.

Kommentar: det totala antalet genererade emigranter kalibreras till det totala antalet emigranter enligt SCB:s demografiska prognos.

Immigration

Riskgrupp: tidigare utvandrade individer/hushåll.

Modelltyp: hazardmodell.

Kovariater: tid sedan utvandring.

Riskgrupp: förstagångsinvandrare.

Algoritm: befintliga hushåll i modellpopulationen klonas och imputeras därefter som förstagångsinvandrare. Selektion av hushåll för kloning görs så att de invandrande hushållen får en lämplig sammansättning med avseende på den högsta ålder som kan observeras i hushållet samt hushållets storlek.

Kommentar: modellens totala immigration kalibreras till total immigration enligt SCB:s prognos genom justering av antalet förstagångsinvandrare inom ramen för den totalt genererade invandringen.

Fertilitet

Riskgrupp: kvinnor mellan 18–49 år som inte bor i föräldrahemmet och som inte tidigare fött barn.

Modelltyp: logistisk regression.

Kovariater: ålder, civilstånd, pensionsgrundande inkomst (kvartiler), indikator för förvärvsarbete, högsta utbildningsnivå.

Riskgrupp: kvinnor mellan 18–49 år som inte bor i föräldrahemmet och som tidigare fött minst ett barn.

Modelltyp: logistisk regression.

Kovariater: åldersgrupp, civilstånd, pensionsgrundande inkomst (kvartiler), indikator för förvärvsarbete, högsta utbildningsnivå, yngsta barnets ålder.

Kommentar: separata modeller estimeras för kvinnor som fött ett, två respektive tre barn. För närvarande modelleras inte fler än fyra födselar.

Kommentar: det totala antalet födselar kalibreras till det totalt antal födda barn per år enligt SCB:s prognos.

Utflyttning från föräldrahemmet

Riskgrupp: individer som är äldre än 17 år och bor kvar i föräldrahemmet.

Modelltyp: logistisk regression.

Kovariater: ålder, högsta utbildningsnivå, indikator för pågående studier, indikator för nollskild taxerad inkomst, nationalitet.⁶⁹

Kommentar: separata modeller estimeras för män och kvinnor. Modellen kalibreras mot observerad andel utflyttade per kön och ålder enligt HEK år 1999.

Bildande av nya hushåll

Riskgrupp: vuxna (ej kvarboende i föräldrahemmet) kvinnor och män som inte är sammanboende.

Modelltyp: logistisk regression.

Kovariater: ålder.

Algoritm: Kvinnor som ska bilda nytt hushåll under året slumpas ut från de skattade sannolikheterna. För varje kvinna genomsöks poolen av valbara män och den första mannen som är tre år äldre än kvinnan väljs ut. Hittas ingen man i rätt ålder upprepas proceduren varvid ett års ytterligare åldersskillnad tillåts för varje upprepad sökning.

⁶⁹ Nationalitet indikerar i detta sammanhang om individen är född i Sverige eller utomlands.

Upplösande av hushåll

Riskgrupp: sammanboende män.

Modelltyp: logistisk regression.

Kovariater: ålder.

Algoritm: män slumpas, utifrån de skattade sannolikheterna, att de lämnar sitt hushåll och bildar ett nytt.

Förtidspension

Riskgrupp: individer mellan 16–29 års ålder som inte är förtidspensionärer.

Modelltyp: logistisk regression.

Kovariater: ålder, högsta utbildningsnivå, indikator för ålder 16, pensionsgrundande inkomst (kvartiler).

Riskgrupp: individer mellan 30–60 års ålder som inte är förtidspensionärer.

Modelltyp: logistisk regression.

Kovariater: ålder, kön, pensionsgrundande inkomst (kvartiler), högsta utbildningsnivå, civilstånd, nationalitet.

Riskgrupp: individer mellan 61–64 års ålder som inte är förtidspensionärer.

Modelltyp: logistisk regression.

Kovariater: ålder, kön, pensionsgrundande inkomst (kvartiler), högsta utbildningsnivå.

Kommentar: under simuleringens första år kalibreras inflödet till förtidspension mot det faktiskt observerade antalet.

Rehabilitering från förtidspension

Riskgrupp: förtidspensionärer.

Modelltyp: logistisk regression.

Kovariater: ålder, tid sedan förtidspensionering, högsta utbildningsnivå.

Utbildning

Händelse/utfall: gymnasiestudier påbörjas.

Riskgrupp: 16-åringar som inte är förtidspensionärer.

Algoritm: samtliga individer antas påbörja gymnasiestudier det år de fyller 16.

Kommentar: maximalt bedrivs tre års studier på gymnasienivå varefter gymnasiekompetens erhålls.

Händelse/utfall: avhopp från gymnasiestudier.

Riskgrupp: gymnasiestudenter.

Modelltyp: logistisk regression.

Kovariater: kön, föräldrars högsta utbildningsnivå, föräldrars högsta ålder, nationalitet, indikator för skilda föräldrar, hushållets inkomst (kvartiler), antal barn i hushållet.

Kommentar: för individer som slumpas till att avbryta sina gymnasiestudier slumpas även en tidpunkt för avbrottet.

Händelse/utfall: universitetsstudier påbörjas direkt efter avslutade gymnasiestudier.

Riskgrupp: 19-åringar som tog gymnasieexamen föregående år och inte är förtidspensionerade eller föräldralediga.

Modelltyp: logistisk regression.

Kovariater: kön, nationalitet, indikator för boende i föräldrahemmet, indikator för egna barn, indikator för boende i storstad, generell arbetslöshetsnivå.

Händelse/utfall: universitetsstudier avbryts innan examen erhållits

Riskgrupp: universitetsstudenter.

Modelltyp: logistisk regression.

Kovariater: kön, nationalitet, civilstånd, ålder, antal barn, indikator för boende i storstad.

Kommentar: universitetsstudier antas pågå tre eller fyra år. Individer som avbryter sina studier efter det tredje året antas då erhålla sin examen. Efter det fjärde året avslutas studierna automatiskt.

Händelse/utfall: transition från arbetsmarknad till universitetsstudier.

Riskgrupp: individer som uppnått gymnasiekompetens och inte deltar i andra studieformer.

Modelltyp: logistisk regression.

Kovariater: kön, nationalitet, BNP-tillväxt, tid sedan gymnasieexamen, ålder, civilstånd, antal barn.

Händelse: transition från arbetsmarknad till komvuxstudier.

Riskgrupp: individer mellan 20–64 års ålder som inte uppnått gymnasiekompetens, inte är förtidspensionerade eller föräldralediga.

Modelltyp: logistisk regression.

Kovariater: kön, nationalitet, BNP-tillväxt, civilstånd, tid sedan gymnasieexamen, antal barn, ålder.

Kommentar: komvuxstudier antas pågå tills tre års studier efter grundskolenivå har uppnåtts. Detta ger individen motsvarande gymnasiekompetens.

Händelse/utfall: universitetsstudier påbörjas direkt efter avslutade komvuxstudier.

Riskgrupp: individer som avslutade komvuxstudier föregående år och inte är förtidspensionerade eller föräldralediga.

Modelltyp: logistisk regression.

Kovariater: kön, nationalitet, ålder, civilstånd, antal barn, indikator för boende i storstad, BNP-tillväxt.

Sysselsättning

Händelse/utfall: sysselsatt/ej sysselsatt under året.

Riskgrupp: individer äldre än 15 år som inte är ålderspensionärer, förtidspensionärer, deltar i studier, är föräldralediga eller är arbetslösa.

Modelltyp: logistisk regression.

Kovariater: indikator för sysselsättning föregående år, ålder, högsta utbildningsnivå, indikator för pågående studier föregående år.

Kommentar: separata modeller estimeras för män respektive kvinnor.

Arbetslöshet

Händelse/utfall: arbetslös/ej arbetslös under året.

Riskgrupp: individer äldre än 15 år som inte är ålderspensionärer, förtidspensionärer, deltar i studier eller är föräldralediga.

Modelltyp: logistisk regression.

Kovariater: indikator för arbetslöshet föregående år, ålder, åldersgrupp, högsta utbildningsnivå.

Kommentar: separata modeller estimeras för män respektive kvinnor.

Arbetsmarknadssektor⁷⁰

Riskgrupp: individer i arbetskraften utan tidigare tilldelning av arbetsmarknadssektor.

Modelltyp: multinomial logit.

Kovariater: ålder, kön, högsta utbildningsnivå, nationalitet.

Riskgrupp: individer i arbetskraften som har en tidigare tilldelad arbetsmarknadssektor.

Modelltyp: multinomial logit.

Kovariater: ålder, kön, högsta utbildningsnivå.

Kommentar: Separata modeller estimeras för varje arbetsmarknadssektor, dvs. individens arbetsmarknadssektor under det innevarande året beror av individens arbetsmarknadssektor föregående år.

Bostadsmarknad/förmögenhet

Händelse/utfall: förekomst av positiv finansiell förmögenhet.

Riskgrupp: individer utan positiv finansiell förmögenhet.

Modelltyp: logistisk regression.

Kovariater: kön, åldersgrupp, högsta utbildningsnivå, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), nationalitet.

⁷⁰ Uppdelning görs mellan privat anställda arbetare, privat anställda tjänstemän, kommunalt anställda, statligt anställda och egna företagare.

Händelse/utfall: finansiell förmögenhet.

Riskgrupp: individer med förekomst av positiv finansiell förmögenhet enligt föregående modell.

Modelltyp: regression, logittransformerad responsvariabel.

Kovariater: kön, åldersgrupp, högsta utbildningsnivå, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), nationalitet.

