



IFAU – INSTITUTET FÖR
ARBETSMARKNADSPOLITISK
UTVÄRDERING

Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?

Anders Björklund
Peter Fredriksson
Jan-Eric Gustafsson
Björn Öckert

RAPPORT 2010:13

Institutet för arbetsmarknadspolitisk utvärdering (IFAU) är ett forskningsinstitut under Arbetsmarknadsdepartementet med säte i Uppsala. IFAU ska främja, stödja och genomföra vetenskapliga utvärderingar. Uppdraget omfattar: effekter av arbetsmarknadspolitik, arbetsmarknadens funktionssätt, arbetsmarknadseffekter av åtgärder inom utbildningsväsendet och arbetsmarknadseffekter av socialförsäkringen. IFAU ska även sprida sina resultat så att de blir tillgängliga för olika intressenter i Sverige och utomlands.

IFAU delar även ut forskningsbidrag till projekt som rör forskning inom dess verksamhetsområden. Forskningsbidragen delas ut en gång per år och sista dag för ansökan är den 1 oktober. Eftersom forskarna vid IFAU till övervägande del är nationalekonomer, ser vi gärna att forskare från andra discipliner ansöker om forskningsbidrag.

IFAU leds av en generaldirektör. Vid institutet finns ett vetenskapligt råd bestående av en ordförande, institutets chef och fem andra ledamöter. Det vetenskapliga rådet har bl.a. som uppgift att lämna förslag till beslut vid beviljandet av forskningsbidrag. Till institutet är även en referensgrupp knuten där arbetsgivar- och arbetstagsidan samt berörda departement och myndigheter finns representerade.

Rapporterna finns även i tryckt format. Du kan beställa de tryckta rapporterna via telefon eller mejl. Se nedanstående kontaktinformation.

Postadress: Box 513, 751 20 Uppsala

Besöksadress: Kyrkogårdsgatan 6, Uppsala

Telefon: 018-471 70 70

Fax: 018-471 70 71

ifau@ifau.uu.se

www.ifau.se

IFAU har som policy att en uppsats, innan den publiceras i rapportserien, ska seminariebehandlas vid IFAU och minst ett annat akademiskt forum samt granskas av en extern och en intern disputerad forskare. Uppsatsen behöver dock inte ha genomgått sedvanlig granskning inför publicering i vetenskaplig tidskrift. Syftet med rapportserien är att ge den ekonomiska politiken och den ekonomisk-politiska diskussionen ett kunskapsunderlag.

Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?

av

Anders Björklund^{*}, Peter Fredriksson[#], Jan-Eric Gustafsson[⊗], Björn Öckert[⊕]

10 juni 2010

^{*} Institutet för social forskning, Stockholms universitet. E-post: anders.bjorklund@sofi.su.se.

[#] Nationalekonomiska institutionen, Stockholms universitet. E-post: peter.fredriksson@ne.su.se.

[⊗] Inst. för pedagogik och didaktik, Göteborgs universitet. E-post: jan-eric.gustafsson@ped.gu.se.

[⊕] Institutet för arbetsmarknadspolitisk utvärdering. E-post: bjorn.ockert@ifau.uu.se.

Innehållsförteckning

1	Inledning.....	7
1.1	Utvecklingen av utbildningspolitiken.....	8
1.1.1	Den obligatoriska skolan.....	9
1.1.2	Förskolan.....	10
1.1.3	Gymnasieskolan.....	11
1.1.4	Högskoleutbildning.....	12
1.1.5	Forskarutbildningen.....	13
1.2	Frågor för denna rapport.....	13
2	Svenska elevers kunskapsutveckling.....	18
2.1	Betyg.....	20
2.1.1	Betygsutvecklingen sedan 1989.....	22
2.1.2	Förändringar i betygsskillnader mellan skolor och kommuner.....	25
2.1.3	Diskussion och slutsatser.....	27
2.2	Nationella utvärderingar.....	28
2.3	Internationella studier.....	30
2.3.1	Utvecklingen av de internationella studierna.....	31
2.3.2	Några metodproblem i de internationella studierna.....	33
2.3.3	Kunskapsutvecklingen sedan 1960-talet.....	36
2.3.4	Diskussion och slutsatser.....	43
2.4	Andra prov.....	45
2.5	Diskussion och slutsatser.....	48
3	Subventioner: Argument och inriktning.....	54
3.1	Argument för utbildningssubventioner.....	56
3.1.1	Externa effekter av utbildning.....	56
3.1.2	Andra former av marknadsmisslyckanden.....	58
3.1.3	Utbildningssubventioner för att korrigera snedvridningar.....	58
3.2	Hur ska subventioner riktas?.....	59
3.3	Sammanfattning och diskussion.....	63
4	Avkastningen på utbildning.....	72
4.1	Begrepp och frågeställningar.....	72
4.2	Internationell jämförelse.....	77
4.3	En historisk jämförelse.....	82
4.4	Sammanfattning och diskussion.....	85
5	Effekter av utbildning på kunskaper och lön.....	89

5.1	Effekter av utbildningsinvesteringar på inkomster	90
5.1.1	Studier som beaktar observerade egenskaper	91
5.1.2	Studier som jämför syskon (tvillingar)	96
5.1.3	Studier som utnyttjar institutionella förhållanden	102
5.1.4	Vad vet vi om avkastningen på utbildning i Sverige?	109
5.2	Effekter av utbildningsinvesteringar på kunskaper	110
5.2.1	Studier som beaktar observerade egenskaper	112
5.2.2	Studier som utnyttjar reglerna för skolstart	115
5.2.3	Studier som utnyttjar studieavbrott	117
5.2.4	Hur stora är effekterna i Sverige?	119
5.3	Effekter av kunskaper på inkomster	120
5.4	Sammanfattning	130
6	Daghem och förskola: vad vet vi om effekterna för barnen?	139
6.1	Vad vi vill veta	143
6.2	Metodproblem	144
6.3	Svenska studier	147
6.4	Studier från andra länder	151
6.5	Studier av tidigare skolstart	156
6.6	Slutsatser	158
7	Lärartäthet och lärarkompetens	164
7.1	Lärartäthet	165
7.1.1	Några grundläggande metodproblem	165
7.1.2	Metaanalyser	167
7.1.3	Studier från andra länder	168
7.1.4	Resultat från svenska undersökningar	172
7.1.5	Diskussion och slutsatser	175
7.2	Lärarkompetens	176
7.2.1	Prestationsrelaterad lärarkompetens	178
7.2.2	Vad påverkar den prestationsrelaterade lärarkompetensen?	182
7.2.3	Lärares kunskaper och färdigheter	183
7.2.4	Den formella kompetensens betydelse	187
7.2.5	Rekrytering till läraryrket	190
7.3	Diskussion och slutsatser	191
8	Organisatoriska frågor	199
8.1	Differentiering	201
8.1.1	Organisatorisk differentiering	201

8.1.2	Nivågruppering.....	204
8.1.3	Specialpedagogiskt stöd	210
8.1.4	Kamrateffekter.....	212
8.1.5	Diskussion och slutsatser.....	215
8.2	Skolval och konkurrens	216
8.3	Decentralisering.....	223
8.4	Inspektioner	225
8.4.1	Skolinspektioner i Sverige.....	225
8.4.2	Erfarenheter av inspektioner från andra länder	226
8.5	Diskussion och slutsatser.....	228
9	Gymnasieskolan och vuxenutbildningen.....	234
9.1	Gymnasiet – en beskrivning	236
9.2	Effekter av ett tredje år i yrkesutbildningen	241
9.3	Konsekvenser av det målrelaterade betygssystemet.....	245
9.4	Kommunal vuxenutbildning.....	254
9.5	Diskussion och slutsatser.....	259
10	Högre studier	264
10.1	Högskolornas utbud av studieplatser	268
10.1.1	Några principiella utgångspunkter.....	268
10.1.2	Dimensioneringen i praktiken	269
10.1.3	Konsekvenser av överutbildning	270
10.2	Studenternas efterfrågan på studieplatser	275
10.2.1	När uppträder selektion till högskolan?.....	277
10.2.2	Studieavgifter och studiemedel	280
10.2.3	Regionalisering av högskolan.....	284
10.3	Antagningssystemets utformning	287
10.3.1	Prediktionsvärdet för olika urvalsinstrument.....	287
10.3.2	Påverkar urvalsinstrumenten vilka som går på högskolan?.....	293
10.3.3	Finns en koppling mellan urvalsinstrument och avkastning?.....	302
10.4	Avkastning på högskolestudier vid olika lärosäten	309
10.4.1	Studier som beaktar observerade egenskaper	309
10.4.2	Studier som jämför syskon (tvillingar).....	312
10.4.3	Studier som utnyttjar begränsningar i utbildningsvalen	313
10.5	Avkastning på högskolestudier med olika inriktning	314
10.6	Sammanfattning.....	316
11	Avslutning	326

11.1	Sammanfattning av forskningsgenomgången.....	326
11.2	Policydiskussion	332
11.2.1	Är avkastningen på utbildning för låg i Sverige?	332
11.2.2	Varför har svenska grundskoleelevers resultat försämrats?	333
11.2.3	Hur ska betygssystemet utformas?	337
11.2.4	Är skolan en central eller kommunal angelägenhet?	339
11.3	Bättre förutsättningar för forskning och utvärdering.....	340

Förord

Huvudsyftet med denna rapport är att försöka besvara frågan: Vilka arbetsmarknadseffekter har svensk utbildningspolitik? För att besvara denna fråga genomför vi både översikter på olika områden och egna empiriska analyser. Vi fokuserar på kvantitativa studier av hur enskilda åtgärder påverkar utfall på individnivå; särskilt fokus ligger på studier som har en tydlig strategi för att mäta orsakssamband.

Ambitionen är att täcka de viktiga åtgärderna inom det reguljära utbildningssystemet, från förskola till grundläggande tertiär utbildning. En konsekvens av denna ambition är att rapporten är mycket omfattande. Rapporten är den första heltäckande genomlysningen av svensk utbildningspolitik, med en kvantitativ utgångspunkt. På grund av rapportens omfattning är det inte meningsfullt att sammanfatta den i ett antal huvudresultat. Sammanfattningen av forskningsgenomgången återfinns i kapitel 11, där vi också diskuterar ett antal policyfrågor som vi uppfattar som särskilt relevanta.

Denna rapport är ett resultat av ett samarbete mellan tre ekonomer och en pedagog. Eftersom vi kommer från två olika discipliner har vi inte alltid haft en samsyn på vad forskning säger. Bristen på samsyn har dock varit anmärkningsvärt liten. En avgörande förutsättning för att samarbetet har fungerat så väl är att vi är empiriska forskare. Vi har alla en stark tilltro till empiriska analyser och en grundläggande förståelse för de krav som ställs om man utifrån empiriska samband vill göra kausala utsagor.

Rapporten är en gemensam produkt. Men som i många andra sammanhang finns det anledning att dra nytta av komparativa fördelar och specialisering. Arbetet med kapitel 2–10 har därför delats mellan oss. Anders Björklund har varit huvudansvarig för kapitel 6, Peter Fredriksson för kapitel 3 och 9, Jan-Eric Gustafsson för kapitel 2 och 7 och Björn Öckert för kapitel 5 och 10. Anvaret för kapitel 4 har delats mellan Anders Björklund och Peter Fredriksson och ansvaret för kapitel 8 mellan Anders Björklund och Jan-Eric Gustafsson.

Under arbetets gång har vi fått många kommentarer. Vi är särskilt tacksamma för Erik Grönqvists, Mikael Lindahls och Jonas Vlachos insiktsfulla kommentarer. Vi tackar även seminariedeltagare på Finansdepartementet och Utbildningsdepartementet och deltagare i ELE-workshopen ”The Economics of Education in Sweden”.

1 Inledning

Avsevärda resurser läggs på utbildning i alla OECD länder. Sverige är inget undantag: De direkta utbildningsutgifterna utgjorde 6,3 % av BNP 2006, vilket kan jämföras med 5,7 % som genomsnitt i OECD vid samma tidpunkt (OECD 2009).¹

Vilka intäkter genererar då dessa utgifter? Ett viktigt motiv för utbildningsinsatserna är att förbättra individens kunskaper vilket i sin tur ska höja deras produktivitet och förbättra deras sysselsättningsmöjligheter, något som ökar den samlade produktionen i landet.

Syftet med denna rapport är att undersöka om utbildningsinsatser har sådana effekter för den enskilde individen. Huvudfrågan i rapporten är: Vilka arbetsmarknadseffekter har svensk utbildningspolitik? Med ”arbetsmarknadseffekter” menar vi huvudsakligen effekter på individens löner, sysselsättning och arbetsinkomster. Men vi kommer också att studera utfall som betyg och testresultat.

Vår ambition med rapporten är att täcka forskningsläget inom i princip hela det reguljära utbildningsväsendet, från förskolan till den grundläggande högskolenivån. Däremot täcker vi inte utbildning som anordnas av Arbetsförmedlingen eller personalutbildning² Översikten har ett kvantitativt fokus och vi lägger särskild vikt vid studier som har en tydlig strategi för att skatta orsakssambandet mellan en utbildningspolitisk intervention och t.ex. ett arbetsmarknadsutfall.

Statistiska utvärderingar är ett nödvändigt hjälpmedel för att utforma en effektiv och rättvis skolpolitik. Eftersom en av utbildningsväsendets huvuduppgifter är att ge ungdomar de färdigheter som är nödvändiga för att nå framgång på arbetsmarknaden, kommer kunskaper såsom de mäts av standardiserade tester att ha en framstående plats i denna översikt. Notera att testresultat i relativt ung ålder har en god prognosförmåga för hur individen klarar sig på arbetsmarknaden; dessa resultat kan därför användas för att få en uppfattning om hur framtida arbetsmarknadsutfall påverkas av en specifik åtgärd.

Den kvantitativa pedagogiska forskningen har en lång tradition i Sverige; svenska forskare var också föregångare i ett internationellt perspektiv. Tidigt gjordes försök att mäta elevers kognitiva förmåga. Viktiga personer på området var Gustav Jäderholm, Siver Hallgren, Torsten Husén och Kjell Härnquist (se

¹ I kostnadskalkylen borde man förstås även beakta all den tid som individer själva lägger ner på studier.

² Den intresserade läsaren kan konsultera översikten av Calmfors m.fl. (2002) då det gäller arbetsmarknadsutbildningen och Ericson (2004) då det gäller personalutbildning.

t.ex. Pearson och Jäderholm 1914, Hallgren 1939, Husén 1951, och Härnquist 1959). Den utbildningsekonomiska forskningen är av senare datum. Internationellt tog den fart runt 1960 med viktiga bidrag av Theodore Schultz (1960) och Gary Becker (1964).

Den utbildningsekonomiska forskningen handlar framförallt om hur formell utbildning och olika typer av utbildningsinterventioner påverkar individers arbetsmarknadsframgång och ekonomin som helhet. Det tar dock en viss tid innan man kan spåra arbetsmarknadseffekterna av enskilda utbildningspolitiska åtgärder – det är ju först när individerna träder in på arbetsmarknaden som det går att mäta effekten av en enskild åtgärd. Därför är det vanligt att man i den utbildningsekonomiska forskningen också studerar mått som är relaterade till kognitiva förmågor, såsom betyg och testresultat.

Det finns ett ytterligare skäl till att studera betyg och testresultat. Även om en given utbildningsintervention påverkar kunskaper i en viss riktning (och därmed borde påverka arbetsmarknadsframgången i samma riktning) är det inte säkert att dessa effekter går att spåra i framtida arbetsmarknadsutfall. Om inriktningen på utbildningspolitiken är kompensatorisk – dvs. inriktad på att stödja individer med sämre studieförframgång – kommer elever som fått ett försteg, till följd av att de fått del av en satsning, sannolikt att få mindre stöd i framtiden. I processen mellan den ursprungliga åtgärden och det slutliga arbetsmarknadsutfallet tunnas effekten ut, vilket gör det svårare att spåra effekten av interventionen på arbetsmarknaden. Även om vi inte kan spåra några skillnader i arbetsmarknadsutfall, är det således inte givet att vi kan dra slutsatsen att den ursprungliga åtgärden var utan effekter.

1.1 Utvecklingen av utbildningspolitiken³

I Sverige liksom i andra länder har det skett en kraftig höjning av befolkningens utbildningsnivå. I levnadsnivåundersökningen (LNU) har man sedan den första undersökningen 1968 ställt frågan: ”Hur många år har Er sammanlagda skol- och yrkesutbildning på heltid varat (från småskolan och uppåt)?” till ett representativt urval av svenska folket. Den genomsnittlige individen i åldern 20–65 uppgav 8,7 år i 1968 års undersökning och 12,7 år i 2000 års undersökning, dvs. en ökning med 4 år eller nära 50 % under 32 år. Detta är en

³ Avsnittet bygger på Richardson (2004).

mycket kraftig ökning under en relativt kort period. Att ökningen är kraftig betyder dock inte att den nödvändigtvis är för stor.⁴

Bakom en sådan höjning av befolkningens utbildningsnivå ligger en expansion av skolpolitiken. Vi ska i detta avsnitt ge en kort översikt av denna expansion med tonvikt på viktiga politiska beslut under de sista 50–60 åren. Vi börjar med att beskriva utvecklingen av den obligatoriska skolan; sedan övergår vi till den frivilliga utbildning som kommer före (förskolan) respektive efter (gymnasium och universitet) detta obligatorium. Den frivilliga utbildningens omfattning bestäms i ett samspel mellan befolkningens efterfrågan och utbudet av studieplatser, där utbudet på kort sikt är politiskt styrt. Även efterfrågan på studieplatser påverkas dock av politiska beslut genom exempelvis studiebidragens och skattesystemets utformning.

1.1.1 Den obligatoriska skolan

1842 brukar anföras som året då Sverige fick en obligatorisk folkskola. Vid denna tidpunkt var det dock snarare frågan om undervisningsskyldighet, dvs. en skyldighet för städer och församlingar på landet att erbjuda undervisning. Först 1882 kom en skolstadga som också innebar skolplikt – vid denna tidpunkt sex år med början vid sju års ålder. År 1937 förlängdes folkskolan till sju år för dem som inte dessförinnan gått över till realskolan, men kommunerna fick 12 år på sig att genomföra denna förlängning.

Nästa förlängning av den obligatoriska utbildningslängden genomfördes som en del av den stora grundskolereformen. Denna reform kom att få ett upplägg av stort intresse från ett utvärderingsperspektiv, något som vi återkommer till i rapporten. År 1950 fattade riksdagen ett beslut som tog sikte på en mer enhetlig nioårig obligatorisk skola – kallad enhetsskolan – som skulle ersätta den tidigare kombinationen av folkskola och realskola. Beslutet innebar dock en försöksverksamhet där olika kommuner kom att införa enhetsskolan vid olika tillfällen med sikte på att dra lärdomar om lämpligheten av reformen. Den första åldersgruppen som kom att omfattas var född 1938. Därefter kom ett ökande antal kommuner att införa enhetsskolan. Erfarenheterna som gjordes under loppet av 1950-talet bedömdes så goda att riksdagen 1962 beslutade att genomföra en obligatorisk nioårig grundskola i hela landet. Varje kommun ålades att genomföra en sådan senast 1967 vilket innebar att åldersgrupperna födda 1957 och senare i sin helhet kom att gå i den nya grundskolan.

⁴ För att ge perspektiv på denna utbildningsexpansion kan man notera att under samma tid ökade den förväntade livslängden med cirka 5,5 år för både män och kvinnor.

Grundskolan var en påtagligt mer sammanhållen obligatorisk skola fram till 16 års ålder i jämförelse med den tidigare kombinationen av relativt tidig uppdelning i realskola och fortsatt folkskola för dem som inte gick till realskolan. Likafullt fanns länge inslag av differentierad undervisning även inom grundskolan. Det var först 1995 – med introduktionen av en ny läroplan (Lpo94) – som en sådan uppdelning formellt försvann. Även om Lpo94 eliminerade det sista formella spåret av differentiering på grundskolenivå har det fria skolvalet (som introducerades 1992) sannolikt bidragit till att skapa en differentiering av ett annat slag inom grundskolan.

1.1.2 Förskolan

Numera har de flesta barn som börjar i första klass tillbringat en längre eller kortare tid på förskola. Förskolan expanderade mycket snabbt från mitten av 1960-talet fram till cirka 1990. Antalet platser ökade under denna period från några tusen till cirka 350 000. Expansionen är självfallet nära förknippad med den ökade förvärvsfrekvensen bland kvinnor med småbarn. Denna utveckling har ökat efterfrågan på barnomsorg, samtidigt som utbyggnaden av förskoleplatser i sin tur stimulerat den kvinnliga förvärvsfrekvensen.

Kopplingen till kvinnors förvärsarbete gör att det är naturligt att tala om barnomsorg. Likafullt används begreppet förskola i politiska sammanhang, något som understryker de pedagogiska ambitionerna. En lag om allmän förskola stadgar också att det finns en rätt till plats i förskolan, även om deltagandet varit och är frivilligt. År 1998 infördes en läroplan för förskolan med tydliga pedagogiska ambitioner. Då infördes också den nya skolformen förskoleklass för 6-åringar, som har en mer skolmässig karaktär än tidigare verksamhet för 6-åringar. Även om förskoleklassen också är frivillig deltar numera de flesta barn i denna verksamhet: läsåret 2008/09 deltog 95 % av alla 6-åringar (se SCB 2009).

Föräldrars efterfrågan på förskoleplatser har säkert också påverkats av de avgifter som utgått. Dessa har förändrats mycket över tiden och varierat en del mellan kommuner, men genomgående har avgifterna täckt klart mindre än hälften av kostnaderna för verksamheten. En omfattande förändring av avgifterna för både förskole- och fritidsverksamheten genomfördes i början av 2000-talet. Denna förändring går under beteckningen maxtaxereformen då den innebär ett statligt direktiv om ett tak för avgifterna. För förskolebarn sattes taket till 3 % av hushållets inkomst för det första barnet, 2 % för det andra, 1 % för det tredje och ingen avgift för det fjärde. Till detta kom ett tak på det absoluta

beloppet för avgiften. Syftet med denna reform var att eliminera de marginal-effekter som skapades av den avgiftsstruktur som kommunerna tidigare tillämpade.

Utbyggnaden av förskolan framstår som en dramatisk förändring i det svenska samhället under en relativt kort tid. Vissa motsvarigheter till dagens förskolor fanns dock även tidigare, om än i mindre skala. Stimulerad av makarna Myrdals inlägg i befolkningsfrågan, ökade andelen barn som deltog i förskoleliknande verksamheter kraftigt fram till 1960-talet då 12,5 % av alla barn i åldrarna 4 till 6 år deltog (Stukát 1966).

1.1.3 Gymnasieskolan

Innan grundskolans införande gick vägen från folkskolan till högre utbildning via urval i två steg. Först från folkskola till realskola, där övergångarna gjordes från årsklasser och åldrar som varierade både över tid och regioner. Därefter skedde en övergång från realskola till gymnasium. Läroverket hade 1905 delats upp i en nedre latinfri realskola och en övre del, gymnasiet, som skulle förbereda för högre studier. Efter folkskolan fanns också ett tämligen heterogent utbud av yrkesinriktade kurser inom yrkesskolan.

En omfattande reform av gymnasieskolan kom 1971 då gymnasium, fackskolan (som bara hade ett par år på nacken), och yrkesskolans heltidskurser sammanfördes till den nya gymnasieskolan.⁵ Nästa stora reform kom i början av 1990-talet och var fullt genomförd läsåret 1995/96. Då ersattes ett stort antal studievägar med 16 nationella program och ett individuellt program, vilka alla blev treåriga. Dessutom övergavs – både på gymnasiet och i grundskolan – det tidigare relativa betygssystemet och ersattes med ett s.k. mål- och kunskapsrelaterat system med fyra nivåer.

1990-talets gymnasiereform kom i stort sett samtidigt som ett antal andra genomgripande reformer och förändringar för den svenska skolan. Det tidigare tämligen detaljerade och styrande statsbidragssystemet förändrades så att varje kommun i stället fick ett samlat statsbidrag att fördela på olika verksamheter. Skolöverstyrelsen och länskolnämnderna avskaffades och ersattes av ett större inslag av målstyrning. Vidare kom 1992 nya bestämmelser som underlättade för fristående skolor att etablera sig genom att godkända s.k. friskolor skulle få ett bidrag per elev som motsvarade minst 85 % av kostnaderna i de kommunala skolorna. Föräldrar fick också ökade möjligheter att välja skola för sina barn. Dessa förändringar kom samtidigt som den ekonomiska krisen tvingade fram

⁵ Denna reform föregicks av en smärre reform av gymnasieskolan 1964.

betydande besparingar för de svenska skolorna. Sammantaget framstår därför första halvan av 1990-talet som en turbulent period för den svenska skolan.

1.1.4 Högskoleutbildning

Utbildningen vid universitet och högskolor uppvisar också en långsiktig ökning. Historiska landmärken är att Uppsala universitet grundades 1477, Lunds universitet 1668, universiteten (först kallade högskolorna) i Stockholm och Göteborg 1878 respektive 1891, samt Umeå universitet 1963. Man kan dock urskilja två perioder då expansionen av antalet studerande var särskilt kraftig. Den första var på 1960-talet: Antalet studerande mer än tredubblades, från 36 000 till 127 000 mellan 1960 och 1970. Den stora expansionen skedde inom samhällsvetenskaplig utbildning (då en del av filosofisk fakultet) där s.k. fri intagning tillämpades. Den andra var 1990-talet: Antalet registrerade i grundutbildning vid universitet och högskolor ökade från 186 000 läsåret 1987/88 till 351 000 läsåret 2001/02, en utveckling som Richardson (2004) betecknar som en ”enorm expansion av högskoleutbildningen”. Båda perioderna innebar således ett kraftigt ökat utbud av högutbildade på den svenska arbetsmarknaden.

När man bedömer siffrorna för de båda perioderna bör man hålla i minnet att 1977 års högskolereform påtagligt breddade innebörden av begreppet högskoleutbildning. Reformen innebar bland annat att flera tidigare icke-akademiska utbildningar, såsom sjuksköterske- och förskolläraryrket, fördes över till högskolesektorn. Enbart denna administrativa förändring innebar en ökning av antalet studerande från cirka 100 000 till 150 000.

Den kraftiga expansionen av antalet studerande sedan 1960 har i stor utsträckning ägt rum vid nya högskole- och universitetsorter. På 1960-talet kom s.k. universitetsfilialer i Karlstad, Linköping, Växjö och Örebro, vilka alla idag har status som självständiga universitet. Dessutom har ett stort antal nya högskoleorter tillkommit, något som torde ha varit nödvändigt för att ta emot den ökande tillströmningen av studenter. Men regionalpolitiska hänsyn, och tanken att kortare geografiskt avstånd gör det lättare att rekrytera studenter från familjer utan tidigare utbildningstradition, har också varit viktiga argument. Den rika floran av universitetsutbildningar och de många utbildningsorterna är självfallet frågor för den utvärderande forskning som vi återkommer till.

Utbyggnaden av svensk universitetsutbildning återspeglas också i befolkningens utbildningsnivå. År 2008 hade 32 % av alla 30-åringar minst 3 års eftergymnasial utbildning, mot 18 % av alla 50-åringar och 12 % av alla

70-åringar (SCB 2009). Bland 70-åringar var andelen lika för män och kvinnor. Men bland 30-åringar hade en påtagligt större andel kvinnor en 3-årig eftergymnasial utbildning: 38 % för kvinnor mot 27 % för män.

Av stor vikt för den högre utbildningens omfattning, liksom för sammansättningen av de studerande, är behörighets- och antagningsregler samt avgifter och studiefinansieringssystem. Det kan här vara värt att nämna att avgifter fanns för praktiskt taget all frivillig utbildning så sent som i slutet av 1950-talet, även om de på slutet troligen hade ringa betydelse. Sedan dess har högskolan i Sverige varit avgiftsfri. Helt nyligen har dock beslut fattats om avgifter för utländska medborgare som studerar vid svenska universitet och högskolor. Det nuvarande svenska systemet för studielån, som har en tämligen generell karaktär, dateras till 1965 då de generella delarna med en låne- och en bidragsdel förstärktes påtagligt. Troligen bidrog denna reform till expansionen av den högre utbildningen under de påföljande åren.

1.1.5 Forskarutbildningen

Forskarutbildningen har också kommit att få en betydande omfattning. Antalet doktorandnybörjare var 2003 uppe i nära 4 000 per år, men sjönk sedan pga. striktare krav på doktorandfinansiering. De allra sista åren har dock tillströmningen stabiliserats kring 3 000 per år, och även antalet doktorsexamina har legat nära 3 000 de sista åren. År 2008 hade 1,6 % av svenska män och 1,1 % av kvinnorna i åldern 35–44 forskarutbildning. Numera har de allra flesta doktorander antingen en doktorandtjänst, stipendium eller en anställning vid en institution med forskningsanknytning.

1.2 Frågor för denna rapport

Denna historiska genomgång reser ett antal frågor som vi helt eller delvis behandlar i denna rapport. I detta avsnitt redogör vi mer precist för de frågor som vi avhandlar i respektive kapitel.

Kapitlen 2 till 5 tjänar som utgångspunkter för resten av rapporten. Kapitel 2 avhandlar svenska elevers kunskaper och hur kunskapsutvecklingen sett ut över tiden. Dessa frågor är högaktuella i skrivande stund och varje publikation av internationella jämförelser röner stor uppmärksamhet i alla deltagande länder. Därför finns ett behov att försöka bringa reda i debatten. Hur förhåller sig svenska elevers kunskaper till elever i andra länder? Och hur har svenska elevers position förändrats över tid?

Kapitel 3 presenterar en modell för hur kunskaper inhämtas. Denna modell, som har hämtat inspiration från olika uppsatser som James Heckman medförfattat (se t.ex. Cunha m.fl. 2006), är naturligtvis inte den enda möjliga modellen, men den förefaller vara en rimlig beskrivning av hur inlärningsprocessen går till. Den grundläggande idén är att kunskaper som inhämtats i ett tidigare skede påverkar effektiviteten på inläringen i ett senare skede. Denna modell har vissa likheter med den s.k. *investment teori* som används i den pedagogiska forskningen. En implikation av modellen är att insatser vid tidig ålder kan vara väl så viktiga för den framtida arbetsmarknadsframgången som vuxenutbildning. Avsnittet behandlar också den grundläggande frågan om varför utbildning subventioneras.

Kapitel 4 är också av grundläggande karaktär. Där avhandlas avkastningsbegreppet, som är det grundläggande verktyg som utbildningsekonomer använder för att värdera olika utbildningsinvesteringar. Vi presenterar även evidens på hur avkastningen på utbildning sett ut i Sverige i jämförelse med andra länder och hur den har förändrats över tid.

Kapitel 5 handlar om sambandet mellan utbildningsinvesteringar, kunskaper, och inkomster. I vilken utsträckning påverkas kunskaper och hur mycket av kunskapsökningen överförs i inkomster? Kapitlet innehåller även en grundläggande diskussion om olika empiriska strategier för att göra effektutvärderingar. Effektutvärderingar syftar till att skatta ett orsakssamband. Vi vill veta hur en enskild åtgärd påverkar studieresultat och framtida inkomster. I många situationer är det inte alldeles lätt att svara på denna fråga, dels för att vi inte vet varför en individ fått del av en åtgärd, och dels för att det inte finns en trovärdig jämförelse som vi kan göra för att skatta effekten av åtgärden.

Efterföljande kapitel – kapitel 6 till 10 – innehåller en genomgång av forskningen om åtgärder inom det reguljära utbildningsväsendet. I grunden har vi utgått från en stadiindelning. Vi börjar med förskolan (kapitel 6) för att avsluta med grundläggande högskoleutbildning (kapitel 10). Med utgångspunkt i modellen som presenteras i kapitel 3 är det naturligt att välja en stadiindelning – modellen bygger ju på att inläring är en kumulativ process. Likväl är många intressanta frågor stadiöverskridande, såsom frågor om skolresursers, lärar-kompetensens och skolorganisationens betydelse. Dessa stadiöverskridande frågor återfinns i kapitel 7 och 8.

I kapitel 6 som handlar om förskolan ställs frågor som: Hur påverkar förskolegång framtida studieutfall? Varierar denna effekt med barnens familjebakgrund?

Kapitel 7 behandlar lärartäthet och lärarkompetens. Några exempel på de frågor som vi tar upp i detta kapitel är: Hur påverkas elevers studieutfall av ett tillskott av resurser? Vilken betydelse har lärare?

Kapitel 8 behandlar olika typer av systemfrågor. En grundläggande fråga är när differentieringen av elever ska ske: Vilken är effekten av att nivågruppera elever vid olika åldrar? Andra frågor rör organisationen av grundskole- och gymnasieutbildning. Flera av 1990-talets reformer i Sverige var systemförändringar och kapitel 8 behandlar därför effekter av skolval och konkurrens. Vi diskuterar även effekterna av decentralisering och de medel – nationella prov och inspektioner – som finns för att säkerställa de nationella målen i ett decentraliserat system.

I kapitel 9 går vi över till gymnasieskolan och den kommunala vuxenutbildningen. De frågor som är unika för gymnasieskolan handlar i mångt och mycket om yrkesutbildningen. En fråga som analyseras är effekterna av att öka det teoretiska inslaget i yrkesutbildningarna. Vuxenutbildningen i Sverige framstår som omfattande i ett internationellt perspektiv. Den grundläggande frågan är hur deltagarnas arbetsmarknadsutsikter påverkas av att delta i den kommunala vuxenutbildningen.

I kapitel 10 kommer vi in på frågor som rör högskolans dimensionering och regionalisering. Vi behandlar också frågan om studiefinansieringens betydelse för rekryteringen till högskolan. Gör existensen av ett studiemedelssystem att individer från en låg socioekonomisk bakgrund i ökad utsträckning väljer högre studier? Andra frågor som berörs är: Finns kvalitetsskillnader mellan högskolor? Påverkar urvalskriterierna studieutfall?

Rapporten avslutas, i kapitel 11, med en sammanfattning av vår syn på forskningsläget: Vad vet vi? Vilka kunskapsluckor finns? Kunskapsluckorna definierar ett tydligt behov av utvärdering och forskning. Det är därför naturligt att ställa frågan: Hur kan man skapa bättre förutsättningar för utvärdering och forskning?

Utifrån sammanfattningen diskuterar vi också fyra större policyfrågor: Är avkastningen på utbildning för låg i Sverige? Varför har svenska grundskoleelevers resultat försämrats sedan mitten av 1990-talet? Vilka kriterier bör ett välfungerande betygssystem tillfredsställa? Är skolan en kommunal eller central angelägenhet?

Referenser

- Becker, G.S. (1964), *Human Capital*. New York: National Bureau of Economic Research,
- Björklund, A., P-A. Edin, P. Fredriksson, och A. Krueger, *Den svenska skolan – effektiv och jämlik?*, Valfärdspolitiska rådets rapport 2003. Stockholm: SNS Förlag.
- Calmfors, L., A. Forslund, och M. Hemström (2002), Vad vet vi om den svenska arbetsmarknadspolitikens sysselsättningseffekter? IFAU rapport 2002:8.
- Cuhna, F., J. Heckman, L. Lochner och D. Materov (2006), Interpreting the evidence on life cycle skill formation, i E. Hanushek och F. Welch (red.), *Handbook of the Economics of Education*. Amsterdam: North-Holland.
- Ericson, T. (2004), Personalutbildning: en teoretisk och empirisk översikt, IFAU rapport 2004:8.
- Hallgren, S. (1939), *Intelligens och miljö. En experimentell undersökning av barn i tredje skolåret vid Malmö folkskolor och privata skolor*, licentiatavhandling, Lunds universitet.
- Husén, T. (1950), *Testresultatens prognosvärde. En undersökning av den teoretiska skolningens inverkan på testresultaten, intelligenstestens prognosvärde och de sociala faktorernas inverkan på urvalet till högre läroanstalter*. Stockholm: Gebers.
- Härnquist, K. (1959), Intelligensutveckling och skolresultat, *Nordisk pedagogisk forskning*, 4, 57–69.
- OECD (2009), *Education at a Glance*. Paris: Organization for Economic Development and Cooperation.
- Pearson, K. och G. Jäderholm (1914), *On the Continuity of Mental Defect*. London: Dulav & Co.
- Richardson, G (2004), *Svensk utbildningshistoria: skola och samhälle förr och nu*. Stockholm: Studentlitteratur.
- SCB (2009), *Utbildningsstatistisk årsbok 2010. Tabeller*, Statistiska Centralbyrån.

Schultz, T. (1960), Capital Formation by Education, *Journal of Political Economy*, 68, 571–583.

Stukát, K-G. (1966), *Lekskolans inverkan på barns utveckling*. Uppsala: Almqvist & Wicksell.

2 Svenska elevers kunskapsutveckling

Syftet med detta kapitel är att ge en beskrivning av huvuddragen i resultatutvecklingen i den svenska skolan under de senaste decennierna. Det empiriska underlaget är dock begränsat, och avser huvudsakligen grundskolan.

Information om elevers kunskaper och färdigheter kan huvudsakligen hämtas ur fyra källor: (1) betyg; (2) nationella utvärderingar; (3) internationella jämförande studier; och (4) prov genomförda i forskningssyfte. De olika informationskällorna kring elevers kunskaper och färdigheter har olika egenskaper, och det är svårt att hävda att någon viss informationskälla skulle ge sannare information än någon annan källa. För att kunna tolka resultaten från olika undersökningar är det dock angeläget att vara medveten om på vilka sätt betyg och prov av olika slag skiljer sig från varandra. Vi inleder därför kapitlet med en kort genomgång av vilka egenskaper som särskiljer de olika informationskällorna.

Betyg används i utbildningssammanhang som en sammanfattande bedömning av en elevs kunskaper och färdigheter. Detta görs med olika syften, och betyg har flera olika funktioner, som att ge grund för urval till nästa nivå i utbildningssystemet, ge information till hemmet om elevens studieresultat, och ge information om utbildningens resultat på olika nivåer inom utbildningssystemet. Den senare funktionen bygger på att betygen återspeglar prestationer som är tolkbara, och kräver att betygen är jämförbara över skolor och över tid. Det är denna utvärderingsfunktion som i detta sammanhang är av störst intresse.

För att det ska vara möjligt att använda betyg i utvärderande syfte måste dessa kunna ge uttryck för uppnådda nivåer av kunskaper och färdigheter. Som närmare beskrivs nedan ersattes vid mitten av 1990-talet det relativa betygssystemet med ett mål- och kunskapsrelaterat betygssystem. En av anledningarna till detta var att det relativa betygssystemet inte kunde användas för att beskriva förändringar i nivåer på kunskaper och färdigheter över tid, eftersom det byggde på att fördelningen av de olika betygsstegen var densamma år från år. I det system som introducerades bygger istället de olika betygsstegen på att eleven skulle ha uppnått de kunskaper och färdigheter som definierats i läro- och kursplaner. Under förutsättning att lärarna kan göra korrekta bedömningar av i vilken utsträckning eleverna uppnått de krav som stipuleras för de olika betygsstegen kan ett sådant system användas som ett utvärderingssystem.

Förtroendet för att lärarna i den svenska skolan skall kunna sätta rättvisa och jämförbara betyg är större än det förtroende lärare i de flesta andra länder

åtnjuter. I många länder sätts slutbetygen i grund- och gymnasieskola på grundval av resultat på olika typer av examensprov, och externa censorer används ofta för bedömningar av elevernas kunskaper och färdigheter. En tänkbar anledning till att lärare i den svenska grund- och gymnasieskolan anförtros mer ansvar att sätta betyg än lärare i andra länder kan vara den långa traditionen av det likvärdiga relativa betygssystemet, med ett strikt regelverk för betygsfördelningen och stöd från standardprov och centrala prov.

Betygssättning bygger på lärarnas mer eller mindre subjektiva bedömningar, och det ofrånkomliga inslaget av subjektivitet är en av betygens svagheter. Betygen har dock fördelar i att betygssättningen bygger på olika typer av observationer gjorda under en längre tidsperiod. Det omfattande och varierade observationsunderlaget ger förutsättningar för att göra bedömningar även avseende sådana kunskaper och färdigheter som är svårfångade med andra metoder, som exempelvis muntlig och skriftlig kommunikationsförmåga, förmåga till samarbete, och engagemang i skolarbetet.

Olika former av prov och test används också som underlag för betygssättning, och som utvärderings- och forskningsinstrument. Prov för att mäta kunskaper och färdigheter kan utformas på en rad olika sätt, men i allmänhet fångar dessa upp ett snävrare spektrum av kognitiva färdigheter än vad betygen kan göra. I vissa prov, som uppsatsprov eller muntliga färdighetsprov, finns omfattande inslag av bedömning, medan andra prov, som exempelvis bygger på uppgifter av flervalstyp, är helt fria från denna typ av subjektiva inslag. En annan viktig skillnad mellan prov är i vilken utsträckning de å ena sidan mäter förmåga att lösa problem med ny information, och å andra sidan mäter förvärvade kunskaper av olika slag, som exempelvis ordförråd och läsförståelse. Den förra typen av prov brukar med en engelsk term betecknas som prov på "fluid abilities" (Gf, formbara förmågor) medan den senare typen brukar betecknas som prov på "crystallized abilities" (Gc, kristalliserade eller formade förmågor). I allmänhet har olika prov varierande omfattning av inslag av dessa två huvudkategorier av förmågor. Sådana prov som traditionellt brukar betecknas som intelligenstag (exempelvis det militära inskrivningsprovet) mäter framförallt Gf medan prov på utvecklade förmågor som används i utbildningssammanhang (t.ex. Höskoleprovet) i störst utsträckning mäter Gc (Carlstedt och Gustafsson 2005).

Det är sålunda möjligt att ordna instrument för att mäta kunskaper och färdigheter längs ett kontinuum där det vid den ena polen finns instrument som mäter förmåga till problemlösning med ny information, och vid den andra

polen finns instrument som mäter grad av uppnående av de kunskaper och färdigheter som anges som mål i läro- och kursplaner. Exempel på den senare typen av instrument är givetvis de nationella proven och betygen. Skillnaden mellan betyg och nationella prov är att de förra i större utsträckning än de senare återspeglar vad som brukar kallas 'Factor X', som är ett komplex av personlighetsegenskaper, intressen, och hembakgrund som förklarar variation mellan elever i betygsresultat (t.ex. motivation, skolanpassning, ansträngning, social förmåga, och muntlig och skriftlig kommunikationsförmåga). 'Factor X' sammanfaller i stort sett med vad som ibland kallas icke-kognitiva förmågor.

Nedan redovisas resultat från studier av resultatutvecklingen i grundskolan baserade på betyg, nationella utvärderingar, internationella jämförande studier, och prov genomförda i forskningssyfte.

2.1 Betyg

Fram till mitten av 1990-talet användes det s.k. relativa betygssystemet i den svenska skolan. Detta normrelaterade system byggde på jämförelse mellan elever, och stipulerade hur stor andel av det totala antalet elever som skulle erhålla vart och ett av de olika betygsstegen. Fördelningen avsåg samtliga elever i landet som läste det aktuella ämnet, medan betygsfördelningen inom den enskilda klassen kunde förväntas vara annorlunda. Som hjälpmedel för att fastställa klassens nivå i relation till den nationella nivån erbjöds lärarna standardprov i vissa ämnen.

Det relativa betygssystemet var dock utsatt för kritik som avsåg flera olika svagheter (t.ex. Skolöverstyrelsen 1970): (1) lärarna tog i alltför hög grad fasta på den stipulerade nationella procentfördelningen av betyg vid den egna betygsättningen; (2) de relativa betygen säger inte något om den faktiska nivån på elevernas prestationer, utan endast hur en elev presterar i förhållande till andra elever, varför förbättringar i kunskaper och färdigheter från ett år till ett annat i landet som helhet inte kan avläsas i form av högre betyg; (3) det relativa betygssystemet inbjuder genom sin konstruktion till konkurrens snarare än till samarbete; (4) det förutsätter i princip att alla elever i landet läser ämnet, eftersom det endast är då som det kan finnas någon grund för det normalfördelningsantagande som den nationella betygsfördelningen bygger på; och (5) det faktum att standardprov endast förekom i tre ämnen (svenska, matematik och engelska), gjorde det svårt att uttala sig om klassens nivå i förhållande till andra klasser för övriga ämnen. Försök gjordes under 1970-talet att inom ramen för

olika projekt och utredningar utveckla ett alternativt betygssystem, utan att några egentliga framsteg gjordes.

År 1994 fattade riksdagen beslut om att införa målrelaterade betyg, som är baserade på bedömning av elevers prestation i förhållande till förväg preciserade mål, i form av vad som betecknades som det mål- och kunskapsrelaterade betygssystemet. För grundskolans del fastställdes en tregradig betygsskala för ämnesbetyg med betygsstegen Godkänd, Väl godkänd, och Mycket väl godkänd. Betyg ges första gången på höstterminen i årskurs 8, varefter terminsbetyg ges varje termin. För slumbetyget från vårterminen i årskurs 9 finns centralt fastställda mål och kriterier, och betygsättningen i årskurs 9 stöds av nationella ämnesprov i svenska, engelska och matematik. Det innebär att betygen kan ge information om prestationer i slutet av grundskolan.