Kommentar: modellen ger en väntevärdesprediktion, dvs. variansen antas vara lika med noll.

Händelse/utfall: finansiell förmögenhet.

Riskgrupp: individer med positiv finansiell förmögenhet (föregående år).

Modelltyp: random walk utan drift. Variansen är proportionell mot förmögenheten. Slumpade värden begränsas till positiva tal.

Kovariater: finansiell förmögenhet föregående år.

Händelse/utfall: bostadsägande.

Riskgrupp: hushåll som bildats under året.

Modelltyp: logistisk regression.

Kovariater: åldersgrupp för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, antal barn under 18 år, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), indikator för hushåll där hushållsföreståndare samt maka/make är svenskfödd, kvantiler för hushållets finansiella förmögenhet.

Kommentar: ett hushåll antas kunna äga maximalt en bostad under sin livstid. Som en konsekvens kan ett bostadsköp och en försäljning äga rum. En individ kan dock genom att byta hushåll vara ägare till fler bostäder under en livstid.

Händelse/utfall: bostadsköp.

Riskgrupp: hushåll som inte äger bostad.

Modelltyp: logistisk regression.

Kovariater: åldersgrupp för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, antal barn under 18 år, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), indikator för hushåll där hushållsföreståndare samt maka/make är svenskfödd, kvantiler för hushållets finansiella förmögenhet.

Händelse/utfall: bostadsförsäljning.

Riskgrupp: hushåll som äger bostad.

Modelltyp: logistisk regression.

Kovariater: åldersgrupp för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, antal barn under 18 år, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), indikator för hushåll där hushållsföreståndare samt maka/make är svenskfödd, kvantiler för hushållets finansiella förmögenhet.

Händelse/utfall: bostadsyta.

Riskgrupp: hushåll som köpt bostad under året.

Modelltyp: regression.

Kovariater: åldersgrupp för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, antal barn under 18 år, kvantiler för skattepliktig inkomst, indikatorer för boende i Stockholm, Göteborg eller Malmö (H-region).

Kommentar: modellen ger en väntevärdesprediktion, dvs. variansen antas vara lika med noll.

Händelse/utfall: bostadens marknadsvärde.

Riskgrupp: hushåll som köpt bostad under året.

Modelltyp: regression, logittransformerad responsvariabel.

Kovariater: ålder för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, antal barn under 18 år, kvantiler för skattepliktig inkomst, indikatorer för boende i tätbefolkad region (H-region), bostadsyta, kvantiler för hushållets finansiella förmögenhet.

Händelse/utfall: förekomst av bostadslån.

Riskgrupp: hushåll som bildats under året och slumpats till bostadsägande.

Modelltyp: logistisk regression.

Kovariater: ålder för hushållets äldsta medlem, åldersgrupp för hushållets äldsta medlem, kvantiler för skattepliktig inkomst, indikatorer för boende i Stockholm, göteborg eller Malmö (H-region), bostadsyta, bostadens marknadsvärde, kvantiler för hushållets finansiella förmögenhet.

Händelse/utfall: bostadslån.

Riskgrupp: hushåll som bildats under året, slumpats till bostadsägande och till att ha bostadslån.

Modelltyp: regression, logtransformerad responsvariabel.

Kovariater: ålder för hushållets äldsta medlem, åldersgrupp för hushållets äldsta medlem, indikator för gift/sammanboende hushåll, kvantiler för skattepliktig inkomst, bostadsyta, bostadens marknadsvärde, kvantiler för hushållets finansiella förmögenhet.

Händelse/utfall: kapitalinkomster.

Riskgrupp: individer med positiv finansiell förmögenhet

Modelltyp: regression.

Kovariater: kön, åldersgrupp, högsta utbildningsnivå, indikatorer för boende i tätbefolkad region (H-region), nationalitet.

Kommentar: responsvariabel är kapitalinkomstens andel av den finansiella förmögenheten.

Arbetsinkomster⁷¹

Riskgrupp: samtliga sysselsatta individer.

Modelltyp: regression (mixed regression), logtransformerad responsvariabel.

Kovariater: arbetserfarenhet, högsta utbildningsnivå, arbetsmarknadssektor, nationalitet, civilstånd, stokastisk interceptterm.

Kommentar: modellen estimeras separat för respektive kön. Separata estimat av modellens varianskomponenter för respektive arbetsmarknadssektor.

Händelse/utfall: förekomst av arbetsinkomster.

Riskgrupp: studenter.

Modelltyp: probitregression.

Kovariater: ålder, kön, nationalitet.

Händelse/utfall: arbetsinkomster.

Riskgrupp: studenter som slumpats i föregående modell.

Modelltyp: regression, logtransformerad responsvariabel.

Kovariater: ålder, kön, nationalitet.

⁷¹ Modellen för inkomstframskrivning beskrivs närmare i avsnitt A.6.

Händelse/utfall: förekomst av arbetsinkomster.

Riskgrupp: individer klassificerade som övriga.

Modelltyp: probitregression.

Kovariater: ålder, kön, nationalitet, högsta utbildningsnivå.

Händelse/utfall: arbetsinkomster.

Riskgrupp: individer klassificerade som övriga som slumpats i föregående modell.

Modelltyp: regression, logtransformerad responsvariabel

Kovariater: ålder, högsta utbildningsnivå.

Sjukdagar

Händelse/utfall: förekomst av positivt antal sjukdagar.

Riskgrupp: individer mellan 16–64 års ålder vilka inte är förtids- eller ålderspensionerade och inte är föräldralediga.

Modelltyp: probitregression (mixed regression).

Kovariater: indikator för positivt antal sjukdagar föregående år, åldersgrupp, högsta utbildningsnivå, kön, generell arbetslöshetsnivå, indikator för boende i storstad, civilstånd, nationalitet, stokastiskt intercept.

Händelse/utfall: antal sjukdagar.

Riskgrupp: individer som slumpats i föregående modell.

Modelltyp: regression, logtransformerad responsvariabel.

Kovariater: åldersgrupp, kön, indikator för barn i hushållet, högsta utbildningsnivå, kön, indikator för boende i storstad, civilstånd, nationalitet, arbetserfarenhet, indikator för sektorkod (statlig/kommunal), indikator för egen företagare, indikator för pågående studier, indikator för arbetslöshet, BNP-tillväxt.

*Take-up av socialbidrag*⁷²

Riskgrupp: hushåll med en vuxen vars disponibla inkomst ligger under riksnormen för tilldelning av socialbidrag.

Modelltyp: logistisk regression.

Kovariater: ålder, kön, antal barn, antal barn under 7 års ålder, högsta utbildningsnivå, indikator för arbetslöshet, skillnad mellan disponibel inkomst och socialbidragsnorm, arbetserfarenhet, indikator för pågående studier, nationalitet, indikator för boende i storstad, indikator för tidigare skilsmässa.

Riskgrupp: hushåll med två vuxna vars disponibla inkomst ligger under riksnormen för tilldelning av socialbidrag.

Modelltyp: logistisk regression.

Kovariater: mannens ålder, kvinnans ålder, antal barn, antal barn under 7 års ålder, mannens högsta utbildningsnivå, kvinnans högsta utbildningsnivå, indikator för arbetslöshet för mannen, indikator för arbetslöshet för kvinnan, skillnad mellan disponibel inkomst och socialbidragsnorm, arbetserfarenhet för mannen, arbetserfarenhet för kvinnan, nationalitet för mannen, nationalitet för kvinnan, indikator för boende i storstad.

Nettosubventioner för offentliga tjänster

Händelse/utfall: subventionsbelopp för skola (förskoleklass, grundskola).

Riskgrupp: alla individer i 6–15 års ålder.

Modelltyp: regression (väntevärden⁷³).

Kovariater: ålder, åldersgrupp.

Händelse/utfall: subventionsbelopp för gymnasiestudier.

Riskgrupp: gymnasiestudenter.

Modelltyp: regression (väntevärden).

Kovariater: kön, nationalitet.

⁷² Termen take-up används för att beskriva den generella benägenheten att ansöka om ett bidrag vilken individen/hushållet är berättigat till. Take-up är alltså relevant för de bidragsformer som inte betalas ut med automatik, till exempel socialbidrag och bostadsbidrag.

⁷³ För dessa modeller görs framskrivningen utifrån modellens prediktion av väntevärden givet sammansättningen av kovariater för individen. Processens varianskomponent ignoreras här då det har visat sig vara svårt att uppskatta dess storlek.

Händelse/utfall: subventionsbelopp för komvuxstudier.

Riskgrupp: komvuxstuderande.

Modelltyp: regression (väntevärden).

Kovariater: kön, nationalitet, ålder.

Händelse/utfall: förekomst av subventionsbelopp för vuxenstudier (inkluderar särvox, statens skola för vuxna, svenska för invandrare, folkbildning/folkhögskola, kvalificerad yrkesutbildning).

Riskgrupp: individer i åldrarna 20–50 år som inte deltar i andra studieformer.

Modelltyp: logistisk regression.

Kovariater: kön, ålder, högsta utbildningsnivå, nationalitet.

Händelse/utfall: subventionsbelopp för vuxenstudier.

Riskgrupp: individer som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: kön, ålder, högsta utbildningsnivå, nationalitet.

Händelse/utfall: subventionsbelopp för universitets/högskolestudier.

Riskgrupp: universitetsstudenter.

Modelltyp: regression (väntevärden).

Kovariater: kön, ålder.

Händelse/utfall: förekomst av subventionsbelopp för barnomsorg.

Riskgrupp: individer i åldrarna 1 till 12 år.