Begreppen *mål* och *kriterier* är centrala i det mål- och kunskapsrelaterade betygssystemet. Målen, om vilka regeringen fattar beslut i form av läroplaner, anger vad eleven skall kunna, och betygskriterierna, vilka fastställs av Skolverket, beskriver de kvaliteter i målluppfyllelsen som svarar mot olika betyg. Kriterierna utgör grunden för en rättvis och likvärdig betygsättning, men dessa kan inte vara så precisa att likvärdighet i betygsättningen kan uppnås endast genom dessa. Istället har det förutsatts att likvärdigheten i tolkningarna av kriterier skall uppnås genom kommunikation i form av professionella samtal mellan dem som planerar och genomför undervisningen, och som bedömer elevens kunskapsutveckling. En stor del av ansvaret för att uppnå en likvärdig betygsättning vilar därför på lärarna och skolorna. En grundfråga vad gäller användbarheten av de mål- och kunskapsrelaterade betygen för utvärderingsändamål är dock om betygen i detta decentraliserade system uppfyller det krav på likvärdighet som enligt skollagen skall gälla all utbildning.

Studier av det målrelaterade betygssystemet visar att det finns brister i betygsättningen med avseende på rättvisa och likvärdighet (t.ex., Klapp Lekholm 2008; Klapp m.fl. 2008, 2009; Riksrevisionen 2004; Skolverket 2007, 2009), liksom att det i synnerhet på gymnasienivå under vissa perioder förelegat betygsinflation (Cliffordson 2004). Resultaten pekar på att bristerna kan ha sin grund både i det sätt på vilket det nya betygssystemet infördes (Tholin 2003), och i systemets konstruktion (SOU 2007:28). En mer utförlig diskussion kring studier av det målrelaterade betygssystemets egenskaper presenteras av Gustafsson och Yang Hansen (2009).

2.1.1 Betygsutvecklingen sedan 1989

De normrelaterade betygen är av intet värde när det gäller att beskriva utvecklingen av kunskaper och färdigheter över tid, eftersom dessa definitionsmässigt hade samma medelvärde för landet som helhet vid varje tidpunkt. Dock är de normrelaterade betygen av stort intresse när de gäller att beskriva skol- och kommunskillnader i resultat liksom skillnader i resultat för olika grupper av elever (pojkar/flickor, högutbildade/lågutbildade föräldrar, svensk/utländsk bakgrund).

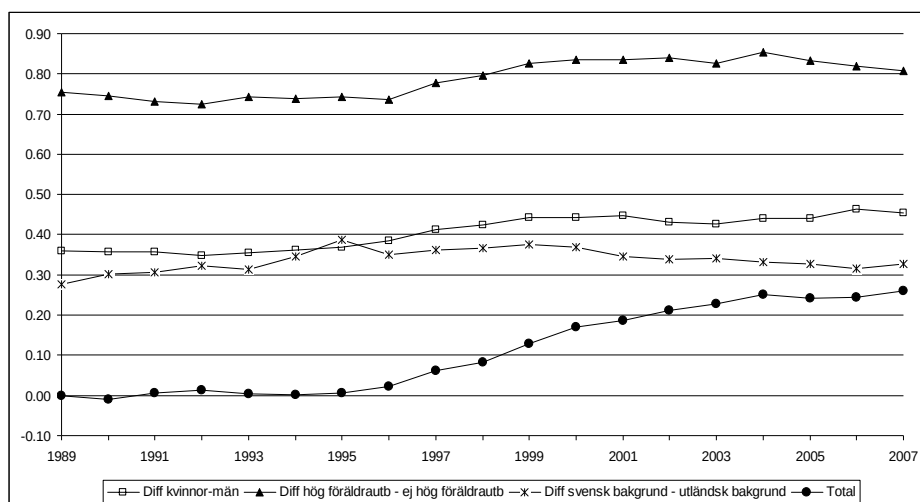
De nedan presenterade analyserna har genomförts av Gustafsson och Yang Hansen (2009) och för en mer utförlig redovisning hänvisas till denna källa.

Vi använder oss av ett enhetligt sätt att beskriva utfall baserat på standardiserade medelvärdesdifferenser, där standardiseringen görs med medeltalet av standardavvikelseerna inom grupperna. Detta standardavvikelseenhetsmått, som börjar bli alltmer allmänt accepterat (se t.ex. Hattie 2009), betecknas ofta Cohens d eller bara d . En stor fördel med d -måttet är att vi kan göra jämförelser mellan variabler som vanligtvis inte direkt låter sig jämföras. Det är denna egenskap hos d som vi här framförallt kan dra nytta av. En annan fördel med d -måttet är att det finns relativt allmänt omfattade tolkningsregler (t.ex. att värden under 0,20 räknas som små, att värden runt 0,50 betraktas som medelstora, och att d -värden över 0,80 är uttryck för stora effekter), liksom att det flesta statistiska mått kan transformeras till d (exempelvis motsvaras $d = 1$ av ungefär 28 steg på percentilskalen). I detta kapitel kommer vi därför genomgående använda oss av d -måttet och i allmänhet kommer vi enbart att presentera resultaten i grafisk form.

I Figur 2.1 presenteras betygsutvecklingen under tidsperioden 1989 till 2007 med avseende på ett sammanfattande betygsmått, som betecknas Total i figuren. Utgångspunkt var medelvärdet över de 16 betyg som alla elever skall erhålla, varvid kodvärdena IG = 0, G = 1, VG = 2, och MVG = 3 användes. Endast elever med 8 eller fler godkända betyg inkluderades i beräkningarna. För det första året (dvs. 1989) har d satts till 0, och för efterföljande år har d -måttet för medelvärdesdifferensen mellan det aktuella året och år 1989 beräknats. Vid bytet av betygssystem 1998 skulle en lösning kunnat vara att sätta värdet för år 1998 till $d = 0$, vilket dock skulle ge det falska intrycket att det detta år inträffade en dramatisk förändring. Istället har utgångspunkten i det nya betygssystemet skattats till att vara värdet för det sista året i det gamla betygssystemet (dvs. år 1997) plus medelvärdet av förändringen mellan de två

åren före (dvs. mellan 1996 och 1997) respektive efter (dvs. mellan 1998 och 1999) införandet av det nya betygssystemet.

Figur 2.1. Betygsutvecklingen 1989–2007



Not. Gruppkillnader avseende kön, utbildningsbakgrund och migrationsbakgrund visas med $d = 0$ år 1989 för män, ej någon högskoleutbildad förälder, respektive utländsk bakgrund. Källa: Gustafsson och Yang Hansen (2009).

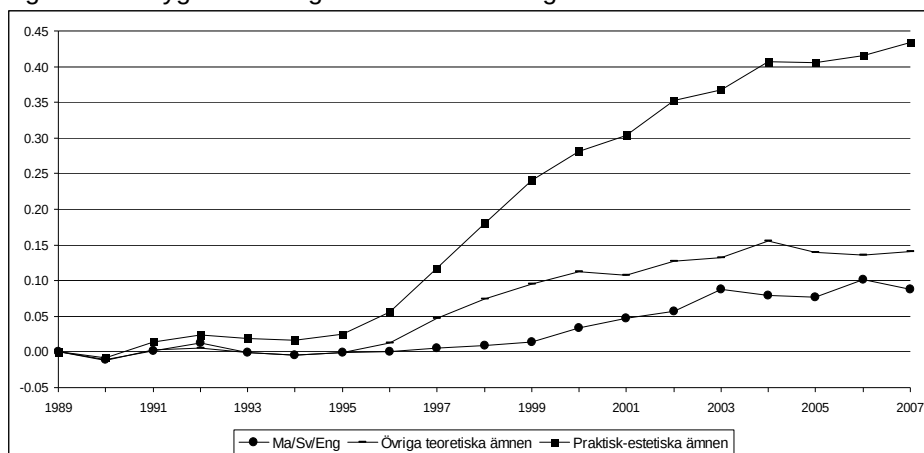
Som framgår av Figur 2.1 ligger Totalkurvan, som förväntat, mycket nära 0 under den tidsperiod då det relativa betygssystemet var i bruk. Från 1998, då det kriterierelaterade betygssystemet infördes, ökar d -värdet successivt över åren. Uttryckt på d -skalan är ökningen 0,26 enheter mellan 1998 och 2007, vilket innebär att meritvärdet under denna tidsperiod ökat med 0,26 standardavvikelser. Om meritvärdet faktiskt återspeglar nivån på elevernas kunskaper och färdigheter innebär detta en betydande resultatförbättring under denna tioårsperiod.

I Figur 2.1 visas också omfattningen av gruppkillnader med avseende på tre aspekter: kön, föräldrarnas utbildningsnivå och migrationsbakgrund. Den översta kurvan anger skillnaden i resultat mellan de elever som har någon förälder som har minst tvåårig postgymnasial utbildning å ena sidan, och övriga elever å den andra. År 1989 var denna differens $d = 0,75$ och var relativt konstant fram till år 1996. Därefter ökade prestationsskillnaden mellan elever med högutbildade föräldrar och övriga elever fram till år 2004 och uppgick då

till $d = 0,85$. Särskilt tydlig var ökningen under åren 1996 till 2000. Efter 2004 har skillnaden minskat något och var år 2007 $d = 0,81$. Även könsskillnaden till förmån för flickor har ökat från mitten av 1990-talet från ca $d = 0,35$ till $d = 0,45$. Skillnaden i resultat för elever med svensk bakgrund och utländsk bakgrund visar även den en trend till ökande skillnader. Det finns dock betydande fluktuationer i storleken på skillnaden över åren, vilket förmodligen delvis förklaras av att olika grupper av invandrare kommit till Sverige vid olika tidpunkter.

Det sammanfattande betygsmåttet är ett mycket heterogent mått som uttrycker kunskaper och färdigheter inom mycket olika fält varför det finns anledning att göra en mer detaljerad beskrivning. Betygen har därför delats upp i tre olika kategorier. Den första kategorin omfattar betygen i svenska, matematik och engelska; dessa tre ämnen har grupperats samman därför att betygsättningen i dessa stöds av nationella prov eller standardprov. Den andra kategorin omfattar betygen i övriga teoretiska ämnen, vilket innebär de natur- och samhällsorienterande ämnena. Den tredje kategorin omfattar de praktisk-estetiska ämnena (bild, musik, idrott och hälsa, hemkunskap och slöjd). För var och en av dessa kategorier har separata d -värden bestämts för åren 1989–2007, på samma sätt som för Totalmåttet i Figur 2.1. Resultaten redovisas i Figur 2.2.

Figur 2.2. Betygsutvecklingen i tre ämneskategorier 1989–2007



Not. Betygsmåttan avser ämneskategorierna matematik/svenska/engelska, övriga teoretiska ämnen och praktisk-estetiska ämnen. Källa: Gustafsson och Yang Hansen (2009).

Utvecklingen inom de tre kategorierna av ämnen är mycket olika. För kategorin Ma/Sv/Eng är skillnaderna mellan de olika åren små, och var i det närmaste obefintliga mellan 1989 och 2001. Därefter har medeltalet ökat något, men var endast $d = 0,09$ år 2007. För kategorin övriga teoretiska ämnen kan vi observera en successiv ökning av medeltalet från 1996, och 2007 uppgick medeltalet till $d = 0,14$. För de praktisk-estetiska ämnena föreligger en kraftig, och i det närmaste linjär, utveckling av d -värdena efter år 1996. År 2007 uppgick d -värdet för denna kategori av ämnen till 0,43.

Den dramatiska skillnaden i utveckling över tid för de tre kategorierna av ämnen pekar på att betygen i det mål- och kunskapsrelaterade systemet inte är okomplicerade och direkta mått på nivån av elevernas kunskaper och färdigheter. Det framstår som orimligt att de redovisade resultaten skulle återspegla en faktisk förändring i elevernas kunskaper och färdigheter. För de tre ämnen där det finns nationella prov är förändringarna över tid minst. Detta kan tolkas så att proven åtminstone delvis kunnat stävja de tendenser till betygsinflation som alltid finns. Även för kategorin övriga teoretiska ämnen kan man tänka sig att de nationella proven haft en viss normerande funktion i betygsättningen. För de praktisk-estetiska ämnena verkar inte de nationella proven haft någon styrande inverkan på betygsättningen, varför den enda vägledningen varit lärarnas och skolornas tolkning av mål och kriterier. De observerade förändringarna över tid är därför förmodligen uttryck för varierande grader av betygsinflation i olika ämnen, möjligen också i kombination med vissa faktiska förändringar av kunskaper och färdigheter, vilka kan vara såväl positiva som negativa.

2.1.2 Förändringar i betygsskillnader mellan skolor och kommuner

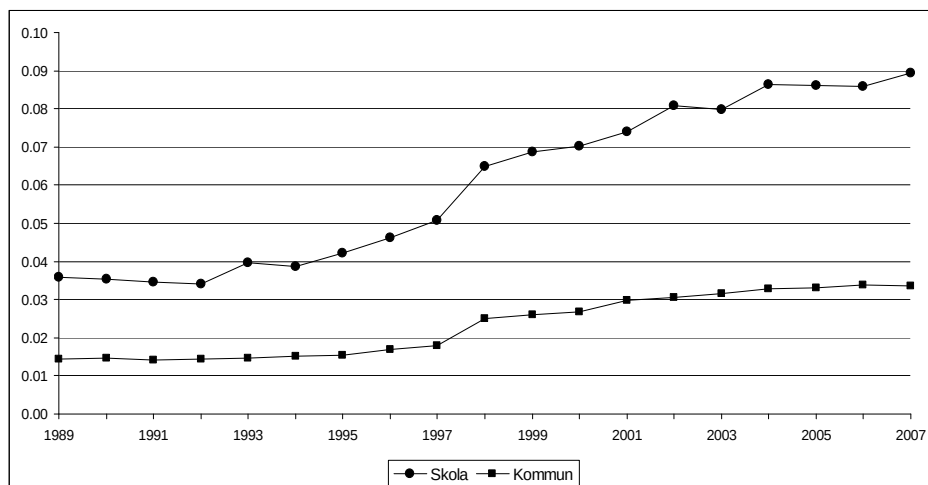
Det är även intressant att göra en beskrivning av utvecklingen av betygsvariationen mellan elever, skolor och kommuner.

Skolverket (2006a) har redovisat en serie analyser av grundskolebetygen i åk 9 med fokus på förändring av resultatskillnader över tid. Meritvärdet uppvisade under perioden 1998–2004 en ökad standardavvikelse, och i synnerhet gällde detta för de fyra första åren. Skolverket visade också att meritvärdet för de 10 % högst presterande eleverna utvecklades positivt under hela den studerade perioden. För de 10 % lägst presterande eleverna hade däremot det genomsnittliga meritvärdet försämrats, och i synnerhet gällde detta under de första fyra åren. I rapporten konstaterar dock också Skolverket att utvecklingen över tid kan tolkas på så sätt att resultatutvecklingen för de högst

presterande eleverna sammanfaller med betygsinflationen, medan de lägst presterande eleverna har halkat efter.

I Figur 2.3 redovisas andelen av den totala betygsvariationen som kan hänföras dels till skolor, dels till kommuner.

Figur 2.3. Variation i betyg mellan skolor och kommuner, 1989–2007



Not. Intraklasskorrelationer, som uttrycker andelen variation av den totala variationen som kan hänföras till skola respektive kommun. Källa: Gustafsson och Yang Hansen (2009).

Kommunvariationen är betydligt mindre än skolvariationen, men mönstret liknar det för skolvariationen. Som framgår av figuren var variationen mellan skolor i huvudsak konstant under perioden 1988 till 1992. År 1993 skedde en viss ökning, som kan vara en effekt av kommunaliseringsreformen, som skedde några år tidigare, och av det riktade statsbidraget till skolan bakades samman med andra statsbidrag till kommunerna detta år. Under perioden 1994 till 1997 ökade betygsvariationen mellan både skolor och kommuner år från år, vilket kan vara en kombinerad effekt av att kommunerna under denna tid fick ökad beslutsrätt avseende bland annat resurstilldelning till skolan, och att tidigare fattade beslut att införa skolpeng successivt implementerades med ökad valfrihet för eleverna mellan både kommunala och fristående skolor som följd.

År 1998 ägde en betydande ökning rum av betygsvariationen mellan skolor. Detta var det första året då eleverna i grundskolan fick betyg enligt det kriterie-relaterade betygssystemet. Som konstaterats av Björklund m.fl. (2003a) är det inte rimligt att anta att den plötsliga ökningen av mellanskolvariationen kan

förklaras av en ökad variation i resurser eller andra faktorer som påverkar resultaten. De menar istället att det nya betygssystemet har ökat betydelsen av den tolkning de olika skolorna gör av vad de olika betygssorgen faktiskt innebär. Det nya betygssystemet återspeglar därför inte lika väl den faktiska variationen mellan skolor i elevernas kunskaper och färdigheter som det relativa betygssystemet.

Mellan 1998 och 2004 förelåg en fortsatt årlig ökning av betygsvariationen mellan skolor, men därefter har ökningen upphört. Det finns flera olika tänkbara förklaringar till den systematiska ökningen av skolvariationen mellan 1998 och 2004.

Skolverket (2006a), liksom Gustafsson (2006), konstaterade att skolsegregationen med avseende på socioekonomisk bakgrund ökat under tidsperioden. Effekten av elevernas socioekonomiska bakgrund på meritvärdet var enligt Skolverkets analyser i huvudsak konstant över tiden när denna analyserades på individnivå (se också Björklund m.fl. 2003b), men den ökade över tid på skolnivå. Den ökade skillnaden i meritvärde mellan elever med hög- och lågutbildade föräldrar under perioden 1998–2004 (se Figur 2.1) skulle enligt Skolverkets resultat ha sin grund i fenomenet på skolnivå, snarare än på individnivå. Detta kan vara en effekt av den ökade skolsegregationen, eller av att social bakgrund på skolnivå har fått en ökad betydelse för betyg utfallet. En förklaring till detta kan vara kontextuella effekter i form av kamrateffekter.

Ett problem när det gäller att finna förklaringar till den ökande mellan-skolvariationen i betygsnivåer under denna period är dock att det inte är klart om variationen är uttryck för skillnader mellan skolor i faktiskt uppnådda kunskaper och färdigheter, eller om den är uttryck för skillnader mellan skolor i bedömningen av elevernas kunskaper och färdigheter. Det är givetvis fullt möjligt att det ökade utrymme för skillnader mellan skolor i grad av stränghet i betygsättningen som infördes år 1998 har fått effekter på skolskillnaderna efter 1998.

2.1.3 Diskussion och slutsatser

Under perioden då det relativa betygssystemet användes är det uppenbart att några uttalanden om den totala resultatutvecklingen inte går att göra eftersom betygsmedelvärdet för nationen som helhet definitionsmässigt var konstant över tid. Då det målrelaterade betygssystemet började användas kan vi notera en stadigt ökande trend i meritvärdet, men vi kan också konstatera att utvecklingen skiljer sig avsevärt mellan olika grupper av ämnen. Ökningen är drama-

tisk för praktisk-estetiska ämnen, och svagt positiv för natur- och samhälls-orienterande ämnen.

Om man tolkar dessa förändringar som uttryck för förändringar i elevernas kunskaper och färdigheter förbättras dessa i huvudsak över tid. Det finns dock mycket starka skäl att ifrågasätta om betygsförändringarna faktiskt återspeglar kunskapsutvecklingen eftersom dessa framstår som orimligt stora och dessutom oförklarliga. En rimligare förklaring är att förändringarna över tid är uttryck för varierande grader av betygsinflation i olika ämnen.

Vi måste därför dra slutsatsen att vare sig de relativa eller de kriterierelaterade betygen ger någon information om kunskapsutvecklingen över tid. Eftersom ett uttalat huvudsyfte med de mål- och kunskapsrelaterade betygen är att de ska kunna användas för utvärderingssyften på systemnivå framstår det som ett misslyckande att det inte kan användas för detta ändamål.

Informationen i de två betygssystemen kan dock användas för att beskriva skillnader i resultat mellan olika grupper av elever. För såväl kön som utbildningsbakgrund framträder ett tydligt mönster av ökande gruppskillnader mellan 1998 och 2004. Även för migrationsbakgrund kan ökande skillnader observeras även om trenden här är något mer komplex. Skillnaderna i resultat mellan olika skolor ökade kraftigt från mitten av 1990-talet till mitten av 2000-talet.

Tidsmässigt sammanfaller de ökande skolskillnaderna med de förändringar av det svenska skolsystemet som genomfördes under början av 1990-talet, som bland annat innebar decentralisering av beslutsfattande till kommunerna, införande av fritt skolval och fristående skolor, ett nytt betygssystem och en ny läroplan. Samtidigt skedde också förändringar i samhället med bland annat minskade resurser till offentlig sektor, en ökande boendesegregation och en förändrad arbetsmarknad. Det är i detta sammanhang omöjligt att hänföra de ökande skillnaderna till specifika förklaringar, men det framstår som sannolikt att de omfattande förändringarna på skolområdet, inte minst vad gäller betygssystemet, har varit betydelsefulla.

2.2 Nationella utvärderingar

Skolverket genomförde våren 2003 en omfattande nationell utvärdering av grundskolelevers kunskaper och färdigheter (NU03). Tidigare nationella utvärderingar har genomförts år 1989, 1992, 1995, och 1998 och det finns vissa möjligheter att göra jämförelser med dessa tidigare undersökningar, och då framförallt med den nationella utvärdering som genomfördes år 1992 (NU92). Gustafsson och Yang Hansen (2009) har gjort en något mer utförlig samman-

ställning av de huvudsakliga resultaten från jämförelser mellan NU03 och NU92 än vad som kan redovisas här, så för en mer omfattande presentation hänvisas till detta arbete.

NU03 omfattade årskurserna 5 och 9 med tyngdpunkt på den senare, och 16 av grundskolans ämnen ingick. Även i NU92 låg tyngdpunkten på åk 9. I båda undersökningarna deltog ca. tio % av landets elever. I vissa ämnen, som svenska, matematik, engelska, natur- och samhällsorienterande ämnen var kunskapsmätningarna i NU03 upprepningar av de mätningar som genomfördes i NU92. I dessa är det möjligt att uttala sig om resultatutvecklingen under den senaste tioårsperioden.

Inom matematik blev dock underlaget för att bedöma kunskapsutvecklingen begränsat, därför att det nationella provet i matematik som var planerat att utgöra en del av det empiriska underlaget i NU03 spreds via Internet strax innan provtillfället. Resultat från NU92 var dock tillgängligt för ett prov som bjöds även i NU03 och det visade på en försämring av medelvärdet. I synnerhet gällde detta taluppfattning avseende hela tal och rationella tal i bråk- och decimalform, och i viss utsträckning även förmåga att utföra beräkningar. Även på uppgifter av problemlösningsskaraktär hade andelen korrekta svar minskat. Den information som behövs för att beräkna d -värden för förändringen finns dock inte tillgänglig i den publicerade rapporten, så det går inte att bedöma nedgångens omfattning. Det måste dock understrykas att underlaget i form av antalet matematikuppgifter är begränsat. Ett annat problem är att bortfallet av elever var betydligt större i NU03 än i NU92.

Inom naturvetenskap fokuserades särskilt på elevernas förståelse av naturvetenskapliga begrepp, modeller och teorier, och även här användes provuppgifter och enkätfrågor från NU92 och NU95. Inom biologi var resultaten i stort sett oförändrade ($d = -0,06$), även om resultaten var bättre på vissa uppgifter och sämre på andra. Även inom fysikområdet var resultatbilden blandad, med förbättringar på vissa uppgifter och försämringar på andra. Den samlade bilden var dock att det skett en nedgång ($d = -0,15$). Inom kemi var utvecklingen över tid mer entydigt negativ ($d = -0,20$).

Läsförmågan mättes med två läsprov vilka hade använts i NU92. Båda läsproven visade på betydande nedgångar i resultaten. För det ena var försämringen $d = -0,61$ och för det andra $d = -0,54$. För båda proven var spridningen i resultat större i NU03 än i den tidigare undersökningen.

För engelska fanns underlag för jämförelse av elevernas resultat på ett läsförståelseprov och ett hörförståelseprov mellan NU92 och NU03. Medel-

värdena på hörförståelseprovet var praktiskt taget identiska i NU92 och NU03 ($d = -0,03$). Dock var spridningen i resultat större i NU03 än i NU92. Medelvärdet på läsförståelseprovet var lägre i NU03 än i NU92 ($d = -0,21$), och även i detta prov var spridningen större i NU03. Det begränsade antalet uppgifter och den splittrade resultatbilden gör det svårt att dra några bestämda slutsatser om utvecklingen i engelska. Resultatbilden kompliceras ytterligare av de resultat som erhöles i en jämförande studie som genomfördes 2002 (Skolverket, 2004b), där Sverige deltog tillsammans med sju andra europeiska länder. I denna undersökning hade Sverige de bästa resultaten av samtliga länder. Inom ramen för denna studie gjordes också vissa jämförelser med resultaten i en tidigare studie som genomfördes 1996. Huvudslutsatsen var att elevernas resultat förbättrats något, och framförallt då vad gäller receptiva färdigheter.

Den huvudsakliga resultatbilden som framträder från den nationella utvärderingen är att resultaten 2003 var sämre än år 1992 i matematik, naturvetenskap och läsning. Dessa resultat framstår som oförenliga med dem som framträder i betygsutvecklingen, vilket ger ytterligare stöd för slutsatsen att betygen är påverkade av inflation i bedömningarna.

2.3 Internationella studier

De första jämförande studierna av resultat från olika utbildningssystem genomfördes under början 1960-talet. Ett huvudsyfte var att försöka använda "världen som ett pedagogiskt laboratorium" för att undersöka effekten av olika organisatoriska och pedagogiska utformningar av skolan. Under de kommande tre decennierna genomfördes sådana undersökningar inom olika ämnesområden som matematik, naturvetenskap och läsning med ungefär 15 års mellanrum. Från mitten på 1990-talet har intresset för internationella jämförelser av utbildningsresultat ökat kraftigt, och flera undersökningar genomförs nu med intervall om tre till fem år, med ett ökande antal deltagande länder. Medias intresse för undersökningsresultaten har också ökat, och inom många länder har utfallen gett upphov till intensiva utbildningspolitiska diskussioner. De internationella studierna har sålunda fått en vidare funktion än att vara redskap för att utvinna generell kunskap om effekter av olika sätt att utforma utbildningssystem, och har alltmer kommit att bli fokuserade på de resultat som uppnås i olika utbildningssystem.

2.3.1 Utvecklingen av de internationella studierna

Pionjärbetet när det gäller utveckling av de internationella studierna har skett inom IEA ("International Association for the Evaluation of Educational Achievement"), där den svenske pedagogen Torsten Husén varit en av de mest drivande. Den första IEA-undersökningen genomfördes år 1964 och hade fokus på matematik. Under åren 1970–72 genomförde IEA den mycket omfattande "Six Subject Survey", som studerade ämnena läsning, litteraturkunskap, naturvetenskap, samhällskunskap, samt engelska och franska som främmande språk i flera åldersgrupper. Under 1980-talet genomförde IEA bland annat nya studier inom matematik ("Second International Mathematics Study", SIMS 1980) och naturvetenskap ("Second International Science Study", SISS 1984).

Uppläggning och genomförande av de första internationella undersökningarna var ett pionjärbete som krävde hantering av svåra metodproblem avseende definitionen av jämförbara populationer, principer för urval och bortfallshantering, utveckling av instrument för mätning av kunskaper och färdigheter med krav på internationell jämförbarhet, och modeller för analys och rapportering av data (Husén 1979). Även om forskarna kom att bemästra många av dessa problem dröjde det till 1990-talet innan IEA-undersökningarna kunde sägas uppfylla högt ställda kvalitetskrav med avseende på urval och mätning.

En av anledningarna till kvalitetshöjningen var, paradoxalt nog, att intressefokus försköts från forskning till resultatmätning på systemnivå. Detta ökade fokus på utbildningssystemens resultat medförde dels att vissa länder, i synnerhet USA, tillsköt betydande finansiering, dels att antalet deltagande länder ökade. Detta gav förutsättningar för att från mitten av 1990-talet genomföra IEA-undersökningarna med högt ställda kvalitetsambitioner.

År 1995 genomfördes den första undersökningen i en serie studier som nu benämns "Trends in International Mathematics and Science Study" (TIMSS) och som omfattade både naturvetenskap och matematik i årskurserna 4 och 8. TIMSS har därefter upprepats vart fjärde år (dvs. 1999, 2003 och 2007), varvid resultaten från de olika undersökningsomgångarna uttryckts på samma skala. Detta gör det möjligt att mäta förändringar i resultat inom länder, vilket gör att kunskapsutvecklingen inom ett land kan studeras oberoende av vilka övriga länder som deltar i undersökningen.

År 2001 genomfördes den första studien i en serie benämnd "Progress in International Reading Literacy Study" (PIRLS) med fokus på elever i årskurs 4. PIRLS är planerad att genomföras vart femte år, och den andra undersökningsomgången genomfördes 2006. Även i PIRLS är det möjligt att undersöka

förändringar i resultat eftersom de olika studiernas resultat är uttryckta på samma skala.

En annan konsekvens av det ökade intresset för utbildningsresultat och dess ekonomiska betydelse var att även OECD vid slutet av 1990-talet påbörjade ett mycket ambitiöst program för internationella jämförelser av utbildningsresultat ("Programme for International Student Assessment", PISA). År 2000 genomförde OECD första omgången av PISA med deltagande av 31 OECD-länder och 10 andra länder. Denna undersökning bygger på samma grundläggande metodik som används inom IEA-undersökningarna, men det finns också viktiga skillnader. En sådan är att urvalen till PISA-undersökningarna baseras på ålder (15-åringar) medan IEA-undersökningarnas urval från 1995 baseras på årskurstillhörighet. En annan skillnad är att IEA-proven utvecklas från de deltagande ländernas läroplaner, medan PISA-proven baseras på kompetenser som bedöms vara viktiga i yrkes- och samhällsliv. I varje PISA-omgång, vilka genomförs vart tredje år, inkluderas uppgifter från de tre områdena läsning, matematik och naturvetenskap. Vid varje tillfälle är ett av dessa områden huvudområde och representeras av ett stort antal uppgifter, medan de övriga två är biområden, och representeras av ett begränsat antal uppgifter. Vid den första mätningen, 2000, var läsning huvudområde; 2003 var matematik huvudområde; och 2006 var naturvetenskap huvudområde.

Medan Sveriges deltagande i de tidiga IEA-undersökningarna var i det närmaste komplett, har det varit mer sporadiskt i TIMSS, i synnerhet vad gäller yngre grundskoleelever. Sedan 2000 har Sverige deltagit i samtliga tre genomförda PISA-undersökningar, vilka dock endast avser 15-åringar. För de yngre eleverna är det därför omöjligt att beskriva resultattrenden, och fram till 1995 finns endast två mätpunkter vardera för matematik, naturvetenskap och läsning.

Gustafsson (2009) har gjort en kritisk granskning av de studier Sverige har deltagit i och har i flera fall kommit fram till en annan tolkning av utfallet än den som varit den officiellt accepterade. Detta har primärt sin grund i att de åldersbaserade urvalsprinciper som använts i de flesta studier missgynnat de svenska eleverna eftersom de haft färre år i skolan än eleverna i de flesta andra länder. Nedan presenteras en sammanfattning av de Gustafssons slutsatser och de överväganden som dessa baseras på.

2.3.2 Några metodproblem i de internationella studierna

Förutsättningen för att de internationella studierna skall kunna användas för att beskriva resultatutvecklingen i den svenska skolan är givetvis att de håller tillräckligt hög kvalitet.

I de första undersökningarna användes delvis olika urvalsmodeller i olika länder, vilka dessutom var mycket ofullständigt dokumenterade. Vissa länder drabbades också av omfattande bortfall. Urvalens representativitet och jämförbarhet måste därför i många fall ifrågasättas i undersökningarna före 1995. Som redan nämnts innebar dock TIMSS 1995 en avsevärd höjning av den tekniska kvaliteten med avseende på urvalsmetoder, bortfallskontroll, samt mät-, analys- och rapporteringsmodeller. De internationella undersökningar som har genomförts av IEA och OECD från 1995 har alla använt en undersöknings-teknik som uppfyller högt ställda kvalitetskrav. Icke desto mindre kvarstår en del grundläggande problem när man vill studera kunskapsutveckling över tid.

En ofta diskuterad svårighet med att genomföra jämförande studier av elevers uppnådda kunskaper och färdigheter är att de prov som används i varierande utsträckning kan vara anpassade efter de deltagande ländernas läroplaner. Eftersom det också skett förskjutningar över tid i fördelningen av uppgifter över olika delområden, kan sådana förändringar i provens utformning ge upphov till vad som verkar vara prestationsförändringar. I de prov som konstruerades för de tidiga undersökningarna användes också ett begränsat antal uppgifter. Vissa områden riskerar att bli underrepresenterade och överrepresenterade, med följden att proven mer eller mindre väl svarar mot olika länders läroplaner.

De studier som genomförts från 1995 är i allmänhet baserade på ett större antal uppgifter, där varje elev endast besvarat ett mindre antal uppgifter i en så kallad matrissamplingdesign. Urval och konstruktion av uppgifter har också skett med större systematik i de senare studierna, under vägledning av explicit formulerade ramverk. Detta gör det möjligt att både åstadkomma en bättre täckning av de olika ämnesdomänerna och att beskriva i vilken utsträckning proven är anpassade till ett visst lands läroplaner. Skolverket (2006b) har gjort jämförelser mellan de svenska läroplanerna och kursplanerna i matematik och naturvetenskap å ena sidan och provuppgifterna i de två stora studierna av kunskaper och färdigheter vid slutet av grundskolan (PISA och TIMSS) å andra sidan. Slutsatsen var att de två undersökningarna tillsammans väl representerar innehåll och mål i den svenska skolan, men att det också finns skillnader i tyngdpunkt mellan PISA och TIMSS (se avsnitt 2.3.4).

Inom ramen för TIMSS- och PIRLS-undersökningarna är det möjligt att undersöka förändringar i resultat mellan olika undersökningstillfällen inom ett visst land, utan jämförelser med andra länder. Även om så är fallet är det dock också i allmänhet av intresse att göra jämförelser med resultaten i andra länder, eftersom det kan finnas generella trendmässiga förändringar i dessa. I de fall det inte är möjligt att undersöka förändringar över tid på samma skala är det nödvändigt att relatera resultaten i ett visst skolsystem till de resultat som uppnåtts i andra skolsystem. Ett sätt att göra detta är att jämföra exempelvis Sveriges resultat med medelvärdet för samtliga deltagande länder, eller för en viss grupp av länder. I synnerhet då det är möjligt att göra jämförelser med en relativt stor grupp av länder vars sammansättning är konstant ger detta förfaringssätt en rättvisande jämförelse.

Ett annat grundläggande metodproblem avser definitionen av de olika populationerna som ingår i undersökningarna, och dessa har också förändrats över tid. Man kan skilja mellan tre olika huvudprinciper: (1) åldersdefinierad population, vilket innebär att populationen omfattar de elever som hör till en viss födelseårskohort (t.ex. 15-åringar) vid teststillfället; (2) årskursdefinierad population, vilket innebär att de elever som vid teststillfället går i en viss årskurs (t.ex. årskurs 8) hör till populationen; och (3) ålderstyrd årskursdefinierad population, vilket innebär att man gör en årskursbaserad definition, men där valet av årskurs bestäms av i vilken årskurs majoriteten av eleverna i viss födelseårskohort går.

Alla dessa huvudprinciper har använts. I korthet kan man sammanfatta utvecklingen med att IEA gått från en åldersbaserad populationsdefinition till en årskursbaserad definition, medan PISA använder en åldersbaserad populationsdefinition.

Valet av princip för populationsdefinition har framförallt betydelse då åldern för skolstart varierar över länder. Bland de länder som deltar i de internationella undersökningarna börjar eleverna i ett mindre antal länder skolan vid fem års ålder (t.ex. England och Skottland), i huvudparten av länderna börjar eleverna skolan vid sex års ålder, och i ett fåtal länder börjar eleverna skolan vid sju års ålder (t.ex. Danmark, Finland och Sverige). Då man använder åldersdefinierade populationer blir konsekvensen att elever i länder med sen skolstart kommer att ha färre skolår bakom sig då undersökningen görs än elever i länder med tidig skolstart. Vid årskursdefinierade populationer kommer elever i länder med sen skolstart att vara äldre vid undersökningstillfället än elever i länder där man börjar skolan vid tidigare år. Effekterna av den ålders-

styrda årskursdefinitionen blir i huvudsak desamma som då man använder åldersdefinierat urval.

Oberoende av vilken modell man väljer uppstår en orättvisa vid jämförelser mellan länder med olika ålder för skolstart. Då man använder åldersprincipen gynnas länder med tidig skolstart eftersom eleverna i dessa kommer att ha fler skolår bakom sig, och därför har hunnit lära mer. Då man använder årskursprincipen gynnas länder med sen skolstart eftersom eleverna i dessa länder kommer att vara äldre, och därför har hunnit lära mer. Om ett år ytterligare i skolan har samma effekt som ytterligare ett levnadsår kan orättvisorna sägas vara lika stora oberoende av vilken princip man väljer. Om däremot effekten av ett skolår skiljer sig från effekten av ett levnadsår skulle man kunna hävda att man bör välja den princip som skapar minst orättvisa.

En betydande mängd forskning har undersökt den effekt som ett års ytterligare skolgång har på utveckling av generella förmågor och på kunskaper och färdigheter i skolsammanhang. (se Gustafsson 2009; Gustafsson och Yang Hansen 2009; samt avsnitt 5.2).

Resultaten visar att effekten av ett skolår är större än effekten av ett kronologiskt år, liksom att effekten av ett skolår är starkare för yngre elever än för äldre elever. Skolårseffekten är starkare för kunskaper och färdigheter inom matematik och naturvetenskap än den är för allmänna kognitiva färdigheter och för läsning. Samtidigt visar resultaten att åtminstone upp till ca. 15 års ålder finns det också en ålderseffekt som inte är försumbar.

Den starka skolårseffekten medför att de tidiga internationella undersökningarna som huvudsakligen genomförts med åldersbestämda populationer underskattat prestationerna i länder med sen skolstart. Det faktum att det även finns en ålderseffekt medför att resultaten i länder med sen skolstart är överskattade då man använder en årskursbaserad populationsdefinition. Detta innebär att det inte finns någon design som gör det möjligt att på ett rättvisande sätt jämföra resultaten i skolsystem med olika skolstartsålder. Genom att utnyttja den kunskap som nu finns om betydelsen som ett skolår och ett kronologiskt år har, är det dock möjligt att i viss utsträckning korrigera för den effekt som skillnader i dessa faktorer medför. Gustafsson (2009) har genomfört sådana korrigeringar för Sveriges resultat i samtliga de internationella undersökningar som Sverige deltagit i.

Det måste understrykas att dessa korrigeringar bygger på starka antaganden, och att den empiriska grunden i form av skattningar av ålders- och skolårseffekter inom olika ämnen för olika åldersgrupper ännu är relativt outvecklad.

Det som i detta komparativa sammanhang framstår som det viktigaste antagandet är att de olika alternativ till formell skolgång som finns i olika länder, t.ex. i form av daghem, förskola och förskoleklass, inte bidrar med starkare effekter på kunskaper och färdigheter än de som representeras av ålders-effekterna. Även om detta antagande sannolikt inte är fullt ut uppfyllt finns det dock mycket som pekar på att en förutsättning för att elever skall utveckla den typ av kunskaper och färdigheter som mäts av de prov som används i de internationella undersökningarna är att de deltar i en läroplansreglerad skol-verksamhet. Vi återkommer till frågan om resultatens pålitlighet i avsnitt 2.3.4.

2.3.3 Kunskapsutvecklingen sedan 1960-talet

Nedan presenteras mycket kortfattat resultaten från de sammanställningar av resultat från de internationella studier med svenskt deltagande som har genomförts avseende matematik, naturvetenskap och läsning. Gustafsson (2009) har i ett första steg inom ramen för varje studie försökt bestämma det svenska resultatet på *d*-skalan i relation till en så jämförbar grupp av länder som möjligt, i huvudsak bestående av sådana länder som hör till de nuvarande EU- och/eller OECD-grupperna. Detta har inneburit att vissa utvecklingsländer uteslutits från de tidiga undersökningarna. För undersökningarna genomförda efter 1995 har det dels inneburit att utvecklingsländer uteslutits från IEA-undersökningarna, och att jämförelser begränsats till den grupp av länder som deltagit i samma undersökningar som Sverige. I nästa steg har Sverige jämförts med medeltalet för övriga länder med avseende på de deltagande elevernas ålder och antal genomgångna årskurser. I det tredje och sista steget har korrekationer gjorts av de svenska resultaten, på så sätt att den förväntade effekten på det svenska resultatet av ålders- och skolårsavvikelsen från det internationella medelvärdet bestämts. Eftersom Gustafsson och Yang Hansen (2009) relativt utförligt redovisar de resultat som framkommer i dessa tre steg, så presenteras de här endast i kortfattat form.

2.3.3.1 Matematik

Sverige har endast deltagit i en undersökning som avser de yngre eleverna i grundskolan, nämligen TIMSS 2007. Några möjligheter att undersöka förändringar över tid för elever i de lägre årskurserna finns därför inte. Dock låg såväl det observerade som det korrigerade resultatet för de svenska eleverna en bit under medelvärdet för jämförelsegruppen av länder.

Sverige har deltagit i fler undersökningar som fokuserar elever i slutet av grundskolan. Den första genomfördes redan 1964 och den senaste år 2007. I de

två första studierna från 1964 respektive 1980 var de svenska elevernas observerade resultat mycket svaga: såväl de korrigerade som de okorrigerade resultaten låg betydligt under det internationella medelvärdet.

I den tredje undersökningen, som genomfördes 1995, hade eleverna i åk 7 relativt sett förbättrat sina resultat, och låg strax under det internationella medelvärdet. I denna undersökning deltog Sverige också med ett inofficiellt urval av elever i åk 8, som inte ingick i den internationella resultatrapporteringen. Eleverna i åk 8 hade mycket goda resultat, en liten bit under den internationella toppnivån. Detta var den första matematikundersökningen i vilken ett urval av svenska elever hade ungefär lika många skolår bakom sig som övriga länders elever. De var dock också äldre än övriga länders elever. När resultaten för eleverna i åk 8 korrigeras för ålders- och årskursskillnader blir den svenska skattningen också något lägre, men ligger fortfarande över det internationella medelvärdet.

Även om Sveriges resultat i TIMSS 1995 var mycket goda är resultatbilden därefter mycket dystert. I de två TIMSS-undersökningar som Sverige deltagit i därefter har Sverige, liksom övriga länder, haft urval från åk 8. Enligt både TIMSS 2003 och TIMSS 2007 har det skett kraftiga försämringar i matematikprestationerna, så att eleverna i åk 8 i dessa undersökningar presterar på en lägre nivå än eleverna i åk 7 gjorde år 1995.

Sverige har deltagit i de tre PISA-undersökningar som genomförts från år 2000. Urvalet till PISA består av elever som är 15 år, vilket innebär att ålderskillnaderna är små, men att de svenska eleverna har gått kortare tid i skolan. De observerade resultaten visar på prestationer något över medelvärdet, med en fallande trend över de tre studierna. Korrektur för ålders- och årskursskillnader medför en höjning av de svenska resultaten, så att Sverige i samtliga tre omgångar hamnar en bit ovanför medelvärdet. Den fallande trenden kvarstår dock i huvudsak.

De svenska matematikresultaten är bättre i PISA än de är i TIMSS, och i synnerhet gäller detta de korrigerade resultaten. Som närmare utreds i avsnitt 2.3.4 har dessa skillnader sin grund i att PISA i huvudsak undersöker problemlösning i vardagssituationer, vilket är en typ av uppgifter där de svenska eleverna vanligtvis har bättre resultat än inom andra matematiska områden. TIMSS har däremot en tyngdpunkt på mer traditionellt matematiskt innehåll, som algebra och geometri, och inom dessa områden är de svenska eleverna traditionellt svagare.

Sammanfattningsvis ger de internationella studierna av förändringar i matematikkunskaper för grundskolans senare år följande bild: Under början av 1960-talet var de svenska matematikresultaten svaga, och detta gällde även år 1980. Vid mitten av 1990-talet hade det skett en betydande förbättring och de svenska elevernas resultat närmade sig då internationell toppnivå. Enligt de undersökningar som har genomförts därefter har resultaten försämrats kraftigt.

2.3.3.2 *Naturvetenskap*

Sverige deltog med urval av elever från grundskolans lägre årskurser i de två första undersökningarna avseende naturvetenskap. Dessa undersökningar visade på resultat över medelvärdet för de svenska eleverna i åk 4. I SISS 1984 deltog Sverige också med ett urval av elever från åk 3, vars resultat låg något under medelvärdet. De flesta övriga länder deltog dock med elever i åk 4 eller åk 5, och korrigeringar för dessa skillnader medför att de redan goda svenska resultaten framstår som än bättre. Med de korrigerade resultaten hörde Sverige resultatmässigt till de främsta länderna i de två första undersökningarna, med resultat som nära sammanföll med dem som erhöles av Japan, som i de tidiga undersökningar var det land som hade högst prestationer.

Efter SISS 1984 har inte Sverige deltagit i någon undersökning avseende naturvetenskap i åk 4 förrän i TIMSS 2007. I denna studie visar både okorrigerade och korrigerade värden på resultat strax under det internationella medelvärdet. Detta pekar på en nedgång i resultaten över tid i naturvetenskap för elever i grundskolan tidiga år.

Sverige har deltagit i flera undersökningar med urval av elever från grundskolans högre årskurser. De observerade resultaten var relativt goda i den första undersökningen från 1970, med ett resultat en bit ovanför det internationella medelvärdet. I den andra undersökningen som gjordes år 1984 deltog Sverige med urval både från åk 7 och åk 8. Den senare gruppen hade resultat något över medelvärdet, medan resultaten för eleverna i åk 7 låg under medelvärdet. De svenska eleverna hade dock gått i skolan kortare tid, och i synnerhet gällde detta urvalet från åk 7 i SISS 1984. Korrigeringarna medför därför kraftiga höjningar av resultaten och för upp de svenska elevernas resultat till internationell toppnivå, på nivå med Japan.

I TIMSS 1995 deltog Sverige som redan nämnts med urval från åk 7 (och åk 6), liksom även med ett inofficiellt urval från åk 8. Det observerade resultatet för åk 7 låg nära det internationella medelvärdet, medan det observerade resultatet för åk 8 låg nära den absoluta toppen, med en placering bland de

östasiatiska länder som i de senare internationella undersökningarna genomgående haft de bästa resultaten. Eftersom de svenska eleverna i åk 8 åldersmässigt låg över det internationella medelvärdet innebar korrigeringen en nedjustering av värdet, dock utan att detta hade mer än marginell effekt. Korrigeringen av resultatet för eleverna i åk 7 förde upp detta till samma höga nivå som det okorrigerade resultatet för åk 8.

Sverige har därefter, som redan framgått, deltagit i två TIMSS-undersökningar (2003 och 2007) med urval från åk 8. I båda dessa har nedgången observerats, och särskilt stark var nedgången mellan 1995 och 2003. Korrekturen medför en ytterligare sänkning av de svenska elevernas resultat, på grund av deras högre ålder. Liksom i matematik innebär nedgången från 1995 att de svenska eleverna i åk 8 år 2007 presterade sämre än eleverna i åk 7 gjorde år 1995.

De svenska resultaten i naturvetenskap i PISA 2000 låg något över medelvärdet, men har i de två senaste omgångarna av PISA gått ner till medelvärdet. Även för naturvetenskap medför korrigeringarna att de svenska PISA-resultaten förbättras relativt andra länder. Skillnaden i resultatnivå mellan PISA och TIMSS är dock mindre inom naturvetenskap än inom matematik.

Sammanfattningsvis framträder följande resultatbild vad gäller utvecklingen inom det naturvetenskapliga området: I de undersökningar där yngre svenska grundskoleelever deltagit har de uppvisat mycket goda resultat, med undantag för den TIMSS-undersökning som genomfördes år 2007. För elever i de senare årskurserna i grundskolan var resultaten mycket goda enligt de undersökningar som gjordes fram till 1995, om hänsyn tas till att de svenska eleverna gått i skolan kortare tid. Därefter har resultaten kraftigt försämrats.

2.3.3.3 *Läsförståelse*

Sverige har deltagit i flera undersökningar av läsförståelse som omfattat urval från grundskolans lägre årskurser. I de första två undersökningarna (1970 och 1991) hade de svenska eleverna mycket goda observerade resultat nära den internationella toppnivån. Korrektur för det faktum att de svenska eleverna gått kortare tid i skolan höjer resultaten ytterligare. I PIRLS 2001 deltog Sverige med urval både från åk 3 och åk 4. Eleverna i åk 4 hade de bästa resultaten bland samtliga deltagande länder, medan eleverna i åk 3 hade resultat nära medelvärdet. Korrekturen medför att de skattade resultaten för de två årskurserna hamnar på samma höga nivå, vilken sammanfaller med resultaten från de två tidiga undersökningarna. Det är intressant att notera att justeringen

för den högre medelåldern för elever i åk 4 endast medför en marginell sänkning av prestationsnivån, som inte påverkar den svenska förstaplaceringen.

De observerade resultaten i PIRLS 2006 visar på en nedgång, och Sveriges har i denna undersökning förlorat sin absoluta tätposition. För PIRLS 2006 har korrigeringen större effekt, och den nedjustering som drabbar Sveriges resultat skulle även innebära förlust av ytterligare fem platser i rangordningen mellan länderna.

I den första undersökningen (1970) av läsförståelse i grundskolans senare årskurser låg de svenska elevernas resultat nära medelvärdet, medan Sverige resultatmässigt i den andra undersökningen (1991) låg bland de främsta länderna. Justeringarna gör att de goda svenska resultaten i denna undersökning blir än mer markerade. De tre PISA-undersökningarna visar att de svenska elevernas resultat ligger något över medelvärdet. Både de observerade och de korregerade resultaten visar dock på en svagt sjunkande trend.

I kort sammanfattning visar genomgången av resultat av studierna av läsförståelse på att de svenska eleverna i både yngre och äldre åldersgrupper haft mycket god läsförståelse jämfört med elever i andra länder. Detta gäller även i de senast utförda undersökningarna, men under 2000-talet kan en nedgående trend noteras.

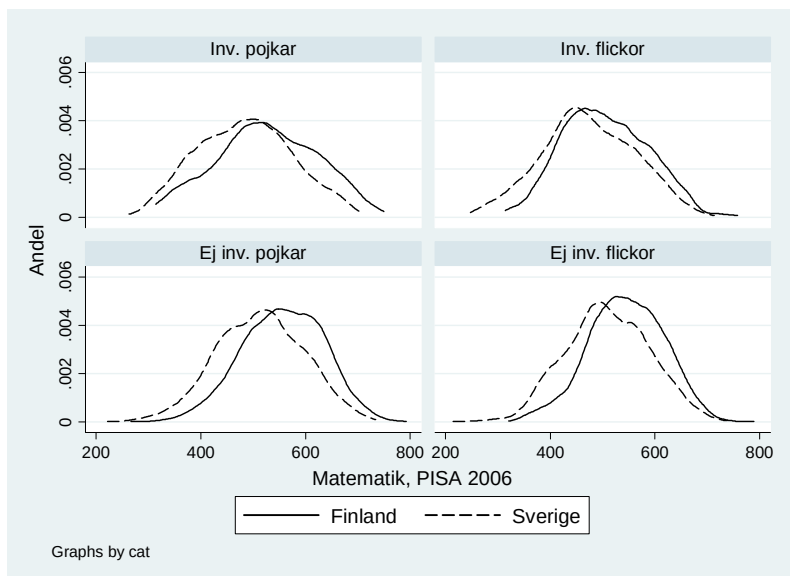
2.3.3.4 Några jämförelser mellan Sverige och Finland

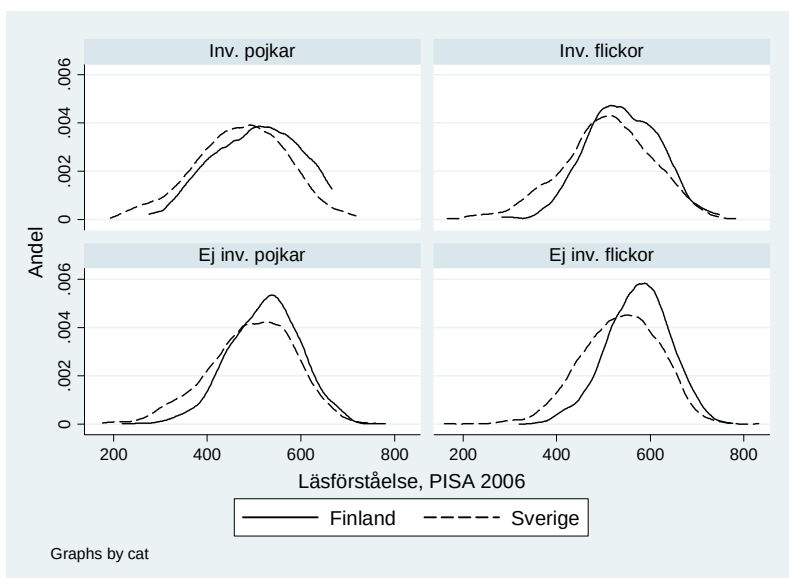
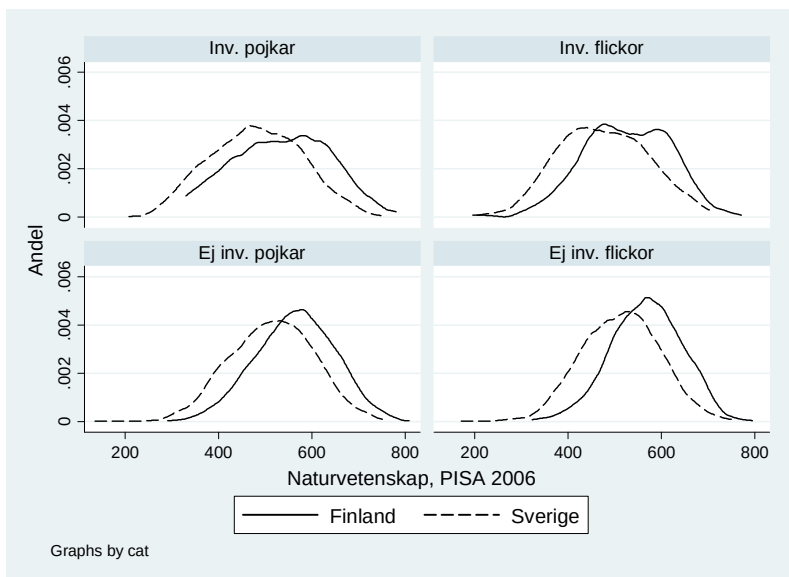
Ett av de mest slående resultaten i de internationella studierna är de utomordentligt goda resultaten som Finland uppnått i PISA-undersökningarna. Inom samtliga områden har de finska eleverna bland de allra bästa resultaten, med en tendens till ett ökat försteg över tid. De finska eleverna hade också de bästa resultaten för såväl äldre som yngre elever i IEAs läsundersökning 1991. I de internationella undersökningar som genomfördes dessförinnan utmärkte sig dock inte de finska eleverna genom särskilt framstående resultat, och i de flesta tidiga studier låg resultaten för de finska och de svenska eleverna mycket nära varandra. Liksom Sverige har Finland skolstart vid sju års ålder, och i samma mån som Sveriges resultat har underskattats har detta även gällt Finlands resultat. Mot denna bakgrund blir den intressanta frågan varför de finska elevernas resultat fortsatte att utvecklas positivt efter början av 1990-talet, medan de svenska elevernas resultat utvecklats i negativ riktning? Denna fråga kommer inte att besvaras här, men det kan ändå noteras att detta resultatmönster utesluter flera av de förklaringar som förts fram för att förklara de framstående finska skolresultaten, som att detta har med den finska kulturen

eller det finska språket att göra. Det utesluter också vissa av de förklaringar som framförts för att förklara den svenska resultatförsämringen, som exempelvis att denna har sin grund i den ekonomiska kris som drabbade Sverige i början av 1990-talet; Finland drabbades ju än hårdare av 1990-tals krisen.

En annan förklaring som framförts är att Finland i jämförelse med Sverige haft en mycket liten invandring, och ytterligare en hypotes är att en större andel av de svaga finska eleverna skulle vara uteslutna ur undersökningen. Det är därför av intresse att göra en mer detaljerad beskrivning av resultatskillnaderna mellan Finland och Sverige på olika prestationsnivåer, och för olika grupper av elever. För läsning, matematik och naturvetenskap har därför täthetsfunktioner skattats på grundval av data från PISA 2006 för elever med olika migrationsbakgrund och kön (se Figur 2.4).

Figur 2.4. Skattade täthetsfunktioner för Sverige och Finland, PISA 2006





Som framgår av figuren finns det stora skillnader i resultat till Finlands fördel för samtliga fyra undergrupper, vilket visar att Finlands bättre resultat inte kan förklaras av en lägre andel elever med utländsk bakgrund. Täthetsfunktionerna visar för både Sverige och Finland på approximativa normalfördelningar, men där täthetsfunktionen för Finland är förskjuten mot högre

värden. För läsförståelse återfinns skillnaderna till förmån för de finska eleverna i första hand för eleverna med lägre prestationer, men för matematik och naturvetenskap finns en prestationsskillnad till förmån för de finska eleverna på alla prestationsnivåer. Detta talar mot hypotesen att de svaga eleverna skulle vara underrepresenterade i det finska urvalet av elever.

2.3.4 Diskussion och slutsatser

Även om de internationella studierna erbjuder en stor mängd information om utvecklingen av kunskaper och färdigheter är det ändå påtagligt hur begränsad informationen är när syftet är att beskriva utveckling under drygt 40 år inom flera ämnesområden för flera åldersgrupper. Fram till mitten av 1990-talet genomfördes de internationella studierna med 10–20 års mellanrum inom varje ämnesområde, vilket gör det svårt att med någon större precision fastställa om och när ett trendbrott ägt rum. Från mitten av 1990-talet har de internationella studiernas frekvens och systematik ökat, men i gengäld har de svenska skolmyndigheternas intresse för att delta i studierna varit vacklande.

Det är också påtagligt i vilken utsträckning utfallet i de olika studierna påverkas av vilken årskurs de deltagande eleverna går i. Detta illustreras med stor tydlighet i de undersökningar där Sverige deltagit med två närliggande årskurser, där skillnaden i resultat för elever i de två årskurserna i allmänhet motsvarar skillnaden mellan ett medelmåttigt resultat och ett resultat bland de högst presterande länderna. Den forskning som delat upp skillnaden i resultat mellan två närliggande årskurser i en komponent som har med effekten av ett kronologiskt år och ett skolår gör det också möjligt att i viss utsträckning korrigera för effekterna av att de svenska elevernas årskurstillhörighet och ålder skiljt sig från många av de övriga länderna.

Antalet undersökningar som ger information om de svenska elevernas resultat i grundskolans tidiga år är begränsat. Dock finner vi för naturvetenskap prestationer på hög internationell nivå under perioden 1970 till 1984. Därefter finns nästa observationspunkt först 2007, då resultatet är genomsnittligt. För matematik visar den enda tillgängliga observationen, som är för 2007, på resultat under medelvärdet. För läsförståelse finns information från perioden 1970 till 2006 och under hela denna period har de svenska elevernas läsresultat legat på hög internationell nivå. Det finns dock en tydlig tendens till sjunkande resultat under tidsperioden.

För eleverna i grundskolans senare år finns information från fler studier. För matematik var resultaten svaga i de två första undersökningarna som genom-

fördes 1964 respektive 1980. TIMSS 1995 visade dock på goda resultat strax under den internationella toppnivån, men enligt de TIMSS-undersökningar som gjorts därefter har de svenska resultaten försämrats till en nivå betydligt under det internationella medelvärdet. PISA-undersökningarna ger en mer positiv bild av de svenska elevernas matematikkunskaper med resultat över medelvärdet, men även PISA visar på en sjunkande trend från 2000 till 2006.

För naturvetenskap hade de svenska eleverna resultat på internationell toppnivå under perioden 1970–95. De har därefter försämrats kraftigt, enligt TIMSS till en nivå under det internationella medelvärdet, och enligt PISA till en nivå över det internationella medelvärdet.

För läsning låg resultaten år 1970 strax ovanför det internationella medelvärdet, men hade internationell toppnivå enligt den undersökning som genomfördes 1991. Enligt PISA var resultaten fortfarande av god internationell klass år 2000 men har därefter försämrats något.

I kort sammanfattning pekar dessa analyser på att den svenska skolan nådde mycket goda resultat fram till mitten av 1990-talet. Ett undantag är matematik, men även inom detta område var resultaten betydligt bättre vid mitten av 1990-talet än vad som tidigare varit fallet. Efter 1995 har resultaten försämrats, och då i synnerhet inom matematik och naturvetenskap där nedgången är dramatisk.

Avslutningsvis finns det anledning att göra några kommentarer till de stora skillnaderna i utfall för Sveriges del mellan TIMSS å ena sidan och matematik och naturvetenskap i PISA å den andra sidan, med bättre resultat i PISA än i TIMSS. Det finns flera möjliga förklaringar till detta. Som framgår av Skolverkets (2006b) analys av PISA- och TIMSS-uppgifternas egenskaper omfattar uppgifterna i PISA betydligt längre texter än uppgifterna i TIMSS. En förklaring är sålunda att PISA-proven i större utsträckning influeras av läsförståelse och läshastighet, och eftersom de svenska eleverna relativt sett är starkare på läsförståelse än matematik och naturvetenskap leder detta till bättre resultat i PISA än i TIMSS. En annan förklaring tar fasta på skillnader i uppgifternas innehåll och art. Medan TIMSS-proven i matematik och naturvetenskap har ett ganska traditionellt ämnesfokus har PISA-proven fokus på sådana kompetenser som man kan förväntas behöva i ett framtida yrkesliv, och uppgifterna har en stark tyngdpunkt på problemlösning i vardagliga kontext. På denna typ av uppgifter har de svenska eleverna traditionellt relativt sett bättre resultat än på uppgifter som avser områden som algebra och geometri. Det är rimligt att tro att denna skillnad mellan uppgiftsinnehållets karaktär i PISA och TIMSS är

den viktigaste förklaringen till de bättre svenska resultaten i PISA. Som påpekats av Skolverket (2006b) är både PISA och TIMSS förenliga med de svenska läroplanerna och båda undersökningarna behövs för att proven skall ge en adekvat täckning av innehållet i läroplanerna.

2.4 Andra prov

En annan typ av information om utvecklingen av kunskaper och färdigheter är mätningar av förändringar över tid i populationers resultat på olika former av psykologiska test av förmågor och färdigheter. Test av olika begåvningsaspekter i samband med obligatoriska inskrivningsprövningar till militärtjänst är i många länder en värdefull källa till information om den trendmässiga utvecklingen inom landet. I allmänhet avser dessa prov dock endast män i 18-19 års ålder. En annan källa till information om förändringar på populationsnivå utgörs av resultat från normeringar av psykologiska test, vilka baseras på slumpmässiga urval av personer i olika åldrar, och vilka görs om med vissa mellanrum.

På grundval av framförallt dessa två informationskällor har Flynn (1987) identifierat det mycket intressanta fenomenet att det under tidsperioden 1930-80 i ett stort antal länder skett en kraftig uppgång i prestationer på de kognitiva testen. Uppgången är starkast för icke-verbala problemslösningsprov (ca. $d = 1,0$), och endast ungefär hälften så stor för verbala prov (ordförråd). Denna s.k. Flynn-effekt har fått många olika förklaringar, där bland annat utbildningens ökade omfattning och kvalitet, ökad spridning av massmedia, och bättre närings- och hälsoförhållanden brukar nämnas. Någon allmänt vedertagen teori om orsakerna till Flynn-effekten finns dock inte (Neisser 1998).

Det är också intressant att notera att under senare tid är den tidigare närmast kontinuerliga ökningen inte längre generellt synlig. I vissa länder har testresultaten fortsatt att förbättras, medan tillväxten stagnerat på andra håll (Flynn 2008). I synnerhet gäller detta Västeuropa och Australien (Lynn och Harvey 2008). Två av de länder där tillväxten ebbat ut är Danmark och Norge. Fram till i början av 1990-talet steg resultaten sakta, men därefter upphörde ökningen i båda länderna (Teasdale och Owen 2000; Sundet m.fl. 2004).

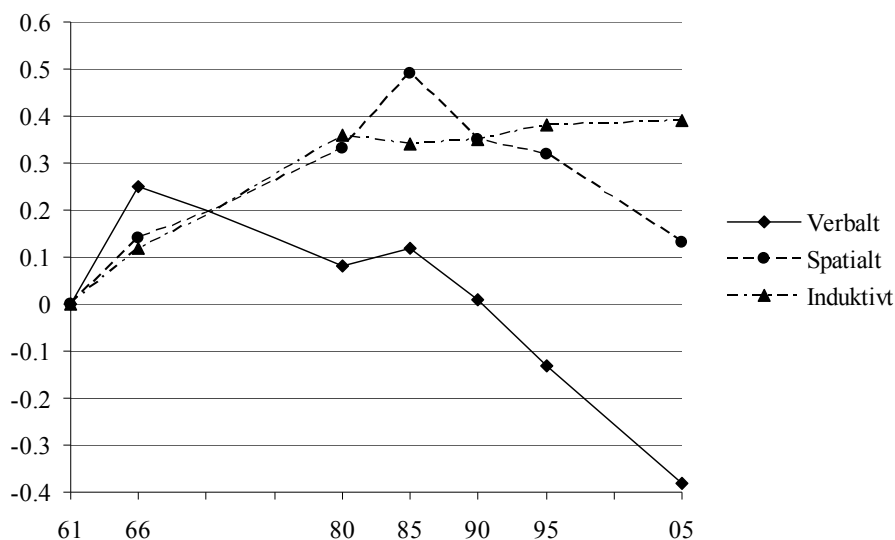
Sverige har lämnat viktiga bidrag till forskningen om Flynn-effekten, bland annat genom det kohortsekventiella UGU-projektet, i vilket successiva urval av elever i åk 6 (13-åringar) under perioden mellan 1961 och 2005 erhållit tre identiska begåvningsprov. Detta material är unikt både genom att det omfattar

urval av båda könen i relativt tidig ålder, och genom den långa tidsperiod som täcks. Svensson (2008) har nyligen analyserat resultat från samtliga sju genomförda mätningar, och fynden från dessa sammanfattas kort nedan.

De tre begåvningsproven mäter verbal förmåga (ordförråd (Motsatser)), induktiv förmåga (Talsrier) och spatial förmåga (Plåtvikning, som innebär att en tredimensionell figur ska skapas från ett tvådimensionellt underlag). Proven utvecklades av Svensson (1964) inför den första insamlingen av data 1961, omfattande personer födda 1948. Datainsamlingar har sedan ägt rum 1966, 1980, 1985, 1990, 1995, och 2005. För de första två insamlingarna användes åldersdefinierade individurval, och för de övriga insamlingarna användes stratifierade klusterurval av klasser. De analyser som Svensson (2008) redovisar omfattar endast normalåriga elever, för att skapa jämförbarhet över de två urvalsmodellerna.

I Figur 2.5 redovisas resultatutvecklingen över de sju mätningarna för de tre testen. Resultaten uttrycks på d -skalan, med värdet för 1961 års mätning satt till $d = 0$.

Figur 2.5. Förskjutningar i testmedeltal från 1961 till 2005 uttryckta i Cohens d



Källa: Svensson (2008).

Av figuren framgår att de tre testen uppvisar olika förändringsmönster över tid. För det verbala provet sker initialt en uppgång, men mellan 1985 och 2005 föreligger en linjär nedgång som uppgår till $d = 0,5$. För det spatiala provet kan vi observera en linjär uppgång mellan 1961 och 1985 ($d = 0,5$) och en linjär nedgång mellan 1985 och 2005 ($d = 0,4$). För det induktiva provet är förändringarna mellan 1961 och 1980 starkt positiva ($d = 0,35$), medan det mellan 1980 och 2005 skett endast en liten positiv förändring (ca. $d = 0,05$).

De försämrade resultaten på det verbala provet framstår vid första anblick som oroväckande. Svensson (2008) visar dock genom en analys av förändringar på uppgiftsnivå att resultatnedgången torde kunna förklaras av att många av de ord som ingår i ordförrådsprovet blivit mindre frekventa, vilket har som konsekvens att svårighetsgraden för uppgifterna ökar. Förklaringen till den observerade nedgången för det verbala provet torde därför vara att uppgifterna blivit svårare, snarare än att elevernas verbala förmåga försämrats.

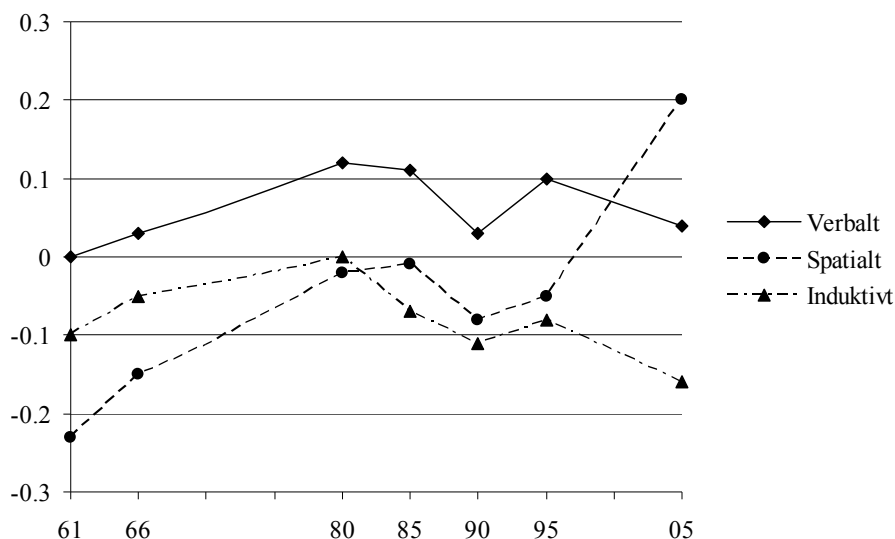
Det dramatiska mönstret av uppgång och fall vad gäller det spatiala provet torde dock inte kunna förklaras i termer av att provuppgifternas egenskaper förändrats. Enligt Svensson (2008) framstår det som lättare att förklara uppgången mellan 1961 och 1985, i termer av utbyggnad av grundskola, TV-nät, och förskola, än det är att förklara nedgången mellan 1985 och 2005. Med det tillgängliga empiriska underlaget finns heller inga ledtrådar till vid vilken ålder förändringen ägt rum, mer än att det måste skett före 13 års ålder.

För det induktiva provet kan uppgången mellan 1961 och 1980 förklaras med hänvisning till samma faktorer som nämndes i anslutning till det spatiala provet. Bristen på ökning under perioden 1980 till 2005 kan möjligen också delvis förklaras med hänvisning till samma faktorer, eftersom dessa inte förändrats i någon större utsträckning efter 1980. Härtill kommer att en ökad invandring efter 1990 förmodligen haft en viss inverkan på testresultaten i negativ riktning, vilket kan ha balanserat förändringar i positiv riktning.

Svensson (2008) redovisar också förändringar i könsskillnader i testresultaten, vilka återges i Figur 2.6. Även i detta avseende uppvisar de tre testen olika utvecklingsmönster. För det verbala provet föreligger en ökning av könsskillnaden till förmån för flickor mellan 1961 och 1980 (ca $d = 0,1$), och därefter en svag minskning fram till 2005 (ca. $d = 0,05$). För det induktiva provet minskar en liten initial könsskillnad till förmån för pojkarna till 0 år 1980, och återkommer därefter för att vara något större än den initiala skillnaden år 2005. Den mest slående utvecklingen av könsskillnaderna uppvisar dock det spatiala testet. År 1961 finns en skillnad till pojkarnas fördel

($d = -0,23$). Denna skillnad minskar sedan successivt, för att vara helt försvunnen år 1985, och är vid de tre efterföljande mätningarna i stort sett konstant. Vid den sista mätningen, år 2005, har emellertid flickorna ett bättre resultat än pojkarna ($d = 0,20$).

Figur 2.6. Könsskillnader 1961–2005 uttryckta i Cohens d



Not: Positiva värden markerar högre värden för flickor.

Källa: Svensson (2008).

Dessa mönster av förändringar i könsskillnaderna framstår som mycket intressanta, men också som svåra att förklara. Inte minst gäller detta förändringen av könsskillnaden i den spatiala förmågan.

Den resultatbild som framträder i analysen av förändringar i de tre kognitiva förmågorna diskuteras nedan i relation till de resultat som framkommit i övriga avsnitt.

2.5 Diskussion och slutsatser

Vi har i detta kapitel beskrivit resultatutvecklingen i den svenska skolan under en tidsperiod som omfattar hela grundskolans existens. En slutsats vi måste dra av genomgången är att det empiriska underlaget är begränsat. Ett skäl för detta är att vi tvingats avvisa betygen som en användbar informationskälla vad gäller

förändring av nivån på kunskaper och färdigheter inom olika områden. De relativa betygen ger pga. sin konstruktion ingen information om förändringar över tid, och de målrelaterade betygen har sådana brister i jämförbarhet över tid att de är oanvändbara som indikatorer på kunskapsutvecklingen. Det har förmodligen skett förändringar av både positiv och negativ art vad gäller kunskaps- och färdighetsnivåer inom olika områden, men de stora förändringar i nivån på de målrelaterade betygen som vi kan observera från 1998 är sannolikt snarare uttryck för varierande grader av betygsinflation än för faktiska förändringar i kunskaper och färdigheter.

Den nationella utvärderingen och de internationella studierna ger ett säkrare underlag för att bedöma kunskapsutvecklingen, men ett huvudproblem med dessa undersökningar är att de endast ger information om utvecklingen inom vissa ämnesområden, och då endast inom delar av dessa. De ger dock en i huvudsak enig bild av att resultaten inom matematik, naturvetenskap och läsförståelse i grundskolans senare år försämrats sedan början/mitten av 1990-talet. De olika studierna ger något olika uppskattningar av storleksordningen av försämringarna, men det samlade intrycket är att resultatnedgången är av betydande omfattning.

De internationella undersökningar som genomfördes till och med år 1995 pekar på att den svenska skolans resultat vid 1990-talets början var utomordentligt goda, och att nedgången därför skett från en hög internationell nivå. De mycket goda resultaten i den svenska skolan framstår som entydiga vad gäller naturvetenskap och läsning. För matematik är resultatbilden mer komplex eftersom resultaten i den IEA-undersökning som genomfördes år 1980 var mycket svaga. Den nationella utvärderingen i matematik visade dock på resultatförbättringar i matematik fram till mitten av 1990-talet. De goda resultaten i matematik för elever i åk 8 i TIMSS 1995 framstår därför som rimliga, i synnerhet som resultatskillnaden mellan elever i åk 8 och åk 7 är av den förväntade storleken i denna undersökning.

En intressant fråga är i vilken utsträckning resultatförändringarna i grundskolan har att göra med förändringar i de kognitiva förmågorna. Förändringen på det spatiala provet, med en kraftig höjning av medelvärdet fram till 1985 och en successiv minskning tillbaka till ett läge nära 1961 års resultat, sammanfaller i sina huvuddrag med mönstret av förändringar i skolprestationerna. Det är dock svårt att hävda att förändringar i den spatiala förmågan skulle vara kausalt relaterade till nedgången i matematik och naturvetenskap, eftersom sambandet mellan spatial förmåga och skolprestation är lågt

(Gustafsson och Balke 1993). En annan möjlighet är att de faktorer som orsakat nedgång i skolprestationerna också orsakat nedgång i den spatiala förmågan. Det är dock i detta sammanhang svårt att spekulera om vilka dessa faktorer skulle kunna vara.

Den förmåga som mest tydligt är kopplad till skolprestationer är den induktiva förmågan, dels därför att denna i stort sammanfaller med den generella intellektuella förmågan (Gustafsson 1984; Valentin Kvist och Gustafsson 2008), dels därför att den här mäts med ett numeriskt prov som innehållsmässigt har särskild relevans för matematikämnet. För denna förmåga finner vi dock ingen nedgång, utan en kraftig ökning fram till 1980, och därefter en svag ökning. Detta resultat pekar på att nedgången i prestationerna i matematik och naturvetenskap inte kan förklaras av försämrade förmågemässiga förutsättningar bland eleverna.

Frågan kvarstår alltså: Vilka faktorer kan förklara den svenska resultatutvecklingen? Vi återkommer till denna fråga i det avslutande kapitel 11.

Referenser

- Björklund, A., P-A. Edin, P. Fredriksson och A. Krueger (2003a), *Den svenska skolan – effektiv och jämlik?* Stockholm: SNS Förlag.
- Björklund, A., M. Lindahl och K. Sund (2003b), Family background and school performance during a turbulent era of school reforms, *Swedish Economic Policy Review*, 10, 111–136.
- Carlstedt, B. och J-E. Gustafsson (2005). Construct validation of the Swedish scholastic aptitude test by means of the Swedish enlistment battery, *Scandinavian Journal of Psychology*, 46, 31–42.
- Cliffordson, C. (2004). Betygsinflation i de målrelaterade gymnasiebetygen. *Pedagogisk Forskning i Sverige*, 9, 1–14.
- Flynn, J. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171–191.
- Flynn, J. (2008). The history of the American mind in the 20th century: A scenario to explain IQ gains over time and a case for the irrelevance of g. I Kyllonen, P., R. Roberts och L. Stankov (red.), *Extending Intelligence. Enhancement and New Constructs*. New York: Lawrence Erlbaum Associates.
- Gustafsson, J-E. (1984), A unifying model for the structure of intellectual abilities. *Intelligence*, 8, 179–203.
- Gustafsson, J-E. (2009). Kunskaper och färdigheter i grundskolan under 40 år: En kritisk granskning av resultat från internationella jämförande studier. Manuskript.
- Gustafsson, J-E. och G. Balke (1993), General and specific abilities as predictors of school achievement. *Multivariate Behavioral Research*, 28, 407–434.
- Gustafsson, J-E. och K. Yang-Hansen (2009), Resultatförändringar i svensk grundskola. Kap 3 i Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer*. Stockholm: Skolverket.
- Hattie, J.A. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.

- Husén, T. (1979). An international research venture in retrospect: The IEA surveys. *Comparative Education Review*, 23, 371–385.
- Klapp Lekholm, A. (2008). *Grades and grade assignment: Effects of student and school characteristics*. Göteborg: Acta Universitatis Gothoburgensis.
- Klapp Lekholm, A., och C. Cliffordson (2008), Discrepancies between school grades and test scores at individual and school level: effects of gender and family background. *Educational Research and Evaluation*, 14, 181–199.
- Klapp Lekholm, A., och C. Cliffordson (2009), Effects of student characteristics on grades in compulsory school. *Educational Research and Evaluation*, 15, 1–23.
- Lynn, R. och J. Harvey (2008), The decline of the world's IQ. *Intelligence*, 36, 112–120.
- Neisser, U. (Red.). *The rising curve. Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association.
- Riksrevisionen (2004). Betyg med lika värde? – en granskning av statens insatser. RIR 2004:11.
- Skolverket (2004). *Internationella studier under 40 år. Svenska resultat och erfarenheter*. Skolverkets aktuella analyser 2004. Stockholm: Skolverket.
- Skolverket (2004b). *Engelska i åtta europeiska länder – en undersökning av ungdomars kunskaper och uppfattningar*. Rapport 242. Skolverket, Stockholm.
- Skolverket (2006a). *Vad händer med likvärdigheten i svensk skola? En kvantitativ analys av variation och likvärdighet över tid*. Rapport 275, Stockholm: Skolverket.
- Skolverket (2006b). *Med fokus på matematik och naturvetenskap. En analys av skillnader och likheter mellan internationella jämförande studier och nationella kursplaner*. Skolverkets aktuella analyser 2006. Stockholm: Skolverket.
- Skolverket (2007). *Provbetyg – Slutbetyg- Likvärdig bedömning? En statistisk analys av sambandet mellan nationella prov och slutbetyg i grundskolans årskurs 9, 1998-2006*. Rapport 300, Stockholm: Skolverket.

- Skolverket (2009). Likvärdig betygssättning i gymnasieskolan? En analys av sambandet mellan nationella prov och kursbetyg. Rapport 338. Stockholm: Skolverket.
- Skolöverstyrelsen (1970). Betygsättning i grundskola och gymnasieskola. Rapport utarbetad av Skolöverstyrelsens arbetsgrupp för betygsfrågor. Stockholm: Skolöverstyrelsen Utbildningsförlaget.
- SOU (2007:28). *Tydliga mål och kunskapskrav i grundskolan – Förslag till nytt mål- och uppföljningssystem*. Stockholm: Utbildningsdepartementet.
- Sundet, J.M., D. Barlaug och T. Torjusssen (2004), The end of the Flynn effect? A study of secular trends in the mean intelligence test scores of Norwegian conscripts during a half century, *Intelligence*, 32, 349–362.
- Svensson, A. (1964). Sociala och regionala faktorerers samband med över- och underprestation i skolarbetet. (Rapport 1964:13). Göteborg: Göteborgs universitet, Institutionen för pedagogik.
- Svensson, A. (2008). Har dagens tonåringar sämre studieförutsättningar? En studie av förskjutningar i intelligenstestresultat från 1960-talet och framåt. *Pedagogisk Forskning i Sverige*, 13, 258–277.
- Teasdale, T.W. och D. Owen (2000), Forty-year secular trends in cognitive abilities. *Intelligence*, 28, 115-120.
- Tholin, J. (2003). *En roligare dans? Svenska skolors första tolkning av innebörden i lokala betygskriterier i tre ämnen för skolår åtta*. Högskolan i Borås.
- Valentin Kvist, A. och J-E. Gustafsson (2008), The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's Investment theory. *Intelligence*, 36, 422–436.

3 Subventioner: Argument och inriktning

Detta kapitel har två syften: dels att illustrera hur vi ser på utbildningsinvesteringar och dels att redogöra för de samhällsekonomiska argument som finns för att subventionera utbildning. Vår utgångspunkt är att utbildning har en positiv effekt på individers produktiva egenskaper och att dessa egenskaper också värderas på arbetsmarknaden.⁶

Utbildning är en investering som är förenad med både intäkter och kostnader. Detta synsätt sammanfattas i den s.k. humankapitalteorin. Humankapitalteorin har många tillämpningar, t.ex. investeringar i hälsa och flyttbeslut, men den vanligaste tillämpningen rör utbildning. Låt oss vara konkreta för att illustrera grundtanken i teorin. Tänk er en individ som precis avslutat sin gymnasieutbildning. Denne står inför valet att fortsätta till högskolan eller att träda in på arbetsmarknaden. Att fortsätta till högskolan och studera i t.ex. 4 år kommer att generera en intäkt som huvudsakligen består av en ökning i den förväntade framtida inkomsten. De kostnader som är förenade med högskolealternativet består framförallt av utebliven inkomst under studieperioden. Alla intäkter och kostnader beräknas i nuvärdestermerna, dvs. de diskonteras till beslutstidpunkten. Precis som för andra former av investeringar fokuserar nationalekonomer ofta på utbildningens avkastning. Avkastningen tar hänsyn till alla typer av intäkter och kostnader (så länge som de kan värderas). Avkastningsbegreppet är centralt eftersom alla förändringar som ökar t.ex. högskoleutbildningens avkastning kommer att göra att fler väljer högskolestudier.⁷ Till exempel kommer ökade löneskillnader mellan högskole- och gymnasieutbildade göra att fler väljer högskoleutbildning. På motsvarande sätt kommer ett förmånligare studiemedelssystem att höja högskoleutbildningens avkastning vilket i sin tur ökar efterfrågan på högre utbildning.

Om den enda effekten av individens utbildning är att den påverkar den enskilde individen och alla marknader fungerar perfekt kommer individens val också att vara samhällsekonomiskt optimalt.⁸ I denna hypotetiska situation

⁶ Man kan ha andra utgångspunkter, t.ex. att utbildning bara signalerar individers inneboende produktivitet (se Spence 1974). Vår uppfattning är dock att evidensen talar för att utbildning ökar individers produktivitet vilket i sin tur genererar ekonomisk tillväxt; se t.ex. Krueger och Lindahl (2001).

⁷ Avkastningen på utbildning är så centralt att vi kommer att ägna kapitel 4 till att reda ut begreppen och beskriva hur avkastningen på utbildning i Sverige har utvecklats över tiden och hur den förhåller sig till avkastningen i andra länder.

⁸ Marknaderna är perfekta under ett antal förutsättningar. Till exempel ska arbetsmarknaden fungera perfekt vilket innebär att full sysselsättning ska råda och att hela produktivitetsoökningen

finns inga argument för staten att subventionera utbildning – att utbildning höjer individens produktivitet är alltså inget skäl till att subventionera utbildning. Detta är naturligtvis ett hypotetiskt scenario, men det har ändå ett viktigt budskap: Det krävs *avsteg* från den perfekta marknadsmodellen för att det ska vara samhällsekonomiskt motiverat att subventionera utbildning.

Vilka avsteg handlar det då om? I grunden rör det sig om avsteg från de förutsättningar som angavs ovan. Följaktligen är ett första argument för utbildningssubventioner att individens utbildning har positiva effekter på resten av samhället. Nationalekonomer pratar här om *externa effekter* – de är externa i meningen att de inte beaktas av individen när denne fattar sitt utbildningsbeslut. Det andra argumentet som kan motivera utbildningssubventioner är andra typer av *marknadssmisslyckanden*, vilket helt enkelt betyder att marknaderna inte är perfekta. Det främsta exemplet på ett marknadssmisslyckande är antagligen att kapitalmarknaden inte fungerar perfekt; individerna kan med andra ord inte låna fullt ut på den privata marknaden för att finansiera utbildningsinvesteringar.

I avsnitt 3.1 redogör vi för några av de argument som finns för att subventionera utbildningsinvesteringar. Det finns naturligtvis ett antal goda argument, vilket kanske inte är så förvånande givet hur mycket offentliga medel som läggs på utbildning inom OECD.⁹ Men genomgången säger inget om hur mycket utbildning ska subventioneras. Och den säger ganska lite om hur subventioner ska riktas.

Generellt sett har forskningen lite att bidra med vad gäller den stora frågan om hur mycket utbildning ska subventioneras. Däremot finns en del forskning som rör frågan om hur subventioner ska riktas. Denna forskning handlar om frågor som: Varierar subventionernas effektivitet med individens ålder? Hur ska utbildningsinvesteringar fördelas över individer med olika ekonomiska förutsättningar?

För att belysa dessa frågor krävs en modell för hur kunskaper förvärvas. Vi skisserar en modell i avsnitt 3.2 och gör en formell genomgång i Appendix. Denna modell är i grunden en humankapitalmodell och bygger på Cunha m.fl.

ska tillfalla individen i form av en inkomstökning. En annan förutsättning är att kapitalmarknaderna ska fungera perfekt. Detta innebär bl.a. att alla individer ska kunna låna hur mycket de vill till en given ränta.

⁹ I genomsnitt spenderade OECD-länderna 5,8 procent av BNP på utbildning 2005; offentliga medel motsvarade 5 procent och privata medel 0,8 procent av BNP (OECD 2008). Även om dessa siffror underskattar det privata åtagandet (eftersom studenternas egen tid inte räknas med) så är subventionsgraden avsevärd. Till och med i USA motsvarade offentliga åtaganden 34 procent av totala utgifter på högskolenivå.

(2006). Utgångspunkten är att kunskaper som inhämtats i ett tidigare skede påverkar subventionernas effektivitet i ett senare skede. I modellen finns ett marknadsmisslyckande inbyggt: barns utbildningsval begränsas av föräldrarnas ekonomiska förutsättningar. För given grundläggande förmåga finns därför ett skäl att rikta subventioner mot barn från familjer med sämre ekonomiska förutsättningar. Sådana subventioner är särskilt effektiva om interventionerna görs tidigt. Trots att tidiga interventioner är effektivast finns under rimliga förutsättningar ett effektivitetsargument för en andrachansutbildning. Detta uppstår om man underskattat barnens grundläggande förmåga eller överskattat föräldrarnas ekonomiska förutsättningar.

Detta var en snabb sammanfattning av modellens grundläggande implikationer. Läsaren som är nöjd med denna kan lämna modellen i avsnitt 3.2 därhän. Modeller skärper tanken men innefattar med nödvändighet ett antal abstraktioner som kan göra materialet mer svårtillgängligt än det övriga materialet i rapporten.

3.1 Argument för utbildningssubventioner

Syftet här är att kort redogöra för de argument som finns för att subventionera utbildningsinvesteringar.¹⁰ Som vi var inne på tidigare krävs externa effekter eller någon annan form av marknadsmisslyckande för att motivera utbildningssubventioner. Men subventioner kan också motiveras för att mildra effekterna av snedvridningar som t.ex. skapas av ett skattesystem och ett socialt skyddsnät med stora omfördelande ambitioner.

3.1.1 Externa effekter av utbildning

Utbildning kan ha positiva externa effekter. Dessa kan ha att göra med att utbildning förbättrar den demokratiska processen, minskar brottslighet och förbättrar hälsa.¹¹ Ett annat exempel på en positiv extern effekt är om utbildning ökar *tillväxttakten* i ekonomin.¹²

Ett av de mest grundläggande motiven till att subventionera grundläggande utbildning torde vara att utbildning kan förbättra den demokratiska processen.