Modelltyp: logistisk regression.

Kovariater: ålder, nationalitet, hushållets högsta utbildningsnivå, hushållets sysselsättningsgrad.

Händelse/utfall: subventionsbelopp för barnomsorg.

Riskgrupp: individer som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: ålder, åldersgrupp, nationalitet, hushållets högsta utbildningsnivå.

Händelse/utfall: förekomst av subventionsbelopp för äldreomsorg (hemtjänst, särskilt boende).

Riskgrupp: individer i åldrarna 65 och äldre.

Modelltyp: logistisk regression.

Kovariater: ålder, åldersgrupp, kön, nationalitet, inkomst från ålderspension (kvintiler), civilstånd, indikator för sysselsättning.

Händelse/utfall: subventionsbelopp för äldreomsorg.

Riskgrupp: individer som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: ålder, inkomst från ålderspension (kvintiler), kön, nationalitet.

Händelse/utfall: förekomst av subventionsbelopp för arbetsmarknadsåtgärder.

Riskgrupp: arbetslösa.

Modelltyp: logistisk regression.

Kovariater: kön, högsta utbildningsnivå, nationalitet.

Händelse/utfall: subventionsbelopp för arbetsmarknadsåtgärder.

Riskgrupp: individer som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: ålder, högsta utbildningsnivå, nationalitet.

Händelse/utfall: subventionsbelopp för vård (primärvård, slutenvård, tandvård).

Riskgrupp: individer som är 19 år och yngre.

Modelltyp: regression (väntevärden).

Kovariater: ålder, kön, nationalitet, hushållets högsta utbildningsnivå.

Kommentar: Samtliga individer som är 19 år och yngre antas erhålla vårdsuventioner.

Händelse/utfall: förekomst av vårdsuventioner.

Riskgrupp: individer i åldrarna 20–64 år.

Modelltyp: logistisk regression.

Kovariater: ålder, kön, indikator för förtidspension, indikator för sysselsättning, nationalitet, civilstånd, taxerad inkomst (kvartil).

Händelse/utfall: subventionsbelopp för vård.

Riskgrupp: individer i åldrarna 20–64 år som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: ålder, åldersgrupp, kön, nationalitet, högsta utbildningsnivå, taxerad inkomst (kvintil).

Händelse/utfall: förekomst av vårdsubventioner.

Riskgrupp: individer som är 65 år eller äldre.

Modelltyp: logistisk regression.

Kovariater: ålder, kön, nationalitet, civilstånd, ålderspension (kvintil).

Händelse/utfall: subventionsbelopp för vård.

Riskgrupp: individer som är 65 år eller äldre som slumpats i föregående modell.

Modelltyp: regression (väntevärden).

Kovariater: ålder, kön, nationalitet, civilstånd, ålderspension (kvintil).

Händelse/utfall: subventionsbelopp för läkemedel.

Riskgrupp: samtliga individer.

Modelltyp: beräknat riksgenomsnitt per kön och åldersgrupper (0–4, 5–9, ..., 90–).

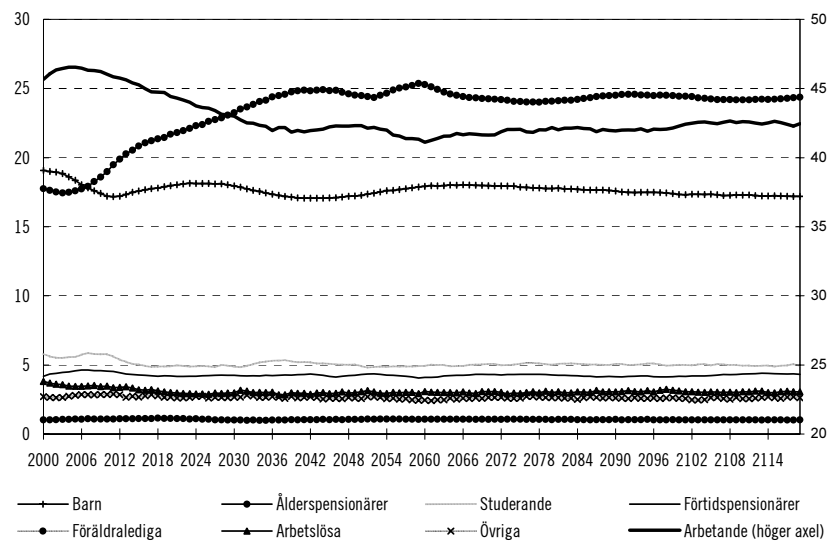
Appendix B: Validering av simulerade inkomster

I detta appendix jämförs de inkomster som genereras i SESIM med verkliga inkomster hämtade från LINDA och HEK. Först beskrivs hur befolkningens sammansättning med avseende på huvudsaklig sysselsättning (status) utvecklas över tiden. Därefter jämförs den fördelning av årsinkomster som SESIM genererar med motsvarande fördelning i de empiriska datamängderna. För att validera livsinkomstfördelningen jämförs fördelningen av 20-årsinkomster från LINDA med motsvarande fördelning simulerad i SESIM. Slutligen jämför den inkomströrlighet som SESIM genererar med den rörlighet som kan observeras i LINDA.

B.1 Status, stockar och flöden

Befolkningens sammansättning beskrivs i SESIM enligt individernas *status*. I diagram B.1 nedan beskrivs hur fördelningen mellan dessa status utvecklas över tiden.

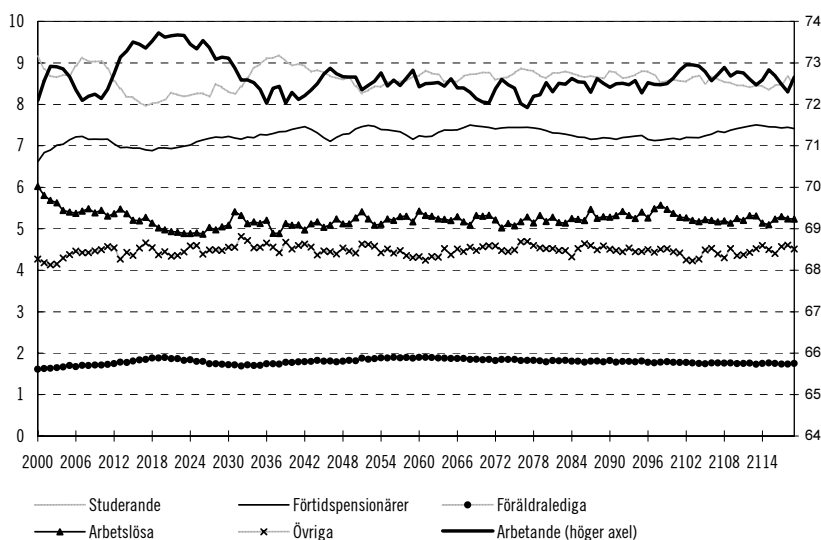
Diagram B.1 Befolkningen uppdelad efter status



Det mest dramatiska som händer är att andelen i arbete, efter en uppgång de första åren i simuleringen, faller kraftigt under en lång period samtidigt som andelen ålderspensionärer ökar. Denna utveckling är en effekt av befolkningens demografiska sammansättning.

För att se vad som händer i den del av befolkningen som är i aktiv ålder visas i diagram B.2 nedan motsvarande andelar för 16–64 åringar.

Diagram B.2 Befolkningen 16–64 år uppdelad efter status



Andelen i olika status är tämligen stabila över tiden, andelen i arbete varierar mellan 72 och knappt 74 procent.

B.2 Inkomst

En förutsättning för att analyser baserade på de hypotetiska inkomster som genereras av SESIM skall vara trovärdiga är att de på ett rimligt bra sätt återspeglar verkliga förlopp. Då hela, verkliga, livsinkomstbanor genererade under konstanta regler och i en stabil överraskningsfri ekonomisk miljö givetvis aldrig har och aldrig kommer att existera måste varje jämförelse med verkliga data vara partiell. Då det huvudsakliga perspektivet i denna bilaga är inkomstfördelning jämförs i första hand den inkomstfördelning som genereras av SESIM med verkliga fördelningar. Såväl årsvisa tvärsnittsfördelningar som fördelningar genererade under längre perioder jämförs.

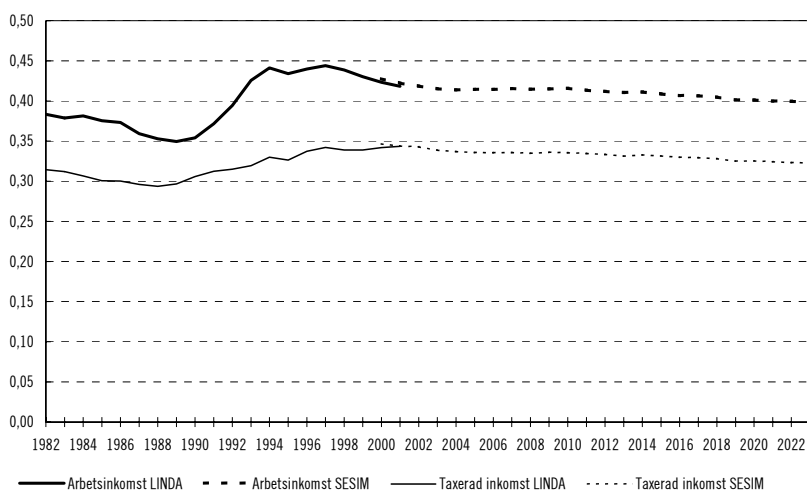
B.2.1 Tvärsnittsfördelning

I ett fördelningsperspektiv betraktas oftast hushållet som en mer intressant inkomstenhet än individen, hushållen antas i någon mån dela på de tillgängliga resurserna. Hushållens inkomster är dock i första hand uppbyggda av de individuella inkomster som hushållsmedlemmarna genererar på marknaden. När den av SESIM genererade inkomstfördelningen skall valideras är det därför naturligt att i ett första steg jämföra fördelningen av individuella inkomster.