¹⁰ Detta avsnitt bygger till stora delar på Björklund m.fl. (2003). En bra och utförlig diskussion om dessa frågor återfinns i Björklund och Lindahl (2005).

¹¹ Till listan av externa effekter skulle man möjligen kunna lägga till intergenerationella effekter av utbildning. Om föräldrarna får högre inkomster tack vare utbildning kan detta på olika sätt också gynna deras barn. Om sådana effekter inte beaktas av föräldrarna då de fattar sina beslut skulle detta kunna vara en extern effekt.

¹² Detta är grundtanken i modeller för endogen tillväxt, se Romer (1990).

Man kan tänka sig att högre utbildade medborgare både deltar i den demokratiska processen i större utsträckning och fattar bättre beslut. Detta är ett argument som framförts av Milton Friedman (1962) – normalt en varm anhängare av den fria marknaden – när han propagerar för en skolplikt eftersom ”ett stabilt och demokratiskt samhälle är omöjligt att upprätthålla utan ett minimum av läskunnighet och kunskap hos majoriteten av medborgarna och en allmän acceptans av en uppsättning av grundläggande värderingar”.

Empiriska studier har visat på ett samband mellan brottslighet, utbildning och ojämlikhet; se t.ex. Ehrlich (1973) och Imrohoroglu m.fl. (2004). Om ojämlikheten är hög finns ett stort antal höginkomsttagare, som utgör potentiella brottsoffer, samtidigt som att alternativkostnaden är låg för individer med begränsade arbetsmarknadsutsikter. Ett sätt att minska brottslighet är därför att minska ojämlikheten i samhället. Riktade utbildningsinsatser för dem med sämst arbetsmarknadsutsikter förbättrar avkastningen på arbete och kan därför förväntas minska brottsfrekvensen.

Ett annat argument för att subventionera utbildningsinvesteringar är externa effekter som verkar via den offentliga budgeten. Ett exempel är om utbildning ökar möjligheterna att få ett jobb. Då genererar högre utbildning dels större skatteintäkter och dels minskade utbetalningar av arbetslöshetsrelaterade ersättningar. Detta är en extern effekt eftersom individer inte beaktar dessa effekter när de söker efter jobb; se t.ex. Fredriksson och Holmlund (2001). Ett annat exempel är om utbildning förbättrar hälsan; se Lleras-Muney (2005). I ett samhälle med offentligt tillhandahållen sjukvård genererar utbildning positiva externa effekter som, återigen, verkar via den offentliga budgeten.

Nelson och Phelps (1966) samt Romer (1990) formulerar modeller där utbildning har externa effekter på tillväxten i ekonomin. Det kan här handla om att andras produktivitet påverkas av att individen utbildar sig eller som i Romers (1990) fall om att tillväxttakten i ekonomin påverkas av utbildning. I båda dessa fall är det alltså fråga om effekter utöver effekten av utbildning på individens egen produktivitet. Det empiriska stödet för att utbildning är en sådan tillväxtmotor är emellertid blandat (se t.ex. Krueger och Lindahl 2001). Enligt Björklund och Lindahl (2005) finns en stark och positiv effekt av utbildning på tillväxttakten i utvecklingsekonomier men inget som tyder på en sådan effekt i OECD-länderna.

Persson och Tabellini (1994) utvecklar ett annat argument som ger ett samband mellan utbildning och tillväxt. I deras analys skapar ojämlikhet politiska krav på omfördelning vilket i sin tur har negativa effekter på till-

växten. En riktad utbildningspolitik som minskar ojämlikhet kan därmed ha positiva effekter på tillväxten.

3.1.2 Andra former av marknadsmisslyckanden

Begränsade möjligheter att låna till studier är det mest framträdande exemplet på ett marknadsmisslyckande som ger ett argument för utbildningssubventioner. Detta är grunden i den modell som vi presenterar i nästa avsnitt och därför diskuterar vi inte dessa här; den empiriska evidensen om betydelsen av sådana kreditrestriktioner avhandlas i kapitel 10.

Ett standardantagande inom nationalekonomi är att individen är rationell. Men det finns en hel del evidens som tyder på att individer, framförallt ungdomar, ibland fattar beslut som inte är rationella. Till exempel värdesätter individer ofta framtida intäkter i alltför liten utsträckning (Warner och Pleeter 2001). Att individer inte agerar i sitt eget intresse kan betraktas som ett marknadsmisslyckande.

I dessa situationer finns argument för offentliga interventioner – inte på grund av externa effekter utan på grund av individernas eget beteende. Det kan vara i studenternas eget bästa att utbilda sig längre, men eftersom de värdesätter framtida intäkter för lite så väljer de att avsluta skolgången. Förlusten för samhället på grund av detta (irrationella) beteende kan vara mycket stor. Detta paternalistiska argument för offentliga interventioner är antagligen viktigt som grund för en allmän skolplikt.

3.1.3 Utbildningssubventioner för att korrigera snedvridningar

Institutionella faktorer som leder till lönesammanpressning och progressiv beskattning snedvrider individers utbildningsval. Orsaken till snedvridningen är att hela produktivitetsökningen inte tillfaller individen i form av en ökad lön; följaktligen investerar individer för lite i utbildning om t.ex. beskattningen är progressiv.¹³ Av dessa skäl finns argument för en utbildningspolitik som korregerar snedvridningar.

Att subventionera utbildning är ett sätt att återställa incitamenten för att investera i utbildning då progressiva skatter används för att omfördela inkomster. Ett exempel på dessa ges i Edin och Holmlund (1995) som visar att avkast-

¹³ När värdet av hemmaproduktion och fritid utvecklas på samma sätt som individens produktivitet och avgifter för utbildning antingen är avdragsgilla eller icke-existerande så påverkar en proportionell skatt inte utbildningsvalet; se t.ex. Jacobs (2005).

ningen på universitetsutbildning stiger avsevärt om man tar hänsyn till studiemedelssystemet.¹⁴

3.2 Hur ska subventioner riktas?

Uppbyggnaden av kunskaper är en dynamisk process. De kunskaper och färdigheter som har förvärvats tidigare är viktiga för hur mycket man lär sig i innevarande period. Vi presenterar en formell modell av denna process i Appendix. Här redogör vi kortfattat för de implikationer som modellen har för utformningen av utbildningspolitiken.

Det finns framförallt två frågor som vi vill belysa. Hur ska investeringar, generellt sett, fördelas över tid (för en given individ)? Och hur ska utbildningsinvesteringar fördelas över individer med olika förutsättningar? För att kunna säga något om dessa frågor krävs några antaganden. För det första, kommer vi att anta att avkastningen (för given grundläggande förmåga) är högre om investeringen görs tidigt snarare än sent. Skälet till detta kan t.ex. vara att man måste ha vissa baskunskaper innan man kan gå in på mer avancerade frågor. Ett annat skäl kan vara att vissa typer av studier är särskilt effektiva vid unga år; en rimlig utgångspunkt är t.ex. att förmågan till språkinläring är betydligt större i yngre åldrar. Samtidigt är det uppenbart att denna utsaga inte gäller för alla typer av utbildningsinvesteringar. Det är rimligen inte produktivt att placera tvååringar i skolbänken till exempel.

För det andra antar vi att barnens framtida inkomspotential inte kan användas som säkerhet för att låna pengar som finansierar studierna.¹⁵ Detta antagande är särskilt rimligt då det gäller utbildningsinvesteringar på lägre nivåer och är antagligen förhållandevis rimligt även på högre nivåer. Antagandet medför att de beslut som fattas av barn och föräldrar (*marknadsjämvikten*) inte kommer att vara samhällsekonomiskt optimala – marknadsjämvikten kommer med andra ord att vara ineffektiv. Implikationen är att utbildningsinvesteringar inte allokeras till dem med högre grundläggande förmåga (den samhällsekonomiskt optimala jämvikten) utan till dem vars föräldrar har högre livsinkomster. Detta innebär per automatik att det finns ett utrymme statliga interventioner som förbättrar marknadsjämvikten. Vi fokuserar nedan på

¹⁴ Edin och Holmlunds kalkyler tyder på att internräntan för män uppgick till knappt 7 % 1991 med hänsyn taget till skatter. Om man också beaktar studiemedel stiger internräntan till 11 %.

¹⁵ Detta kan kanske verka som ett udda antagande. Innebörden är bara att det inte är möjligt att övertyga en privat långivare om att man ska låna ut pengar för att finansiera ett barns utbildning med argumentet att barnet har en stor inkomspotential (och därmed stor återbetalningsförmåga) i vuxen ålder.

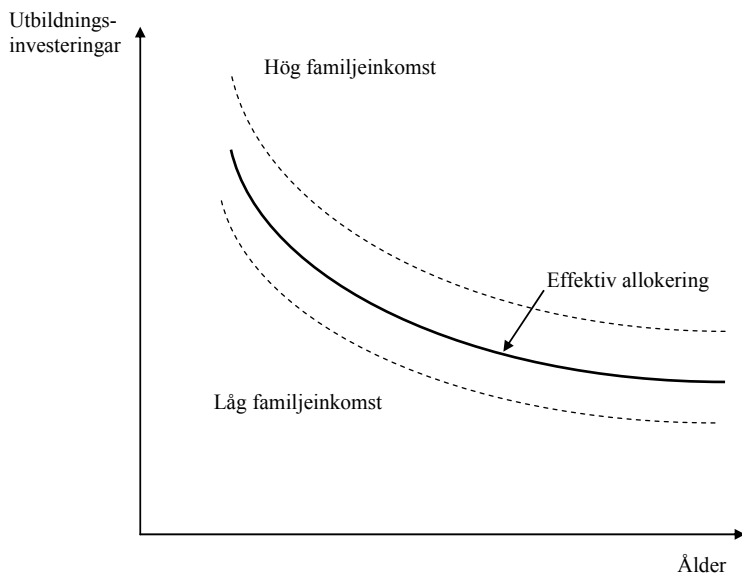
skillnader mellan barn i termer av föräldrarnas livsinkomster. Föräldrar med olika livsinkomster skiljer sig högst sannolikt åt även i andra avseenden, t.ex. utbildning och förmåga, vilket rimligen innebär att uppväxtmiljön kommer att vara systematiskt bättre för barn med högre livsinkomster. Om vi också skulle beakta sådana skillnader skulle argumentet för statliga interventioner för att hjälpa de sämst ställda att stärkas än mer.

De streckade linjerna i Figur 3.1 illustrerar utbildningsinvesteringar efter ålder för två barn med samma förmåga men olika familjebakgrund. Utbildningsinvesteringarna har en avtagande profil vilket följer av antagandet att avkastningen på tidiga investeringar är högre än på investeringar senare i livet. Den avtagande lutningen är densamma för de två familjerna, vilket har att göra med att vi antagit att föräldrarna inte är begränsade av restriktioner på kapitalmarknaden.

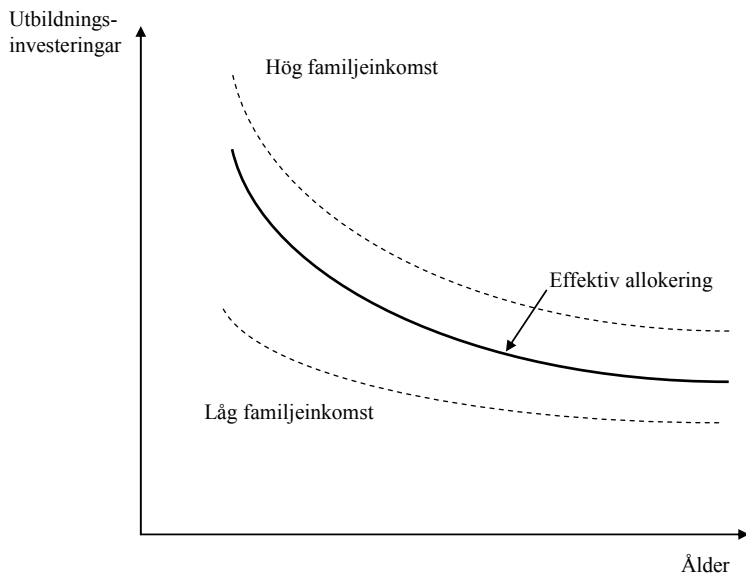
Skillnaden mellan de två familjerna är att föräldrar med lägre inkomst investerar mindre i sina barn, totalt sett. Denna lösning kan dock inte vara samhällsekonomiskt optimalt. Barnen har ju samma förmåga – vilket i modellen innebär att avkastningen på utbildningsinvesteringar är densamma – och därför borde lika mycket investeras i dem. Den heldragna linjen i figuren visar den samhällsekonomiskt optimala jämvikten, vilken följaktligen innebär att lika mycket resurser ska investeras i barn med samma förmåga.

En viktig fråga är på vilket sätt lånerestriktioner påverkar optimala utbildningsval och hur existensen av sådana restriktioner påverkar värdet av samhälleliga interventioner. Under de förutsättningar som gäller här (tidiga investeringar är produktivare) kommer familjer som inte kan låna hur mycket de vill investera för lite till en början; se Figur 3.2. Om lånerestriktioner är relevanta är argumentet för tidiga utbildningssubventioner riktade mot låginkomstfamiljer särskilt starkt.

Figur 3.1. Utbildningsinvesteringar efter ålder, samma förmåga olika familjeinkomst



Figur 3.2 Utbildningsinvesteringar efter ålder, samma förmåga olika familjeinkomst. Låginkomstfamiljen möter lånerestriktioner.



En samhällsekonomiskt optimal utbildningspolitik kommer att vara omfördelande *i förhållande till marknadsjämvikten*. Detta följer så länge som förmågor och familjeinkomster inte är perfekt positivt korrelerade. Om det inte finns något samband alls mellan familjeinkomsten och barns förmågor, ska utbildningspolitiken medföra att lika mycket investeras i barn med samma förmåga oberoende av familjeinkomst. Notera att det torde vara omöjligt att implementera en sådan lösning. Skälet är att staten aldrig kan förhindra familjer att investera mer än vad som är samhällsekonomiskt optimalt.

Den samhällsekonomiskt effektiva lösningen som illustreras i t.ex. Figur 3.1 ska tolkas som en optimal plan – den visar hur staten skulle vilja agera givet full information om barns förmågor och familjers livsinkomster. Men det är knappast rimligt att beslutsfattare besitter en sådan mängd information. En intressant fråga är om det finns argument för kompensation om det visar sig att man inte investerat tillräckligt tidigt i livet. Ibland har modellen som skisseras ovan tagits som intäkt för att avkastningen på investeringar i vuxen ålder är väldigt låg och att man därför inte bör subventionera denna typ av investeringar. Vår uppfattning är att detta inte följer av modellen. Det är förvisso sant att det med fullständig information är bättre att investera tidigt än sent, men utan perfekt information kan det i efterhand visa sig avkastningen på investeringar vid en vuxen ålder trots allt är hög. Den kan vara hög om det visar sig att familjen hade sämre ekonomiska förutsättningar än man från början trodde eller om barnet hade en högre förmåga än man från början trodde.

I Figur 3.3 illustrerar vi detta resonemang. Vi tänker oss två perioder. Den horisontella axeln visar faktiska investeringar i period 1 som andel av optimala investeringar i samma period; den vertikala axeln visar optimala investeringar i period 2 som andel av vad som hade varit optimalt i period 2 om man gjort den optimala investeringen i period 1. Den streckade 45-graderslinjen visar fallet där man inte kompenserar alls, utan bara fortsätter på den inslagna banan.

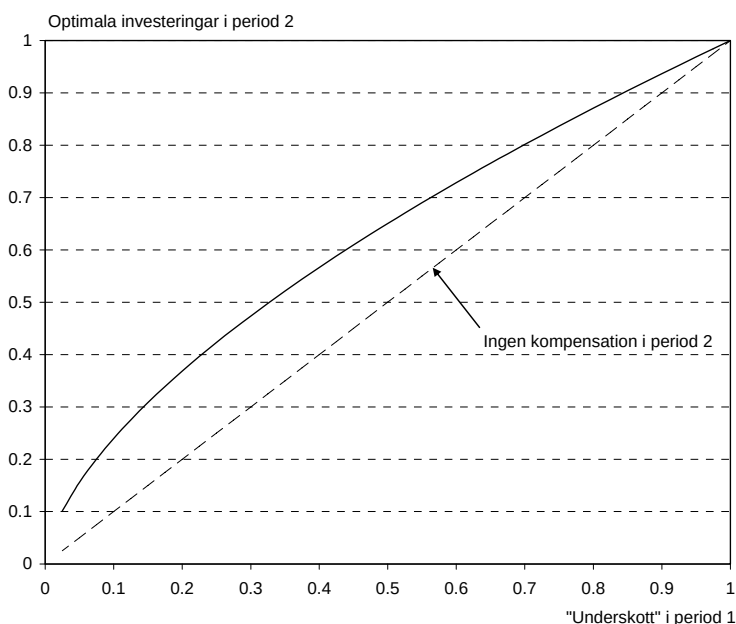
I modellen finns två olika krafter som verkar i motsatt riktning. För det första är investeringar i period 1 och 2 är komplement. Detta innebär att marginalavkastningen av att investera i period 2 är ökande i investeringarna i period 1. Denna mekanism talar mot kompensation i period 2. För det andra är investeringarnas skalavkastning avtagande. Detta betyder att marginalavkastningen av att investera i period 2 att vara fallande i investeringar i period 1. Denna mekanism talar alltså för kompensation i period 2.

Eftersom den heldragna linjen i Figur 3.3 ligger ovanför 45-graderslinjen är det optimalt att kompensera för underinvesteringar i period 1. Det finns alltså

ett argument för att fördela om utbildningsresurser så att mer investeras i barn som fått för lite till en början; om man så vill kan man säga att modellen innehåller ett effektivitetsargument för en andra chans i utbildningssystemet. Kompensationen blir dock aldrig fullständig, vilket har att göra med att investeringar i de två perioderna är komplement.

Exakt hur mycket omfördelning som ska ske mellan individer beror på hur utbildningsinvesteringar omsätts till kunskaper; se Appendix. Tyvärr har vi väldigt lite empirisk kunskap som hjälper till att avgöra denna fråga.

Figur 3.3 Optimala investeringar i period 2 som en funktion av "underskottet" i period 1



Not: Investeringar i period 1 och 2 anges som andel av optimala investeringar i respektive period.

3.3 Sammanfattning och diskussion

I detta kapitel har vi bl.a. presenterat en modell som har implikationer för inriktningen på utbildningsinsatserna. Modellen är grovt stiliserad men har likväl ett antal intressanta policyimplikationer. Modellen visar t.ex. att arbetsmarknadseffekterna av tidiga utbildningsinvesteringar kan vara väl så stora som effekterna av senare investeringar. Detta följer av att det man lärt sig under tidigare perioder påverkar ens förkovran i senare perioder. Det finns också ett

argument för offentliga interventioner. Talangfulla barn från fattiga familjer antas vara begränsade av föräldrars inkomst. En samhällsekonomiskt effektiv politik fördelar utbildningsinvesteringar efter avkastningen på ytterligare investeringar, inte efter föräldrars inkomst. I modellen finns alltså ett argument för utbildningssubventioner som är riktade mot barn från familjer med sämre ekonomiska förutsättningar. En sådan politik bidrar både till att öka effektiviteten och jämlikheten. Utbildningssubventionerna är särskilt effektiva om interventionerna görs tidigt.¹⁶

Argumentet för investeringar i tidig ålder innebär dock inte att avkastningen på sena investeringar behöver vara låg. Om det i efterhand visar sig att inriktningen på utbildningsinsatserna varit felaktig, så att alltför litet investerats i ett givet barn, kan avkastningen på sena investeringar vara hög. Det hade förstås varit än bättre om man insett detta på förhand, men trots allt fattas beslut om inriktningen på offentliga interventioner under osäkerhet, dvs. utan full kännedom om barn förmåga och familjers ekonomiska förutsättningar.

Det finns förstås ett stort antal andra argument, utöver det som presenterats i modellen, för offentliga interventioner. Vi har kort redovisat också dessa. Men att det finns argument för utbildningssubventioner betyder dock inte att mer alltid är bättre än mindre. Det är därför viktigt att kontinuerligt övervaka och utvärdera systemet så att resurserna används på ett effektivt sätt.

¹⁶ Att en politik som mildrar dessa s.k. kreditrestriktioner kan vara en politik som både är effektiv och jämlik noterades redan av Arthur Okun i den klassiska boken, *Equality and Efficiency: The Big Tradeoff* (Okun 1975).

Appendix: En modell för utbildningsval

Syftet med denna appendix är att formulera en enkel modell som vi kan använda för att diskutera utformningen av utbildningspolitiken. Modellen, som i flera avseenden liknar modellen i Cunha m.fl. (2006), kommer att vara mycket stiliserad; likväl fångar den många nyckelfrågor i den utbildningspolitiska diskussionen.

Vi tänker oss att barnens humankapital ackumuleras under två perioder; i period 3 träder barnen in på arbetsmarknaden. Barnen kan skilja sig åt med avseende på förmåga (α) och vilka ekonomiska resurser som deras föräldrar har till förfogande (Y). Humankapitalet för ett barn med förmågan α och familjeinkomsten Y ges av

$$h_{\alpha Y} = \alpha[\gamma I_1^\varphi + (1-\gamma)I_2^\varphi]^\rho/\varphi$$

där I_1 och I_2 betecknar utbildningsinvesteringar i respektive period, γ är ett mått på hur effektiva investeringar är i period 1 i förhållande till period 2. Om de kunskaper som förvärvas i period 1 är en nödvändig förutsättning för kunskapsinhämtande i period 2 skulle vi förvänta oss att γ är förhållandevis stort. φ mäter hur lätt det är att substituera utbildningsinvesteringar mellan perioder. Ju högre φ är desto lättare är det att kompensera för låga investeringar i period 1 med investeringar i period 2. Ett extremfall är när $\varphi=1$ då investeringar i de båda perioderna är perfekt substituerbara. Detta extremfall verkar dock inte realistiskt; typiskt kommer vi att anta att $\varphi \leq 0$. Skalavkastningen mäts av ρ . Denna parameter anger med andra ord hur den totala mängden utbildningsinvesteringar $I = I_1 + I_2$ påverkar humankapitalet. Vi kommer att anta att $\rho \leq 1$, vilket innebär att skalavkastningen är avtagande.

De ekonomiska resurser som finns till förfogande för utbildningsinvesteringar ges av

$$Y = Y_1 + \frac{Y_2}{1+r} = I_1 + \frac{I_2}{1+r}$$

Denna budgetrestriktion är skriven i nuvärdestermerna. Detta innebär i sin tur att vi smugit in ett viktigt antagande, nämligen att föräldrar fritt kan låna och spara för att fördela om utbildningsinvesteringar över tid. Vi antar också att

barnen inte kan belåna sitt framtida humankapital för att göra utbildningsinvesteringar. Det senare antagandet innebär att marknadssjämvikten inte kommer att vara effektiv vilket vi illustrerar nedan.

Vi bortser alltså från att föräldrar kan möta restriktioner på kapitalmarknaden; om föräldrar inte kan låna eller spara skulle investeringsproblemet ha den triviala lösningen $I_t = Y_t, t = 1, 2$. Notera dock att antagandet att familjer fritt kan låna och spara till en given ränta, r , inte är så restriktivt som det kan tyckas vid en första anblick; man kan gott tänka sig att familjer möter olika räntesatser – de som har svårare att få lån kan bara få dem om de betalar en högre ränta.

Syftet är nu att beskriva hur familjer som agerar på egen hand skulle investera i sina barn och sedan jämföra hur det offentliga skulle vilja fördela utbildningsinvesteringar över tid och mellan familjer. Vi kommer först att studera dessa beslut ur ett ex ante perspektiv, vilket betyder att det handlar om att lägga upp en optimal plan utan full kännedom om förmågor och inkomster. Vi antar att både familjer och det offentliga maximerar inkomspotentialen (humankapitalet) vid det framtida arbetsmarknadsinträdet. Inkomspotentialen antas vara proportionell mot humankapitalet; arbetsmarknaden fungerar med andra ord perfekt.

Följande ekvationer beskriver en optimal (privat) plan för en given familj

$$\frac{I_1}{I_2} = \left[\frac{\gamma}{(1-\gamma)(1+r)} \right]^{\frac{1}{1-\varphi}} \equiv \delta$$

$$I = \frac{1+\delta}{(1+r)^{-1} + \delta} Y$$

Dessa ekvationer innebär att fördelningen av utbildningsinvesteringar över de två perioderna är oberoende av familjeinkomsten. De styrs bara av humankapitalproduktionens ”teknologi” (γ, φ) och diskonteringsräntan (r). Under rimliga förutsättningar kommer $(I_1/I_2) > 1$, vilket följer om $\gamma > (1+r)/(2+r)$. Givet definitionen av γ är det naturligt att anta att $\gamma > 1/2$. Det är också lätt att övertyga sig om att ju högre diskonteringsräntan är, desto lägre är I_1/I_2 . Familjer som begränsas av restriktioner på lånemarknaden kommer alltså att välja att allokera en större mängd investeringar till period 2.

Den totala mängden insatta resurser (I) kommer att variera med familjens livstidsinkomst (Y). Om vi jämför två barn med samma förmåga (α) och samma diskonteringsränta (r) men med olika familjeinkomst $Y_H > Y_L$, får vi

$$\frac{I_{aH}}{I_{aL}} = \frac{Y_H}{Y_L} > 1 \Rightarrow \frac{h_{aH}}{h_{aL}} > 1 \Rightarrow \frac{\partial h_{aH} / \partial I_{aH}}{\partial h_{aL} / \partial I_{aL}} < 1$$

Mindre resurser kommer alltså att investeras i barn med lägre familjeinkomst, även då deras förmågor är desamma. I själva verket spelar inte barnens förmåga någon roll för hur mycket resurser som investeras i dem, bara familjeinkomsten är relevant.

Låt oss nu istället fundera på hur det offentliga skulle designa en optimal plan för olika familjer. För "staten" kommer inte den enskilda familjens inkomst, eller den diskonteringsränta som familjen möter, att vara relevant; för staten är den enda relevanta aspekten individens förmågor.

I en optimal plan för staten ska utbildningsinvesteringar allokeras över tid enligt

$$\left(\frac{I_1}{I_2} \right)_s = \left[\frac{\gamma}{(1-\gamma)(1+r_s)} \right]^{\frac{1}{1-\phi}} \equiv \delta_s$$

där r_s är den ränta som staten möter. Den samhällsekonomiskt optimala allokeringen sammanfaller alltså med den privatekonomiska om $r = r_s$. Alla familjer som har en högre diskonteringsränta än r_s investerar för lite i period 1 relativt period 2.

Låt oss nu undersöka samma fråga som ovan: Hur ser den samhällsekonomiskt optimala lösningen ut för barn med samma förmåga, men olika familjeinkomst? Svaret är att den inte ska skilja sig åt – lika mycket resurser ska investeras i barn med samma förmåga. Villkoret för en samhällsekonomiskt optimal allokering av resurser över barn är

$$\frac{\partial h_{\alpha Y} / \partial I_{\alpha Y}}{\partial h_{\alpha' Y'} / \partial I_{\alpha' Y'}} = 1$$

En omskrivning av detta villkor ger

$$\frac{I_{\alpha H}}{I_{\alpha' L}} = \frac{I_{\alpha}}{I_{\alpha'}} = \left[\frac{\alpha}{\alpha'} \right]^{\frac{1}{1-\rho}}$$

Från detta villkor följer att barn med högre förmåga ska ges mer resurser – om $\alpha > \alpha' \Rightarrow I_{\alpha} > I_{\alpha'}$ – och att barn med samma förmåga ska ges lika stora resurser.

Vad kan man då lära sig av detta? En lärdom är att en samhällsekonomiskt optimal utbildningspolitik kommer att vara omfördelande i *förhållande till marknadsjämvikten*. Detta följer så länge som förmågor och familjeinkomster inte är perfekt positivt korrelerade. Om det inte finns något samband alls mellan familjeinkomsten och barns förmågor, ska utbildningspolitiken medföra att lika mycket investeras i alla barn oberoende av familjeinkomst. Det torde vara omöjligt att implementera en sådan lösning. Skälet är att staten aldrig kan förhindra familjer att investera mer än vad som är samhällsekonomiskt optimalt.

En fråga som rönt stor uppmärksamhet i litteraturen är betydelsen av kreditrestriktioner. Frågan är om, och i så fall när, de är relevanta. I den modell som vi har formulerat här, medför dessa restriktioner framförallt att familjer investerar för lite i den första perioden. Detta ligger i linje med resultat i t.ex. Carneiro och Heckman (2002).

Vi har betonat att implikationerna från denna modell ska tolkas som en ”plan”; de gäller under förutsättning att man (separat) kan identifiera barns förmågor och familjers ekonomiska förutsättningar. Men planer slår ofta fel. Den relevanta frågan är vad man ska göra om planen slog fel: Vad är marginalavkastningen på investeringar i period 2 om man investerat för litet i period 1? Denna fråga har genererat en hel del kontrovers i litteraturen; se t.ex. debatten mellan Alan Krueger och James Heckman i Friedman (2002). Ibland har modellen som presenteras här tagits som intäkt för att marginalavkastningen på sena investeringar är låg och därför bör investeringar huvudsakligen ske tidigt i livet. Detta behöver inte vara fallet om planen slagit fel. Marginalavkastningen på sena investeringar kan vara hög om man investerat alltför lite tidigt i livet.

För att illustrera denna poäng beräknar vi den optimala ex post allokeringen givet att man investerat för litet i period 1. Vi gör detta för ett specialfall då $\varphi = 0$; humankapitalet produceras då enligt en s.k. Cobb-Douglas teknologi. Vi

tänker oss att staten kan allokera om resurser marginellt från individer (med samma förmåga) vars investeringar i period 1 var optimala till personer där man underinvesterat i period 1. Om vi löser detta problem får vi följande uttryck

$$\ln\left(\frac{I_2}{I_2^*}\right) = \frac{\rho\gamma}{1-\rho(1-\gamma)} \ln\left(\frac{\bar{I}_1}{I_1^*}\right)$$

Detta uttryck visar att om investeringar i period 1 var optimala ska investeringarna också följa den optimala banan. Vi ser också att avtagande skalavkastning är avgörande för att någon ”kompensation” i period 2 ska komma tillstånd; om $\rho = 1$ underinvesterar man lika mycket i period 2 som man gjorde i period 1. Till sist så är det värt att notera att det aldrig är optimalt att ”återställa” underskottet i period 1 – $(I_2/I_2^*) < 1$. I Figur 3.3 i huvudtexten visas beräkningar av den optimala ex post investeringen givet olika värden på ρ , γ , och $(\bar{I}_1/I_1^*)$. Vi har här satt $\rho = \gamma = 0.7$.

Referenser

- Björklund, A., P-A. Edin, P. Fredriksson och A. Krueger (2003), *Den svenska skolan – effektiv och jämlik?* Stockholm: SNS Förlag.
- Björklund, A. och M. Lindahl (2005), *Utbildning och ekonomisk utveckling – Vad visar den empiriska forskningen om orsakssambanden?*, Rapport till ESS 2005:4. Stockholm: Fritzes.
- Carneiro, P. och J. Heckman (2002), The evidence on credit constraints in post-secondary schooling, *Economic Journal*, 112, 705–734.
- Cuhna, F., J. Heckman, L. Lochner och D. Masterov (2006), Interpreting the evidence on life cycle skill formation, i E Hanushek och F Welch (red.), *Handbook of the Economics of Education*. Amsterdam: North-Holland.
- Edin, P-A. och B. Holmlund (1995), The Swedish wage structure: The rise and fall of solidarity wage policy? i R. Freeman and L. Katz (red.) *Differences and Changes in Wage Structures*. Chicago: University of Chicago Press.
- Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation, *Journal of Political Economy*, 81, 521–565.
- Fredriksson, P. och B. Holmlund (2001), Optimal unemployment insurance in search equilibrium, *Journal of Labor Economics*, 19, 370–399.
- Friedman, B. (red) (2002), *Inequality in America. What Role for Human Capital Policies?* Cambridge MA: MIT Press.
- Friedman, M. (1962), *Capitalism and Freedom*. Chicago: University of Chicago Press.
- Imrohoroglu, A., A. Merlo och P. Rupert (2004), What accounts for the decline in crime? *International Economic Review*, 45, 707–729.
- Jacobs, B. (2005), Optimal income taxation with endogenous human capital, *Journal of Public Economic Theory*, 7, 295–315.
- Krueger, A. och M. Lindahl (2001), Education and growth: Why and for whom?, *Journal of Economic Literature*, 39, 1101–1136.
- Lleras-Muney, A. (2005), The relationship between education and adult mortality in the United States, *Review of Economic Studies*, 72, 189–221.

- Nelson, R. och E. Phelps (1966), Investment in humans, technological diffusion, and economic growth, *American Economic Review*, 56, 69–75.
- Okun, A. (1975), *Equality and Efficiency: The Big Tradeoff*. Washington DC: Brookings Institution.
- Persson, T. och G. Tabellini (1994), Is inequality harmful for growth?, *American Economic Review*, 84, 600–621.
- Romer, P. (1990), Endogenous technological change, *Journal of Political Economy*, 98, 71–102.
- Spence, M. (1974), *Market Signaling: Informational Transfer in Hiring and Related Processes*. Cambridge MA: Harvard University Press.
- Warner, J. och S. Pleeter (2001), The personal discount rate: Evidence from military downsizing programs, *American Economic Review*, 91, 33–53.

4 Avkastningen på utbildning

Avkastningen på utbildning är ett centralt begrepp, särskilt för ekonomer men även inom andra discipliner. Måttet sammanfattar utbildningens intäkter och kostnader på ett effektivt sätt. Likväl finns ett antal olika avkastningsbegrepp som cirkulerar i litteraturen. Man kan tala om utbildningens *lönepremie*, dess *inkomstpremie*, eller dess *internränta*. Man skiljer även på den *privatekonomiska* och den *samhällsekonomiska* avkastningen. Dessa begrepp är nära relaterade, men det finns såklart skillnader dem emellan. I avsnitt 4.1 preciserar vi dessa begrepp och förklarar hur de hänger ihop.

I avsnitt 4.2 redovisar vi skattningar av hur avkastningen varierar över länder och i avsnitt 4.3 hur avkastningen varierar över tid i Sverige. Vi fokuserar då på skattningar av lönepremien, men redovisar även försök att beräkna internräntor utifrån ett privatekonomiskt och samhällsekonomiskt perspektiv.

4.1 Begrepp och frågeställningar

I kapitel 3 presenterade vi en modell där utbildningsval betraktas som investeringar. Härav följer också att vi vill veta vilken avkastningen är på olika utbildningsinvesteringar. Det finns efterfrågan på sådan kunskap både från politiskt håll – politiker vill veta vad man får ut av utbildningssatsningar av olika slag – och från forskare som vill förstå hur utbildningsval påverkar (exempelvis) arbetsmarknaden på olika sätt. Det finns numera en omfattande litteratur som på olika sätt belyser avkastningen på utbildningen. Innan vi går igenom denna litteratur vill vi reda ut en del begrepp. Framför allt vill vi betona att det finns många aspekter på avkastningen på utbildning och därmed olika avkastningsbegrepp.

En stor del den empiriska litteraturen – särskilt den ekonomiska men också en del av den sociologiska litteraturen – har haft en mycket enkel statistisk modell som utgångspunkt. Denna modell går under namnet Mincer-ekvationen:

$$\ln W_i = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \beta_4 X_i + \varepsilon_i \quad (1)$$

där $\ln W_i$ står för logaritmen av lön per arbetat timme för individ i , S_i för antalet utbildningsår, E_i för arbetslivserfarenhet (i år) och E_i^2 arbetslivserfarenhet i kvadrat, X_i representerar en uppsättning av kontrollvariabler (t.ex. kön och invandrarbakgrund) som forskaren önskar beakta för att få fram den

rena kausala effekten av utbildning, ε_i fångar upp alla andra faktorer som påverkar individens lön.

Låt säga att vi har en skattning av koefficienten för utbildning, β_1 , från ett datamaterial som utgör ett tvärsnitt av den svenska arbetsmarknaden. Eftersom den beroende variabeln är logaritmerad kan β_1 approximativt tolkas som den procentuella effekten på lönen av ytterligare ett utbildningsår. Det är naturligt att kalla detta *lönepremien* för ett extra utbildningsår på den svenska arbetsmarknaden.

Denna koefficient kan under vissa antaganden tolkas som *internräntan* på ytterligare ett utbildningsår. Detta är ett mer renodlat avkastningsmått och anger den räntesats med vilken framtida intäkter och kostnader kan diskonteras för att nuvärdet av framtida intäkter respektive kostnader ska vara lika stora. Sådana avkastningsmått är värdefulla eftersom de kan beräknas även för andra investeringsprojekt än utbildning. Det går således att jämföra utbildningens internränta med motsvarande avkastningsmått för investeringar i byggnader, vägar med mera.

Det ska dock understrykas att den skattade utbildningskoefficienten bara kan tolkas som en internränta under vissa förutsättningar. För det första antas de enda kostnaderna vara studenternas uteblivna inkomster under studietiden. Detta kan vara en rimlig approximation för den enskilde individen och därmed för den *privatekonomiska avkastningen*, men i en beräkning av den *samhällsekonomiska lönsamheten* måste man också beakta kostnaderna, i form av lärare, lokaler och utrustning, för att bedriva undervisningen. För det andra bör man även i en strikt privatekonomisk kalkyl beakta inkomstskatter (som om de är progressiva reducerar individens avkastning av utbildningen) och studiebidrag (som entydigt höjer utbildningsavkastningen).¹⁷ För det tredje beaktar inte ekvation (1) att utbildningen kan ha kausala effekter på arbetstiden, vilket påverkar både den privat- och samhällsekonomiska kalkylen. Slutligen gäller tolkningen av β_1 som en internränta bara under ett antal tekniska antaganden som vi inte fördjupar oss i här.¹⁸

Det finns en klar samvariation mellan arbetstider och utbildningslängd. Högre utbildade arbetar både längre tider per vecka, större delar av året och pensionerar sig senare. Det svåra är dock att avgöra vilken del av variationen i

¹⁷ I länder med studieavgifter skall sådana beaktas och de höjer självfallet individens (men inte samhällets) kostnader för utbildningen.

¹⁸ Björklund och Kjellström (2002) visar dock att rimliga avvikelser från dessa antaganden inte betyder så mycket.

arbetstid som är en kausal effekt av utbildning. Karriärinriktade personer har sannolikt längre utbildning och jobbar fler timmar per vecka. Men detta är en effekt av dessa personers inneboende egenskaper snarare än en effekt av utbildning. Samtidigt är det svårt att komma ifrån att utbildning rimligen har en effekt på individers sysselsättningsmöjligheter. Om man tror att den senare effekten huvudsakligen driver variationen i arbetstid är det naturligt att använda den logaritmerade årsarbetsinkomsten som den beroende variabeln i stället för den logaritmerade timlönen; skattningen av koefficienten på utbildningslängd i en sådan specifikation för benämns lämpligen utbildningens *inkomstpremie*.

Konventionen, som är starkt influerad av studier på amerikanska data, är dock att använda timlönen som den beroende variabeln. Vi kommer huvudsakligen att följa den konvention. Samtidigt vill vi betona att det inte finns något rätt eller fel här, så länge som man talar om exakt vad det är för ekvation som man skattat.

Ytterligare en precisering följer av att ekvation (1) normalt är skattad från ett tvärsnitt av arbetande personer vid en viss tidpunkt. Det betyder att en skattningen avspeglar hur *tidigare genomförd utbildning* belönas på arbetsmarknaden *vid observationstidpunkten*. Men utbildningval handlar om framtida investeringsbedömningar. Man bör således skilja mellan *realiserad avkastning* på tidigare utbildning och *förväntad avkastning* på framtida utbildning. Ett exempel tydliggör distinktionen mellan dessa två begrepp ytterligare. Om man skattar Mincer-ekvationen på data från slutet av 1960-talet är den skattade avkastningen tämligen hög (se avsnitt 4.3). Men de som genomförde sin utbildning vid denna tid fick se sin realiserade avkastning falla på grund av fallande realiserade lönepremier under de efterföljande åren. I princip är det den förväntade avkastningen som är relevant för att förstå utbildningsbeslut. Samtidigt påverkar rimligen den realiserade avkastningen presumtiva studenters förväntningar om den realiserade avkastningen.

Man kan också fråga sig hur man rent teoretiskt kan försvara en så enkel modell som ekvation (1). Faktum är att den har en starkare teoretisk förankring än vad man kanske kan tro. Man kan visa att en stiliserad modell över arbetsmarknaden och utbildningsbeslut, på lång sikt, ger upphov till ett samband mellan löner och utbildningslängd (och arbetslivserfarenhet) av just den typ som ekvationen anger.¹⁹ I modellen skiljer sig yrken åt enbart genom att de kräver olika lång utbildning. Ett linjärt samband med logaritmen på lönen och

¹⁹ Se Willis (1986) för den mest utvecklade modellen som i långsiktig jämvikt ger upphov till detta samband mellan lön och utbildningens längd.

utbildningsår är liktydigt med att alla yrken ger samma livsinkomster. Det är således en modell med löneskillnader som kompenserar för skillnader i utbildningens längd precis så mycket att den diskonterade livsinkomsten blir densamma i alla yrken.

Även om modellen är kraftigt stiliserad pekar den på vissa starka mekanismer som på sikt leder till en lönestruktur som i ekvation (1). Den kan därför tjäna som en första utgångspunkt när man ska undersöka sambandet mellan löner och utbildning. Behovet av ett enkelt samband är dessutom särskilt starkt när man med ett mindre datamaterial önskar komprimera sambandet mellan löner och utbildning. De mikrodata som fanns tillgängliga när humankapitalteorin etablerades för cirka 50 år sedan kom ofta från små surveyundersökningar. Ekvationen är dessutom lätt att tolka, varför vi kommer att använda den i kommande kapitel då vi diskuterar ett antal grundläggande metodproblem som skattningen av lönepremier är förknippad med.

Utbildningskoefficienten i ekvation (1) kan således ses som ett sammansatt mått på hur utbildning i genomsnitt belönas av arbetsmarknaden vid en given tidpunkt. Men den ger ingen vägledning för hur en utbildningspolitiker bör prioritera mellan olika delar av utbildningssystemet. För detta krävs uppenbarligen studier av mer detaljerade samband. En första rättfram utvidgning är att göra funktionsformen mer flexibel så att effekten av ett ytterligare utbildningsår tillåts variera. En mycket flexibel funktionsform är att använda dummyvariabler för varje enskilt utbildningsår så att varje utbildningslängd (mätt i antalet år) har sin egen lönepremie. En sådan ansats är särskilt informativ för ett land som USA där de olika delarna av utbildningssystemet följer en tydlig sekvens utan några parallella delar.

Litteraturen är numera tämligen rik på studier som använder en flexiblere funktionsform. De flesta explicita tester av den linjära funktionsformen mot en mer flexibel funktionsform ger resultatet att den linjära formen statistiskt förkastas till förmån för den mer flexibla.²⁰ Avvikelserna är däremot inte alltid särskilt stora i ekonomiska termer och är ofta inte systematiska över tid. Därför framstår ekvation (1) som ett anmärkningsvärt stabilt allmänt samband mellan utbildningslängd och lön. Däremot har lönepremien visat sig variera mycket över tiden och skilja sig mellan länder.

Det svenska utbildningssystemet har parallella delar. Efter en inledande obligatorisk skolgång på nio år, delas de som läser vidare upp i praktiska och

²⁰ Se Edin och Holmlund (1995) och Björklund och Kjellström (2002) för sådana tester på svenska data.

teoretiska linjer. Likaså sker en uppdelning i yrkesinriktade och teoretiska studier efter fullföljt gymnasium. Detta gör Mincer-ekvationen inte ger så nyanserad information som de utbildningspolitiska avvägningarna i Sverige ställer krav på.²¹

En önskedröm för en utbildningsminister vore att ha en komplett lista över lönepremier för varje enskild utbildning i landet. Eftersom lönepremien för en given utbildning dessutom kan förväntas bero på elevernas förutsättningar skulle listan också visa hur lönepremien varierar beroende på dessa förutsättningar. Detta är dock en helt orealistisk önskedröm, bl.a. beroende på ett antal metodproblem förknippade med att skatta avkastningen på utbildning (se kapitel 5). Vi vill dock formulera ett antal olika utbildningspolitiska avvägningsproblem som ställer krav på olika typer av information i form av lönepremier. Längre fram i rapporten kommer vi att se att det finns studier som belyser vissa av dessa problem men inte alla. I det avslutande kapitel 11 diskuterar vi möjligheterna att med svenska data och befintliga metoder förbättra det politiska beslutsunderlaget med vidare forskning.

Grunden för alla utbildningssystem är en *obligatorisk skola* av en längd som varierat över tiden och som varierar mellan länder. En central fråga är hur lång den obligatoriska utbildningen bör vara. Vad vore avkastningen på att förlänga skolplikten, antingen genom att den påbörjas tidigare eller förlängs med något år? En annan frågeställning är vilka effekterna vore av att förlägga en obligatorisk utbildning av given längd till en tidigare eller senare ålder.

Utöver den obligatoriska utbildningen finns *frivillig utbildning* både *före* – främst i form av förskola – och *efter* i form av olika påbyggnadsutbildningar. En avgörande fråga är hur omfattande dessa bör vara och vilken den ekonomiska lönsamheten är av att på marginalen öka eller minska omfattningen på dessa utbildningar. En relaterad fråga är hur attraktiva de frivilliga utbildningarna bör vara för den enskilde (och dennes föräldrar) i form av avgifter och bidrag. Inom alla utbildningsformer kan man också ställa frågan om vilka anordnare som är effektivast i meningen att de generar störst utbyte för samhället. Det kan gälla kommunala skolor i jämförelse med friskolor; men motsvarande frågeställning kan också tillämpas på olika lärosäten på högskolenivå.

Vi kommer att gå igenom vad man kan säga om effekterna av utbyggnaden av den svenska förskolan i kapitel 6. I kapitel 9 och 10 behandlar vi gymnasiet och högskolan. Och i kapitel 8 redogör vi för vad man kan säga om effekten av konkurrens mellan olika typer av utbildningsanordnare.

²¹ Se Card (1999) för vidare diskussion.

4.2 Internationell jämförelse

I likhet med övriga nordiska länder har Sverige en internationellt sett mycket sammanpressad lönestruktur. Löneskillnaderna mellan olika grupper, t.ex. olika utbildningsgrupper, är alltså små. Små löneskillnader mellan utbildningsgrupper antyder också att den privatekonomiska avkastningen på utbildningsinvesteringar är förhållandevis låg. Detta följer dock inte med nödvändighet eftersom skatter, eventuella utbildningsavgifter samt studiemedel också har effekt på den privatekonomiska kalkylen.²²

Som vi poängterat i avsnitt 4.1 finns det många skäl till att den samhällsekonomiska avkastningen kan skilja sig från den privatekonomiska avkastningen. De samhällsekonomiska intäkterna involverar aspekter som det i många fall är svåra att mäta, t.ex. effekter på hälsa och sannolikheten att valmanskåren fattar välinformerade beslut (se kapitel 3). Frånvaron av fullödiga internationella jämförelser av den samhällsekonomiska avkastningen är därför inte förvånande. Men även någorlunda heltäckande jämförelser av den privatekonomiska avkastningen lyser i stort med sin frånvaro. Det finns naturligtvis en uppsjö av skattningar av löneekvationer eller inkomstekvationer. Men dessa är inte alltid fullt jämförbara pga. att data skiljer sig åt och att skattningarna är gjorda på data från olika tidsperioder. Eftersom det finns internationella trender i hur lönespridningen varierar över tid, leder en jämförelse av utbildningens lönepremie skattad på data från 1970 med en lönepremie skattad på data från 2000 obönhörligen fel.

Typiskt mäts den privatekonomiska avkastningen i termer av löner (åtminstone i forskningslitteraturen). Skälet till att välja löner snarare än årsarbetsinkomster är att löneskattningarna inte direkt avspeglar skillnader i arbetad tid. Det underliggande argumentet är att skillnader i arbetad tid över utbildningsgrupper avspeglar individens val av arbetsutbud snarare än att de orsakas av utbildningsnivån. Men hur ska man se på skillnader i årsarbetsinkomster som har att göra med skillnader i arbetslöshet? Det är inte orimligt att argumentera för att skillnader i sysselsättningsmöjligheter orsakas av utbildning och i sådant fall borde dessa skillnader också beaktas i den privatekonomiska kalkylen.

²² Edin och Holmlund (1995) visar att skatter och studiemedel påverkar den privatekonomiska kalkylen mycket. De beräknar internräntan för universitetsutbildning utan hänsyn till skatter och studiemedel till 6 procent 1991. När de även beaktar skatter och studiemedel stiger den till 11 procent. Björklund m.fl. (2003) uppdaterade kalkylerna i Edin och Holmlund. De finner att internräntan med hänsyn taget till skatter och studiemedel uppgår till ca. 10 procent 2000. Den marginella minskningen har att göra med att studiemedlen är mindre förmånliga 2000 än 1991.

I detta avsnitt presenterar vi skattningar av utbildningens löne- och inkomst-premier som gjorts för olika länder. Skattningarna av lönepremierna kommer från enskilda studier och har sammanställts av Psacharopoulos och Patrinos (2004). Från denna studie väljer vi de skattningar som för oss synes passera ett minimikrav på jämförbarhet. Skattningar av utbildningens inkomstpremier kommer från OECD-publikationen *Education at a Glance* som varje år presenterar internationella jämförelser av avkastningen på utbildning. I denna publikation finns även skattningar som har ambitionen att fånga skattesystemets utformning och skattningar som, under vissa antaganden, kan betraktas som approximationer av den samhällsekonomiska avkastningen.

Genomgående är det viktigt att komma ihåg att de skattningar som presenteras för varje enskilt land inte gör anspråk på att fånga de intäkter och kostnader som *orsakas* av ett års ytterligare utbildning (vi diskuterar orsaks-sambandet ytterligare i kapitel 5). De gör inte heller anspråk på att fånga alla monetära och icke-monetära aspekter av olika utbildningsval. Även om sambanden är beskrivande är det inte orimligt att anta att den eventuella snedvridning som finns är likartad över länder (tid), vilket gör att jämförelser över länder (tid) ändå är meningsfulla.

Tabell 4.1 presenterar olika skattningar av Mincers löneekvation, dvs. en regression av den typ som visas i ekvation (1).

Tabell 4.1. Utbildningens lönepremie, procent

Land	År	Skattning (%)
Belgien	--	--
Danmark	1990	4,5
Finland	1991	8,8
Frankrike	--	--
Italien	1987	2,7
Kanada	1989	8,9
Nederländerna	1989	7,3
Norge	1991	5,4
Schweiz	1990	7,5
Spanien	1991	7,2
Storbritannien	1987	6,8
Sverige	1991	5,0
Tyskland	1988	7,7
USA	1991–95	10,0

Källa: Psacharopoulos och Patrinos (2004), Tabell A4.

Anm. Sammanställningen bygger på en rad ursprungskällor som redovisas i Psacharopoulos och Patrinos. Flertalet av skattningarna (Danmark, Finland, Italien, Nederländerna, Norge, Schweiz och Sverige) kommer ursprungligen från en bok med en sammanställning av avkastningen på humankapital i Europa; se Asplund och Pereira (1999). Skattningen för USA är baserad på data med information om löner och utbildning för tvillingpar.

I möjligaste mån redovisar vi skattningar som är gjorda med enklast möjliga och likartade metoder (dvs. vanlig linjär regression snarare än sofistikerade instrumentalvariabelansatser) och med data från ungefär samma tidpunkt, ca. 1990.

Tabell 4.1 visar att utbildningens lönepremie är förhållandevis låg i Sverige. I likhet med Norge och Danmark ligger lönepremien per utbildningsår på ungefär 5 %. ²³ Denna premie är förvisso högre än i Italien men avsevärt lägre än i USA (även om USA är den ”udda fågeln”, både i termer av nivå på avkastningen och i termer av skattningsmetod).

Sammanställningen i Psacharopoulos och Patrinos (2004) innehåller skattningar för en stor uppsättning länder som skiljer sig åt i termer av ekonomisk utveckling och genomsnittlig utbildningsnivå. Avkastningen på utbildning tenderar att vara högre i utvecklingsländer och i länder där utbildningsnivån är låg.

²³ Det är värt att notera att lönepremien för Sverige är baserad på data där man direkt frågat individer om deras utbildningslängd. Med sådana data skattar man i allmänhet en lägre avkastning på utbildning än med data där man utnyttjat registerdata med information om den typiska utbildningslängden för olika utbildningsnivåer; se kapitel 5. Detta är ett exempel på de jämförbarhetsproblem som man ofrånkomligen stöter på vid internationella jämförelser av denna typ.

Den verkar också vara marginellt högre för kvinnor än för män och högre för primär utbildning än för högre utbildning.

De uppgifter som ligger till grund för sammanställningen i Psacharopoulos och Patrinos (2004) är förhållandevis gamla. De uppgifter som publiceras av OECD ger en något aktuellare bild. Vi är då hänvisade till årsinkomster, vilket antagligen försvårar länderjämförelserna.

Tabell 4.2 sammanställer information från OECD (2005, 2006, 2007) vad gäller avkastningen på högskoleutbildning. OECD redovisar dels relativa skillnader i årsarbetsinkomster och dels regelrätta internräntekalkyler där hänsyn tas till skattesystemen och de direkta utbildningskostnader som faller på individen (dock tar de inte hänsyn till studiemedel). Man gör även ett försök att ta hänsyn till hur ett års ytterligare utbildning påverkar den offentliga budgeten. Intäkterna utgörs här av de ökade skatteintäkter som följer av att individer med högre utbildning har högre arbetsinkomster. Den stora kostnadsposten är offentliga utgifter för att finansiera utbildningssystemet (t.ex. kostnader för lärarlöner). I ett proportionellt skattesystem avspeglar den fiskala intäkten (i princip bara) bruttoskillnaden i årsarbetsinkomster mellan universitets- och gymnasieutbildade. Den huvudsakliga kostnaden utgörs av den offentliga sektorns kostnader för universitetsutbildning. Ett progressivt skattesystem kommer att öka den fiskala internräntan för givna utbildningskostnader vilket är ett problem om vi skulle vilja tolka denna ”budgetkalkyl” som den samhälls-ekonomiska avkastningen av utbildningsinvesteringar. Även i andra avseenden är budgetkalkylen ett långt steg från den samhälls-ekonomiska avkastningen på utbildningen då svårfångade aspekter som utbildningens effekter på hälsa inte beaktas i kalkylen. Likväl tillförs viss information i förhållande till den privata kalkylen eftersom budgetkalkylen också beaktar offentliga utgifter på utbildningssystemet.

Kolumn (1) i Tabell 4.2 redovisar en grov kalkyl som är baserad på årsarbetsinkomster för alla individer i åldern 25–64. För att konvertera den relativa inkomstskillnaden mellan universitets- och gymnasieutbildade till en årsvis avkastning har vi antagit att universitetsutbildningen i alla länder är 4 år; rimligheten i detta antagande varierar naturligtvis över länder.

Tabell 4.2. Avkastningen på högskoleutbildning 2003, internräntor (procent)
beräknade separat för män och kvinnor

	Privat avkastning				"Fiskal" avkastning	
	Arbetsinkomster (brutto)		Arbetsinkomster + skatter och direkta utbildningskostnader			
	(1)	(2)	(3)			
	Män	Kvinnor	Män	Kvinnor	Män	Kvinnor
Belgien	7,2	7,2	10,7	15,2	12,2	17,9
Danmark	7,6	6,2	8,3	8,1	7,8	6,9
Finland	12,5	9,9	16,7	16,0	13,6	11,3
Frankrike**	11,4	9,7	--	--	--	--
Italien**	16,3	7,6	--	--	--	--
Kanada	9,4	9,5	--	--	--	--
Nederländerna*	9,4	11,6	--	--	--	--
Norge	6,6	6,8	12,1	15,7	9,5	9,9
Schweiz	8,4	11,8	10,0	9,8	6,3	5,8
Spanien**	7,2	9,2	--	--	--	--
Storbritannien	10,9	15,8	16,8	19,6	13,7	16,1
Sverige	8,2	6,4	8,9	9,2	7,5	6,3
Tyskland	10,7	9,7	--	--	--	--
USA	17,2	15,3	14,3	13,1	14,1	13,0

Källa: OECD (2005, 2006, 2007).

Anm.: Siffrorna i kolumn (1) är beräknade utifrån relativa skillnaden i årsinkomster mellan universitets- och gymnasieutbildade 25-64 åringar. "Internräntan" är beräknad som internränta = $\exp(\ln(\text{relativ inkomstskillnad})/4)$. Kolumn (2) och (3) är baserade på antagandet att individen övergår direkt från gymnasie- till universitetsutbildning. När individen väl kommer ut på arbetsmarknaden möter de den inkomstprofil efter ålder som råder vid mättdpunkten.

* Data från 2002; ** data från 2004.

Variationer i årsarbetsinkomster avspeglar löneskillnader, skillnader i sysselsättningsmöjligheter, och skillnader i arbetsutbud. Variationen som skapas av skillnader i arbetad tid försvårar tolkningen av den internationella jämförelsen för kvinnor avsevärt. Det är tydligt att i länder där kvinnligt arbetskraftsdeltagande är lågt är det stora skillnader i arbetad tid mellan universitets- och gymnasieutbildade. Vi fokuserar därför på resultaten för män.

Grovt sett konfirmerar kolumn (1) bilden som ges av Tabell 4.1. Avkastningen är högst i USA; Sverige hamnar på den nedre halvan i denna jämförelse, tillsammans med länder som Danmark, Norge, Schweiz och Spanien.

Kolumn (2) presenterar resultaten från OECD:s lite mer sofistikerade kalkyler. Dessa kalkyler tar hänsyn till skillnader i utbildningslängd mellan länder och hur inkomsten varierar över olika åldrar. I tillägg beaktas skatter och direkta utbildningskostnader, men däremot inte studiemedel. Att studiemedelsystemet inte beaktas kommer att underskatta avkastningen mer i ett land som Sverige där studiemedlen är förhållandevis generösa.

Även med hänsyn taget till att det är färre länder som ingår faller Sverige i denna jämförelse. Internräntan i Sverige ligger runt 9 % och det är bara Danmark som har en lägre skattad avkastning på universitetsutbildning. Den fiskala kalkylen i kolumn (3) kommunicerar ungefär samma budskap. I flertalet länder har utbildningsinvesteringar en större positiv inverkan på den offentliga budgeten än i Sverige.

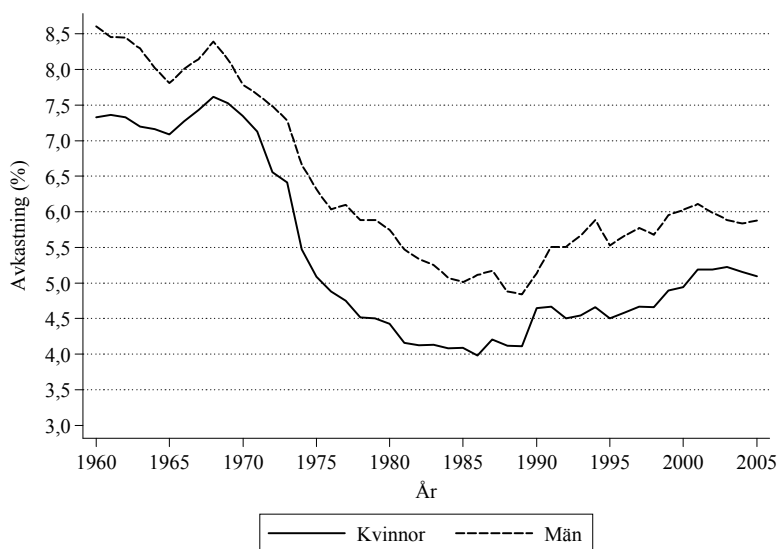
I detta sammanhang är det viktigt att notera att siffrorna i kolumn (2) och (3) bygger på antagandet att individen går direkt från gymnasie- till universitetsutbildning. OECD redovisar också internräntor under antagande att individen börjar studera vid 40 års ålder. Detta alternativa antagande har marginell betydelse för den privatekonomiska avkastningen, men har stor betydelse för budgetkalkylen. För många länder halveras den uppmätta fiskala avkastningen, vilket har att göra med att den offentliga utbildningskostnaden är densamma, oberoende av vilken ålder som investeringen görs, medan intäktssperioden halveras då individen väljer att börja universitetsstudier vid 40 års ålder snarare än direkt efter gymnasiet.

De internräntor som redovisas i Tabell 4.2 tyder på att avkastningen på universitetsutbildning är jämförelsevis låg i Sverige. Frågan är i vilken utsträckning vi ska ta detta budskap på allvar. Skattningar som är baserade på löner och förhållandevis jämförbara data kommunicerar i stort samma budskap som kolumn (1) för män: avkastningen på utbildning i Sverige ligger under genomsnittet. Som vi varit inne på tidigare är det vanskligt att tolka skillnaderna för kvinnor över länder. Och utan vidare insikt i hur OECD gjort sina internräntekalkyler är det svårt att förhålla sig till siffrorna i kolumn (2) och (3). Varför skulle t.ex. internräntan i Norge vara 3 procentenheter högre än i Sverige när bruttointkomstskillnaden mellan utbildningskategorier i Norge är lägre än i Sverige? Detta kräver skillnader i utbildningslängd, skattesystem, eller ålderinkomstprofilerna som det knappast finns stöd för i data. Internräntekalkylerna som presenteras av OECD bör nog tas med en del nypa salt.

4.3 En historisk jämförelse

Hur har avkastningen på utbildning i Sverige utvecklats över tid? För att sprida ljus över denna fråga presenterar vi i Figur 4.1 skattningar av utbildningens ”lönepremie” under perioden 1960–2005.

Figur 4.1. Utbildningens "lönepremie", procent, 1960–2005



Anm.: Skattningar baserade på LINDA – se Edin och Fredriksson (2000). I skattningarna har vi exkluderat individer vars årsarbetsinkomster understiger ett basbelopp. Dessutom har vi exkluderat individer vars inkomster är så låga att de tillhör den lägsta kvintilen av den resulterande inkomstfördelningen.

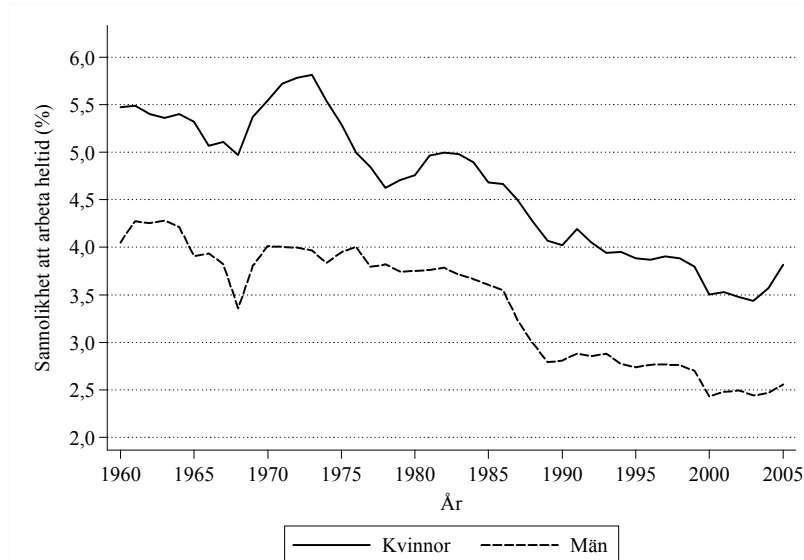
Underlaget för dessa skattningar är pensionspoängsdata för perioden 1960–67 och årsinkomstdata från 1968 och framåt. Eftersom vi vill att dessa skattningar ska motsvara en lönepremie har vi skurit bort den nedre delen i inkomstfördelningen. Det är värt att notera att skattningar för perioden 1960–67 inte har presenterats tidigare och att utvecklingen efter 1968 följer mer konventionella skattningar baserade på lön per timme i data från t.ex. Levnadsnivåundersökningen mycket väl.

Utbildningspremien följer i stort svängningarna i lönespridning över tid. I slutet av 1960-talet när lönespridningen var relativt hög var också utbildningspremien hög. Under de därpå följande 20 åren minskade utbildningspremien (och lönespridningen) avsevärt. I mitten av 1980-talet vände utvecklingen och sedan dess kan vi observera en kontinuerlig ökning i utbildningspremien.

När vi skattat dessa utbildningspremier har vi tagit bort individer vars årsarbetsinkomst är lägre än ett basbelopp. Därtill har vi tagit bort de lägsta 20

procenten i inkomstfördelningen för varje enskilt år.²⁴ Skälet till dessa begränsningar är som sagt att vi har ambitionen att skatta en ”lönepremie”. Men utbildningens effekter på sysselsättningsmöjligheter kan vara väl så intressanta. Därför visar vi i Figur 4.2 hur ett års ytterligare utbildning påverkar sannolikheten att ha inkomster som överstiger denna nedre inkomstgräns.

Figur 4.2. Utbildningens ”sysselsättningspremie”, procent, 1960–2005



Anm.: Skattningar baserade på LINDA. Figuren visar hur ett års ytterligare utbildning påverkar sannolikheten att ha en arbetsinkomst som överstiger den nedre gränsen som definieras i texten.

Figur 4.2 visar att den relativa sysselsättningsutvecklingen för dem som har ett års ytterligare utbildning har försämrats trendmässigt; mellan 1975 och 2005 minskar sysselsättningspremien med ungefär 1,5 procentenheter för både kvinnor och män.

²⁴ Skälet till att vi genomgående tar bort individer vars inkomster understiger ett basbelopp är att dessa inkomster inte finns med i pensionspoängsdata. Skälet till att lägga på ytterligare begränsningar är att Antelius och Björklund (2000) har visat att om man tar bort den lägsta kvintilen i arbetsinkomstfördelningen så får man väldigt likartade resultat om man skattar utbildningspremien på arbetsinkomster och regelrätta löner.

4.4 Sammanfattning och diskussion

Länderjämförelser tyder på att avkastningen på utbildning är relativt låg är Sverige. Den ligger under genomsnittet och är därmed i nivå med avkastningen i Danmark och Norge. Då den internationella jämförelsen är baserad på gamla data har vi också visat hur utbildningens lönepremie varierar över tid i Sverige. Lönepremien minskade kraftigt i Sverige mellan slutet av 1960-talet och mitten av 1980-talet. Därefter sker en trendmässig ökning. Våra skattningar tyder på att lönepremien ligger mellan 5 och 6 procent 2005.

Vad förklarar variationen i utbildningsavkastningen mellan länder och över tid? Det är naturligt att söka förklara förändringarna med variationer i efterfrågan och utbud på kvalifikationer. Särskilt utbudet torde vara relevant, i synnerhet på längre sikt.

Som vi har antytt ovan finns mönster i data som pekar på att variationer i utbudet är en viktig faktor. Avkastningen på utbildning är högre i utvecklingsländer vilket rimligen har att göra med att utbudet av kvalifikationer är lägre i dessa länder än i den industrialiserade världen. Ett annat exempel på utbudets betydelse är att avkastningen på utbildning i Sverige (och USA) föll kraftigt efter 1960-talets utbildningsexpansion.

Formella test av den enkla efterfråge- och utbudsmodellen ger dock något varierande resultat. Edin och Holmlund (1995) visade att den enkla modellen fungerar väl för att förklara utvecklingen i högskolelönepremien från slutet av 1960-talet till början av 1990-talet. Gustavsson (2006) uppdaterade analysen i Edin och Holmlund (1995) med data för perioden 1992–2001. Det är tydligt att mönstret under 1990-talet avviker från tidigare perioder. Under 1990-talet steg högskolelönepremien trots att det relativa utbudet av högskoleutbildade ökade avsevärt. Detta är inte förenligt med en modell där den relativa efterfrågan på högskoleutbildade är konstant. För att efterfråge- och utbudsmodellen ska kunna förklara 1990-talsutvecklingen krävs att den relativa efterfrågan på högskoleutbildade accelererade under 1990-talet och tillväxten i den relativa efterfrågan var kraftigare än ökningen i det relativa utbudet. Det faktum att universitetslönepremien ökat i många andra länder under 1990-talet är förenligt med denna hypotes. En acceleration i relativ efterfrågan, som drivs på av teknologiförskjutningar som gynnar högutbildade eller ökad handel, borde vara gemensam över länder och följaktligen borde vi se samma mönster i flera länder vilket också observeras i data.

En annan förklaring till att utbuds- och efterfrågemodellen fungerar sämre under 1990-talet är att lönebildningsinstitutionerna har förändrats. I Sverige kan

vi observera en trend mot en individualiserad lönebildning; s.k. sifferlösa avtal blev mer frekventa under 1990-talet, särskilt i tjänstemannakollektivet (se Fredriksson och Topel 2006). Allmänt sett är skillnader länder i lönebildningsinstitutioner en potentiellt viktig förklaring till variationen i lönepremien mellan länder (se Autor och Katz 1999).

Referenser

- Antelius, J. och A. Björklund (2000), How reliable are register data for studies of the return on schooling? An examination of Swedish data, *Scandinavian Journal of Educational Research*, 44, 341–355.
- Asplund, R. och P. Pereira (red.), *Returns to Human Capital in Europe*. Helsingfors: ETLA.
- Autor, D. och L. Katz (1999), Changes in the Wage Structure and Earnings Inequality, i O. Ashenfelter och D. Card (red.), *Handbook of Labor Economics*, vol. 3A, Amsterdam: North-Holland.
- Björklund, A., P-A. Edin, P. Fredriksson, och A. Krueger (2003), *Den svenska skolan – effektiv och jämlik?* Valfärdspolitiska rådets rapport 2003. Stockholm: SNS Förlag.
- Björklund, A. och C. Kjellström (2002), Estimating the returns to schooling in Sweden: How useful is the Mincer-equation? *Economics of Education Review*, 21, 195–210.
- Card, D. (1999), The causal effect of education on earnings, i Ashenfelter, O och D Card (red.), *Handbook for Labor Economics*, Vol. 3. Amsterdam: North-Holland.
- Edin, P-A. och P. Fredriksson (2000), LINDA – Longitudinal INDividual DATA for Sweden, Working Paper 2000:19, Nationalekonomiska institutionen, Uppsala universitet.
- Edin, P-A. och B. Holmlund (1995), The Swedish wage structure: The rise and fall of solidarity wage policy? I R. Freeman och L. Katz, *Differences and Changes in Wage Structures*. Chicago: University of Chicago Press.
- Fredriksson, P. och R. Topel (2006), Förändrade förutsättningar för svensk lönebildning, i R. Freeman, B. Swedenborg, och R. Topel (red.), *Att reformera välfärdsstaten – Amerikanskt perspektiv på den svenska välfärdsstaten*. Stockholm: SNS Förlag.
- Gustavsson, M. (2006), The evolution of the Swedish wage structure: New Evidence for 1992–2001, *Applied Economics Letters*, 13, 279–286.
- OECD (2005), *Education at a Glance*. Paris: Organization for Economic Cooperation and Development.

- OECD (2006), *Education at a Glance*, Paris: Organization for Economic Cooperation and Development.
- OECD (2007), *Education at a Glance*. Paris: Organization for Economic Cooperation and Development.
- Psacharopoulos, G. och H. Patrinos (2004), Returns to investment in education: A further update, *Education Economics*, 12, 111–134.
- Willis, R. (1986), Wage determinants: A survey and reinterpretation of human capital earnings functions, i O. Ashenfelter och R. Layard (red.), *Handbook of Labor Economics*, Vol 1. Amsterdam: North-Holland.

5 Effekter av utbildning på kunskaper och lön

Det finns stora löneskillnader mellan personer med olika lång utbildning. År 2006 tjänade exempelvis de som slutfört en högskoleutbildning i genomsnitt nästan 30 % mer än de med gymnasiekompetens. Frågan är vad dessa löneskillnader beror på. En möjlig förklaring är att personer som läst på högskolan har förvärvat produktiva kunskaper som värdesätts på arbetsmarknaden. Den högre lönen skulle därmed ha orsakats av den längre utbildningen. En annan förklaring, är att personer som läst på högskolan redan från början hade bättre produktiva egenskaper än de som inte läst vidare. Personer med en högskoleutbildning har exempelvis högre gymnasiebetyg, vilka värdesätts på arbetsmarknaden, än personer som inte läst vidare. Den högre lönen för högskoleutbildade skulle därmed, åtminstone delvis, ha orsakats av ursprungliga skillnader mellan människor med olika lång utbildning. Den empiriska forskningen om utbildningens löneeffekter försöker fastställa hur stor del av löneskillnaderna mellan individer som har orsakats av skillnader i utbildningslängd, och hur stor del som beror på andra skillnader mellan personer med olika lång utbildning.

I avsnitt 5.1 går vi igenom forskningen om avkastningen på utbildning i Sverige.²⁵ Särskilt fokus läggs på att beskriva de metoder som används för att fastställa den *kausala* effekten av utbildning, dvs. hur stor del av de observerade löneskillnaderna som har *orsakats* av skillnader i utbildningslängd. Vi diskuterar styrkor och svagheter med olika ansatser och presenterar vad vi anser vara de trovärdigaste skattningarna av avkastningen på utbildning i Sverige.

Det kan ta lång tid innan det är möjligt att mäta arbetsmarknadseffekterna av utbildningsinvesteringar. Vår utgångspunkt i den här översikten är att utbildning leder till produktiva kunskaper som i sin tur värdesätts på arbetsmarknaden (se Kapitel 3). I avsnitt 5.2 sammanfattar vi därför forskningen om sambandet mellan utbildningslängd och olika kunskaper. Även här ges företräde åt studier som försöker mäta *kausala* effekter av utbildning.

I avsnitt 5.3 belyser vi effekterna av kunskaper på arbetsmarknaden. Vi studerar betydelsen av såväl kognitiva som icke-kognitiva färdigheter och visar hur de värdesätts på arbetsmarknaden. Avsnitt 5.2 och 5.3 illustrerar också att kunskaper kan användas som intermediära (mellanliggande) utfall; effekterna av utbildning på kunskaper är med andra ord informativa om dess effekter på

²⁵ Det här avsnittet behandlar huvudsakligen effekten av utbildningsinvesteringar i genomsnitt. Effekter av utbildning på olika nivåer och med olika inriktning diskuteras i avsnitt 9.1 (gymnasieutbildning) och avsnitt 10.4 och 10.5 (högskoleutbildning).

arbetsmarknaden. Man kan således skaffa sig en god uppfattning om hur utbildningsåtgärder påverkar individers situation på arbetsmarknaden, redan innan de har passerat genom utbildningssystemet. Det är därmed också möjligt att förändra utbildningspolitiken innan det är "för sent".

5.1 Effekter av utbildningsinvesteringar på inkomster

Det finns en lång nationalekonomisk tradition av att mäta den ekonomiska avkastningen på utbildningsinvesteringar (se kapitel 4). Det uppmätta sambandet mellan lön och utbildning kan i de flesta fall inte tolkas som en kausal effekt av utbildning. Utbildningsval är inte slumpmässiga, utan bestäms av en rad faktorer, såsom familjebakgrund, förmåga och motivation. Problemet är att vi, dels inte vet exakt vilka faktorer som påverkar utbildningsbeslut, dels ofta saknar information om många faktorer som vi tror är betydelsefulla. Eftersom personer med längre utbildning troligen redan från början hade bättre produktiva egenskaper (såsom förmåga och familjebakgrund) än personer med kortare utbildning, är det troligt att det beskrivande sambandet mellan utbildning och lön *överskattar* den kausala effekten av utbildning.²⁶

Förutom svårigheten att avgöra hur mycket av löneskillnaderna som orsakats av skillnader i utbildningslängd, så tampas forskningen med ett annat problem: mätfel i uppgifterna om utbildningslängd. Mätfelens egenskaper varierar mellan olika källor. En del studier använder sig av svar på enkät- eller intervjuuppgifter, medan andra baseras på registeruppgifter. Vid enkät- eller intervjuundersökningar kan individer ha svårt att minnas exakt hur länge de gått i skolan, vilket (i bästa fall) ger upphov till slumpmässiga fel. Administrativa uppgifter kan också innehålla fel, alltifrån rena felkodningar till utebliven registrering av uppnådd utbildningsnivå. Även om mätfelet är helt slumpmässigt, så leder det till att sambandet mellan utbildning och lön *underskattas*. Ju större mätfel, desto svagare blir sambandet mellan utbildning och lön.²⁷

Ytterligare en omständighet att beakta är att avkastningen på utbildning kan skilja sig åt mellan individer. Det kan bero på skillnader i ursprungliga produktiva egenskaper – och därmed omfattningen av uteblivna inkomster under studietiden – men också på olika förutsättningar att tillgodogöra sig undervisningen i skolan. Om utbildningseffekten varierar mellan individer, kan

²⁶ Man brukar säga att sambandet mellan utbildning och lön lider av "ability bias", vilket är ett resultat av obeaktade skillnader i "förmågan att uppnå utfallet" (såsom förmågan att erhålla en hög lön) mellan personer med olika lång utbildning.

²⁷ I litteraturen används begreppet "attenuation bias", vilket innebär att slumpmässiga mätfel "tunnar ut" sambandet mellan utbildning och lön.

resultaten från studier som baseras på olika urval av personer också komma att variera. Jämförbarheten mellan studier försvåras ytterligare av att vanligtvis bara en del av utbildningsskillnaderna faktiskt används för att skatta utbildningseffekten. Det är därför viktigt att fråga sig ”för vem” det uppmätta sambandet mellan utbildning och lön gäller.

I litteraturen används tre olika metoder för att mäta den kausala effekten av utbildning på lönen:

- 1 Jämför lönerna för personer som har samma observerade egenskaper men olika lång utbildning.
- 2 Jämför lönerna för syskon (tvillingar) som har olika lång utbildning.
- 3 Jämför lönerna för personer som har olika lång utbildning på grund av institutionella förhållanden.

Gemensamt för alla tre ansatser är att de försöker isolera och mäta effekterna av skillnader i utbildningslängd som kan antas vara mer eller mindre slumpmässiga. Rimligheten i detta antagand är avgörande för olika studiers trovärdighet. Det har också visat sig att flera av metoderna för att beakta ursprungliga skillnader mellan personer med olika lång utbildning, samtidigt förvärrar mätfelsproblemet. Vidare använder metoderna olika delar av skillnaderna i utbildningslängd för olika personer. Det uppmätta sambandet mellan utbildning och lön från olika metoder, kommer därför att avse olika delar av befolkningen. Vi kommer nedan att redogöra för de tre metoderna var för sig, och kritiskt granska de studier som kan hänföras till respektive metod.

5.1.1 Studier som beaktar observerade egenskaper

Det kanske mest uppenbara sättet att bemöta problemet med att personer med olika lång utbildning kan skilja sig åt också i andra avseenden, är att försöka mäta och beakta dessa skillnader. Problemet är att vi inte vet exakt vilka faktorer som påverkar utbildningsval och att flera av de faktorer som vi tror är viktiga kan vara svåra att mäta.²⁸ I litteraturen har i huvudsak två typer av ursprungliga egenskaper beaktats: (1) förmåga och (2) familjebakgrund. Dessa faktorer har visat sig vara betydelsefulla för individers utbildningsval. Sambandet mellan tidig förmåga och utbildningslängd är starkt, inte minst därför att

²⁸ Det är bara ursprungliga skillnader mellan individer med olika lång utbildning som ska beaktas, inte skillnader som kan tänkas ha påverkats av utbildningen. Annars riskerar man att *underskatta* den kausala löneeffekten av utbildning.

möjligheten att avancera i utbildningssystemet huvudsakligen beror på tidigare studieprestationer. På samma sätt finns en omfattande litteratur som dokumenterar sambandet mellan familjebakgrund (såsom föräldrarnas utbildning, yrke och inkomst) och utbildningsnivå.²⁹ I praktiken har dock valet av faktorer kommit att bero på tillgången av data.

Det finns en rad internationella studier som försöker mäta den kausala effekten av utbildning på lön, genom att beakta skillnader i familjebakgrund och förmåga (t.ex. Griliches 1977; Blackburn och Neumark 1993 och 1995; Murnane m.fl. 1995). Det finns också några svenska studier på området.³⁰ Kjellström (1999) använde UGU-undersökningen och registerinformation från SCB för att mäta effekten av utbildning på inkomster 1993 för män födda 1948 och 1953.³¹ UGU-undersökningen innehåller en rad olika mått på förmåga (resultat från kunskaps- och begåvningstester samt betyg i sjätte klass), samt på uppväxtmiljö (föräldrarnas yrke och utbildning). Utan att ta hänsyn till faktorer som kan påverka individers utbildningsval, uppmätte Kjellström (1999) avkastningen på utbildning till 5,0 %. Effekten sjönk till omkring 4,7 % när uppgifter om föräldrarnas utbildning och yrke fördes in i analysen. När hänsyn också togs till resultat från kunskaps- och begåvningstester i årskurs 6, minskade avkastningen till omkring 3,8 %. Både arv- och miljöfaktorer tycks således påverka utbildningsval, och därmed snedvrider det observerade sambandet mellan utbildnings och lön.

Vi har uppdaterat Kjellströms (1999) studie, dels med fler kohorter (födda 1948, 1953, 1967, 1972, 1977 och 1982), dels med färskare inkomstuppgifter

²⁹ Se exempelvis den sociologiska litteraturen om social selektion (Erikson och Jonsson 1993) eller den ekonomiska litteraturen om intergenerationell rörlighet (Holmlund m.fl. 2008).

³⁰ Fägerlind (1975) är en av de allra första studierna både i Sverige och internationellt. Han använde mätte inkomsteffekter av utbildning under olika år mellan 1953 och 1971 för personer födda 1928. Dessutom beaktades skillnader i resultat från ett begåvningsprov som genomfördes i tredje klass. Något förvånande fann han i princip inga inkomsteffekter av utbildning. Resultaten är dock mycket svårtolkade, eftersom utbildning och yrke beaktas samtidigt i analyserna. Studien mäter således bara inkomsteffekter av skillnader i utbildningslängd för personer med samma yrke. Eftersom utbildningslängden påverkar möjligheten att välja yrke – och därmed också vilken lön man får – är det troligt att studien kraftigt underskattar effekten av utbildning.

³¹ Utvärdering genom uppföljning (UGU) drivs av institutionen för pedagogik och didaktik vid Göteborgs universitet och består av ett representativt urval av personer födda 1948, 1953, 1967, 1972, 1977, 1982, 1987, 1992 och 1998, som har följts från tredje eller sjätte klass. Undersökningen innehåller registerbaserad information, skoladministrativa uppgifter, enkätuppgifter samt resultat från kunskaps- och begåvningsprov. För en mer detaljerad beskrivning av UGU-undersökningen, se <http://www.ipd.gu.se/Forskning/forskningsprojekt/UGU>.

(2006).³² Till skillnad från Kjellström studerar vi avkastningen för både män och kvinnor. Resultaten presenteras i Tabell 5.1. Kolumn (1) återger det beskrivande sambandet mellan utbildningsår och logaritmerade inkomster, där ytterligare ett års studier är förknippat med 4.6 % högre inkomster. I kolumn (2) beaktar vi skillnader i föräldrarnas utbildning, vilket leder till en marginellt lägre utbildningseffekt. I kolumn (3) tar vi också hänsyn till resultat från begåvningsprov i årskurs 6. Precis som andra studier visat, så leder kontroll för tidig förmåga till en försvagning av sambandet mellan utbildning och inkomst. Avkastningen på utbildning uppmäts nu till 3,8 %. Läger man dessutom in kontroll för skolbetyg, se kolumn (4), så minskar avkastningen till 3,5 %.

Frågan är om resultaten från denna och liknande studier ska tolkas som kausala effekter av utbildning. Det finns två huvudsakliga invändningar. Den första är att det uppmätta sambandet mellan utbildning och lön bara kan ges en kausal tolkning under förutsättning att *alla* faktorer som påverkar både utbildningsval och lön har beaktats. Vanligtvis har man bara tillgång till grundläggande uppgifter om personers uppväxtmiljö, såsom föräldrarnas utbildning, yrke eller inkomst. Rimligtvis kan även andra miljökomponenter vara betydelsefulla för utbildningsbeslut, t.ex. stöd och hjälp med läxläsning, vilken skola man gick i, vilka vänner man umgicks med etc. Samma sak gäller uppgifterna om individers förmåga. I bästa fall har tidigare studier haft tillgång till tidiga mått på kognitiv förmåga, såsom resultat från kunskaps- och förmågetester eller betyg. Information om icke-kognitiva förmågor, såsom studieintresse, motivation och uthållighet, är betydligt mer sällsynt. Det är dessutom tveksamt om testresultat och betyg ger en fullständig bild av en persons kognitiva förmågor.

Mycket talar således för att många viktiga faktorer har utelämnats. Eftersom högutbildade personer i allmänhet har bättre *observerade* egenskaper – såsom föräldrabakgrund, testresultat och betyg – än lågutbildade personer, verkar det också rimligt att de skulle ha bättre *icke-observerade* egenskaper. Problemet med utelämnade variabler tenderar därför att överskatta sambandet mellan utbildning och lön.

³² Vi har valt att använda årliga arbetsinkomster, eftersom löneuppgifterna inte är heltäckande. Analyserna begränsas dock till personer som har en arbetssinkomst över 135 000, vilket ungefär motsvarar kollektivavtalsreglerade minimilöner på årsbasis (se Skedinger 2006). Med denna begränsning blir skattningarna mer eller mindre identiska med att använda lön som utfall.

Tabell 5.1. Avkastning på utbildning och kontroll för observerade egenskaper

	(1)	(2)	(3)	(4)	(5)	(6)
Utbildningslängd	0,046 (0,001)	0,044 (0,001)	0,038 (0,001)	0,035 (0,001)	0,033 (0,001)	0,034 (0,001)
Föräldrarnas utbildningslängd	.	0,008 (0,001)	0,006 (0,001)	0,005 (0,001)	0,005 (0,001)	0,004 (0,001)
Begåvningsprov Åk 6	.	.	0,045 (0,001)	0,026 (0,002)	0,025 (0,002)	0,024 (0,002)
Betyg Åk 6	.	.	.	0,032 (0,003)	0,026 (0,002)	0,023 (0,003)
Utbildningsaspiration Åk 6	.	.	.	.	0,023 (0,002)	0,022 (0,002)
Korrigerigering av mätfel						X
Antal personer	35 634	35 634	35 634	35 634	35 634	35 634

Noter: Tabellen visar regressionskattningar av sambandet mellan olika egenskaper och inkomsten. Beroende variabel är logaritmen av årsinkomsten 2006. Resultaten baseras på ett urval av personer födda 1948, 1953, 1967, 1972, 1977 och 1982. Endast personer med en årsinkomst på mer än 135 000 kronor samt med fullständiga värden på alla variabler ingår. Föräldrarnas utbildningslängd avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Begåvningsprovet i Åk 6 är summan av tre delprov som avser att mäta verbal, logisk och spatiell förmåga. Skolor hade möjlighet att avskaffa betygen i årskurs 6 för personer födda 1967. I de fall betyg i årskurs 6 saknas, har betyg i årskurs 7 använts. För personer födda 1977 (1982) har betyg i årskurs 8 (9) använts. Utbildningsaspiration avser antalet år som personerna tror att de ska gå i skolan, antingen som svar på en direkt fråga eller som de antal utbildningsår det yrke som de tror sig ha i framtiden motsvarar. För att göra alla variabler jämförbara med varandra och mellan kohorter har begåvningsprov, betyg och ambition standardiserats för varje kohort. Estimatet motsvarar effekten att röra sig en standardavvikelse högre upp i fördelningen. Förutom de variabler som återges i tabellen innehåller regressionerna ett intercept samt kontroller för kön och kohort.

Kolumn (5) i Tabell 5.1 illustrerar hur avkastningen på utbildning påverkas av att beakta ytterligare skillnader, i det här fallet elevernas utbildningsaspirationer.³³ I enlighet med resonemanget ovan, leder ytterligare kontroller till en något lägre avkastning på utbildning.³⁴ Skillnaden kan tyckas liten, men motsvarande effekt är att vänta om ännu fler egenskaper skulle kunna beaktas. Den första invändningen mot studier som försöker beakta skillnader i observerade egenskaper mellan personer med olika lång utbildning är således att de troligen *överskattar* utbildningens löneeffekter.

Den andra invändningen mot en kausal tolkning är att kontroller för observerade skillnader mellan individer förvärrar mätfelsproblemet. Som nämnts tidigare så leder mätfel till att sambandet mellan utbildning och lön underskattas. Kontroller för observerade egenskaper gör att sambandet blir ännu svagare. Detta skulle tala för att studier som beaktar skillnader i egenskaper mellan personer med olika lång utbildning *underskattar* den kausala effekten av utbildning. Den centrala frågan är hur stort problemet med mätfel egentligen är. Flera studier tyder på att mätfelet i enkät- och intervjuundersökningar vanligen utgör ungefär 10 % av observerade skillnader i utbildningslängd (Card 1999).