Individuella inkomster

Den inkomstgenererande processen i SESIM baseras på ett inkomstbegrepp som ligger nära det arbetsinkomstbegrepp som används i inkomststatistiken. Det första testet blir därför att jämföra fördelningen av arbetsinkomster i SESIM med motsvarande fördelning i LINDA. Då endast individer i aktiv ålder kan förväntas ha någon arbetsinkomst begränsas jämförelsen till individer i åldrarna 20–64 år. I diagram B.3 nedan visas hur Gini-koefficienten för arbetsinkomst utvecklats mellan 1982 och 2001 enligt LINDA samt mellan 1999 och 2023 i simulerade data. Vi ser att fördelningen var som jämnast mot slutet av 1980-talet och att ojämnheten ökade markant i början av 1990-talet. Ökningen berodde delvis på den skattereform som genomfördes 1990–1991 men framförallt var det den ökade arbetslösheten som drev upp ojämnheten. Arbetslösa erhåller förvisso arbetsmarknadsstöd men detta stöd räknas inte in i arbetsinkomsten. Fördelningen enligt simulerade data ligger väldigt nära det verkliga utfallet för de år som överlappar, därefter sker en långsam minskning av ojämnheten. Detta skall inte ses som någon prognos av hur fördelningen av arbetsinkomster kommer att utvecklas. En sådan prognos skulle kräva ett detaljerat scenario avseende utveckling av utbildningsnivåer, arbetsmarknadens sektorer, relativlöner mellan utbildningsnivåer och sektorer mm. En sådan prognos faller utanför ramen för syftet med denna bilaga, fördelningsutfallet framåt i tiden är helt enkelt det utfall som erhålls om de samband som bygger upp SESIM får verka över tiden. Det viktiga för bilagans syften är att utvecklingen är stabil och inte drar iväg på ett orimligt sätt åt något håll.

Diagram B.3 Gini-koefficienten för individuella inkomster enligt LINDA respektive SESIM. 20–64 år



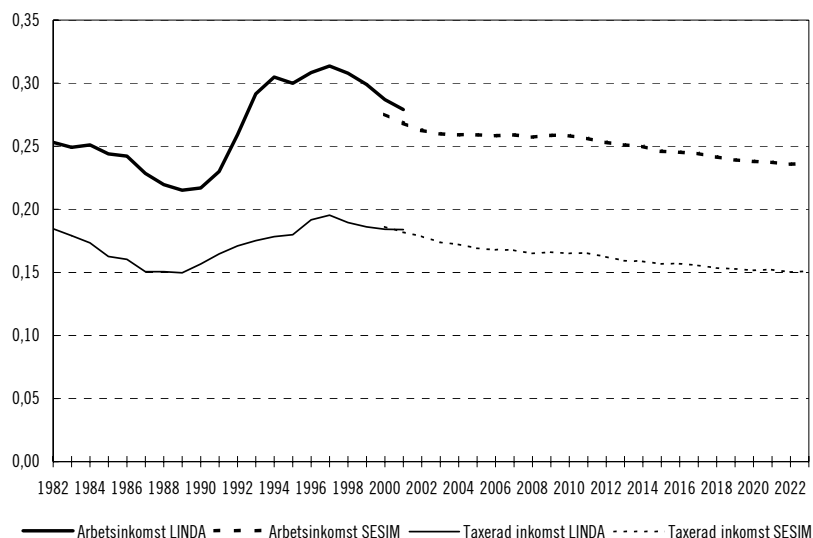
Källa: LINDA respektive SESIM.

I diagrammet visas även fördelningen av taxerad inkomst, i denna ingår förutom arbetsinkomst också arbetsmarknadsstöd samt pensioner. Utvecklingen av denna fördelning följer i stort fördelningen av arbetsinkomst men på en lägre nivå. Den ökade ojämnheten i början av 1990-talet framstår dock som mindre dramatisk då taxerad inkomst används, detta beror naturligtvis på att arbetsmarknadsstödet ingår.

Av särskild vikt för analyserna i denna bilaga är hur väl den nedre delen av fördelningen efterbildas. De system som har ett uttalat syfte att omfördela mellan individer på kort sikt, såsom t.ex. bostads- och socialbidragen, avräknas mot övriga inkomster. Om förekomsten av låga inkomster är för hög eller för låg så kommer omfattningen av dessa bidrag över- respektive underskattas. Detta kan i sin tur påverka de slutsatser som dras angående hur stor del av den totala omfördelningen som är omfördelning mellan individer.

I diagram B.4 nedan redovisas andelen med inkomster under halva medianinkomsten respektive år.

Diagram B.4 Andel med låg inkomst i LINDA respektive SESIM. Individuella inkomster 20–64 år

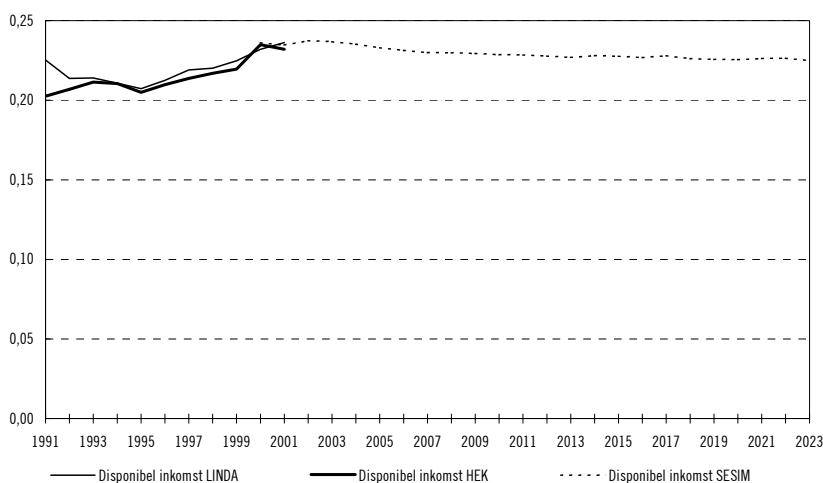


I SESIM tvingas alla individer in i en status hela året, är man arbetslös så är man det hela året och har således ingen arbetsinkomst. I verkligheten är många arbetslösa endast under en del av året. I LINDA, och all annan inkomststatistik, finns även individer med "oklar försörjningssituation" (svart arbete, företagare etc.) och dessa framstår i LINDA som låginkomsttagare. I SESIM finns inga som har låg inkomst på grund av mätproblem utan de som redovisas med låg inkomst är i någon mening genuina låginkomsttagare. Mot denna bakgrund är det förväntat att andelen låginkomsttagare blir lägre i simuleringen än i registerdata. Skillnaderna mellan SESIM och LINDA är därför förvånansvärt små och tvärsnittsfördelningen av arbets- och taxerade inkomster modelleras på ett tillfredsställande sätt såväl vad gäller fördelningen i sin helhet som den nedre delen.

Hushållsinkomster

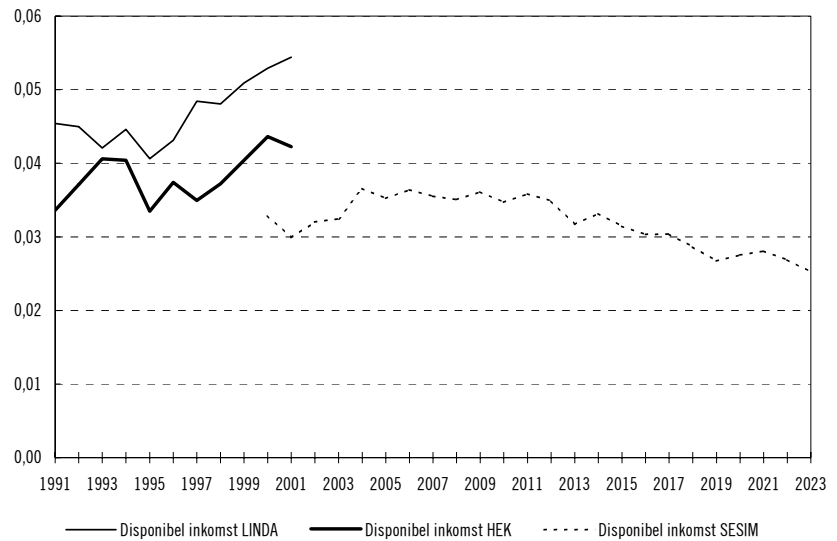
Fördelningen av de inkomster som SESIM genererar på individnivå överensstämmer förhållandevis väl med jämförbara verkliga fördelning. Detta medför inte automatiskt att fördelningen av hushållsinkomster hamnar "rätt". Korrelationerna mellan makars inkomster och hushållens sammansättning i övrigt kan påverka resultatet. Vidare beräknas vissa inkomstslag på hushållsnivå. Samtliga individer ingår i analysen, inkomsterna sammanräknas på hushållsnivå och justeras för försörjningsbörd. Redovisning sker på individnivå.

Diagram B.5 Gini-koefficienten för hushållsinkomster enligt LINDA, HEK och SESIM



Ojämnheten i den av SESIM genererade inkomstfördelningen tar vid på ungefär samma nivå som de sista observationerna av verkligt utfall. Nivån utvecklas därefter stabilt.