Många svenska studier baseras på uppgifter om högsta utbildningsnivå från registret över befolkningens utbildning (utbildningsregistret). Isacsson (2004) uppskattade mätfelet i utbildningsregistret till 12–13 % av den totala variationen i utbildningslängd. Denna skattning bygger dock på relativt starka antaganden.³⁵ Vi har därför valt att göra en ny skattning av mätfelets storlek. Vi använder oss då av SCB:s (1997) utvärdering av utbildningsregistret. I utvärderingen jämfördes uppgifterna om högsta utbildningsnivå i utbildningsregistret med motsvarande uppgifter i Arbetskraftsundersökningen (AKU). Med hjälp av detta material kan vi skatta mätfelet i utbildningsregistret genom att studera hur mycket ett utbildningsår i utbildningsregistret motsvarar i AKU.³⁶

Våra resultat visar att mätfelet i utbildningsregistret uppgår till omkring 5 %, vilket står i skarp kontrast till Isacssons (2004) skattningar på 12–13 %.³⁷

³³ Utbildningsaspiration avser antalet år som eleverna tror att de ska gå i skolan, antingen som svar på en direkt fråga eller som antalet utbildningsår det det förväntade yrket motsvarar.

³⁴ Skillnaden i avkastning på utbildning mellan kolumn (4) och (5) är dock inte signifikant.

³⁵ Isacsson (2004) antog att SCB (1997) har tagit fram ett sant utbildningsregister.

³⁶ Vi har vi inte tillgång till datamaterialet som ligger till grund för SCB:s utvärdering. Med hjälp av en korstabell över utbildningsregistret och AKU, är det dock möjligt att "återskapa" materialet. Varje utbildningsnivå tillskrivs därefter det genomsnittliga antalet utbildningsår enligt Levnadsnivåundersökningen (LNU) 1991. Under antagandet att mätfelet är slumpmässigt, kan vi skatta mätfelets storlek genom en regression av Utbildningsregistret mot AKU.

³⁷ Holmlund m.fl. (2008) skattar också mätfelets storlek i utbildningsregistret till omkring 5 %.

Förekomsten av mätfel är alltså ett betydligt mindre problem för studier som baseras på registerinformation än vad som tidigare hävdats. I kolumnen (6) i Tabell 5.1 presenterar vi den mätfelskorrigerade avkastningen på utbildning. Eftersom mätfelet är förhållandevis litet, blir avkastningen på utbildning bara något högre. Underskattningen av utbildningseffekten på grund av mätfel är således blygsam. Sammantaget är det därför troligt att studier som beaktar observerade skillnader mellan människor med olika utbildningslängd över-skattar den kausala effekten av utbildning på lön.

5.1.2 Studier som jämför syskon (tvillingar)

Ett alternativ till att försöka mäta alla skillnader i människors arv och uppväxtmiljö, är att studera personer som delar flera av dessa egenskaper. Man kan sedan eliminera gemensamma (icke-observerade) egenskaper genom att jämföra skillnaden i utbildningslängd med motsvarande skillnader i lön. Ett antal studier mäter därför avkastningen på utbildning för syskon. De delar vanligen uppväxtmiljö och i genomsnitt halva genuppsättningen. Ett flertal undersökningar försöker mäta kausala effekter av utbildning genom att studera tvillingar; helst enäggstvillingar eftersom de är genetiskt identiska. Man kan därmed eliminera alla ärftliga egenskaper som påverkar både utbildning och lön.

Det finns flera internationella studier som mäter avkastningen på utbildning genom att ta hänsyn till gemensamma icke-observerade produktiva egenskaper hos tvillingar (t.ex. Griliches 1979; Ashenfelter och Krueger 1994; Behrman m.fl. 1994; Miller m.fl. 1995; Rouse 1999; Ashenfelter och Rouse 1998). Gemensamt för flertalet av dessa studier är att antalet tvillingar är litet och att urvalet inte är särskilt representativt för befolkningen i helhet. Isacson (1999, 2004) använde det svenska tvillingregistret och registeruppgifter från SCB, för att mäta avkastningen på utbildning för samkönade tvillingar födda 1926–58. Förutom registerinformation om utbildning och inkomster 1989–93, fanns enkätuppgifter om huruvida personerna var enäggs- eller tvåäggstvillingar. Urvalet var riksrepresentativt och bestod av totalt 4 984 enäggstvillingar och 6736 samkönade tvåäggstvillingar.

Om ingen hänsyn togs till gemensamma arvs- och miljöfaktorer, uppmätte Isacson (1999) avkastningen på utbildning till ungefär 4,5 % för både enäggs- och tvåäggstvillingar. När han istället mätte effekten av skillnader i utbildningslängd *inom* tvillingparen, sjönk den uppmätta effekten till 3,9 % för tvåäggstvillingar och 2,2 % för enäggstvillingar. Sambandet mellan utbildning och inkomster mer än halveras således för enäggstvillingar, när man beaktar ge-

mensamma arvs- och miljöfaktorer. För tvåäggstvillingar är nedgången blygsammare och är i nivå förändringen som Kjellström (1999) fann när han beaktade skillnader i observerade faktorer. Skillnaden i uppmätt avkastning mellan enäggs- och tvåäggstvillingar verkar rimlig, eftersom enäggstvillingar har identisk genupsättning medan tvåäggstvillingar i genomsnitt delar hälften av generna. En tolkning av detta resultat är att kontroll för ärftliga faktorer är viktigt för att kunna mäta kausala effekter av utbildning.

I Tabell 5.2 presenterar vi nya mätningar av avkastningen på utbildning för tvillingar 2006. Vi har tillgång till ett större tvillingmaterial,³⁸ men till skillnad från Isacsson (1999) saknar vi information om huruvida tvillingarna är enäggs- eller tvåäggstvillingar. Vi har därför valt att dela upp tvillingarna i olikkönade och samkönade tvillingar. Olikkönade tvillingar är alltid tvåäggstvillingar. Av samkönade tvillingar är ungefär hälften tvåäggstvillingar och hälften enäggstvillingar (Karolinska institutet, 2005). Genom att jämföra resultaten från de olika tvillinggrupperna, kan vi under vissa förutsättningar beräkna hur stor effekten skulle vara för enbart enäggstvillingar.

Kolumn (1) och (2) visar avkastningen på utbildning för olikkönade tvillingar, utan respektive med kontroller för gemensamma faktorer. Avkastningen sjunker från 4,8 till 3,7 %, när vi mäter inkomstskillnader *inom* tvillingpar istället för inkomstskillnader mellan tvillingar i allmänhet.³⁹ I kolumnerna (4) och (5) återges resultaten för samkönade tvillingar. För denna grupp sjunker den uppmätta avkastningen på utbildning med nästan 40 % när icke-observerade skillnader mellan tvillingarna beaktas. Minskningen är mindre än vad Isacsson (1999) fann för enäggstvillingar, vilket förklaras av att samkönade tvillingar består av både enäggs- och tvåäggstvillingar.

³⁸ Urvalet består av alla svenska tvillingpar födda 1942–1990, där båda tvillingarna har en inkomst på minst 135 000 kronor år 2006.

³⁹ Vi har också studerat avkastningen på utbildning mellan syskon. Resultaten är i princip identiska för syskon som för olikkönade (tvåäggs)tvillingar. Det är kanske inte så förvånande med tanke på att tvåäggstvillingar inte är mer genetiskt lika än syskon i allmänhet. Notera också att förändringen av sambandet mellan utbildning och lön för olikkönade tvillingar är i nivå med vad Isacsson (1999) fann för *samkönade* tvåäggstvillingar. Det skulle kunna tyda på att samkönade och olikkönade tvåäggstvillingar ungefär har lika mycket produktiva egenskaper gemensamt, något vi kommer att utnyttja nedan.

Tabell 5.2. Avkastning på utbildning för tvillingar

	Olikkönade tvillingar			Samkönade tvillingar		
	(1)	(2)	(3)	(4)	(5)	(6)
Utbildningslängd	0,048 (0,001)	0,037 (0,002)	0,038 (0,002)	0,047 (0,001)	0,030 (0,001)	0,033 (0,001)
Föräldrarnas utbildningslängd	0,007 (0,001)	. .	. .	0,009 (0,001)	. .	. .
Kontroll för tvillingpar		X	X		X	X
Korrigering av mätfel			X			X
Antal personer	13 614	13 614	13 614	28 756	28 756	28 756

Noter: Tabellen visar regressionsskattningar av sambandet mellan olika egenskaper och logaritmerad inkomst 2006. Resultaten baseras på personer födda 1942-90. Urvalet begränsas till tvillingpar där båda har en årsinkomst på mer än 135 000 kronor samt fullständiga värden på alla variabler. Föräldrarnas utbildningslängd avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Förutom de variabler som återges i tabellen innehåller regressionerna ett intercept samt kontroller för kön, ålder och ålder i kvadrat.

Under förutsättning att samkönade och olikkönade tvillingar i övrigt är jämförbara, kan man använda de båda skattningarna för att beräkna avkastningen för enäggstvillingar.⁴⁰ Avkastningen (standardfelet) på utbildning för enäggstvillingar uppskattas i så fall till 2,3 (0,3) procent. Det är nästan exakt vad Isacsson (1999) fann och den korregerade skattningen tycks därför ge en god approximation på avkastningen på utbildning för enäggstvillingar.

Frågan är om syskonstudierna, och kanske framförallt tvillingstudierna, mäter kausala effekter av utbildning. Invändningarna mot studier som beaktar icke-observerade skillnader är desamma som för studier som beaktar observerade skillnader: (1) obeaktade skillnader och (2) mätfel i utbildningsvariabeln. Ett flertal forskare har kritiserat syskon- och tvillingstudierna, för att många viktiga faktorer som påverkar människors utbildningsval och lön inte har beaktats (se t.ex. Neumark 1999; Bound och Solon 1999). Tvillingstudier tampas också med en svårlöst paradox. Å ena sidan är fördelen med att studera tvillingar att de är *lika*, å andra sidan kan vi bara mäta effekten av utbildning för tvillingar som väljer *olika*.⁴¹ För att avkastningen på utbildning i tvillingstudier ska kunna tolkas kausalt, måste man anta att alla skillnader i utbildning inom ett tvillingpar är slumpmässiga. Frågan är förstås varför lika tvillingar väljer olika, och om det finns obeaktade skillnader i produktiva egenskaper som ger upphov till skillnaderna i utbildningsval.

Black m.fl. (2007) visade att det finns skillnader i födelsevikt mellan tvillingar, och att dessa skillnader påverkar både utbildningsdeltagande och inkomster. Även Behrman och Rosenzweig (2004) har liknande resultat. Tabell 5.3 redovisar avkastningen på utbildning när vi kombinerar ansatserna som kontrollerar för observerade respektive icke-observerade egenskaper. Mer precist studerar vi hur avkastningen på utbildning påverkas av att ta hänsyn till skillnader i uppmätt begåvning i Åk 6.⁴² Vi använder återigen UGU-materialet, men studerar bara de tvillingar som ingår i studien. Detta medför dessvärre att analysen baseras på ett fåtal personer, även om urvalsstorleken inte är mindre än i många internationella tvillingstudier. Kolumn (1) och (2) respektive (5)

⁴⁰ Avkastningen för enäggstvillingar beräknas enligt följande formel: $b_{MZ} = (b_{MZ,DZ} - \theta \cdot b_{DZ}) / (1 - \theta)$, där b_{MZ} är avkastningen för enäggstvillingar, $b_{MZ,DZ}$ är avkastningen för (samkönade) enägg- och tvåäggstvillingar, b_{DZ} är avkastningen för (olikkönade) tvåäggstvillingar och θ är andelen tvåäggstvillingar av samkönade tvillingar.

⁴¹ Ungefär hälften av de samkönade tvillingarna i våra data har samma utbildningsnivå.

⁴² En liknande analys finns i Sandewall m.fl. (2009). De använder dock ett mått på kognitiva egenskaper som uppmäts vid 18 års ålder. Eftersom utbildning påverkar individers kognitiva förmågor (se avsnitt 5.3) är det svårt att veta om det är skillnader i förmågor som medfört skillnader i utbildningslängd eller tvärtom. Deras resultat är därför svårtolkade.

och (6) replikerar resultaten för olikkönade och samkönade tvillingar i Tabell 5.2. Generellt sett tycks den uppmätta avkastningen på utbildning vara ungefär densamma i UGU-materialet som för hela tvillingbefolkningen. Effekten av att beakta gemensamma arv- och miljöfaktorer är dock något större.

I kolumn (3) och (7) beaktar vi skillnader i begåvning i åk 6. Detta leder till ett försvagat samband mellan utbildning och inkomst för både olikkönade och samkönade tvillingar. Eftersom urvalen är små är skillnaden dock inte statistiskt säkerställd. Avkastningen på utbildning är faktiskt inte ens signifikant skild från noll för samkönade tvillingar. Notera också att begåvningsmåttet har en direkt effekt på inkomsten för både olikkönade och samkönade tvillingar, även vid samma utbildningsnivå. Om vi gör samma korrigering som ovan, skulle avkastningen för enäggstvillingar faktiskt vara negativ (även om osäkerheten är stor). Dessa resultat ifrågasätter antagandet att tvillingar är identiska i alla relevanta avseenden, och tyder på att den kausala avkastningen på utbildning i själva verket kan vara lägre än vad tvillingstudierna finner.⁴³

Mätfel medför potentiellt problem i tvillingstudier. Eftersom utbildningslängd skiljer sig mycket mellan familjer, och därmed förklarar en stor del av variationen i utbildning, är mätfelsproblemet större i syskonstudier än i studier som beaktar andra observerade skillnader.⁴⁴ Isacsson (1999) korregerade den uppmätta avkastningen på utbildning under olika antaganden om mätfelets storlek. Våra skattningar ovan tyder på att mätfelet i utbildningsregistret bara uppgår till omkring 5 %. Det innebär att avkastningen på utbildning för enäggstvillingar (tvåäggstvillingar) i hans studie uppgår till 2,6 (4,2) %.⁴⁵ Om vi gör motsvarande mätfelskorrigeringar av våra skattningar, vilket sker i kolumn (3) och (6) i Tabell 5.2, så uppgår avkastningen på utbildning för samkönade (olikkönade) tvillingar till 3,3 (3,8) %. Den beräknade avkastningen för enäggstvillingar uppgår till 2,8 %. Våra skattningar är därmed i nivå med vad Isacsson (1999) fann. Om vi mätfelskorregerar skattningarna som beaktar skillnader i begåvning mellan tvillingar (Tabell 5.3), är motsvarande inkomsteffekter 1,2 (3,4) %. Avkastningen för enäggstvillingar är i princip noll.

⁴³ Isacsson (2004) mäter avkastningen på utbildning för olika utbildningsnivåer. Vi har gjort motsvarande skattningar för samkönade tvillingar. Resultaten är dock lite svårtolkade, eftersom selektionen till olika utbildningsnivåer – och därmed risken för att personer med olika lång utbildning skiljer sig från varandra – är olika för olika nivåer.

⁴⁴ Som exempel kan nämnas att resultat från begåvningstest, betyg, mått på utbildningsaspirationer samt föräldrarnas utbildningsnivå förklarar omkring 30 % av variationen i utbildning mellan slumpmässigt valda personer, medan skillnader mellan familjer förklarar ungefär 60 %.

⁴⁵ Dessa siffror är något lägre än vad Isacsson (1999) rapporterade, vilket beror på att mätfelskorrigeringen blir något mindre om antalet utbildningsår har predicerats från en annan källa.

Tabell 5.3. Avkastning på utbildning för tvillingar och kontroll för observerade egenskaper

	Olikkönade tvillingar				Samkönade tvillingar			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Utbildningslängd	0,051 (0,008)	0,042 (0,014)	0,034 (0,014)	0,034 (0,014)	0,047 (0,007)	0,018 (0,012)	0,011 (0,013)	0,012 (0,014)
Föräldrarnas utbildningslängd	-0,002 (0,009)	.	.	.	0,007 (0,008)	.	.	.
Begåvningsprov i Åk 6	.	.	0,084 (0,044)	0,084 (0,044)	.	.	0,064 (0,035)	0,064 (0,035)
Kontroll för tvillingpar		X	X	X		X	X	X
Korrigerig av mätfel				X				X
Antal personer	154	154	154	154	326	326	326	326

Noter: Tabellen visar regressionskoefficienter av sambandet mellan olika egenskaper och logaritmerad inkomst 2006. Resultaten baseras på ett urval av personer födda 1948, 1953, 1967, 1972, 1977 och 1982. Endast tvillingar med en årsinkomst på mer än 135 000 kronor samt med fullständiga värden på alla variabler ingår. Föräldrarnas utbildningslängd avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Begåvningsprovet i Åk 6 består av tre delar, och avser att mäta verbal, logisk och spatiell förmåga. Begåvningsprovet har standardiserats för varje kohort. Estimatens motsvarar effekten att röra sig en standardavvikelse högre upp i fördelningen. Förutom de variabler som återges i tabellen innehåller regressionserna ett intercept samt kontroller för kön och kohort.

Sammantaget visar tvillingstudierna att det är viktigt att beakta skillnader i ärftliga egenskaper mellan personer som har olika lång utbildning. Vi har också visat att det finns skillnader i begåvning vid 13 års ålder mellan tvillingar som förklarar både skillnader i utbildningslängd och inkomster. Eftersom urvalet av tvillingar med har uppgift om begåvning är litet, medför inte dessa kontroller en signifikant lägre avkastning på utbildning. Resultaten tyder dock på att det kan finnas kvarstående skillnader variabler som både påverkar utbildningslängd och lön även mellan enäggstvillingar. Vi har också visat att förekomsten av mätfel i utbildningsvariabeln är ett betydligt mindre problem i studier som baseras på utbildningsregistret. Den mätfelskorrigerade avkastning på utbildning för enäggstvillingar uppgår till omkring 2,8 %, trots att det antagligen finns kvarstående skillnader mellan tvillingarna. Den verkliga avkastningen på utbildning är sannolikt lägre.

5.1.3 Studier som utnyttjar institutionella förhållanden

Det huvudsakliga problemet med ansatserna som försöker beakta skillnader mellan personer med olika lång utbildning – antingen genom att ta hänsyn till observerade egenskaper eller genom att jämföra syskon – är att många viktiga skillnader troligen ändå består. Ett alternativt angreppssätt är att försöka hitta situationer då skillnader i utbildningslängd kan antas ha uppkommit mer eller mindre av en slump. Individer möter olika institutionella förhållanden när de gör sina utbildningsval, såsom skillnader i när utbildningssystemen byggs ut i olika delar av landet eller platsbegränsningar på gymnasiet eller högskolan. Om det kan antas vara slumpmässigt vem som ges en viss valmöjlighet, borde personer med olika begränsningar i genomsnitt inte skilja sig åt i relevanta avseenden. De uppmätta skillnaderna i inkomster för personer som möter olika institutionella begränsningar, kan därmed relateras till motsvarande skillnader i utbildningslängd.⁴⁶ Dessa typer av ansatser brukar benämnas kvasiexperiment eller instrumentalvariabelansatser (IV) i litteraturen.

För att begränsningar i utbildningsvalen ska kunna utnyttjas för att mäta avkastningen på utbildning måste två förutsättningar vara uppfyllda. För det första måste de institutionella förhållandena påverka individers utbildningsval. Det är oftast lätt att påvisa. Man jämför helt enkelt utbildningslängden för personer som exempelvis bott i kommuner som infört nioårig grundskola vid olika tidpunkter, eller som precis klarat respektive precis inte klarat betygs-

⁴⁶ Den uppmätta avkastningen på utbildning i studier som utnyttjar institutionella begränsningar i utbildningsvalen påverkas inte av eventuella (slumpmässiga) mätfel i utbildningsuppgifterna.

kraven vid antagningen till gymnasiet eller högskolan. För det andra måste individer som möter olika institutionella förhållanden i genomsnitt vara jämförbara i alla andra relevanta avseenden. Det är inte möjligt att påvisa, utan är ett grundläggande (identifierande) antagande. Man måste således anta att exempelvis personer som bott i kommuner som introducerade nioårig grundskola vid olika tillfällen eller som har betyg som ligger precis över respektive precis undre betygsgränsen vid antagningen till gymnasiet eller högskolan, i genomsnitt skulle ha haft samma inkomster i det hypotetiska fall att det inte funnits några institutionella begränsningar.

Tolkningen av studier som utnyttjar institutionella förhållanden kan skilja sig något från tolkningen av andra studier. De uppmätta inkomsteffekterna av utbildning gäller bara personer vars utbildningsval har påverkats av de institutionella begränsningarna. Det är möjligt att avkastningen för dessa personer skiljer sig från avkastningen för en genomsnittlig person. Det betyder å ena sidan att skillnader i utbildningslängd och inkomster direkt kan ses som en utvärdering av en genomförd åtgärd, t.ex. utbyggnaden av skolsystemet eller antalet platser på gymnasiet eller högskolan. Man kan därmed tolka effekterna som ett resultat av olika politiska beslut. Å andra sidan kan inkomsteffekten av utbildning vara avsevärt annorlunda för personer som har haft olika begränsningar i utbildningsvalen än för personer i allmänhet. Man måste därför vara försiktig när man generaliserar resultaten till hela befolkningen.

En rad internationella studier utnyttjar institutionella förhållanden för att mäta avkastningen på utbildning. En del studier jämför personer som möter olika institutionella begränsningar över tid (Angrist och Krueger 1991; Harmon och Walker 1995; Oreopoulos 2006); andra utnyttjar skillnader i valmöjligheter mellan regioner (Kane och Rouse 1993; Card 1995; Kling 2001); en sista grupp av studier utnyttjar både tidsmässiga och regionala skillnader (Duflo 2001; Ichino och Ebmer-Winter 2004; Lemieux och Card 2001, Webbink 2007). En del forskare har dock ifrågasatt det grundläggande antagandet i en del av dessa studier (Bound m.fl. 1995; Carneiro och Heckman 2002).

Det finns ett fåtal svenska studier som utnyttjar institutionella förhållanden. Meghir och Palme (2005) studerade utbildnings- och inkomsteffekterna av införandet av nioårig grundskola. Före 1950 hade Sverige ett relativt selektivt grundskolsystem. Eleverna gick tillsammans under de första sex (tre) åren i folkskolan. I sjunde (fjärde) klass kunde eleverna välja att gå kvar i folkskolan (1–2 år), eller att gå över till realskolan (3 år). Realskolan var också en förutsättning för vidare studier. Urvalet av sökande till realskolan baserades på skol-

betyg.⁴⁷ 1950 beslutade Riksdagen att introducera en enhetlig nioårig grundskola med en nationell läroplan. Det skedde i form av en storskaligt försök, där fler och fler kommuner gick över till det nya grundskolesystemet under perioden 1949–62.⁴⁸ Uppväxtkommun och födelseår kom därmed att avgöra om en elev fick gå i folkskolesystemet eller i grundskolesystemet. Meghir och Palme (2005) utnyttjade dessa institutionella begränsningar för att skatta effekterna av grundskolereformen.

Meghir och Palme (2005) studerade personer födda 1948 och 1953 i UGU-undersökningen, tillsammans med registerinformation från SCB, för att mäta effekten av grundskolereformen på utbildning och inkomster åren 1985–96. I många kommuner (reformkommuner) kom de två kohorterna att gå i olika skolsystem: personer födda 1948 gick i det gamla folkskolesystemet och personer födda 1953 gick i det nya grundskolesystemet. I andra kommuner (jämförelsekommuner) kom båda kohorterna att gå i samma skolsystem; antingen gick båda kohorter i det gamla skolsystemet eller så gick båda i det nya. Meghir och Palme (2005) jämförde skillnaden i utbildning och inkomster mellan personer födda 1948 och 1953 i reformkommuner, med motsvarande skillnad mellan personer födda 1948 och 1953 i jämförelsekommuner. Metoden kallas Difference-in-differences och tar hänsyn till både tidskonstanta skillnader mellan kommuner och trendmässiga skillnader över tid.

Resultaten visade att grundskolereformen påverkade både deltagarnas utbildning och inkomster. Elever i det nya skolsystemet skaffade sig i genomsnitt 0,3 års längre utbildning, huvudsakligen beroende på att den enhetliga grundskolan i genomsnitt var längre än folk- och realskolan. Det fanns dock vissa ”dynamiska” effekter, i och med att de som gick i det nya grundskolesystemet hade högre sannolikhet att läsa vidare på gymnasiet och högskolan än de som gick i det gamla skolsystemet. Grundskolereformens effekter på utbildning hade dock ingen motsvarighet i inkomster. I genomsnitt fanns ingen statistiskt signifikant inkomsteffekt av grundskolereformen.⁴⁹

⁴⁷ Andelen som gick vidare till realskolan ökade med tiden. På 1930-talet påbörjade omkring 10 % av en kohort realskolan. Motsvarande siffra på 1940-talet var ca. 15 %. Omkring hälften av dem som började blev utkuggade (Husén 2002). I UGU-materialet var det 35–40 % av personer födda 1948 som gick över till realskolan.

⁴⁸ Reformen påverkade framförallt personer födda 1945–55. Bara 6 % av personer födda 1945 gick i det nya systemet. Motsvarande siffra för personer födda 1955 är 95 % (Holmlund 2007).

⁴⁹ Holmlund (2007) mätte effekten av grundskolereformen på utbildningsdeltagande på ett större urval personer. Hon fann genomgående lägre utbildningseffekter än Meghir och Palme (2005), vilket delvis kan förklaras av att uppgifterna om när olika kommuner införde den nya grundskolan innehöll en del mätfel.

Meghir och Palme (2005) visade att grundskolereformen hade intressanta fördelningspolitiska effekter. Hela den genomsnittliga ökningen i utbildningsdeltagande berodde på att barn till lågutbildade föräldrar gick längre i skolan. Effekten för övriga barn – som gick länge i skolan redan i det gamla folkskole-systemet – var försumbar. Därmed kom reformen att utjämna en del av skillnaderna i utbildning som fanns mellan grupperna. På samma sätt skilde sig inkomsteffekterna med avseende på familjebakgrund. För barn med lågutbildade föräldrar medförde den nya grundskolan inkomstökningar på drygt 3 %, medan barn med högutbildade föräldrar i genomsnitt fick drygt 5 % lägre inkomster av att gå i den nioåriga grundskolan. Även i detta avseende var grundskolereformen således utjämnande mellan grupperna.

Det finns en del omständigheter som gör att man måste tolka resultaten i Meghir och Palme (2005) med viss försiktighet. För det första beaktade Meghir och Palme skillnader i betyg och begåvning i årskurs 6 samt förvärvsregion. Eftersom grundskolereformen kan ha haft effekter på både skolprestationerna och val av förvärvsort, är det möjligt att detta snedvrider resultaten.⁵⁰ För det andra uteslöt de personer som saknar inkomster 1985–96 från analysen (16 % av personer födda 1948 respektive 9 % av födda 1953).⁵¹ Eftersom grundskolereformen kan tänkas påverka sannolikheten att förvärvsarbete blir resultaten därmed svåra att tolka. För det tredje kan man ifrågasätta om kommuner som införde reformen vid olika tillfällen hade utvecklats likartat även i det hypotetiska fall att reformen införts alls. Holmlund (2007) visade att utvecklingen i kommuner som introducerade reformen vid olika tidpunkter såg olika ut. Det skulle kunna bero på att reformen hade olika effekt i olika kommuner, men indikerar också på att kommuner skulle kunna ha haft olika utveckling även utan reformen.⁵²

Vad säger då resultaten i Meghir och Palme (2005) om avkastningen på grundskoleutbildning? Faktum är att Meghir och Palme inte använde de institutionella begränsningarna för att mäta effekten av utbildning på inkomster, utan bara för att utvärdera effekten av grundskolereformen på utbildning och inkomster.⁵³ Orsaken är att reformen inte enbart förlängde grundskoleutbildningen, utan också innehöll flera andra komponenter. Det är därför svårt att

⁵⁰ Av samma skäl är det svårt att tolka effekterna som särredovisas med avseende på uppmätt begåvning i årskurs 6.

⁵¹ Meghir och Palme (2005) redogjorde inte för orsakerna till bortfallet. När vi analyserade materialet fann vi att knappt 4 % av personerna faller bort på grund av mortalitet och migration.

⁵² I praktiken verkar dock detta problem ha liten betydelse.

⁵³ I tidigare versioner mätte de dock utbildningsavkastningen (Meghir och Palme 1999).

avgöra om de uppmätta inkomsteffekterna beror på (1) skillnader i utbildningslängd, (2) senarelagd uppdelning av elever eller (3) en ny läroplan.

Låt oss för ett ögonblick anta att tidpunkten för elevernas gruppindelning och skillnaden i läroplanernas utformning har försumbara inkomsteffekter. Hur stor skulle då avkastningen på utbildning vara om man använder resultaten i Meghir och Palme (2005)? I genomsnitt medförde reformen 0,3 års längre utbildning, medan inkomsteffekten var 1,4 %. Genom att dividera reformeffekten för inkomster med motsvarande effekt för utbildning, erhålls en genomsnittlig avkastning på grundskoleutbildning på 4,8 %. Effekten ligger således mycket nära det beskrivande sambandet mellan utbildning och inkomster (se ovan), även om den skattade avkastningen inte är statistiskt skild från noll. Grundskolereformen hade dock större effekter för barn med lågutbildade föräldrar än för övriga barn. Avkastningen på grundskoleutbildning för barn med lågutbildade föräldrar uppmäts till 8,3 %. Eftersom inte utbildningslängden påverkades för barn till högutbildade föräldrar är det inte möjligt att beräkna ett avkastningsmått för dem. Notera dock återigen att deras inkomster sjönk med drygt 5 %.

Hur ska våra beräkningar av grundskoleutbildningens inkomsteffekter från Meghir och Palme (2005) tolkas? Resultaten från studier som utnyttjar institutionella begränsningar mäter effekterna för personer vars utbildningsval har påverkats av begränsningarna. Eftersom grundskolereformen bara påverkade utbildningslängden för barn till lågutbildade föräldrar, bör resultaten tolkas som avkastningen på utbildning för personer i denna grupp som gick längre i skolan pga. grundskolereformen. Det är möjligt att denna grupp har högre avkastning på grundutbildning än befolkningen som helhet, vilket skulle kunna förklara den replativt höga avkastningen på 8,3 %.⁵⁴

Det finns några omständigheter som talar för att våra beräkningar av avkastningen på utbildning är missvisande. För det första påverkade reformen antagligen kvaliteten på grundskoleutbildningen. Före reformen fanns stora kvalitetsskillnader mellan skolor. Den nioåriga grundskolan hade en nationell läroplan, vilket troligen minskade kvalitetsskillnaderna mellan skolor. För det andra innebar övergången från ett selektivt till ett uniformt utbildningssystem att signalvärdet av en genomgången nioårig grundläggande skolutbildning försvann. Tidigare var det bara de ”bästa” eleverna som kunde läsa nio år, medan i princip alla elever gjorde detta efter reformen. Det innebär att personer

⁵⁴ Studier som utnyttjar institutionella begränsningar finner ofta att skattningar som överstiger det beskrivande sambandet mellan utbildning och inkomster (se Card 1999).

som annars skulle ha gått i realskolan – och vars utbildningslängd därmed inte påverkades – kom att förlora på grundskolereformen. För det tredje är det möjligt att grundskolereformens omfattning påverkade lönesättningen för utbildad arbetskraft. Eftersom utbudet av personer med minst nioårig skolgång ökade samtidigt som utbudet av personer med högst 7 (8) års skolgång minskade, är det troligt att löneskillnaderna mellan dessa grupper minskade. Sammanfattningsvis är det svårt att veta hur förekomsten av skillnader i utbildningskvalitet, signaleffekter och relativa utbudseffekter påverkar den skattade avkastningen på utbildning.

En andra studie som utnyttjar institutionella förhållanden är Öckert (2010), som använde platsbegränsningarna till högskolan för att mäta avkastningen på grundläggande högskoleutbildning. Han jämförde sökande till högskolan som precis blivit antagna till en viss utbildning med sökande som precis blivit utan en plats vid samma utbildning. Dessa personer hade likvärdiga formella meriter, men fick olika chans att läsa på högskolan. I vissa fall lottades de sista platserna ut till personer med lika meriter. Under förutsättning att det inte finns några andra skillnader mellan personer som konkurrerar om de sista utbildningsplatserna, kan platsbegränsningarna utnyttjas för att mäta avkastningen på utbildning. Ansatsen brukar benämnas ”Regression-Discontinuity design”, och utnyttjar i det här fallet brytpunkter (diskontinuiteter) i sannolikheten att läsa på högskolan vid den lägsta antagningspoängen.

I början av 1980-talet var Sverige inne i en lågkonjunktur och söktrycket till högskolan var högt. Platsbegränsningarna till högskolan kom därmed att spela en stor roll för de sökandes chanser att bli antagna. Öckert (2010) studerade sökande till högskolan via universitets- och högskoleämbetet (UHÄ) hösten 1982. Han hade tillgång till detaljerad information från antagningen, såsom sökta utbildningar, kvotgruppstillhörighet, meritpoäng och antagningsstatus. Som referensgrupp använde han ett slumpmässigt urval av avgångna från gymnasiet. Individernas utbildningskarriär och arbetsmarknadssituation följdes under perioden 1981–96. För att kunna utnyttja platsbegränsningarna, avgränsade Öckert sig till sökande till utbildningar som var övrefterfrågade.

Öckert (2010) fann att personer som antagits till högskolan hösten 1982 i genomsnitt hade 0,2 års längre utbildning i mitten av 1990-talet, än sökande som inte antagits. Den blygsamma utbildningseffekten beror på att (1) studien avgränsade sig till två utbildningsalternativ och två kvotgrupper, (2) många sökande som antagits hoppade av sina utbildningar, samt (3) många av dem som inte antagits 1982 blev antagna senare. Antagningsstatus hade också –

åtminstone inledningsvis – stor betydelse för de sökandes arbetsinkomster. Årsinkomsten minskade med mellan 8 000 och 15 000 kronor under de tre efterföljande åren. Det var förstås en naturlig följd av de antagnas större sannolikhet att läsa på högskolan under denna period. Mycket överraskande kunde Öckert (2010) dock inte påvisa några långsiktiga inkomsteffekter av att bli antagen till högskolan. De positiva effekterna på utbildningsdeltagande hade således ingen motsvarighet i ökade årsinkomster.

Öckert utnyttjade de detaljerade uppgifterna om antagningsprocessen för att skatta avkastningen på utbildning under 1990-talet på två olika sätt. För det första beaktade han observerade skillnader mellan sökande med olika lång utbildning. För det andra utnyttjade han skillnader i utbildningslängd som orsakats av platsbegränsningarna till högskolan. I den första analysen visade Öckert att den uppmätta avkastningen på utbildning var känslig för val av jämförelsegrupp och kontrollvariabler. I det slumpmässiga urvalet av avgångna från gymnasiet, uppmättes avkastningen på högskoleutbildning till omkring 3,5 % 1991–96. För personer som sökt till högskolan – som därmed troligen också delar många egenskaper – skattades avkastningen till omkring 2,5 %. När Öckert tog hänsyn till skillnader i sökt utbildning och meritpoäng sjönk avkastningen till knappt 1,5 %.

Analysen som utnyttjade platsbegränsningarna för att mäta avkastningen på utbildning fann negativa inkomsteffekter av högskoleutbildning. Skattningarna hade dock genomgående dålig precision. Inte i något fall gick det att påvisa att avkastningen på högskoleutbildning var skild från noll eller från det observerade sambandet mellan utbildning och inkomster. Det går däremot att utsluta att avkastningen på utbildning är högre än 7,5 %, vilket har uppmätts i många anglosaxiska länder. En möjlig tolkning av skattningarna som utnyttjar platsbegränsningar är att de i alla fall inte motsäger de låga skattningarna på knappt 1,5 % när observerade egenskaper beaktas.

Frågan är hur resultaten i Öckert (2010) ska tolkas. Som nämns ovan mäter studier som utnyttjar institutionella förhållanden bara avkastningen på utbildning för personer som påverkas av begränsningarna. Med andra ord avser avkastningen på högskoleutbildning de sökande som kom in på högskolan hösten 1982 men som annars inte skulle ha läst vidare. Varför började inte alla de som blev utan utbildningsplats 1982 att läsa på högskolan senare? Det finns två möjliga förklaringar. För det första kanske de inte sökte till högskolan igen, antingen för att de blev avskräckta eller för att de fått ett (bra) arbete. För det andra kanske de sökte till högskolan senare, men lyckades inte bli antagna då

heller. En möjlig farhåga är således att avkastningen på högskoleutbildning bara avser personer som annars skulle ha fått ett bra jobberbjudande? Även under de efterföljande åren var dock konkurrensen om platserna till högskolan hård. Det enda sättet att förbättra sina chanser var att arbeta, eftersom arbetslivserfarenhet var meriterande vid antagningen. Det innebär att refuserade sökande som hittat ett arbete hade högre sannolikhet att komma in på högskolan senare. I enlighet med resonemanget visade Öckert (2010) att inkomstillväxten 1981–82 ökade sannolikheten för att de som inte antagits 1982 kom in 1983 istället. Sammanfattningsvis tycks inte förekomsten av en hög andel goda jobberbjudanden bland refuserade sökanden vara en trolig förklaring till den låga avkastningen på högskoleutbildning.

5.1.4 Vad vet vi om avkastningen på utbildning i Sverige?

Effekten av utbildning på arbetsmarknaden är ett av de mest skattade sambanden i nationalekonomiska mikrostudier. Metodproblemen är dock betydande. Och antalet studier som verkligen försöker hantera problemen är klart färre, men ändå betydande. Vi har i det här avsnittet kritiskt behandlat olika metoder för att mäta den kausala effekten av utbildning på löner och inkomster. För varje angreppssätt finns minst en svensk studie representerad. I flera fall måste de svenska studierna betraktas som högklassiga. Det borde därmed finnas goda förutsättningar att uttala sig om avkastningen på utbildning i Sverige.

Hur ska vi då sammanfatta evidensen? Isacssons (1999) och våra tvillingstudier utgör troligen en övre gräns för den avkastningen på utbildning i Sverige. Den genomsnittliga avkastningen skulle därmed vara högst 2,8 %.

Resultaten från studierna som utnyttjar institutionella förhållanden skulle kunna tolkas som att avkastningen på grundläggande skolutbildning är långt högre än avkastningen på högskoleutbildning. Vi tror dock att avkastningen på mer än 8 % för personer till lågutbildade föräldrar (Meghir och Palme 2005) överstiger den genomsnittliga avkastningen på grundskoleutbildning i befolkningen. Att låta hela befolkningen gå ytterligare ett år i grundskolan har antagligen blygsammare inkomsteffekter.

Öckert (2010) utnyttjar bl.a. detaljerad information från antagningen för att skatta avkastningen på högskoleutbildning. Hans resultat tyder på att avkastningen är knappt 1,5 %. Dessa skattningar har dock inte justerats för mätfel (vilket troligen är ett mindre problem) och det kan finnas kvarvarande skillnader mellan personer som har olika lång utbildning. Eftersom den genomsnittliga avkastningen på utbildning är högst 2,8 %, och eftersom avkastningen på

grundskoleutbildning antagligen är högre, verkar det rimligt att avkastningen på högskoleutbildning inte är högre än 1,5–2,0 %.⁵⁵

I kapitel 4 konstaterade vi att det observerade sambandet mellan utbildning och löner innebär att utbildningens lönepremie ligger på 5–6 %. Analysen i detta kapitel tyder på att ungefär hälften av denna lönepremie är kausal och att hälften har att göra med icke-observerade skillnader som gör att individer både har längre utbildning och högre inkomster.

Vad är orsaken till den relativt låga avkastningen på utbildning i Sverige?⁵⁶ Vi diskuterade två möjliga förklaringar i kapitel 4: den ena relaterar till utbudet av utbildad arbetskraft och den andra till lönebildningsinstitutionerna i Sverige.

Utbudsförklaringen handlar om att ett relativt högt utbud av välutbildad arbetskraft, i förhållande till efterfrågan, leder till en låg lönepremie. Utbudet av utbildad arbetskraft påverkas av utbildningspolitiska åtgärder. Till exempel är subventionsgraden av den svenska högskoleutbildningen hög: avgifter förekommer inte och studiemedlen är relativt förmånliga i Sverige. Dessa faktorer ökar det relativa utbudet av utbildad arbetskraft.

Statsmakterna påverkar även utbudets struktur genom dimensioneringen av olika studievägar. Det är möjligt att dessa dimensioneringsbeslut leder till en större omfattning på lågavkastande studievägar än vad som annars skulle vara fallet. Formerna för lönebildningen påverkar sannolikt också avkastningen på utbildning. Även om lönebildningen decentraliserats under 1990-talet bestäms lönerna för en stor andel av arbetskraften fortfarande i centrala förhandlingar mellan arbetsgivarorganisationer och fackföreningar. Centrala löneavtal är särskilt vanliga i låglönesektorer. Detta komprimerar lönestrukturen och kan ge upphov till en lägre avkastning på utbildning i Sverige än i länder utan centrala förhandlingar.⁵⁷

5.2 Effekter av utbildningsinvesteringar på kunskaper

Vi har i kapitel 3 diskuterat en modell för ackumulering av humankapital. De förmågor som människor ärver från sina föräldrar eller förvärvar under uppväxten eller senare i livet värdesätts i sin tur på arbetsmarknaden. Det innebär att människors förmågor och kunskaper – såsom de exempelvis mäts i

⁵⁵ Detta förutsätter dock att avkastningen på gymnasieutbildning inte är extremt låg.

⁵⁶ Notera att de skattningar som presenteras här försöker fastställa orsakssamband. Det är därför inte möjligt att avgöra om avkastningen på utbildning är lägre i Sverige än i andra länder, genom att relatera avkastningen på högst 2,8 % till de länderjämförelser som görs i avsnitt 4.2.

⁵⁷ Se Edin och Topel (1997) och Fredriksson och Topel (2010) för en utförligare diskussion om lönepolitiken och dess konsekvenser för avkastningen på utbildning i Sverige.

standardiserade tester eller betyg – är informativa om hur de senare klarar sig på arbetsmarknaden.⁵⁸ Man behöver därför inte vänta tills personerna kommer ut på arbetsmarknaden, för att utvärdera arbetsmarknadseffekterna av förd utbildningspolitik. Studier av effekterna på testresultat eller betyg utgör därför ett viktigt komplement till studier av rena arbetsmarknadsutfall. Därmed ges beslutsfattare också möjlighet att snabbare utöka lyckade satsningar eller avsluta mindre lyckade åtgärder.

Precis som ekonomer varit intresserade av löneeffekterna av ytterligare ett års utbildning, finns en omfattande pedagogisk litteratur om hur ett år i skolan påverkar elevernas förmågor och kunskaper. Dessa studier fyller två viktiga funktioner. För det första är studier som försöker isolera orsakssamband mellan utbildningslängd och förmågor också informativa om avkastningen på utbildning, pga. den nära relationen mellan testresultat och arbetsmarknadsframgång. För det andra utgör studier av utbildningslängdens kunskapseffekter en naturlig jämförelsenorm för studier som utvärderar andra delar av utbildningssystemet. De effekter som uppmäts i andra studier kan därmed relateras till effekten av att gå i skolan överhuvudtaget.

De empiriska problemen med att mäta effekten av utbildning på olika testresultat beror på vilken population som studeras. För personer i skolåldern är skillnader i utbildningslängd synonyma med skillnader i ålder. Eftersom mognad och erfarenhet i sig kan tänkas påverka testresultaten, är det nödvändigt att försöka separera effekten av skolgång från effekten av ålder. För personer i arbetsför ålder – när skolplikten avslutats – är skillnader i utbildningslängd ett resultat av individers val. Precis som i studierna som försöker mäta avkastningen på utbildning, är det därför nödvändigt att separera effekten av utbildningslängd från andra skillnader mellan personer med olika lång utbildning.

Studierna som försöker man mäta orsakssambandet mellan utbildning och testresultat använder tre olika angreppssätt:

- 1 Jämför testresultaten för personer som har samma observerade egenskaper (vanligen tidigare testresultat) men olika lång utbildning.
- 2 Jämför testresultaten för personer som är ungefär lika gamla, men som har olika lång utbildning på grund av reglerna för skolstart.
- 3 Jämför förändringen i testresultaten för personer som går i skolan med personer som inte går i skolan (antingen på grund av skollov eller för att studierna avslutats).

⁵⁸ Vi ska i avsnitt 5.3 visa att så också är fallet.

Vi kommer i det här avsnittet kritiskt gå igenom evidensen om skolans förmåge- och kunskapseffekter.

5.2.1 Studier som beaktar observerade egenskaper

Svenska pedagoger började tidigt försöka mäta effekterna av utbildning på elevernas förmågor med hjälp av omfattande representativa urval av befolkningen. Redan 1950 studerade Torsten Husén effekten av skolutbildning på de militära mönstringsresultaten med hjälp av den s.k. Malmöundersökningen. Kjell Härnqvist (1968) genomförde en motsvarande analys på UGU-materialet. De svenska studierna har också rönt stor internationell uppmärksamhet. När Winship och Korenman (1997) sammanfattade forskningen om effekterna av utbildning på elevernas förmågor, var två av sex studier utförda av svenska pedagoger baserade på svenska data.

De svenska pedagogerna var inte bara bland de första att studera effekter av utbildning på testresultat, de insåg också tidigt problemen med att fastställa orsakssamband. Vanligen syftade studierna till att förklara resultaten på den militära mönstringen med skillnader i utbildningslängd vid testtillfället. För att ta hänsyn till att personer med olika lång utbildning troligen skiljer sig åt också i andra avseenden, beaktades resultat från tester som genomförts före den obligatoriska skolans avslutande. Detta upplägg är därmed identiskt med de studier som försöker skatta avkastningen på utbildning genom att beakta observerade skillnader (avsnitt 5.1.1). Skillnaden ligger i att studierna av utbildningens effekter på testresultat kan ta hänsyn till tidigare skillnader i det utfall man är intresserad av, vilket gör dessa studier trovärdigare än studierna av avkastningen på utbildning. Det beror på att kontroll för skillnader i ett tidigare testresultat antagligen hanterar mer av (de för utfallet relevanta) skillnaderna mellan personer med olika lång utbildning, om man studerar utbildningens effekter på testresultat än om man studerar effekter på inkomster.

Husén (1950, 1951), Härnqvist (1959), Tuijnman och Husén (1991) samt Falch och Sandgren (2006) använde alla Malmöundersökningen för att mäta effekterna av skolutbildning på mönstringstestet för omkring 650 män födda 1928. Som kontroll för tidigare förmåga användes ett begåvningsprov som genomfördes i tredje klass (när eleverna vanligen var 10 år). Vid den här tiden var den obligatoriska skolgången 7 år. Variationen i utbildningslängd uppkom därför genom att en del elever gick ytterligare 1 eller 2 år i realskolan, samt att ett fåtal läste vidare på gymnasiet. Falch och Sandgren (2006) visade att ytterligare ett års utbildning var förknippat med 0,33 *d* bättre testresultat vid 20 års

ålder, om inte skillnader i tidigare förmåga beaktades. När testresultat vid 10 års ålder lades in i analysen sjönk effekten av ett skolor till 0,21 *d*. Resultaten i Husén (1950, 1951), Härnqvist (1959) samt Tuijnman och Husén (1991) är något mer svårtolkade eftersom de studerar effekten av att ha uppnått olika utbildningsnivåer snarare än utbildningslängd.⁵⁹ Eftersom det principiella upp-lägget i studierna var detsamma – även om metoderna skilde sig åt – är det troligt att resultaten ligger i paritet med vad Falch och Sandgren (2006) fann.

Härnqvist (1968) samt Balke-Aurell (1982) använde UGU-materialet för att mäta effekten av utbildning på mönstringsresultaten. Härnqvist (1968) stude-rade män födda 1948, medan Balke-Aurell (1982) studerade män födda 1948 och 1953. Som mått på tidigare förmåga användes resultat från begåvningsprov i sjätte klass. Dessutom beaktades skillnader i föräldrarnas utbildningsnivå. Variationen i utbildningslängd uppkom huvudsakligen på grund av skillnaden i längd mellan folkskolan och realskolan, men även av att många kommuner infört den nioåriga grundskolan. Därtill kom variation i andelen som läst vidare. Båda studierna analyserade skillnader i testresultat för män som uppnått olika utbildningsnivåer. Balke-Aurell (1982) lade också speciellt fokus på stu-diernas inriktning samt inriktningen på arbetslivserfarenheten för de som arbetat. Om man översätter resultaten för olika utbildningsnivåer till skillnader i testresultat med avseende på utbildningslängd – vilket inte är helt enkelt – uppmäts effekten av ett års skolutbildning i båda studierna till 0,13–0,20 *d* (Härnqvist 1977). Balke-Aurell (1982) visade dessutom att effekten av utbild-ning berodde på studiernas inriktning.

Gustafsson (2001) studerade betydelsen av gymnasieutbildningens inrikt-nings på elevernas testresultat. Han använde skolbetyg för att ta hänsyn till att personer med olika utbildningsinriktning kan skilja sig åt också i andra avseen-den. Betygen inhämtades när personerna avslutade grundskolan. Studien base-rades på män födda 1976 som mönstrade vid 18 års ålder, och begränsades till personer som läst på gymnasiet.⁶⁰ Varken skillnader i utbildningslängd eller ålder vid testtillfället beaktades i analyserna. Gustafsson (2001) skiljer sig där-med från andra studier på området, eftersom effekterna av utbildning också fångar upp ålderseffekter.

⁵⁹ Ett problem i Malmöundersökningen är att ca. 15 % av urvalet – huvudsakligen personer som gått om en eller flera klasser – mönstrade tidigare och skrev därmed ett annat mönstringstest. Sandgren och Falch (2006) standardiserar testerna separat, samt kontrollerar för om personerna var överåriga, vilket borde tendera att underskatta effekterna av utbildning. Det är oklart om och hur Husén (1951), Härnqvist (1959) samt Husén och Tuijnman (1991) hanterar problemet.

⁶⁰ Bara personer som skrev det nya mönstringsprovet (I-94) inkluderades.

Analysen delades upp med avseende på typ av gymnasielinje/program. Under antagandet att personerna i genomsnitt gått 2 år på gymnasiet vid mönstringen, medförde ett års teoretiska gymnasiestudier 0,17 *d* förbättrade testresultat. Effekten för yrkesinriktade utbildningar (vanligen 2-åriga) var i allmänhet lägre än för teoretiska inriktningar (vanligen 3-åriga). Det skulle dock delvis kunna förklaras av skillnader i utbildningslängd vid mönstringstillfället, eftersom de 2-åriga (men inte de 3-åriga) utbildningarna ibland kunde ha avslutats vid mönstringen. Han fann ingen skillnad i mönstringsresultat mellan personer som gick i det gamla linjesystemet och de som gick i den nya gymnasieskolan.

Det finns, precis som i fallet med studierna som försöker mäta avkastningen på utbildning, två invändningar med ansatsen att beakta observerade skillnader mellan personer som har olika lång utbildning. För det första är det möjligt att det kan finnas kvarstående skillnader även om man tar hänsyn till resultatet från ett tidigare test. Det kan bero på att tidigare testresultat inte är den enda bestämningsfaktorn av utbildningslängd och senare testresultat. Man kan exempelvis tänka sig att personernas familjebakgrund, motivation eller utbildningsaspirationer skulle kunna påverka båda utfallen. Eftersom de flesta tänkbara utelämnade variabler påverkar både utbildningslängden och testresultaten i samma riktning, är det troligt att effekten av utbildningslängd överskattas.⁶¹

Det är dessutom svårt att hitta tillförlitliga mått på personers förmåga. Testresultat innehåller många slumpelement, såsom elevernas dagsform och motivation. Frågor kan också vara för vagt formulerade, vilket ökar tendensen till att eleverna chansar. Dessutom kan frågorna vara för enkla eller för svåra, vilket begränsar variationen på grund av golv- och takeffekter. Sammantaget är det troligt att tidigare testresultat är oprecisa mått på individers nuvarande förmåga. Ju större mätfelet är, desto sämre kan tidigare testresultat beakta skillnader i förmågor mellan personer med olika lång utbildning. Detta leder också till överskattningar av utbildningseffekten.

Den andra invändningen mot studier som beaktar observerade skillnader, är förekomsten av mätfel i utbildningsvariabeln. Vi har tidigare framhållit att omfattningen av mätfel i de svenska utbildningsregistren är relativt liten. Även om flera tidigare studier använder gamla datakällor, kommer uppgifter om skolutbildning huvudsakligen från administrativa register. Bara i enstaka fall har information om utbildningsdeltagande samlats in från enkäter. Vår bedöm-

⁶¹ Falch och Sandgen (2006) visade att effekten av utbildning sjunker från 0,21 *d* till 0,17 *d* när skillnader i familjebakgrund, betyg och lärarens bedömning av eleverna prestationer beaktades.

ning är därför att problemet med mätfel i studier som beaktar observerade skillnader mellan personer med olika lång utbildning är relativt litet. Sammantaget är det därför troligt att dessa studier överskattar effekten av utbildning på testresultat.

5.2.2 Studier som utnyttjar reglerna för skolstart

Flera internationella studier utnyttjar reglerna för när barn ska börja skolan för att mäta effekten av utbildningslängd på testresultat. I de flesta länder börjar barn skolan det kalenderår de uppnår en viss ålder (vilken varierar mellan länder). Det innebär att barn som är födda i slutet på året vanligen börjar skolan ett år tidigare än barn födda i början av påföljande år. Om födelsetidpunkten i övrigt inte har någon direkt effekt på testresultatet, kan man utnyttja det faktum att barn födda på var sin sida om nyåret är ungefär lika gamla men har gått i skolan olika länge. Det är därmed inte nödvändigt att försöka mäta alla skillnader mellan personer med olika lång utbildning. Dessutom kan studier som utnyttjar reglerna för skolstart också mäta effekten av utbildningslängd på lägre utbildningsnivåer.

Studier som utnyttjar reglerna för skolstart för att mäta effekten av utbildning baseras vanligen på tester som skrivs av elever i två efterföljande årskurser vid samma tidpunkt (Baltes och Reinert 1969; Cahan och Davis 1987; Cahan och Cohen 1989; Bentin m.fl. 1991; Ferreira och Morrison 1994; Varnhagen m.fl. 1994; Bisanz m.fl. 1995; Morrison m.fl. 1995; Morrison m.fl. 1997; Crone och Whitehurst 1999; Luyten 2006; och Luyten m.fl. 2008). Skillnaden i testresultat för barn i samma årskurs födda under olika delar av året tillskrivs effekten av ålder, medan skillnaden i resultat för barn i olika årskurser födda på var sin sida om nyåret tillskrivs effekten av utbildningslängd.

Det finns två svenska studier som utnyttjar denna ansats för att mäta effekterna av grundskoleutbildning på elevernas prestationer. Gustafsson (2008) använder läsförståelsetestet i PIRLS 2001 för att separera effekten av ålder och skolgång. Det svenska urvalet av PIRLS omfattar drygt 10 000 elever i åk 3 och 4, och kan därför användas för att mäta effekten av ett fjärde grundskoleår. Eftersom vissa elever börjat skolan tidigare/senare än normalt, och då detta troligen är nära relaterat till deras prestationer, avgränsades studien till normalåriga barn (96 % av eleverna). Gustafsson fann att ett ytterligare år i grundskolan förbättrade läsförståelsen med 0,26 *d*. Effekten av ett års ålder uppmättes till 0,16 *d*.

Cliffordson (2008) utnyttjade matematikproven i TIMSS 1995 för att mäta effekten av grundskoleutbildning. Undersökningen omfattar närmare 9 000 elever i åk 6 till åk 8, vilket gör det möjligt att mäta effekterna av sjunde respektive ett åttonde grundskoleår. När hon studerade normalåriga elever fann hon att ett år i åk 7 förbättrade matematikprestationerna med hela 0,36 *d*. Motsvarande effekt för åk 8 var 0,25 *d*. Ålderseffekterna uppskattades till 0,17 samt 0,14 *d* i åk 7 respektive åk 8.

Cliffordson och Gustafsson (2008) studerade effekterna av tid på gymnasiet som uppkommit på grund av reglerna för skolstart. De studerade män födda 1972–75 som gick på en teoretisk gymnasieutbildning vid månstringen. Bara personer med ”normal ålder” beaktades, dvs. de som börjat skolan det år de fyllde sju, som inte gått om eller hoppat över en årskurs och som påbörjade sina gymnasiestudier direkt efter grundskolan. Dessutom begränsades studien till män som månstrade det kalenderår de fyllde 18 år. Under den studerade perioden skedde månstringen vanligen i relation till personernas 18-årsdag. De som var födda i början på året – och vanligen började skolan när de var något äldre – hade därmed gått kortare tid på gymnasiet när de månstrade än personer födda senare på året. Cliffordson och Gustafsson (2008) utnyttjar således att personer som var lika gamla vid månstringen – på grund av skolstartsreglerna – hade gått i skolan olika länge när deras förmågor testades. Eftersom alla män inte månstrar exakt på sin födelsedag, fanns också betydande åldersvariation även för personer med lika lång utbildning.

Cliffordson och Gustafsson (2008) visade att ett års teoretiska gymnasiestudier förbättrade eleverna färdigheter med i genomsnitt 0,18 *d*. Gymnasiestudierna har dock inte samma effekt på all förmågor. Störst effekt av utbildning uppmättes på personernas logiska och verbala förmåga, medan effekterna på spatial förmåga och teknisk förståelse var betydligt mindre. Effekten av personernas ålder på testresultaten var genomgående liten, ibland till och med negativ. Cliffordson och Gustafsson (2008) fann de största skillnaderna mellan olika utbildningsinriktningar vad gäller testet för teknisk förståelse, där personer som läste teknisk eller naturvetenskaplig linje presterade bättre ju längre de gått på gymnasiet.

Ett problem med studier som utnyttjar reglerna för skolstart är att föräldrar, lärare och rektorer har viss möjlighet att tidigare- eller senarelägga skolstarten för en del barn. Dessutom är det i vissa länder vanligt att barnen går om eller hoppar över en årskurs. Många studier – däribland de svenska – väljer därför att ta bort elever som inte har för årskursen ”normal” ålder. Eftersom lågpres-

terande barn födda sent på året (ska börja skolan tidigt) har större sannolikhet att senarelägga skolstarten eller att gå om ett år, tenderar det att underskatta effekten av utbildningslängd. Några internationella studier använder därför födelsedatum (vanligen födelsekvartal) som instrument för utbildningslängd (Neal och Johnson 1996; Hansen m.fl. 2004; Gormley och Gayer 2005; Cascio och Lewis 2006).

Ett annat – och kanske allvarigare – problem med studier som utnyttjar reglerna för skolstart, är att de antar att åldern för skolstart i sig inte har någon effekt på elevernas skolresultat. Så är dock inte fallet. Det finns en överväldigande evidens från nästan alla världens länder att barn som är äldre vid skolstarten klarar sig bättre i skolan än sina yngre kamrater (t.ex. Bedard och Dhuey 2006; Fredriksson och Öckert 2008). Eftersom skolstartsreglerna medför att de som är födda i början av året har gått kortare i skolan vid mättillfället, än personer med samma ålder som är födda senare på året, kommer utbildningslängden att vara negativt relaterad till skolstartsåldern. Med andra ord: de som vid testtillfället har längre utbildning kommer i genomsnitt att ha varit yngre när de började skolan. Eftersom effekterna av skolstartsålder är stora, tenderar dessa studier att kraftigt underskatta effekten av utbildning på testresultaten.⁶² Eftersom Gustafsson (2008), Cliffordson (2008) och Cliffordson och Gustafsson (2008) inte tar hänsyn till skillnader i skolstartsålder, är det rimligt att deras studie underskattar effekten av utbildning.

5.2.3 Studier som utnyttjar studieavbrott

Det finns en rad internationella undersökningar som försökt utnyttja olika typer av studieavbrott eller förkortad skolutbildning för att identifiera effekterna av utbildning på skolprestationerna. En del av dessa uppsatser utnyttjar dock relativt drastiska avbrott i skolstudierna, såsom krigsutbrott (DeGroot 1948) eller nedstängning av skolor under längre perioder (Green m.fl. 1964). Det är troligt att sådana händelser har direkta effekter på elevernas resultat, vilket gör att studierna kan vara missvisande. Det finns dock några undersökningar som utnyttjar mindre dramatiska studieuppehåll, såsom skollov eller frivilliga avbrott (t.ex. Hansen m.fl. 2004; Leuven m.fl. 2010).

Heynes (1978) och Entwisle m.fl. (1997) visade att amerikanska elever lär sig mer under terminerna än under sommarlovet, och att kunskapsförändringen under sommaren är negativt relaterad till familjeinkomsten. Lindahl (2001)

⁶² Fredriksson och Öckert (2008) visade att svenska elever som är ett år äldre vid skolstarten hade ungefär 0,17 *d* bättre betyg i årskurs nio.

utnyttjade skillnaden i kunskapstillväxt under skolåret respektive sommarlovet för att mäta effekterna av skolgång på kunskaper för elever i sjätte klass. Han samlade in test- och enkätinformation för 556 elever vid 16 slumpvis valda stockholmskolor. Eleverna gavs samma prov i matematik vid tre tillfällen: (1) i slutet av vårterminen i femte klass; (2) i början av höstterminen i sjätte klass; samt (3) i slutet av vårterminen i sjätte klass. Under förutsättning att kunskapsutvecklingen under sommarlovet enbart fångar upp effekten av mognad och uppväxtmiljö (annan än skolan), kommer skillnaden i kunskapsförändring under skolåret och sommarlovet att mäta effekten av att gå i skolan. Lindahl (2001) visade att skillnaden i inläring under skolåret och under sommarlovet motsvarade en effekt av skolgång på ungefär 0,21 *d*.

Det finns två problem med Lindahls studie. För det första genomfördes testen inte precis före respektive precis efter sommarlovet, utan lite längre in i terminerna. Prestationsförändringen under sommarlovet fångar därmed till viss del upp effekter av skolgång. Lindahl är medveten om detta problem och försöker justera skattningarna under olika antagande om hur inläring påverkas av tid i skolan. Om inläring är linjär, d.v.s. att prestationerna ökar lika mycket över hela skolåret, medförde ett år i femte klass en förbättring av matematikprestationerna med hela 1,44 *d*. Under antagande att inläringen är konkav, dvs. att elevernas prestationer avtar med tid i skolan, uppskattades effekten till 0,94 *d*. För det andra finns risk att eleverna lär sig provet med tiden, eftersom samma test gavs vid tre tillfällen. Lindahl beaktar detta problem genom att utnyttja det faktum att inte alla elever skrev (hela) provet vid alla tillfällen. Under antagande om icke-linjär inläring, uppskattades då ett års skolestudier förbättra matematikprestationerna med 1,07 *d*. Sammantaget verkar Lindahls (2001) resultat vara mycket känsliga för olika antagande om inläring och effekter av att skriva samma prov flera gånger.

Öckert (2009) utnyttjade reglerna för skolstart samt sommarloven för att separera effekterna av tid på gymnasiet, ålder och skolstartsålder på skolprestationerna. Han använde resultaten från inskrivningsprovet vid den militära mönstringen för män födda 1965–75. Effekten av skolstartsålder bestämdes genom att jämföra personer födda precis före respektive precis efter nyåret. De var födda ungefär samtidigt men började – på grund av reglerna för skolstart – vanligen skolan vid olika åldrar (jfr. Fredriksson och Öckert 2008). För att separera effekterna av utbildning och ålder på prestationerna, utnyttjade han skolloven samt att en del personer avslutat sin utbildningskarriär vid mönstringstillfället. Förändringen i prestationer för personer som mönstrade under

olika delar av skollåret (då eleverna både gick i skolan och åldrades) relaterades till förändringen i prestationer för personer som mönstrade under olika delar av skolloven eller när gymnasiet avslutat (då eleverna bara åldrades).

Öckert (2009) fann att ett år på gymnasiet förbättrade resultaten på inskrivningsprovet med i genomsnitt 0,23 *d*. Ålderseffekterna var genomgående mycket små (i flera fall till och med negativa). Störst effekt hade gymnasieutbildning på logisk och verbal förmåga (0,25 *d*), medan effekterna på spatial förmåga var relativt små (0,10 *d*). Öckert (2009) fann dock inga större skillnader mellan teoretiska och yrkesinriktade gymnasieprogram.

5.2.4 Hur stora är effekterna i Sverige?

Precis som i fallet med avkastning på utbildning finns en omfattande svensk forskning om hur mycket elever lär sig i skolan. Svenska pedagoger var tidigt ute och försökte mäta sambandet mellan skolgång och förmågor. På senare år har också en rad nationalekonomiska studier tillkommit. Flera olika metoder för att separera skoleffekter från ålderseffekter respektive selektionseffekter har använts. Vidare täcker studierna effekter av skolgång stora delar av utbildningssystemet, från fjärde klass till och med gymnasiet. Vilka lärdomar kan vi då dra från dessa studier?

Även om de svenska studierna använder olika ansatser och avser olika tidsperioder är resultaten ändå – med något undantag – förvånansvärt samstämmiga. Effekterna av ett år i skolan ligger inom intervallet 0,15–0,25 *d* i nästan alla studier. Det är värt att betona att en skattning på omkring 0,20 *d* av ett års skolgång är mycket. Svenska skolor tycks därför i allmänhet kunna förmedla mycket kunskaper till sina elever.

Vi tycker oss också kunna se ett mönster med något större effekter av skolgång i början av utbildningskarriären, och något mindre effekter för senare utbildningsdeltagande. Gustafsson (2008) uppmätte effekten av att gå fjärde klass till 0,26 *d*. Studierna som baseras på Malmöundersökningen och UGU-undersökningen mäter effekterna av folkskola (åk 7–8), realskola (åk 7–9) och till viss del gymnasium. Där tycks effekterna i genomsnitt ligga på 0,20 *d*. Studierna av effekter av gymnasieutbildning visar på något lägre skattningar: 0,17 *d* för teoretiska utbildningar och 0,14 *d*, eller betydligt lägre, för yrkesinriktade utbildningar. Trots en tendens till sjunkande kunskapseffekter av skolgång ju längre upp i utbildningssystemet man kommer, är det värt att notera att även relativt sena utbildningsinvesteringar (gymnasiet) medför stora kunskapsförbättringar för eleverna.

Det är viktigt att betona att litteraturen om effekter av skolgång – både den svenska och den internationella – nästan helt uteslutande studerar olika mått på kognitiv förmåga. Vad vi vet saknas det i princip helt studier som på ett trovärdigt sätt försöker mäta effekter utbildning på elevernas icke-kognitiva förmåga.⁶³ Detta trots att – som vi ska visa i nästa avsnitt – både kognitiva och icke-kognitiva förmågor är viktiga för individers arbetsmarknadsframgång. Studier av effekten av skolgång på elevernas icke-kognitiva förmågor är ett viktigt område för framtida forskning.

5.3 Effekter av kunskaper på inkomster

Föregående avsnitt visade att det finns en överväldigande evidens för att skolor producerar kunskaper. Syftet med det här avsnittet är att utröna om – och i så fall hur mycket – dessa kunskaper värdesätts på arbetsmarknaden. Det finns en stor internationell litteratur som visar på starka samband mellan kognitiva förmågor å den ena sidan och arbetsmarknadsframgång å den andra (t.ex. Murnane m.fl. 1995; Neal och Johnsson 1996; Murnane m.fl. 2001; Zax och Rees 2002; Hanushek och Woessmann 2008). I det här avsnittet kommer vi att belysa den svenska forskningen på området, både vad gäller arbetsmarknads-effekter av kognitiva såväl som av icke-kognitiva förmågor.

Låt oss först slå fast att det är mycket svårt att mäta kausala effekter av olika kunskaper. Det finns flera orsaker till det. För det första påverkas personers kunskaper av en rad faktorer, såsom arv, uppväxtmiljö, skola och vänner. Detta tenderar att överdriva effekten av kunskaper på arbetsmarknaden. Vanligtvis saknar forskaren alla dessa faktorer i de data som används. För det andra är det troligt att testresultat utgör oprecisa mätningar av människors kunskaper – testresultat innehåller mätfel. Som vi tidigare diskuterat leder detta till underskattningar av effekten av kunskaper på arbetsmarknaden. För det tredje är kunskaper mångdimensionella. Forskare har vanligen bara uppgifter om en del av undersökningspersonernas färdigheter. Eftersom olika kunskaper samvarierar med varandra är det troligt att detta också leder till att enskilda kunskapers betydelse på arbetsmarknaden överdrivs.

Sammantaget gör dessa problem att det i princip saknas studier – både i Sverige och internationellt – som på ett trovärdigt sätt mäter orsakssamband mellan kunskapsmått och arbetsmarknadsutfall. Man får därför nöja sig med beskrivande studier om sambandet mellan olika förmågor eller kunskaper på

⁶³ Luyten m.fl. (2008) studerade betydelsen av utbildning för elevernas läsaktiviteter och läsengagemang, men hittade inga signifikanta effekter.

exempelvis sysselsättning eller löner. I bästa fall har det varit möjligt att beakta skillnader i uppväxtmiljö samt flera olika kunskapsmått samtidigt. Det är dock viktigt att komma ihåg att sambanden som uppmäts under rimliga antaganden bara är beskrivande och att den kausala effekten av olika egenskaper på arbetsmarknaden kan vara annorlunda. Det är därför inte helt enkelt att översätta kunskapseffekterna av skolgång eller förd utbildningspolitik till arbetsmarknadseffekter. Det beskrivande sambanden kan ändå vara informativa om olika kunskapers relativa värde på arbetsmarknaden.

Föregående avsnitt visade att utbildning påverkar skolprestationer. Detta har viktiga implikationer för hur man ska mäta effekter av kunskaper på arbetsmarknaden. För att kunskapseffekterna inte ska spegla skillnader i utbildningslängd är det nödvändigt att använda mått på skolprestationer som insamlats när personerna gått i skolan lika länge. Detta sker enklast vid en given tidpunkt i grundskolan. Detta innebär dock inte att effekterna av kunskaper på arbetsmarknaden inte fångar upp skillnader i utbildningslängd. Skolprestationer, inte minst betyg, har stor betydelse för människors fortsatta utbildningskarriär. Det är precis av den anledningen som olika förmågetest används som kontrollvariabler i studier som försöker mäta kausala effekter av utbildning på arbetsmarknaden (se avsnitt 5.1.1). Eftersom utbildning snarast ska ses som ett utfall i det här sammanhanget, kommer dock studier som beaktar både utbildning och skolprestationer att underskatta effekter av kunskaper på arbetsmarknaden. Vi kommer därför att avgränsa oss till studier som bara kontrollerar för egenskaper som bestämdes för grundskolan avslutades.

Falch och Sandgren (2008) använde Malmöundersökningen för att mäta effekten av kognitiv förmåga vid tio års ålder på livsinkomsterna. De avgränsade sig till personer som har positiv inkomst och kontrollerar för utbildningsnivå eller utbildningslängd, vilket gör resultaten lite svårtolkade. Som utgångspunkt presenterade de dock utvecklingen av sambandet mellan tidig kognitiv förmåga och inkomster för män. Det är resultaten som vi väljer att återge här.

I slutet på 1950-talet, när personerna i undersökningen var 20 år, var inkomsteffekterna av en standardavvikelse bättre kognitiva förmågor omkring 5 %. Effekterna kom sedan att öka kraftigt under de efterföljande 20 åren. I slutet av 1960-talet, när personerna var i 40-årsåldern, var inkomsteffekten hela 25 %. Det är dock värt att komma ihåg att resultaten potentiellt fångar, dels en förändring av avkastningen på kognitiva förmågor över tiden, dels en förändring av avkastning över livscykeln. Från slutet av 1960-talet till och med mitten av 1980-talet sjönk avkastningen på de kognitiva förmågorna till omkring

15 %. Därefter kom inkomsteffekterna att öka fram till mitten av 1990-talet, då personerna närmade sig pensionsåldern, till 20 %. Utvecklingen av sambandet mellan inkomster och kognitiva förmågor sedan slutet av 1960-talet är snarlik utvecklingen av sambandet mellan inkomster och utbildning (se avsnitt 4.3).

Lindahl (2001) använde UGU-undersökningen för att mäta inkomsteffekter av matematikkunskaper och förmåga i årskurs 6 för personer födda 1948 och 1953. Analysen begränsas till dem som tjänade minst 100 kronor år 1987, 1990 eller 1993, och genomsnittet av årsinkomsterna användes som utfall. Lindahl kontrollerade för familjebakgrund samt analyserade kunskaps- och förmågemåtten samtidigt, vilket gjorde att deras separata inkomstbidrag kunde mätas. Han fann att en standardavvikelse bättre matematikkunskaper i sjätte klass var förknippad med ca. 8,5 % högre inkomster. Motsvarande inkomstpremie för uppmätt förmåga i åk 6 uppskattades till 5,5 %.

Zetterberg (2005) studerade inkomsteffekter av kognitiva och icke-kognitiva förmågor i åk 6 för personer i UGU-undersökningen födda 1967. Analysen begränsades till män (kvinnor) med högre årsinkomster än 154 000 (102 000) kronor. Som mått på kognitiv förmåga använde Zetterberg verbal, spatial och induktiv förmåga i åk 6. Måtten på icke-kognitiv förmåga baserades på svar på enkätfrågor till eleverna. Han skapade ett mått på relativt självförtroende, som utgick från frågor om elevens bedömning om sina prestationer i förhållande till klasskamraternas. Därtill skapades ett mått på absolut självförtroende baserat på elevernas bedömning av sina prestationer i största allmänhet. I de resultat som återges nedan analyseras alla förmågemått samtidigt.

Zetterberg fann att en standardavvikelse bättre kognitiva förmågor var förknippad med 7,6 (6,6) % högre inkomst för män (kvinnor). Han fann inga inkomsteffekter för måttet på absolut självförtroende, men en inkomstpremie på 4,4 (3,6) % för en förbättring av relativt självförtroende med en standardavvikelse för män (kvinnor). Han fann också positiva interaktionseffekter mellan de kognitiva och icke-kognitiva förmågorna, vilket tyder på att dessa egenskaper är komplement snarare än substitut på arbetsmarknaden. Zetterberg visade också att inkomsteffekterna för både kognitiva och icke-kognitiva förmågor ökade ju längre upp i inkomstfördelningen personerna befann sig.

Lindqvist och Westman (2009) använde data från den militära mönstringen för att mäta betydelsen av både kognitiv och icke-kognitiv förmåga på arbetsmarknaden. Studien baserades på män födda 1965–74 och utfallen på arbetsmarknaden uppmättes 2006. Som mått på kognitiv förmåga användes det militära inskrivningsprovet, som är avsett att mäta verbal, logisk, spatial och tek-

nisk förmåga. Måttet på icke-kognitiv förmåga kom från en strukturerad intervju med en certifierad militärpsykolog, och skulle fånga uthållighet, social förmåga, ledarskapsegenskaper, etc. I de flesta fall analyserades effekten av båda måtten samtidigt, vilket gör det möjligt att renodla effekten av varje enskild förmåga. Författarna korregerade också för förekomsten av mätfel i de båda måtten.

Lindqvist och Vestman (2009) visade att både kognitiva och icke-kognitiva egenskaper är mycket viktiga för att förklara individers arbetsmarknads-situation. Icke-kognitiva förmågor visade sig vara speciellt viktiga för att förklara ogynnsamma arbetsmarknadsutfall, såsom arbetslöshet och låga inkomster. Personer med en standardavvikelse bättre icke-kognitiv förmåga hade exempelvis 2,4 procentenheter lägre risk att vara arbetslös. Båda egenskaperna var också starkt relaterade till inkomster, och en standardavvikelse bättre kognitiv (icke-kognitiv) förmåga var förknippad med omkring 8,3 (8,6) % högre lön. Detta gällde trots att de kontrollerade för om personerna gått på gymnasiet eller inte, vilket troligen fångar upp en del av framförallt personernas kognitiva förmågor.

Vi har valt att göra egna skattningar av effekten av kognitiva och icke-kognitiva förmågor över livscykeln. Liksom Lindahl (2001) och Zetterberg (2005) använder vi oss av UGU-undersökningen, dock studerar vi alla kohorter födda 1948, 1953, 1967, 1972, 1977 och 1982. Uppgifter om årsinkomst finns tillgängliga för perioden 1968, 1971, 1973, 1976, 1979, 1982 och 1985–2006, vilket gör det möjligt att studera sysselsättnings- och inkomsteffekter för åldersintervallet 20–58 år. Som mått på kognitiva förmågor använder vi medelvärden av resultaten från tester av verbal, spatial och induktiv förmåga i årskurs 6. De icke-kognitiva egenskaperna representeras av individernas utbildningsaspirationer i åk 6.

De flesta tidigare studier av arbetsmarknadseffekterna av kunskaper och färdigheter har begränsats till personer som arbetar. Det är dock mycket möjligt att människors förmågor och färdigheter också samvarierar med deras sysselsättningsstatus. Tabell 5.4 visar skillnader i sannolikheten att vara sysselsatt för personer som befinner sig en standardavvikelse från varandra i fördelningen.

Tabell 5.4. Sysselsättningseffekter för olika egenskaper över livscykeln (20-58 år)

	Alla	Kvinnor	Män	Lågutbildade föräldrar	Högutbildade föräldrar
Begåvningsprov Åk 6	0,025 (0,002)	0,033 (0,004)	0,016 (0,003)	0,027 (0,003)	0,022 (0,004)
Utbildningsaspiration Åk 6	0,019 (0,002)	0,029 (0,003)	0,010 (0,003)	0,022 (0,003)	0,017 (0,004)
Föräldrarnas utbildningslängd	-0,002 (0,001)	-0,000 (0,001)	-0,004 (0,001)	-0,011 (0,002)	-0,001 (0,002)
Antal personer	53 488	24 747	25 674	32 110	21 378

Noter: Tabellen visar regressionsskattningar av sambandet mellan olika egenskaper och sannolikheten att ha en inkomst över 135 000 (i 2006 års penningvärde) för respektive år. Beroende variabel är en sysselsättningsindikator för åren 1968, 1971, 1973, 1976, 1979, 1982 samt 1985-2006. Resultaten baseras på ett urval av personer födda 1948, 1953, 1967, 1972, 1977 och 1982. Föräldrarnas utbildningslängd avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Begåvningsprovet i Åk 6 är summan av tre delprov som avser att mäta verbal, logisk och spatiell förmåga. Utbildningsaspiration avser antalet år som personerna tror att de ska gå i skolan, antingen som svar på en direkt fråga eller som de antal utbildningsår det yrke som de tror sig ha i framtiden motsvarar. För att göra alla variabler jämförbara med varandra och mellan kohorter, har föräldrarnas utbildningslängd, begåvningsprov, betyg och ambition standardiserats för varje kohort. Estimatet visar skillnaden i sannolikheten att vara sysselsatt för personer som befinner sig en standardavvikelse från varandra i fördelningen. Förutom de variabler som återges i tabellen innehåller regressionerna ett intercept samt fixa effekter för interaktionen mellan år, kohort och kön. Klusterjusterade standardfel (individer) finns inom parantes.

Första kolumnen visar att personer med högre kognitiv eller icke-kognitiv förmåga har större sannolikhet att vara sysselsatt. En förflyttning i fördelningen av begåvning med en standardavvikelse är förknippad med i genomsnitt 2,5 procentenheter större sannolikhet att vara sysselsatt under livscykeln. Motsvarande effekt för utbildningsaspirationer är 1,9 procentenheter. Både kognitiva och icke-kognitiva förmågor är således viktiga för individers sysselsättnings-sannolikheter. Det är vidare värt att notera att givet måtten på kognitiv och icke-kognitiv förmåga finns i princip inga sysselsättningsskillnader med avseende på föräldrarnas utbildning.

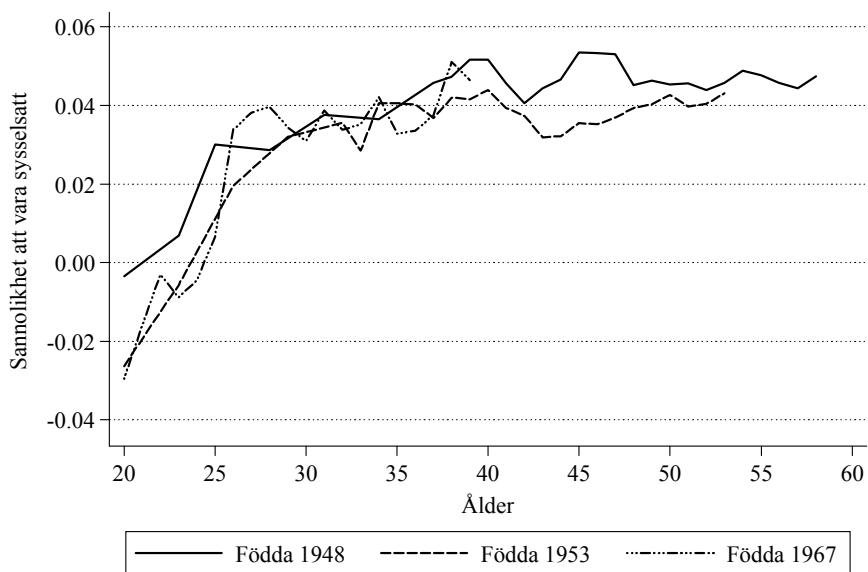
De två nästkommande kolumnerna delar upp analysen på män och kvinnor. Båda förmågemåtten har ungefär dubbelt så stor betydelse för kvinnor än för män. De två sista kolumnerna delar upp analysen med avseende på föräldrarnas utbildningsnivå, dock utan att hitta några statistiskt säkerställda skillnader mellan grupperna. Det tycks dock som om både kognitiva och icke-kognitiva förmågor är förknippade med något högre sysselsättning för barn till lågutbildade föräldrar än för barn till högutbildade föräldrar.

Figur 5.1 och Figur 5.2 återger åldermönstret för sysselsättningseffekterna över livscykeln för hela urvalet. Personer med hög kognitiv eller icke-kognitiv förmåga har lägre sysselsättningssannolikhet i början av livet. Det beror antagligen på att de har större sannolikhet att studera vid dessa åldrar. När personerna träder in på arbetsmarknaden vid 20–30 års ålder ökar betydelsen av både kognitiva och icke-kognitiva förmågor för sannolikheten att arbeta kraftigt. Störst är dock ökningen för kognitiva förmågor. Från och med 30 års ålder ligger sysselsättningseffekten av kognitiv förmåga relativt stabilt kring 4 procentenheter. Effekten av icke-kognitiv förmåga ligger också relativt stabilt på ungefär samma nivå.⁶⁴

Tabell 5.5 presenterar effekten av olika förmågor på livsinkomsten (20–58 år). Analysen är begränsad till de som arbetar (dvs. har en real årsinkomst överstigande 135 000 kronor respektive år). Om man rör sig en standardavvikelse högre upp i fördelningen av kognitiva förmågor ökar livsinkomsten med 5,4 %. Motsvarande resultat för icke-kognitiva förmågor är ungefär 3,8 %. Även föräldrabakgrund spelar roll för livsinkomsten, med omkring 1,2 % högre inkomst för barn till föräldrar med ett års längre utbildning.

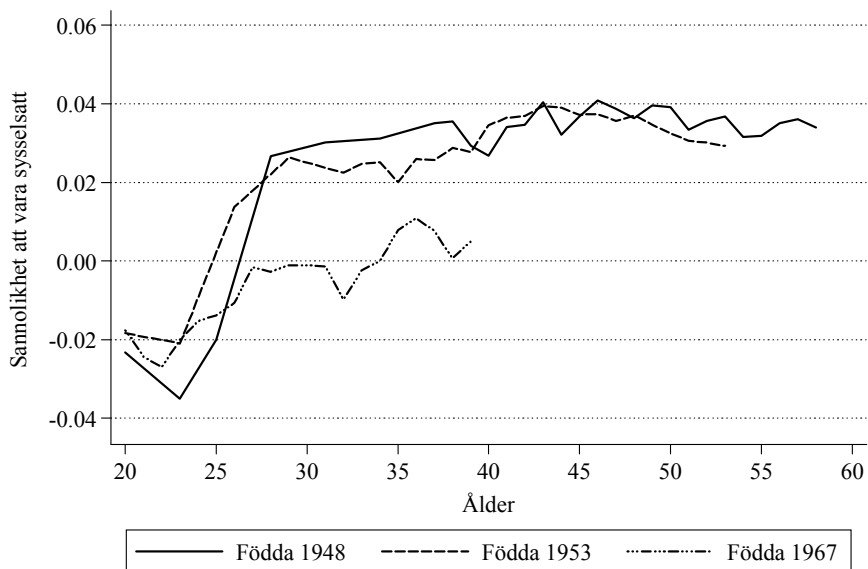
⁶⁴ De svagare sambandet mellan icke-kognitiva förmågor och sysselsättning för personer födda 1967 beror antagligen på att måttet har tagits fram på ett något anorlunda sätt.

Figur 5.1. Begåvning i Åk 6 och sysselsättning över livsrykeln



Not: Figuren visar skillnaden i sysselsättningssannolikheter för personer som befinner sig en standardavvikelse från varandra i fördelningen. Skillnader i aspirationer, föräldrabakgrund och kön har beaktats.

Figur 5.2. Utbildningsaspirationer i Åk 6 och sysselsättning över livsrykeln



Not: Figuren visar skillnaden i sysselsättningssannolikheter för personer som befinner sig en standardavvikelse från varandra i fördelningen. Skillnader i begåvning, föräldrabakgrund och kön har beaktats.

Tabell 5.5. Inkomsteffekter för olika egenskaper över livscykeln (20-58 år)

	Alla	Kvinnor	Män	Lågutbildade föräldrar	Högutbildade föräldrar
Begåvningsprov Åk 6	0,054 (0,002)	0,048 (0,003)	0,059 (0,003)	0,053 (0,002)	0,054 (0,004)
Utbildningsaspiration Åk 6	0,038 (0,002)	0,031 (0,002)	0,044 (0,003)	0,035 (0,002)	0,038 (0,003)
Föräldrarnas utbildningslängd	0,012 (0,001)	0,011 (0,001)	0,014 (0,001)	0,005 (0,001)	0,015 (0,002)
Antal personer	42,450	19,632	21,764	26,959	16,191

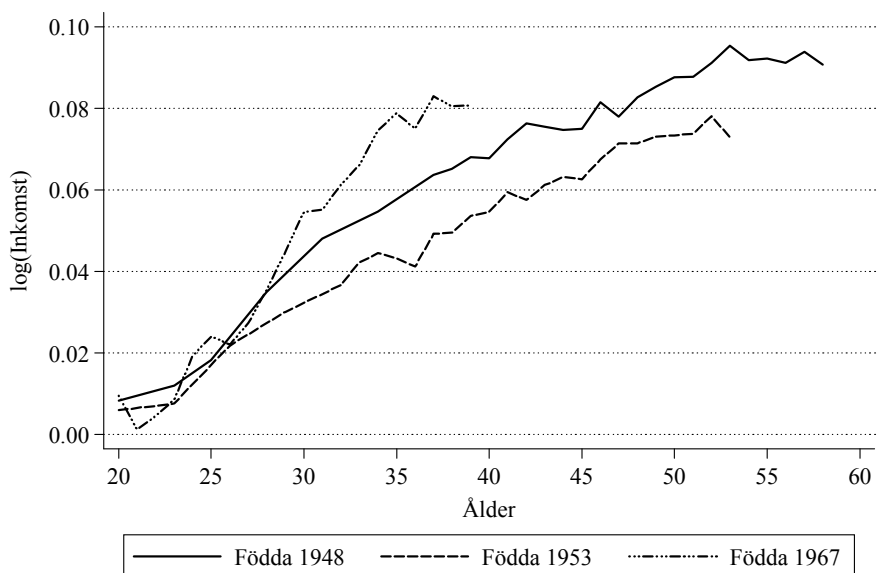
Noter: Tabellen visar regressionsskattningar av sambandet mellan olika egenskaper och inkomsten. Beroende variabel är logaritmen av årsinkomsten för åren 1968, 1971, 1973, 1976, 1979, 1982 samt 1985-2006 (i 2006 års penningvärde). Resultaten baseras på ett urval av personer födda 1948, 1953, 1967, 1972, 1977 och 1982. Föräldrarnas utbildningslängd avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Begåvningsprovet i Åk 6 är summan av tre delprov som avser att mäta verbal, logisk och spatiell förmåga. Utbildningsaspiration avser antalet år som personerna tror att de ska gå i skolan, antingen som svar på en direkt fråga eller som de antal utbildningsår det yrke som de tror sig ha i framtiden motsvarar. För att göra alla variabler jämförbara med varandra och mellan kohorter, har föräldrarnas utbildningslängd, begåvningsprov, betyg och ambition standardiserats för varje kohort. Estimaten visar skillnaden i sannolikheten att vara syselsatt för personer som befinner sig en standardavvikelse från varandra i fördelningen. Förutom de variabler som återges i tabellen innehåller regressionerna ett intercept samt fixa effekter för interaktionen mellan år, kohort och kön. Klusterjusterade standardfel (individer) finns inom parantes.

De två nästkommande kolumnerna delar upp analysen på män och kvinnor. Till skillnad från sysselsättningseffekterna så är inkomsteffekterna av båda förmågemåtten större för män än för kvinnor. Både kognitiva och icke-kognitiva förmågor har således stor betydelse för både män och kvinnors arbetsmarknadssituation, men effekten verkar i större utsträckning via sysselsättning än via löner för kvinnor i förhållande till män. De två sista kolumnerna i Tabell 5.5 visar att skillnaden i inkomstpremier med avseende på familjebakgrund är ganska liten, dock med en tendens till större effekter för barn till högutbildade än till lågutbildade föräldrar. Det beror antagligen på att begåvade barn till högutbildade föräldrar utbildar sig i högre utsträckning än begåvade barn till lågutbildade föräldrar (se avsnitt 10.2).