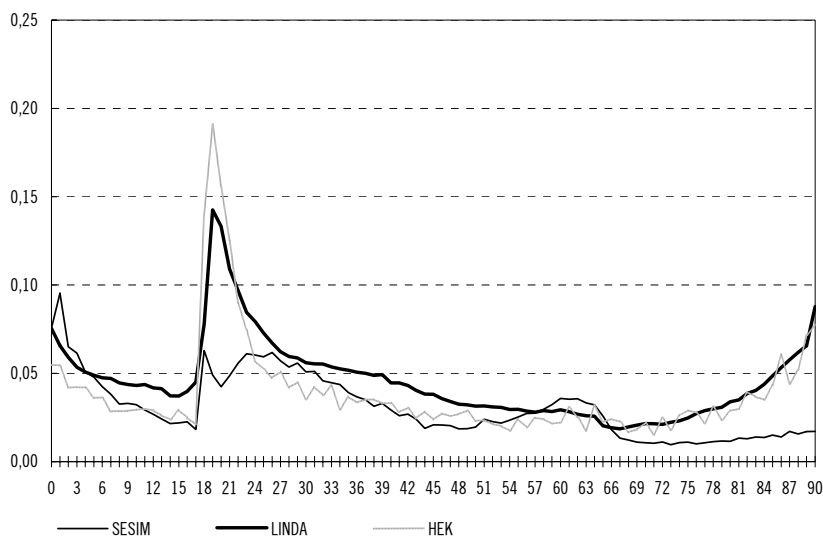
Diagram B.6 Andel med låg inkomst i LINDA, HEK och SESIM. Hushållsinkomster



Andelen med låg inkomst, här uttryckt som andelen med en inkomst som understiger halva medianinkomsten, underskattas i SESIM. Den nedre delen av fördelningen modelleras sämre än fördelningen i sin helhet. Som redan nämnts finns det i statistik över inkomstfördelningen alltid individer vars konsumtionsmöjligheter inte fångas av den registrerade disponibla inkomsten, många har en relativt god levnadsstandard men det går utifrån statistiken inte att avgöra av vad de lever. Motsvarande mätproblem finns inte i SESIM, alla är väldefinierade i en status. Detta i sig ger en lägre nivå på andelen med låga inkomster. Till detta kommer att regelsystemen i SESIM efterlevs exakt, inga bidrag betalas ut till de som inte är berättigade och alla betalar korrekt skatt. För social- och bostadsbidragen finns förvisso modeller för ansökningsbenägenheten men träffsäkerheten i bidragssystemen torde ändå bli betydligt högre i SESIMS stiliserade värld än i verkligheten.

Av diagram B.7 nedan framgår att det framförallt är bland ungdomar som precis flyttat hemifrån och bland de äldsta som andelen underskattas.

Diagram B.7 Andel med låg disponibel inkomst per ålder i LINDA HEK och SESIM



Anm.: Då antalet individer i respektive åldersklass är låg har andelarna beräknats som genomsnitt över flera år, för HEK och LINDA 1991–2001, för SESIM 1999–2023. Alla som är 90 år eller äldre slås samman till en grupp.

Källa: LINDA, HEK och SESIM.

Bland ungdomarna vet vi att både LINDA och HEK är behäftade med mätfel⁷⁴ avseende hushållsstrukturen som leder till en överskattning av andel med låg inkomst. I ungdomsgruppen finns förstås många som har flyttat hemifrån, saknar jobb och som inte är kvalificerade för a-kassa och som därmed har en låg faktisk inkomst. Andelen med låg inkomst bland ungdomarna är även korrekt mätt högre än för andra åldersgrupper men nivån överdrivs i HEK och LINDA. Den nivå som SESIM anger är å andra sidan förmodligen något för låg på grund av ovan nämnda skäl.

Bland de äldre beror den låga andelen på att SESIM tilldelar de äldre förmögenheter och kapitalinkomster i en något överdriven omfattning.

Vilka konsekvenser detta får för resultatet i kapitel 4 är inte helt lätt att förutsäga. Å ena sidan kan underskattningen av andelen

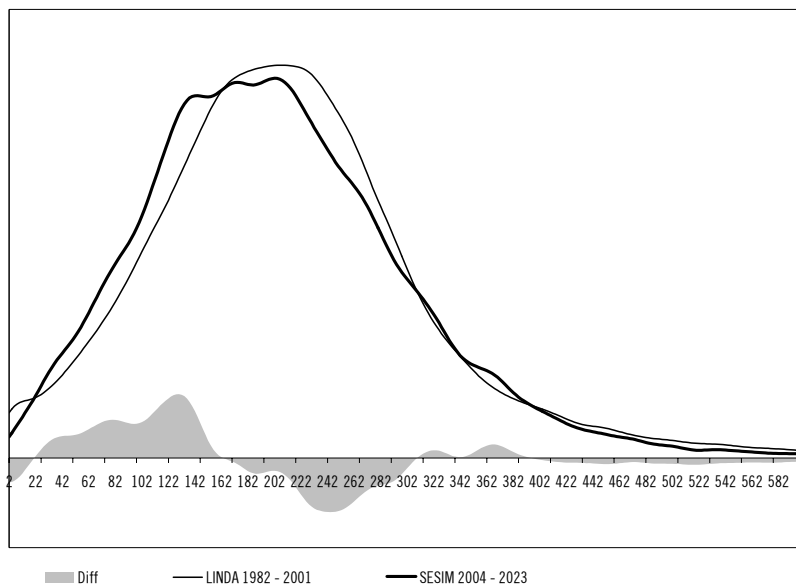
⁷⁴ Ungdomar som bor hos sina föräldrar betraktas ofta felaktigt som ensamstående, i HEK bygger förvisso hushållsbegreppet på enkätuppgifter men för bortfallet (25–30 procent av urvalet) används endast registeruppgifter.

med låg inkomst leda till en underskattning av behovet av omfördelade transfereringar och som en följd av detta en underskattning av den interpersonella omfördelningen. Å andra sidan kan den överdrivna träffsäkerheten i bidragssystemen leda till en överskattning av densamma.

B.2.2 Livsinkomstfördelning

I brist på data över hela livsinkomster att jämföra med används inkomster över en 20-årsperiod hämtade från LINDA. Den senast tillgängliga 20-årsperioden i LINDA är 1982–2001, denna period jämförs med simulerade data avseende perioden 2004–2023. När årsinkomsterna summeras till periodinkomster trendrensas data från LINDA med medianinkomsten 2001. I diagram B.8 nedan visas fördelningen av 20-åriga arbetsinkomster för de som var i aktiv ålder (20–64 år) i LINDA respektive SESIM. Endast de som har en arbetsinkomst samtliga 20 år ingår i redovisningen.

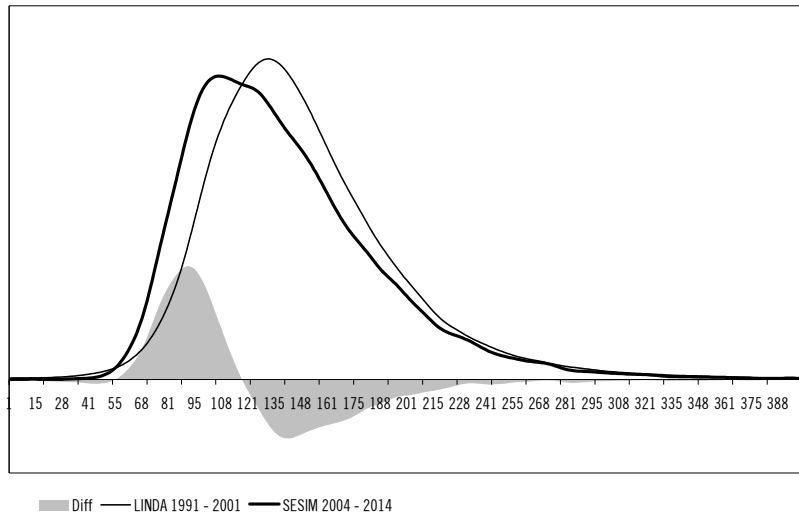
Diagram B.8 Fördelning av arbetsinkomst beräknad som genomsnitt under en 20-årsperiod



Den simulerade fördelningen av individuella arbetsinkomster sammanfaller överraskande väl med den verkliga fördelningen. Andelen med mycket låga inkomster (< 25 000) underskattas något och fördelningen är "förskjutet" bakåt. Den simulerade inkomstfördelningen är något ojämnare än vad som uppmäts i LINDA under den aktuella perioden, Gini-koefficienten för 20-årsinkomsten blir 0,295 respektive 0,284.

I diagram B.9 nedan visas motsvarande fördelningar av justerad disponibel inkomst över en 11-årsperiod.

Diagram B.9 Fördelning av justerad disponibel inkomst beräknad som genomsnitt under en 11-årsperiod



Även denna fördelning är förskjuten bakåt men fördelningens form påminner i allt väsentligt om motsvarande fördelning i LINDA. Den simulerade fördelningen är även här något ojämnare, Gini-koefficienten för 11-årsinkomsten är 0,189 respektive 0,181.

B.2.3 Inkomströrlighet

När olika inkomstslags omfördelningseffekter skall analyseras i ett tvärsnittsperspektiv kommer resultatet i stor utsträckning att vara beroende av hur fördelningen av övriga inkomster ser ut i tvärsnittet. Om marknadsinkomsterna är jämnt fördelade är möjligheterna att ytterligare utjämna med skatter och transfereringar begränsade. Om marknadsinkomsterna däremot är ojämnt fördelade finns stort utrymme för utjämning. När analysen av inkomstslagen däremot analyseras i ett livscykelperspektiv minskar tvärsnittsfördelningens betydelse för resultatet, i stället kommer rörligheten att vara avgörande. Om rörligheten är stor och individerna under sina liv befinner sig på olika platser i tvärsnittsfördelningen så kommer livsinkomstfördelningen att vara betydligt jämnare än årsinkomstfördelningarna. Möjligheterna att utjämna livsinkomst-

erna med skatter och transfereringar kommer då att vara begränsade. Det kan t.o.m. vara så att transfereringar som verkar utjämnande i tvärsnittet har motsatt effekt över livet. Om det däremot är samma individer som återfinns i botten av fördelningen under större delen av sina liv så kan omfördelning via skatter och transfereringar vara av stor betydelse även sett i ett livscykelperspektiv. Det är alltså av största vikt att den rörlighet som genereras av modellen på ett rimligt bra sätt återspeglar den rörlighet som kan observeras i verkligheten. För att validera inkomströrligheten jämförs simulerade inkomstbanor från SESIM med historiska inkomstbanor från LINDA. Valideringsperioden har satts till 20 år, detta motsvarar på intet sätt någon *livsinkomst* men periodlängden bedöms ändå vara tillräcklig för att utvärdera den inkomstgenererande processen. Genom att välja en grupp som är mellan 20–64 år (20–44 år startåret) kan rörligheten under en stor del av den aktiva perioden studeras. Före 20 års ålder lever många fortfarande på sina föräldrar och inkomströrligheten för dessa är mindre intressant. Efter pensionering är rörligheten låg och inkomstnivåerna bestäms i stort av de inkomster som genererades under den yrkesverksamma delen av livet. De inkomstbegrepp som analyseras är arbetsinkomst samt justerad disponibel inkomst. Arbetsinkomst beskriver väl den inkomstgenererande processen och rörlighet som beror på arbetslöshet fångas upp. Med disponibel inkomst fångas den rörlighet som beror på varierande familjesituation upp. Förutom åldersvillkoret krävs också att individerna finns med i de respektive panelerna samtliga år, de som dör eller utvandrar under perioden exkluderas.

Two olika sätt att mäta rörlighet används, först analyseras den totala rörligheten med *Shorrocks rigiditetsindex*,⁷⁵ därefter analyseras fördelningen av perioder med låg inkomst.

Shorrocks index

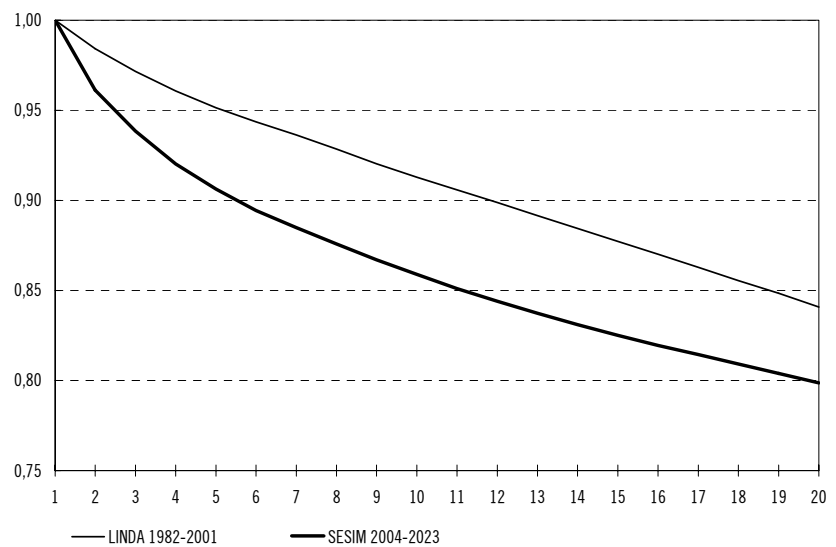
Shorrocks rigiditetsindex R definieras som kvoten mellan ojämnheten i den totala inkomsten över en längre period (mer än ett år) och det viktade medelvärdet av ojämnheten i de årliga inkomsterna under samma period. Formellt skrivs R som:

⁷⁵ Shorrocks [1978].

$$R = \frac{G_T}{\sum_{k=1}^T w_k G_k} \quad \text{där } w_k = \frac{\bar{Y}_k}{\sum_{i=1}^T \bar{Y}_i}$$

G_T avser Gini-koefficienten för den sammanlagda inkomsten under hela 20-årsperioden, G_k de årsvisa Gini-koefficienterna och \bar{Y}_i genomsnittsinkomsterna för motsvarande perioder. Rörligheten M definieras som $M=1-R$. Enligt denna definition är rörlighet det samma som den grad av utjämning som uppkommer av att mätperioden förlängs. Ojämnheten kan mätas med valfritt ojämnhetsindex, här används Gini-koefficienten.⁷⁶ I diagram B.10 nedan visas resultaten för arbetsinkomst.

Diagram B.10 Shorrocks R avseende arbetsinkomst



Källa: LINDA, SCB, Finansdepartementets beräkningar, SESIM.

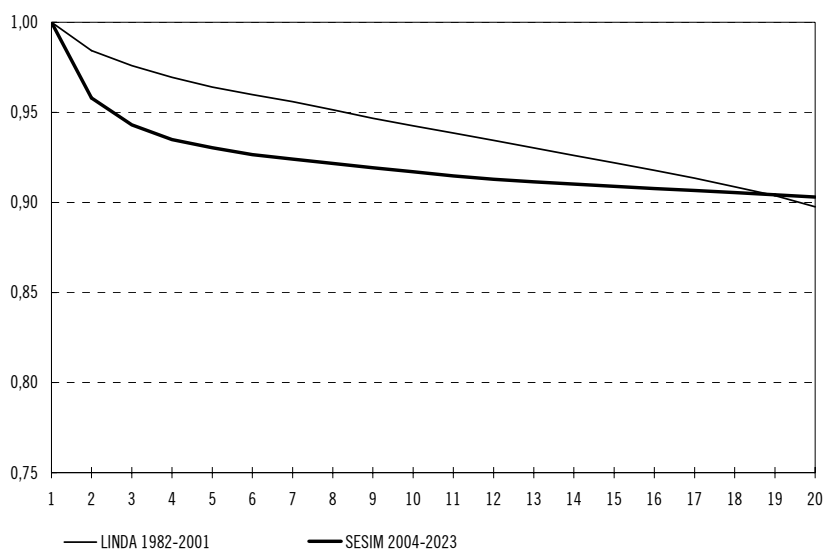
Rörligheten är större i SESIM och detta beror på att den årsvisa variationen är större. Detta kan ses i form av en snabbare minskning i R över tiden för SESIM än för LINDA. Spridningen i

⁷⁶ När rörlighet enligt Shorrocks index baseras på Gini-koefficienten så måste en individ byta plats så att den inbördes rankingen förändras för att någon rörlighet skall uppstå. Rörlighet som inte påverkar rangordningen registreras inte.

5-årsinkomster är i SESIM 10 procent lägre än spridningen i årsinkomster. I LINDA är dessa endast 5 procent lägre. När mätperioden förlängts till 6–7 år så ökar inte avståndet mellan kurvorna mer. Mot slutet antyds att skillnaden minskar, detta skulle i såna fall betyda att SESIM genererar mer kortsiktig variation men rörligheten i *permanent inkomst* är större i LINDA. Om perioden förlängs mer är det sannolikt att kurvorna möts.

I SESIM avser alla *status* helår, detta får till följd att den som blir arbetslös är det hela året och således inte får någon arbetsinkomst alls. I verkligheten, och följaktligen även i LINDA, så är många arbetslösa endast delar av året. En arbetslöshetsperiod ger därför upphov till en större förändring av arbetsinkomsten i SESIM jämfört med i LINDA. För att isolera en grupp som i huvudsak försörjer sig via arbete har analysen upprepats för den grupp som, under varje år i 20-årsperioden, har en arbetsinkomst som överstiger 75 000 kronor.

Diagram B.11 Shorrocks R avseende arbetsinkomst för de vars arbetsinkomst överstiger 75 000 om året



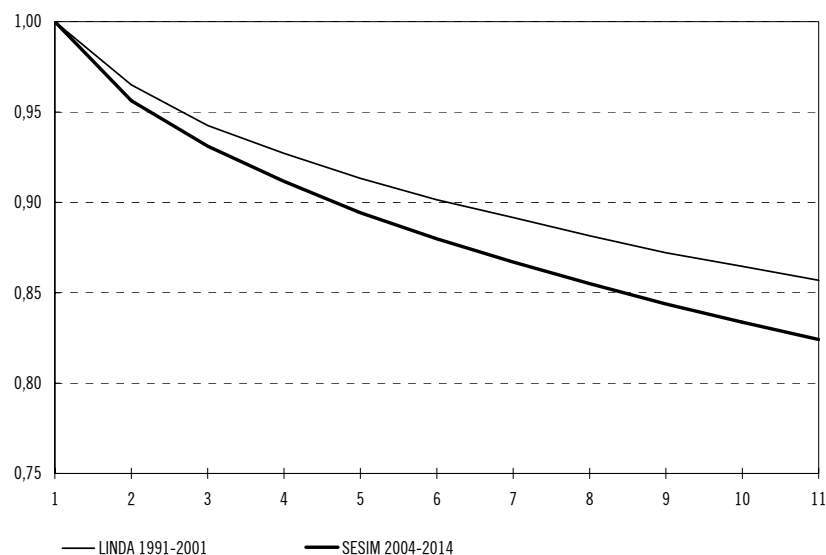
Källa: LINDA, SCB, Finansdepartementets beräkningar. SESIM.

När populationen begränsas på detta sätt (diagram B.11) bekräftas att SESIM genererar mer kortsiktig men mindre långsiktig rörlighet.

En förklaring till detta kan vara att 90-talskrisen, med exceptionellt hög arbetslöshet, ingår i den undersökta perioden (i LINDA). Sådana kriser genererar sannolikt en ökad rörlighet i permanent inkomst på ett sätt som inte återskapas i SESIM.

Analysen av rörlighet i disponibel inkomst baseras på kortare perioder beroende på att längre perioder med jämförbara LINDA-data är svåra att skapa. Av avgörande betydelse för analys av rörlighet i hushållsinkomster är hushållsbegreppet. Från 1991 och framåt finns LINDA-data med ett utvidgat hushållsbegrepp som kopplar ihop barn och föräldrar även när barnen blivit myndiga. Från LINDA används därför 11-årsperioden 1991–2001, från SESIM används åren 2004–2014.

Diagram B.12 Shorrock's R avseende justerad disponibel inkomst



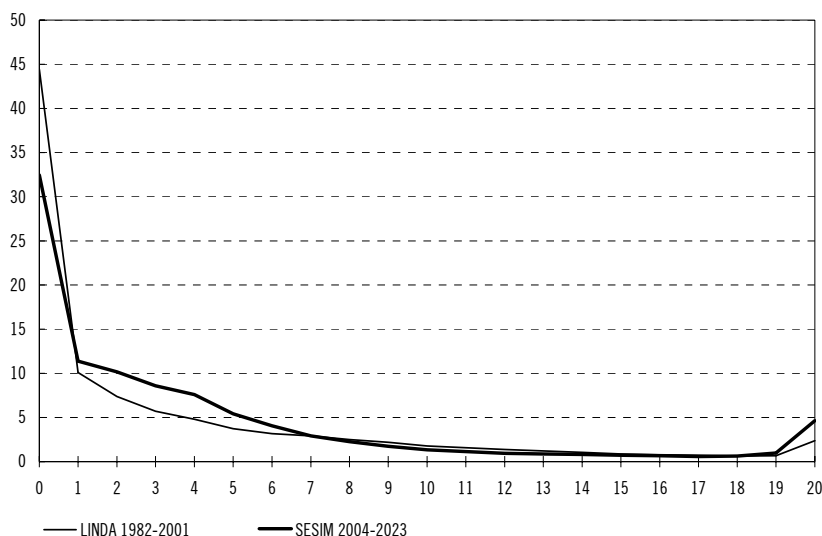
SESIM ger en högre rörlighet än LINDA men skillnaderna är mindre än när arbetsinkomsten studerades. Rörlighet i justerad disponibel inkomst kan uppkomma på en mängd olika sätt. Effekter som beror på karriär, arbetslöshet, sjukskrivning, eller förtids-

pension blandas med effekter av förändrade familjeförhållanden. Det är därför svårt att uttala sig om vad som orsakar de observerade skillnaderna i denna rörlighet mellan SESIM och LINDA. Vi nöjer oss därför med att konstatera att skillnaderna är förhållandevis små och att de därför troligen inte i någon större utsträckning kommer att störa analyserna i bilagan.

Låg inkomst

Shorrocks index fångar upp all typ av rörlighet och skiljer inte på rörlighet i toppen och i botten av fördelningen. Rörligheten i botten av fördelningen är dock extra intressant i livscykelperspektivet. Om samma personer ofta återfinns i botten av fördelningen kommer transfereringar som riktas mot dessa att verka omfördelande även över livet men om många individer befinner sig i botten under kortare perioder och på detta sätt delar på de svåra perioderna så kommer de transfereringar som de mottar under dessa perioder inte ha så starka omfördelande effekter i ett livscykelperspektiv. I diagram B.13 nedan visas fördelningen av antal perioder med låg arbetsinkomst under en 20-årsperiod i LINDA respektive SESIM. Låg inkomst definieras här som en inkomst understigande hälften av medianinkomsten under respektive år.

Diagram B.13 Fördelning av antal år med låg arbetsinkomst under en 20-årsperiod enligt LINDA respektive SESIM

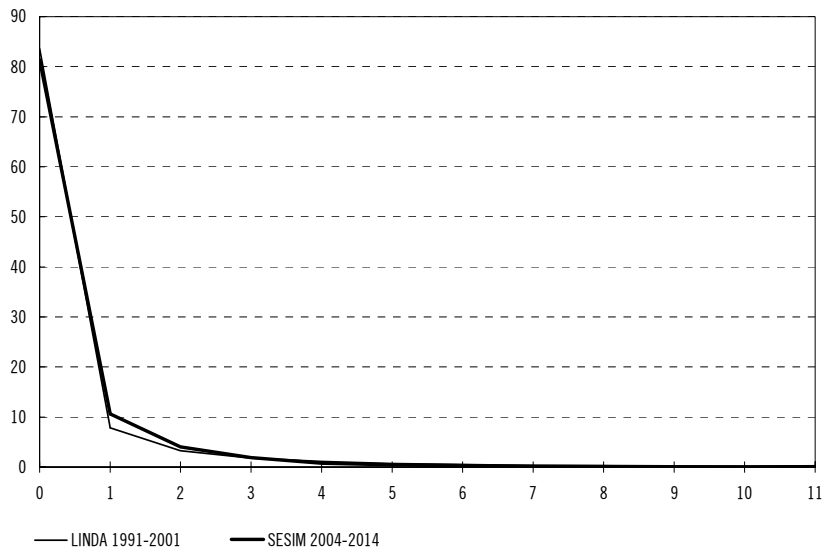


Drygt 55 procent av de som ingår i analysen upplevde minst ett år med låg arbetsinkomst enligt LINDA, motsvarande för SESIM var 67,5 procent. Att förekomsten av låg inkomst är högre i SESIM än i LINDA kan, mot bakgrund av vad som redovisades avseende tvärsnittsfördelningen i diagram B.4 framstå som något överraskande. Vi studerar här en grupp av individer som var 20–44 år det första året i 20-årsperioden och 40–64 år det sista året. Andelen med låg inkomst i tvärsnittet kommer att variera över tiden på grund av att åldersfördelningen förändras. I början dras andelen upp av ungdomarna och mot slutet av de äldre, lägst andel fås i mitten av perioden när åldrarna varierar mellan ca 30–55 år. I LINDA sammanfaller detta med högkonjunkturen i decennieskiftet mellan 1980 och 1990-talet då andelen med låga arbetsinkomster var väldigt liten. Detta medför att antalet perioder med låg inkomst totalt sett blir lägre i LINDA än i SESIM. Fördelningen av antalet perioder bland de som har låg inkomst i minst ett år är ganska lika i de bägge materialen. Överskottet av låginkomstperioder i SESIM ger en högre andel med 1–6 år samt 20 (eller fler) år. SESIM tycks alltså inte underskatta förekomsten av ”permanent bad luck”

De flesta upplever inte någon period med låg disponibel inkomst (83 procent i LINDA och 81 procent i SESIM) och fördelning av

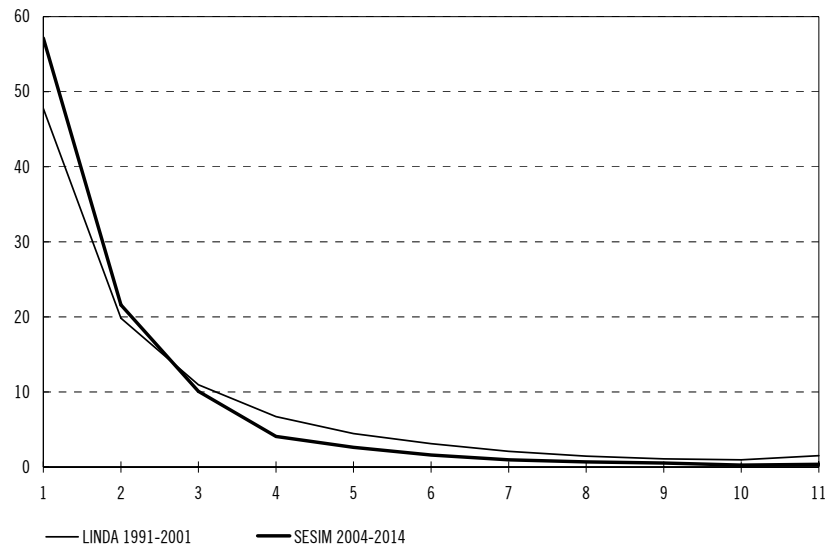
antal perioder för de som har minst ett år med låg inkomst är ganska lika.

Diagram B.14 Fördelning av antal år med låg disponibel inkomst under en 11-årsperiod enligt LINDA respektive SESIM



För att tydliggöra de skillnader som faktiskt finns visas i diagram B.15 nedan samma fördelning som i diagram B.14 men utan de som saknar låginkomstperioder.

Diagram B.15 Fördelning av antal år med låg disponibel inkomst under en 11-årsperiod enligt LINDA respektive SESIM bland de som har minst ett år



Förekomsten av låg inkomst är som väntat lägre i SESIM än i LINDA, den konjunkteffekt som gav det omvända resultatet för arbetsinkomst syns inte här. Detta kan till viss del förklaras av att andelen med låg disponibel inkomst inte är lika konjunkturkänslig, de som blir arbetslösa erhåller a-kassa samt eventuellt bostads- och socialbidrag. Huvudförklaringen är dock att populationen inte avgränsas lika hårt som i analysen av arbetsinkomst, perioden är kortare och en större del av befolkningen omfattas.

Fördelningen av antal perioder med låg inkomst simuleras så pass väl i SESIM att avvikelserna jämfört med LINDA knappast kan ha någon större effekt på analyserna i kapitel 3, 4 och 5.

Begrepp och definitioner

Nedan presenteras och förklaras kort några av de begrepp som används i bilagan. Förklaringarna avser främst de sätt på vilka begreppen används här. I andra sammanhang kan de olika begreppen ha en mer generell betydelse.

Beteendeeckvation: Statistisk modell där utfallet inte bara beror av individens/hushållets egenskaper utan även av omvärldens egenskaper. I till exempel en modell för prediktion av ett hushålls arbetsutbud påverkas detta inte bara av hushållets egenskaper utan kan även påverkas av förändringar i skatte- och transfereringssystemen.

Dynamisk process: Process som påverkas av historisk information. Till exempel en statistisk modell där den beroende variabeln är en funktion av värden på den beroende variabeln från tidigare tidpunkter.

Ekonomisk miljö: De antaganden som görs avseende den ekonomiska utvecklingen (löner, priser, tillväxt etc.).

Ekvivalensskala: Skala som beskriver hushållens relativa försörjningsbörda. Används för att justera inkomstnivåerna så att de blir jämförbara mellan hushåll av olika storlek och sammansättning.

Endogen: Variabel vars värde bestäms av modellutfallet. Jämför med exogen.

Estimera, skatta: En statistisk modell förklarar värdet av en ”beroende variabel” genom ”förklarande variabler” och parametrar. De olika variabelernas värden är givna i datakällan och parametrarnas värden estimeras, eller skattas, genom att hitta parametervärden sådana att modellen förklarar så mycket som möjligt av den variation som kan observeras i den beroende variabeln.

Exogen: Variabel vars värde bestäms (antas) utanför modellen. Jämför med endogen.

Imputering, imputeringsmodell: Påförande av information (variabler) till ett datamaterial genom användning av observerade samband, i ett annat datamaterial, mellan den påförda variabeln och variabler som finns tillgängliga i datamaterialet. Sambanden kan specificeras med hjälp av en imputeringsmodell.

Interceptterm: Term som anger den generella nivån (givet att samtliga förklarande variabler är lika med noll) för responsvariabeln i en statistisk modell.

Kalibrering: Metod för att generera ett modellutfall som sammanfaller med ett exogent givet antagande.

Kosthushåll: Utgörs av personer som bor i samma bostad och har gemensam "hushållning".

Kovariater: Förklarande variabler i en statistisk modell.

Monte Carlo - simulering: Metod för generering av slumpmässiga data.

Nettosubvention: Totalt (brutto-)subventionsvärde, till exempel för äldreomsorg, efter avdrag för erlagda avgifter.

Paneldata, longitudinella data: Data som insamlats vid minst två tidpunkter för hela, eller delar av, stickprovet. Data innehåller alltså information om förändringar mellan tidpunkter och inte bara om tillståndet vid en viss tidpunkt som för tvärsnittsdata.

Progressiv: Inkomstslag som bidrar till en utjämning av det studerade inkomstbegreppet.

Regelanalys: Analys där effekterna av ett eller flera regelsystem (för t.ex. skatter eller transfereringar) isoleras från andra effekter.

Registerdata: Data baserade på administrativa register till skillnad från data baserade på enkätundersökningar.

Regressiv: Inkomstslag som bidrar till en ökad ojämnheter i det studerade inkomstbegreppet.

Rekursiv struktur: Struktur för sekventiellt system av modeller där varje modell endast kan vara beroende av variabler som är antingen exogena eller som är responsvariabler i modeller som föregår den aktuella modellen. Motsatsen är en simultan struktur.

Riskaversion: Graden av riskbenägenhet, en riskbenägen individ har låg riskaversion medan en person som är obenägen att ta risker har hög riskaversion.

Riskgrupp: En grupp av objekt som löper risk att utsättas för en händelse. Till exempel individer som är "under risk" att bli förtidspensionerade.

Simulera: Generera data med hjälp av en matematisk funktion eller modell.

Självfinansierad: Transferering eller subvention som, vid någon tidpunkt i livet, finansierats av individen själv genom skatter eller avgifter.

Stokastisk: Styrdd av slumpen. I ett stokastiskt experiment som upprepas kan utfallet variera även om förutsättningarna är identiska.

Syntetisk datamaterial: Data som har genererats eller simulerats fram med hjälp av en modell eller funktion. Motsatsen är empiriska data som beskriver verkliga förhållanden.

Tvärsnittsdata: Data som insamlats vid en specifik tidpunkt, vanligen den 31 december ett visst år. Jämför paneldata/longitudinella data.:

Referenser

- Björklund, A. [1992], "Långsiktiga perspektiv på inkomstfördelningen", i *Inkomstfördelningens utveckling*, Bilaga 8 till Långtidsutredningen 1992.
- Blomquist, N. S. [1976], *The Distribution of Lifetime Income – A Case Study of Sweden*, doktorsavhandling, Princeton University.
- Davison, A. C. & Hinkley, D. V. [1997], *Bootstrap Methods and Their Application*, Cambridge University Press.
- Diggle, P. J., Liang, K. Y. & Zeger, S. L. [1994], *Analysis of Longitudinal Data*, Oxford University Press.
- Ds 1994: 86, *Fördelningseffekter av offentliga tjänster*, Rapport till Expertgruppen för studier i offentlig ekonomi (ESO) av Johan Fritzell.
- Ds 1994:135, *Skatter och socialförsäkringar över livscykeln – en simuleringsmodell*, Rapport till Expertgruppen för studier i offentlig ekonomi (ESO) av Joakim Hussénus & Jan Selén.
- Ds 2000:19, *Studiebidragen i det långa loppet*, Rapport till Expertgruppen för studier i offentlig ekonomi (ESO) av Peter Ericson och Joakim Hussénus.
- Edin, P. A. & Fredriksson, P. [2000], "LINDA – Longitudinal Individual Data for Sweden", Working Paper 2000:19, National-ekonomiska institutionen, Uppsala Universitet.
- Falkingham, J. & Harding, A. [1996], *Poverty Alleviation vs Social Insurance Systems: A Comparison of Lifetime Redistribution*, Contributions to Economic Analysis, vol. 232, North Holland, Amsterdam.
- Falkingham, J. & Lessof, C. [1992], "Playing god: The construction of LIFEMOD, a Dynamic Cohort Microsimulation Model", in Hancock, R. & Sutherlands, H. (red.) *Microsimulation Models for Public Policy Analysis: New Frontiers*, STICRED Occasional Paper No. 17, London School of Economics, London.

- Galler, H. P. [1997], "Discrete-Time and Continuous-Time Approaches to Dynamic Microsimulation reconsidered", NATSEM Technical Paper No. 13, October, National Centre for Social and Economic Modelling, University of Canberra.
- Harding, A. [1993], *Lifetime Income Redistribution: Applications of a Dynamic Cohort Microsimulation Model*, North Holland, Amsterdam.
- Harding, A. m.fl. [2000], "The Lifetime Distributional Impact of Government Health Outlays", Discussion Paper No. 47, National Centre for Social and Economic Modelling, University of Canberra.
- Kakwani, N. C. [1977], "Applications of Lorenz Curves in Economic Analysis, *Econometrica*, Vol. 45, No. 3.
- Klevmarken, A. [1997], "Behavioral Modeling in Micro Simulation Models. A survey", Working Paper 1997:31, Nationalekonomiska institutionen, Uppsala Universitet.
- Klevmarken, A. [1998], "Statistical Inference in Micro Simulation Models: Incorporating External Information", Working Paper 1998:20, Nationalekonomiska institutionen, Uppsala Universitet.
- Nelissen, J. H. M. [1998], "Annual Versus Lifetime Income Redistribution by Social Security", *Journal of Public Economics*, 68. pp. 223–249.
- O'Donoghue, C. [2001a], "Dynamic Microsimulation: A Methodological Survey", *Brazilian Electronic Journal of Economics*, Vol. 4.
- O'Donoghue, C. [2001b], *Redistribution in the Irish Tax-Benefit System*, doktorsavhandling, London School of Economics, London.
- Pylkkänen, E. [2003], *Studies on Household Labor Supply and Home Production*, doktorsavhandling, Nationalekonomiska institutionen, Göteborgs universitet.
- Rake, K. & Zaidi, A. [2001], "Dynamic Microsimulation Models: A Review and Some Lessons for SAGE", SAGE Discussion Paper No. 2 (SAGEDP/02).
- Regeringens proposition [1999/00:1], *Budgetpropositionen för 2000*, Bilaga 4, Fördelningspolitisk redogörelse.
- Regeringens proposition [2001/02:100], *2002 års Vårpropositionen*, Bilaga 3, Fördelningspolitisk redogörelse.
- SCB [2002], *Sveriges framtida befolkning 2002–2050*, Statistiska meddelanden, BE18SM0201.
- SCB [2003], *Inkomstfördelningsundersökningen 2001*, Statistiska meddelanden, HE 21 SM 0301.

- Shorrocks, A. [1978], "Income Inequality and Income Mobility", *Journal of Economic Theory* 19, pp. 376–393.
- Smeeding, T. m.fl. [1993], "Poverty, Inequality, and Family Living Standards Impacts Across Seven Nations: The Effect of Noncash Subsidies for Health, Education and Housing, *The Review of Income and Wealth*, Series 39, No. 3.