Figur 5.3 och Figur 5.4 visar hur inkomsteffekterna för kognitiv och icke-kognitiv förmåga utvecklas över livsrymden. Inkomstskillnaderna för personer med olika förmågor är mycket små i början av arbetsmarknadskarriären, vilket kan förklaras av större sannolikhet för personer med bättre förmågor att läsa vidare. Ingen av måtten är förknippade med inkomstskillnader. Inkomsteffekterna för båda förmågorna stiger brant under större delen av yrkeslivet, även om det finns en tendens till svagare tillväxttakt vid högre åldrar. Måttet på kognitiva förmågor värderas högre än måttet på icke-kognitiva förmågor. Vid slutet av yrkeskarriären ligger avkastningen på kognitiva förmågor på omkring 10 %, medan icke-kognitiva förmågor har en inkomstpremie på ungefär 6 %.

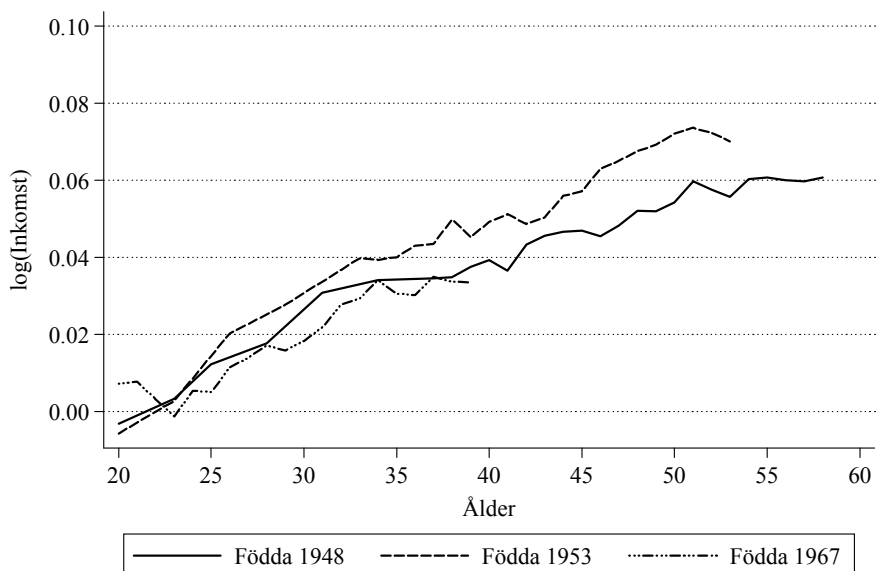
Resultaten från tidigare studier och våra egna beräkningar visar att det finns ett starkt samband mellan mått på både kognitiva och icke-kognitiva förmågor och senare arbetsmarknadsframgång. Det är dock viktigt att återigen poängtera att de samband som presenterats i litteraturen – och i vår översikt – knappast kan göra anspråk på att vara kausala effekter av olika förmågor på arbetsmarknaden. Det finns problem med både utelämnade variabler och mätfel. Vi tycker dock genomgången av litteraturen på området ger stöd för att kunskaper kan ses som intermediära utfall: effekterna av utbildning på kunskaper är informativa om dess effekter på arbetsmarknaden. Vi kommer därför att fortsättningsvis att rapportera både effekter av utbildningspolitik på kunskaper och färdigheter, samt omväxlande presentera effekterna på sysselsättning, löner och inkomster.

Figur 5.3. Begåvning i Åk 6 och inkomster över livsrykeln



Not: Figuren visar inkomstskillnaderna för personer som befinner sig en standardavvikelse från varandra i fördelningen. Skillnader i aspirationer, föräldrabakgrund och kön har beaktats.

Figur 5.4. Utbildningsaspirationer i Åk 6 och inkomster över livsrykeln



Not: Figuren visar inkomstskillnaderna för personer som befinner sig en standardavvikelse från varandra i fördelningen. Skillnader i begåvning, föräldrabakgrund och kön har beaktats.

5.4 Sammanfattning

Vi har i det här kapitlet diskuterat effekten av utbildning på kunskaper och inkomster, samt arbetsmarknadens värdering av de kunskaper som utbildnings-systemet producerar. Den mest trovärdiga evidensen tyder på att avkastningen på utbildning i Sverige är låg. En kvalificerad gissning är att den kausala avkastningen är ungefär hälften så stor som det beskrivande sambandet mellan utbildning och inkomst. Det skulle innebära att ett års utbildning medför omkring 2–3 procents högre lön. Vi tycker oss ha funnit stöd för en något högre avkastning på utbildning vid lägre utbildningsnivåer, och en något lägre avkastning vid högre nivåer. Dessutom tycks personer med svagare familjebakgrund ha mer att vinna på att utbilda sig än personer med bättre bakgrund.

Trots det relativt svaga sambandet mellan utbildning och arbetsmarknadsframgång finns evidens som pekar på att svenska skolor är effektiva på att producera kunskaper. Ett år i skolan förbättrar olika mått på kognitiv förmåga med omkring 0,2 *d*. Även här tycks effekterna vara något större i de första årskurserna, men kunskapseffekterna är betydande även relativt sent i utbildningskarriären.

Vi har också visat att det finns ett starkt samband mellan personers kognitiva och icke-kognitiva förmågor och olika mått på arbetsmarknadsframgång. Både kognitiva och icke-kognitiva förmågor är viktiga för att förklara både sysselsättning och inkomster. I genomsnitt är en standardavvikelse bättre kognitiv förmåga förknippad med 2,5 % större sannolikhet att arbeta för personer i arbetsför ålder samt 5,4 % högre livsinkomster för de som arbetar. Motsvarande samband för vårt mått på icke-kognitiv förmåga är 1,9 procents större sannolikhet att arbeta respektive 3,8 % högre livsinkomster.

Om vi för ett ögonblick antar att sambanden mellan uppmätta kognitiva förmågor och inkomster skulle kunna tolkas kausalt, kan en utbildningseffekt på 0,2 *d* motsvara en inkomsteffekt på 1,1 %. Kanske är skolorna lika bra på att också producera icke-kognitiva kunskaper (vad vi vet saknas helt sådana studier) och man kanske borde lägga till en ytterligare en inkomsteffekt på 0,78 % per utbildningsår. Vi skulle då landa på en avkastning på utbildning på omkring 2 %, vilket är mycket nära den inkomsteffekt som de bästa studierna redovisar. Sammantaget tycks svenska skolor således vara effektiva i att producera kunskaper – i varje fall historiskt sett – men kunskaperna värderas inte fullt ut på arbetsmarknaden.

Referenser

- Angrist, J., och A. Krueger (1991), Does compulsory school attendance affect schooling and earnings?, *Quarterly Journal of Economics*, 106, 979–1014.
- Ashenfelter, O. och A. Krueger (1994), Estimates of the economic return to schooling from a new sample of twins, *American Economic Review*, 84, 1157–1173.
- Ashenfelter, O. och C. Rouse (1998), Income, schooling and ability: evidence from a new sample of twins, *Quarterly Journal of Economics*, 113, 253–284.
- Balke-Aurell, G. (1982). Changes in ability as related to educational and occupational experience. Göteborg: Göteborg studies in educational sciences, 40., Acta Universitatis Gothoburgensis.
- Baltes, P. och G. Reinert (1969), Cohort effects in cognitive development of children as revealed by cross-sectional sequences, *Developmental Psychology*, 1, 169–177.
- Bedard, K. och E. Dhuey (2006), The persistence of early childhood maturity: International evidence of long-run age effects, *Quarterly Journal of Economics*, 121, 1437–1472.
- Behrman, J., M. Rosenzweig och P. Taubman (1994), Endowments and the allocation of schooling in the family and in the marriage market: the twins experiment. *Journal of Political Economy*, 102, 1131–1174.
- Bentin, S. R. Hammer och S. Cahan (1991), The effects of aging and first grade schooling on the development of phonological awareness, *Psychological Science*, 2, 271–274.
- Bisanz, J., F. Morrison och M. Dunn (1995), Effects of age and schooling on the acquisition of elementary quantitative skills, *Developmental Psychology*, 31, 221–236.
- Björklund, A. och M. Jäntti (2009), Intergenerational income mobility and the role of family background, i *Oxford Handbook of Economic Inequality*. Oxford: Oxford University Press.
- Björklund, A. och C. Kjellström (2002), Estimating the returns to schooling in Sweden: How useful is the Mincer equation?, *Economics of Education Review*, 21, 195–210.

- Black, S., P. Devereux och K. Salvanes (2007), From the cradle to the labor market: The effect of birth weight on adult outcomes, *Quarterly Journal of Economics*, 122, 409–439.
- Blackburn, M. och D. Neumark (1993). Omitted-ability bias och the increase in the return to schooling, *Journal of Labor Economics*, 11, 521–544.
- Blackburn, M. och D. Neumark (1995), Are OLS estimates of the return to schooling biased downward? Another look, *Review of Economics and Statistics*, 77, 217–230.
- Bound, J., D. Jaeger och R. Baker (1995), Problems with instrumental variables estimation when the correlation between the instrument och the endogenous explanatory variable is weak, *Journal of the American Statistical Association*, 90, 443–450.
- Bound, J., och G. Solon (1999), Double trouble: on the value of twin-based estimation of the return to schooling. *Economics of Education Review*, 18, 169–182.
- Cahan, S., och N. Cohen (1989), Age versus schooling effects on intelligence development, *Child Development*, 60, 1239–1249.
- Cahan, S., och D. Davis (1987), A between-grades-level approach to the investigation of the absolute effects of schooling on achievement, *American Educational Research Journal*, 24, 1–12.
- Card, D. (1995), Using geographic variation in college proximity to estimate the return to schooling. Ingår i L. Christofides, K. Grant och R. Swidensky (red.), *Aspects of labour market behaviour: Essays in honour of John Vanderkamp*. Toronto: University of Toronto Press.
- Card, D. (1999), The causal effect of education on earnings. Ingår i *Handbook for Labor Economics*, Vol. 3, Ashenfelter, O. och D. Card (red). Amsterdam: Elsevier.
- Carneiro, P. och J. Heckman (2002), The evidence on credit constraints in post-secondary schooling, *Economic Journal*, 112, 705–734.
- Cascio, E och E. Lewis (2006), Schooling och the Armed Forces Qualifying Test: Evidence from School-Entry Laws, *The Journal of Human Resources*, 41, 294–318.

- Cliffordson, C. (2008), Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 95 data. Manuskript, Pedagogiska Institutionen, Göteborgs universitet.
- Cliffordson, C. och J-E. Gustafsson (2008), Effects of age och schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age och length of schooling, *Intelligence*, 36, 143–152.
- Crone, D. och G. Whitehurst (1999), Age och Schooling Effects on Emergent Literacy och Early Reading Skills, *Journal of Educational Psychology*, 91, 604–614.
- DeGroot, A. (1948), The Effects of War upon the Intelligence of Youth, *Journal of Abnormal Social Psychology*, 43, 311–317.
- Duflo, E. (2001), Schooling och labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment, *American Economic Review*, 91, 795–813.
- Edin P-A. och R. Topel (1997), Wage Policy och Restructuring: The Swedish Labor Market since 1960, i R. Freeman, R. Topel och B. Swedenborg (red.), *The Welfare State in Transition*. Chicago: University of Chicago Press.
- Entwisle, D., K. Alexander och L. Olson (1997), *Children, schools, and inequality*, Boulder, CO: Westview Press.
- Erikson, R. och Jonsson, J. (1993), *Ursprung och utbildning: social snedrekrytering till högre studier*, SOU 1993:85.
- Falch, T och S Sandgren (2006), The effect of education on cognitive ability, WP 9/2006, Department of Economics, Norwegian University of Science och Technology.
- Falch, T och S Sandgren (2008), The effect of early cognitive ability on earnings over the life-cycle. WP 1/2008, Department of Economics, Norwegian University of Science och Technology.
- Ferreira, F. och F. Morrison (1994), Children's metalinguistic knowledge of syntactic constituents: Effects of age and schooling, *Developmental Psychology*, 30, 663–678.

- Fredriksson P. och R. Topel (2010), Wage Determination and Employment in Sweden Since the Early 1990s, i R. Freeman, R. Topel och B. Swedenborg (red.), *Reforming the Welfare State*. Chicago: University of Chicago Press.
- Fredriksson, P. och B. Öckert (2008), Är det bättre att börja skolan tidigare?, *Ekonomisk debatt*, 36, 17–32.
- Furu, M. (2000), The Malmö Study, ingår i Seven Swedish longitudinal studies in the behavioral sciences, Rapport 2000:8, Forskningsrådsnämnden, Stockholm.
- Fägerlind, I. (1975). *Formal education and adult earnings: A longitudinal study on the economic benefits of education*. Stockholm: Almqvist & Wiksell.
- Gormley, W. och T. Gayer (2005), Promoting school readiness in Oklahoma: An evaluation of Tulsa's pre-K program, *Journal of Human Resources*, 40, 533–558.
- Green, R., L. Hoffman, R. Morse, M. Hayes och R. Morgan (1964), *The Educational Status of Children in a District without Public Schools*, Office of Education, U.S. Department of Health, Education and Welfare, Washington, D.C.
- Griliches, Z., (1977), Estimating the returns to schooling: some econometric problems. *Econometrica* 45(1), 1–22.
- Griliches, Z. (1979), Sibling models and data in economics: beginnings of a survey. *Journal of Political Economy*, 87, S37–S64.
- Gustafsson, J-E. (2001), Schooling och intelligens: Effects of track of study on level och profile of cognitive abilities, *International Education Journal*, 2, 166–186.
- Gustafsson, J.-E. (2008), Strukturell ekvationsmodellering, ingår i I. Djurfeldt och M. Barmark (red). *Statistisk verktygslåda II*. Lund: Studentlitteratur.
- Hansen, K., J. Heckman and K. Mullen (2004), The effect of schooling and ability on achievement test scores, *Journal of Econometrics*, 121, 39–98.
- Hanushek, E. och L. Woessmann (2008), The role of cognitive skills in economic development, *Journal of Economic Literature*, 46, 607–668

- Harmon, C. och I. Walker (1995), Estimating the Economic Return to Schooling for the United Kingdom, *American Economic Review*, 85, 1278–1289.
- Heyns, B. (1978). *Summer learning and the effects of schooling*. New York: Academic Press.
- Holmlund, H. (2007), A researcher's guide to the Swedish compulsory school reform, Swedish Institute for Social Research, Stockholm University, Working paper 9/2007.
- Holmlund, H., M. Lindahl och E. Plug (2008), The causal effect of parent's schooling on children's schooling: A comparison of estimation methods, IZA DP No. 3630.
- Husén, T (1950), *Testresultatens prognosvärde. En undersökning av den teoretiska skolningens inverkan på testresultaten, intelligenstestens prognosvärde och de sociala faktorernas inverkan på urvalet till högre läroanstalter*. Stockholm: Gebers.
- Husén, T (1951), The influence of schooling upon IQ. *Theoria*, 17, 61–88.
- Husén, T (2002), *Bokslut – essäer om utbildning*. Stockholm: Carlsson Bokförlag.
- Husén, T., och A. Tuijnman (1991). The contribution of formal schooling to the increase in intellectual capital. *Educational Researcher*, 20, 17–25.
- Härnqvist, K. (1959), Intelligensutveckling och skolresultat. *Nordisk Pedagogisk Forskning*, 4, 57–69.
- Härnqvist, K. (1968), Relative changes in intelligence from 13 to 18. *Scandinavian Journal of Psychology*, 9, 50–82.
- Ichino, A. och R. Winter-Ebmer (2004), The long-run educational cost of World War II, *Journal of Labor Economics*, 22, 57–86.
- Isacsson, G. (1999), Estimates of the return to schooling in Sweden from a large sample of twins. *Labour Economics*, 6, 471–489.
- Isacsson, G. (2004), Estimating the economic return to educational levels using data on twins, *Journal of Applied Econometrics*, 19, 99–119.

- Kane, T., och C. Rouse (1993), Labor-market returns to two- och four-year college: Is a credit a credit och do degrees matter? Working Paper no. 311, Industrial Relations Section, Princeton University.
- Karolinska institutet (2005), Svenska tvillingregistret.
- Kling, J. (2001), Interpreting instrumental variables estimates of the return to schooling, *Journal of Business och Economic Statistics*, 19, 358–364.
- Kjellström C. (1999) Omitted ability bias och the wage premium for schooling: new Swedish evidence. I *Essays on investments in human capital*, Ph.D. thesis. Swedish Institute for Social Research, Stockholm University.
- Lemieux, T. och D. Card (2001), Education, earnings och the 'Canadian G.I. bill', *Canadian Journal of Economics*, 34, 313–344.
- Leuven, E., M. Lindahl, H. Oosterbeek och D. Webbink (2010), Expanding schooling opportunities for 4-year-olds, *Economics of Education Review*.
- Lindahl, M. (2001), Summer learning and the effect of schooling: Evidence from Sweden, IZA DP No. 262.
- Lindqvist, E. och R. Westman (2009), The labor market returns to cognitive och noncognitive ability: Evidence from the Swedish enlistment, Working Paper No. 794, Institutet för näringslivsforskning.
- Luyten, H. (2006), An emperical assessment of the absolute effect of schooling: regression-discontinuity applied to TIMSS-95. *Oxford Review of Education*, 32(3), 397–429.
- Luyten, H., J. Peschar och R. Coe (2008), Effects of schooling on reading performance, reading engagement and reading activities of 15-year-olds in England. *American Educational Research Journal*, 45, 319–342.
- Meghir C. och M. Palme (2005), Educational reform, ability och family background, *American Economic Review*, 95, 414–424.
- Miller P., C. Mulvey och N. Martin (1995), What do twins studies reveal about the economic returns to education? A comparison of Australian and U.S. findings. *American Economic Review*, 85, 586–599.
- Mincer, J. (1974), *Schooling, Earnings and Experience*. New York: Columbia University Press.

- Morrison, F., L. Smith och M. Dow-Ehrensberger (1995), Education and cognitive development: A Natural experiment, *Developmental Psychology* 31, 789–799.
- Morrison, F., E. Griffith, och D. Alberts (1997), Nature-nurture in the classroom: Entrance age, school readiness, and learning in children, *Developmental Psychology*, 33, 254–263.
- Murnane, R., J.B. Willett och F. Levy, (1995), The growing importance of cognitive skills in wage determination, *Review of Economics and Statistics*, 77, 251–266.
- Murnane, R. J.B. Willett, M.J, Braatz och Y. Duhaldeborde, (2001), Do different dimensions of male high school students' skills predict labour market success a decade later? Evidence from the NLSY, *Education Economic Review*, 20, 311–320.
- Neal, D. och W. Johnson (1996), The role of pre-Market factors in black-white wage differences, *The Journal of Political Economy*, 104, 869–895.
- Neumark, D. (1999), Biases in twin estimates of the return to schooling: a note on recent research, *Economics of Education Review*, 18, 143–148.
- Nordin, M. (2008), Ability and the rates of return to schooling – making use of the Swedish enlistment battery test, *Journal of Population Economics*, 21, 703–717.
- Oreopoulos, P. (2006), Estimating average och local average treatment effects in education when compulsory schooling laws really matter, *American Economic Review*, 96, 152–175.
- Rouse, C. (1999), Further estimates of the economic return to schooling from a new sample of twins, *Economics of Education Review*, 18, 149–157.
- Sandewall, Ö., D. Cesarini och M. Johannesson (2009), The co-twin methodology och returns to schooling – testing a critical assumption, IFN Working Paper, No. 806.
- SCB (1997), Kvalitetsdeklaration Utbildningsregistret 1997-01-16.
- Skedinger, P. (2006), Svenska minimilöner i den globaliserade ekonomin, *Ekonomisk debatt*, 34, 63–77.

- Varnhagen, C. K., F.J. Morrison, och R. Everall (1994), Age and schooling effects in story recall and story production, *Developmental Psychology*, 30, 969–979.
- Webbink, D. (2007), Returns to university education: Evidence from a Dutch institutional reform, *Economica*, 74, 113–34.
- Winship, C., och Korenman, S. (1997). Does staying in school make you smarter? The effect of education on IQ in the Bell curve. I B. Devlin, S. E. Fienberg, D. P. Resnick, och K. Roeder (Red.), *Intelligence, genes, & success. Scientists respond to the bell curve* New York: Springer-Verlag.
- Zax, J. och D. Rees (2002), IQ, academic performance, environment, and earnings, *Review of Economics och Statistics*, 84, 600–616.
- Zetterberg, J. (2005), Swedish evidence on the impact of cognitive och non-cognitive ability on earnings – an extended pre-market factor approach, WP 16/2005, IFAU.
- Öckert, B (2009), Does going to high school make you smarter?, manuskript, IFAU.
- Öckert, B (2010), What's the value of an acceptance letter? Using admissions data to estimate the return to college, *Economics of Education Review*, 29, 504–516 .

6 Daghem och förskola: vad vet vi om effekterna för barnen?

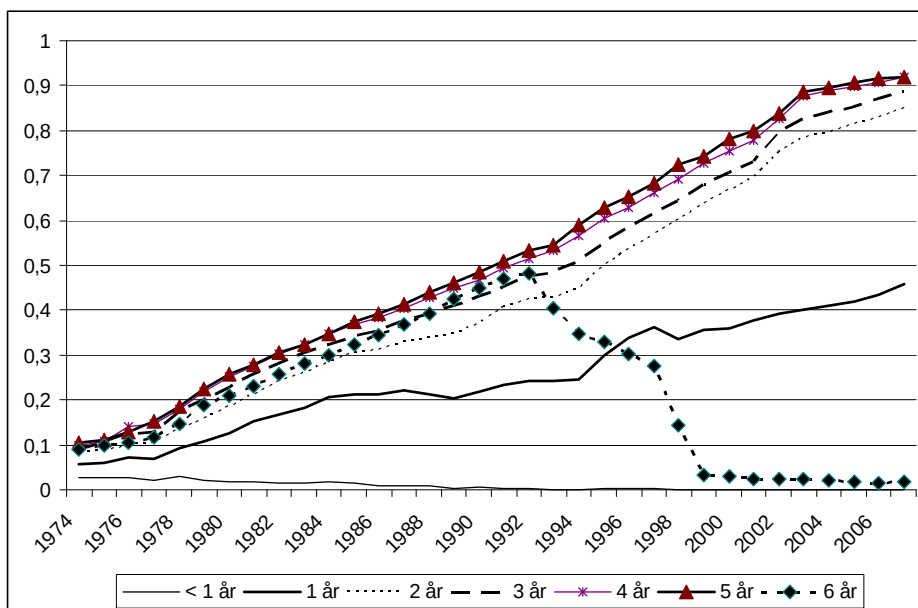
Från slutet av 1960-talet och fram till början av 1990-talet skedde en snabb utbyggnad av den svenska förskolan, eller av barnsomsorgen som det omväxlande kallas. Under en period av ca. 25 år ökade antalet platser från ett fåtal tusen till 350–400 000. Expansionen skedde i huvudsak i kommunal regi och innehöll en blandning av platser på familjedaghem ("dagmamma") och daghem, med successivt ökande vikt för det senare. Under samma tidsperiod förlängdes föräldrapenningen till ca. ett år varför barnen under sitt första år vanligen kom att vårdas av en förälder. Sedan mitten av 1990-talet har antalet platser stagnerat just under 400 000. Samtidigt infördes dock skolformen förskoleklass för sexåringar år 1998, varför man sedan dess bör relatera de nära 400 000 platserna till fem årskullar av barn från ett till fem år.

Andelen barn i olika åldrar som varit inskrivna i förskolan under åren 1974 till 2007 presenteras i Figur 6.1. Vi ser att för barn i åldern två till fem år har andelen stigit från omkring 10 % i mitten av 1970-talet till nära 90 % cirka 30 år senare. För 6-åringar noteras en nedgång i samband med införandet av förskoleklass. För 1-åringar når talen bara upp till knappt 50 % medan andelen barn under ett år varit mycket lågt. Det senare sammanhänger med att den betalda föräldraledigheten varit omkring, eller just över, ett år.

Ett annat sätt att belysa utvecklingen av barnsomsorgen, liksom av lekskolan, är att använda data från Levnadsnivåundersökningarna (LNU) där personer födda mellan 1966 och 1981 fått frågan om hur länge de som barn deltog i olika barnomsorgsformer inklusive lekskola.⁶⁵ Figur 6.2a visar den förväntade uppgången av andelen med erfarenhet av daghem, familjedaghem eller någon av dessa båda former. Figur 6.2b visar också att de som faktiskt deltagit i dessa förskoleformer har deltagit under successivt allt längre tid. Det är dock intressant att talen för deltagande i lekskola är så pass höga. För alla åldersgrupperna är andelen som deltagit i sådan kring 80 % (Figur 6.2a) och i genomsnitt har tiden varit ett drygt år (Figur 6.2b). Det är därför en tämligen låg andel barn födda från 1966 och framåt som helt saknar erfarenhet av någon förskoleform.

⁶⁵ Jonsson (2004) innehåller både information om LNU och utbyggnaden av barnsomsorgen.

Figur 6.1. Andelen barn i olika åldrar som är inskrivna i förskola, 1974–2007

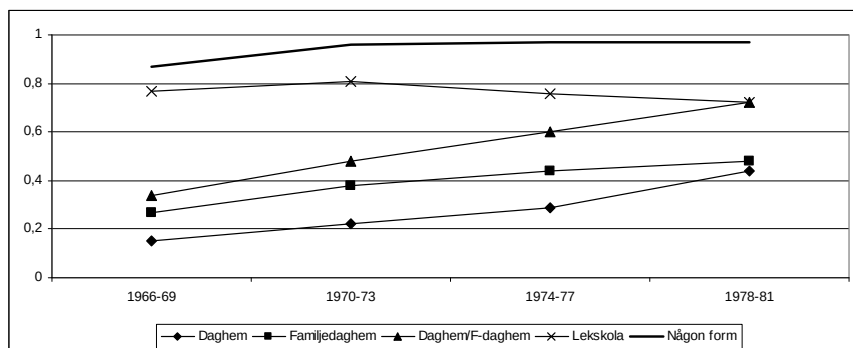


Källa: Statistiska centralbyrån.

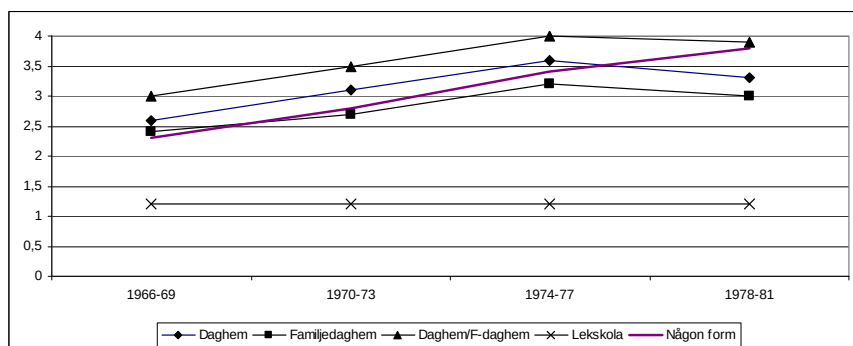
Förskolans utbyggnad innebär en stor samhällsförändring under en relativt kort tid. Även i ekonomiska termer är det frågan om en kraftfull satsning. Föräldrarnas avgifter har som regel bara täckt cirka 10 % av kommunernas kostnader för verksamheten. Utbyggnaden har motiverats med olika argument. Ett viktigt syfte har varit att underlätta kvinnors förvärvsarbete. Detta har lett till att barn med högutbildade föräldrar kommit att delta i verksamheten mer än andra. Ett annat syfte har varit socialpolitiskt, nämligen att underlätta för ensamstående föräldrar att försörja sig med hjälp av förvärvsarbete. Detta har lett till att barn med ensamstående föräldrar också tenderat vara överrepresenterade inom barnomsorgen. Båda dessa mönster framgår av Figur 6.2c, som dock också visar att dessa skillnader minskat över tiden. Ibland har även argumentet att subventionerad barnomsorg leder till högre födelsetal framförts.⁶⁶

⁶⁶ Se Mörk m.fl. (2009) för en analys, som tyder på förekomsten av sådana effekter, och en litteraturgenomgång.

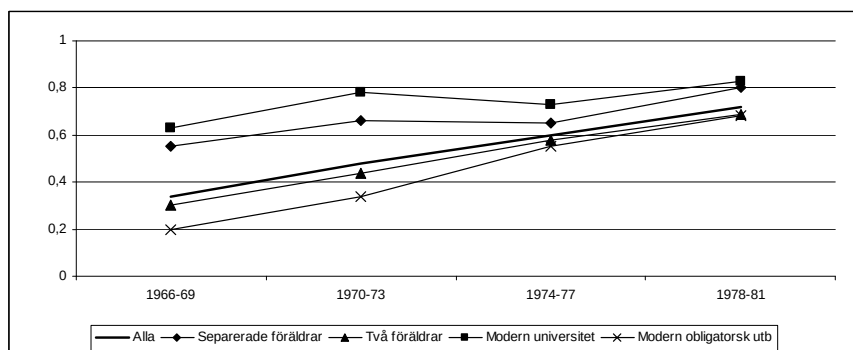
Figur 6.2a. Andelen barn med någon erfarenhet av barnomsorg, efter typ av barnomsorg och födelseår



Figur 6.2b. Genomsnittligt antal år i barnomsorg, efter typ av barnomsorg och födelseår



Figur 6.2c. Andelen barn med någon erfarenhet av daghem eller familjedaghem, efter födelseår och föräldrabakgrund.



Källa: Levnadnivåundersökningarna 1991 och 2000.

Men framför allt har ambitionerna varit höga utifrån barnens perspektiv. Ett tecken på detta är att det centrala politiska ansvaret 1996 överfördes från Socialdepartementet till Utbildningsdepartementet. År 1998 kom den första Läroplanen för förskolan (Lpfö 98). Denna är visserligen tämligen allmänt hållen och inleds med en betoning av förskolans värdegrund, men likafullt betonas också lärandet och att förskolan ska ge en bra grund för det fortsatta skolarbetet. Följande stycken från läroplanen visar på sådana ambitioner:

”Förskolan skall lägga grunden för ett livslångt lärande”

”Förskolan skall sträva efter att varje barn utvecklar sitt ord- och begreppsförråd... och sitt intresse för skriftspråk”

”Förskolan skall sträva efter att varje barn utvecklar sin förmåga att upptäcka och använda matematik i meningsfulla sammanhang”

”Vid övergången till nya verksamheter (dvs. skolan, vår anm.) skall särskild uppmärksamhet ägnas åt de barn som behöver särskilt stöd.”

Mot denna bakgrund är det naturligt att försöka utvärdera förskolan efter dess effekter på de fortsatta prestationerna i skolan, även om andra utfall också är av intresse. Det är också naturligt att undersöka om förskolan bidragit till en social utjämning genom att kompensera barn från en ur studiesynvinkel mindre gynnsam hemmiljö.

Även om det i allt väsentligt rått politisk enighet om utbyggnaden av den kommunala barnomsorgen har kritiska röster då och då höjts mot denna satsning. Ett exempel är Per Kågessons bok från 2006. Detta är en personligt färgad bok där författaren driver tesen att det finns många hälsorisker med tidig placering av barn i omsorg utanför det egna hemmet, särskilt före tre års ålder. Författaren går igenom en stor mängd litteratur som han menar ger underlag för denna bedömning. De flesta svenska studier som han åberopade kommer vi också att granska i detta kapitel.

Vi börjar kapitlet med att diskutera vilka frågor som vi idealt skulle vilja ha svar på när det gäller förskolans effekter på barnens utveckling i skolan och andra relaterade utfall. Därefter diskuterar vi i avsnitt 6.2 de metodproblem som uppstår då man vill beräkna dessa effekter. I avsnitt 6.3 går vi igenom de svenska studier som gjorts på området, vi diskuterar både de empiriska resultaten och studiernas trovärdighet. Därefter följer i avsnitt 6.4 en översikt av resultaten från studier från andra länder. I avsnitt 6.5 går vi igenom studier av en besläktad fråga, nämligen effekten av tidig respektive sen skolstart. Vi avslutar med en något nedslående sammanfattning i den meningen att vi anser

att våra kunskaper är mycket bräckliga i dessa viktiga frågor, särskilt vad avser Sverige. Vi diskuterar också vilken typ av studier som skulle ha kunnat genomföras på området för att ge oss svar på en del av de centrala frågorna.

6.1 Vad vi vill veta

En första fråga är: effekter på vad? Med tanke på de mål som finns för den svenska förskolan är det rimligt att undersöka effekterna på kognitiva färdigheter som att kunna läsa, räkna och skriva, eller olika indikatorer på förmågan att tillgodgöra sig sådana färdigheter som ordförråd, kommunikationsförmåga och förståelse för enkla matematiska begrepp. En utvärderingsstudie måste därför ha tillgång till mått på sådana färdigheter. Man kan också argumentera för att förskolan kan ha effekter på sådana kognitiva färdigheter både på kort och lång sikt. Kortsiktiga utfall är naturliga att mäta i samband med skolstarten då daghemsperioden normalt avslutas. Men det finns också anledning att undersöka om det finns långsiktiga effekter på skolresultaten längre fram i åldrarna. Även det ”slutliga” utfallet i form av framgång på arbetsmarknaden är naturligt att undersöka. Sådana långsiktiga effekter ligger i linje med den dynamiska modell för inläring som vi diskuterade i kapitel 3.

Även om kognitiva färdigheter kan sägas vara ett slutligt mål för en väl fungerande förskola, kan det vara svårt att mäta sådana på ett tillförlitligt sätt redan kring skolstarten. Därför framstår andra utfall som skolmognad och social förmåga att fungera tillsammans med skolkamrater som rimliga utfallsvariabler om man ska studera kortsiktiga utfall. Även hälsovariabler och infektionskänslighet är relevanta då det i debatten hävdats att daghemsvistelse kan påverka sådana utfall. Skolmognad och hälsa torde sedan i ett längre perspektiv påverka skolresultaten och de kognitiva färdigheterna.

En andra metodfråga är vilket jämförelsealternativ vi ska ha. Med effekt menar vi skillnaden i utfall mellan två väldefinierade alternativ, i detta fall daghem och något annat alternativ där barnet skulle ha vistats fram till skolstart om inte daghem funnits eller föräldrarna gjort ett annat val. Det är långt ifrån klart vilket detta alternativ är. I vissa fall kan alternativet vara att vistas i hemmet med en förälder (vanligen mamman) eller en ”barnflicka”, i andra fall privat barnomsorg, och i ytterligare andra fall kan det vara frågan om att någon släkting tar hand om barnet. För en övergripande utvärdering av den svenska förskolan är det rimligt att tro att alternativet är en blandning av dessa omsorgsformer. De alternativ som står mot varandra är sannolikt, å ena sidan, två arbetande föräldrar och att barnet går på förskola, och, å den andra sidan, både

en hemmavarande mor och en (eventuellt deltidsarbetande) arbetande mor i kombination med någon annan barnomsorg än förskola.

Ett fall där alternativen kan framstå som klarare är när man i det svenska fallet vill jämföra förlängd föräldraledighet med senarelagd rätt till daghem. Ett exempel, som vi återkommer till, är en reform som genomfördes 1988 och som innebar att den betalda föräldraledigheten förlängdes med tre månader och rätten till daghemsplats i praktiken flyttades lika långt fram i tiden.

Även när man begrundar effekter av tidigarelagd skolstart är det viktigt att precisera sina jämförelsealternativ. Ett tänkbart alternativ till tidigare skolstart är just daghem, men det är också möjligt att privat barnomsorg eller vistelse i hemmet med en förälder är alternativ för många 5- eller 6-åringar som skulle börja skolan vid en tidigare skolstart.

En tredje fråga är: effekter för vilka barn? Som framgick av inledningen har förskolan som uppgift att ägna särskild uppmärksamhet åt de barn som behöver extra stöd och underlätta övergången till skolan för dessa. Mot denna bakgrund är det angeläget att undersöka om förskola har särskilt gynnsamma effekter för barn med olika problem eller som kommer från en resursmässigt svagare socioekonomisk bakgrund. Det finns således ingen anledning att utgå från att en förskolesatsning har en och samma effekt på alla barn. Vi diskuterar detta närmare från ett metodperspektiv i nästa avsnitt.

För det fjärde finns många delfrågor kring utformningen av förskoleverksamhet och barnomsorg. Hit hör vid vilken ålder det är lämpligt att börja i någon form av förskola eller barnomsorg utanför hemmet. Hur många timmar per dag eller per vecka är lämpligt? Vad betyder familjedaghem i förhållande till daghemsgrupp? Vad betyder personalens kompetens och personaltätheten?

Sammantaget finns det således många delfrågor kring effekterna av förskola och det gäller framför allt att precisera vilka alternativ man jämför med. När vi senare går igenom olika studier kommer vi att diskutera vilka delfrågor som studierna i bästa fall ger oss svar på.

6.2 Metodproblem

Även med ett väldefinierat alternativ till daghem och med en tydlig utfallsvariabel är det svårt att skatta effekten av daghem. Ett antal metodproblem aktualiseras och de beror väsentligen på vad vi tror om karaktären på de effekter som vi är intresserade av och på hur selektionen av barn till daghem har gått till. Vi ska nu först allmänt diskutera dessa problem och återkommer sedan till dem när vi går igenom olika specifika studier.

Anta först att de effekter vi är intresserade av är tämligen homogena i den meningen att alla barn påverkas ungefär lika, med undantag för att effekterna kan variera något beroende på observerade egenskaper som barnets kön och föräldrarnas utbildning. Säg nu att vi vill skatta effekterna genom att jämföra utfallet för en grupp barn som deltagit i daghemsverksamhet med en grupp som inte deltagit. Vårt metodproblem består då av att beakta att daghemsbarnen kan ha andra egenskaper som påverkar utfallet än jämförelsegruppen. Det är då naturligt att försöka hantera detta problem med hjälp av regressionsanalys eller matchningstekniker så att man kontrollerar för skillnader i form av exempelvis föräldrars utbildning, yrke och inkomster, familjestruktur (såsom antal syskon och förekomst av separation i familjen), region med mera. Skillnaderna i effekter kan hanteras med hjälp av så kallade interaktionseffekter mellan daghemsvariabeln och de aktuella variablerna. Denna ansats låter sig genomföras med variabler som är mätbara och som forskaren har tillgång till i sin studie. Det är möjligt att dessa variabler kan räcka ganska långt som kontroller, men det är alltid svårt att argumentera för att de beaktar alla relevanta skillnader i förutsättningar mellan barn som gått på daghem och de som inte gjort det.

Anta nu i stället att effekterna av daghem inte är homogena, utan varierar mellan barn på ett sätt som inte fångas upp av några mätbara variabler, dvs. att även bland barn med samma uppmätta familjebakgrund så varierar effekterna en hel del. Anta också att föräldrarna (åtminstone delvis) känner till dessa skillnader och väljer omsorgsform för sitt barn på basis av denna kunskap. Föräldrar som känner på sig att daghem passar för deras barn tenderar att välja detta alternativ och föräldrar som gör den motsatta bedömningen väljer att själva ta hand om barnet.⁶⁷ Om dessa förutsättningar är uppfyllda förvärras metodproblemet betydligt. Intuitivt är det lätt att inse varför. Vi vill använda utfallet för barn som *inte* gått på daghem för att bedöma hur det *skulle* ha gått för daghemsbarnen om daghem inte hade funnits. Men vi tvingas då skatta effekten av daghem genom att jämföra utfallen mellan barn som gått på daghem, och för vilka föräldrarna gjorde bedömningen att daghem var bäst, och barn som inte gått på daghem, och för vilka föräldrarna gjorde bedömningen att just detta alternativ var det bästa.

Ytterligare en konsekvens av sådan heterogenitet i effekterna i kombination med selektion efter effekternas storlek är att begreppet ”genomsnittlig effekt av

⁶⁷ I metodlitteraturen kallas detta problem ”essential heterogeneity”. Se Heckman m.fl. (2006) för mer ingående analys av de problem som uppkommer och de möjligheter som finns att med hjälp av olika typer av information skatta relevanta effekter.

vistelse på daghem” måste nyanseras. Utgångspunkten är ju att det finns påtagliga skillnader i effekternas storlek mellan olika barn. Vi måste då vara noga med för vilken grupp vi definierar genomsnittet. Ett begrepp är den genomsnittliga effekten av daghem för dem som gått där. Ett annat begrepp är den genomsnittliga effekten för dem som inte gått på daghem om de hade valt att delta (exempelvis till följd av lägre avgifter). Det är rimligt att den förra effekten är större än den senare. I en sådan situation blir det också högst relevant att skatta effekter för dem som på marginalen skulle välja att delta i förskola om t.ex. antalet platser skulle öka eller avgifterna skulle sänkas. En sådan analys kräver dock information som sällan finns tillgänglig, nämligen information om faktorer som på marginalen bidragit till att vissa barn gått på daghem och vissa inte gjort detta.

Frågan är då om det finns stor anledning att förvänta sig sådan heterogenitet i effekterna i kombination med val baserade på effekternas storlek. Vi menar att så är fallet när det gäller barnomsorg. Betänk vad vi menar med effekt av daghem: det är skillnaden i utfall mellan alternativen att delta i verksamheten och att vara hemma med en av föräldrarna. Det betyder att även om själva daghemsverksamheten är likartad för de flesta barn så kan alternativet skilja sig betydligt. Alternativet med föräldraomsorg kan fungera bra om en av föräldrarna är motiverad att stanna hemma och vårda barnet, medan det kan fungera illa om ingen av föräldrarna är utrustad med det tålamod som krävs härvidlag.

Ytterligare ett metodproblem uppstår då man vill studera effekter på längre sikt, dvs. längre fram i tiden än vid övergången från daghem till skola. Som vi noterade kapitel 3 finns anledning att tro att exempelvis en lyckosam förskoleperiod kan påverka inlärandet längre fram i livet vilket innebär att effekterna av daghem fortsätter långt efter barnen har lämnat verksamheten. Under vissa omständigheter kan det vara svårt att skatta sådana effekter med hjälp av individdata över barn som gått respektive inte gått på daghem. Ett exempel är om det finns så kallade kamrateffekter i skolarbetet. Sådana kan vara både positiva och negativa, men om vi diskuterar fallet med positiva kamrateffekter är innebörden att de som fått en bra start genom förskola hjälper övriga elever som förebilder eller på annat sätt. Ett annat exempel är ifall skolan verkar kompensatoriskt så att lärarna inte behöver lägga ned lika mycket tid på elever som är väl förberedda för skolan genom en bra förskola och att de i stället kan fokusera på övriga barn. I så fall kommer skillnaderna mellan förskolebarnen och övriga att minska över tiden trots att det föreligger gynnsamma effekter. Ett botemedel i detta fall kan vara att genomföra analysen på en annan aggre-

gationsnivå än på individnivå. Om det exempelvis finns påtagliga skillnader mellan kommuner i hur mycket de har satsat på förskola kan dessa skillnader användas för att spåra de sammantagna effekterna av förskola. Denna analysstrategi kan dock vara förknippad med andra problem.

6.3 Svenska studier

En tidig studie av lekskolans betydelse genomfördes av Stukát (1966). Hans urval var relativt litet, cirka 250 barn i Göteborg som började skolan höstterminen 1960. Dessa var lika fördelade mellan barn som deltagit i lekskola och en jämförelsegrupp av barn som inte gjort detta. Syftet var att genom en jämförelse mellan dessa grupper skatta effekten av att ha gått i lekskola. Stukát kunde välja sin jämförelsegrupp bland dem som stod i Göteborgs kommuns kölista för att få delta i lekskola. Vid denna tid var det betydligt fler som anmälde sitt intresse för lekskola än som fick plats att delta. Kötiden användes tämligen strikt för att välja ut vilka barn som fick delta. Detta betyder att studiens jämförelsegrupp består av barn vars föräldrar velat att barnen skall gå i lekskola men förhindrats till detta av bristen på platser och kösystemet. Detta är en klar styrka med studiens upplägg. Man kan dock inte utesluta att de föräldrar som anmälde sina barn tidigt har barn med andra förutsättningar än övriga. Därför genomfördes också en parvis matching med avseende på barnets ålder, kön, antal syskon, bostadsort samt faderns yrke. Huruvida den lika fördelningen av dessa variabler mellan lekskolegruppen och jämförelsegruppen är tillräcklig för eliminera skillnaderna mellan tidigt och sent anmälda barn är dock en öppen fråga.

Utfallen var av flera slag, såsom ”skolberedskap”, ordförråd, språklig förmåga, orienterings- och räkneförmåga, variabler som samlades in i samband med skolmognadsprov våren 1960 och under det första skolarbetet. Skillnaderna var överlag mycket små och icke signifikanta, men i några fall till lekskolebarnens fördel. I slutet av åk 3 genomfördes en uppföljningsundersökning genom att be lärarna bedöma elevernas prestationer. Analysen av dessa data visade inga eller små kvarstående skillnader.

Den första vetenskapliga studien av den från 1960-talets slut utbyggda daghemsverksamheten och andra omsorgsformer genomfördes av Andersson (1989, 1992). Han följde upp 128 barn födda 1975 och 1976 från åtta låg- och medelklassområden i Stockholm och Göteborg fram till 8 respektive 13 års ålder. Urvalen gjordes vid 3–4 års ålder. Med hjälp av intervjuer av föräldrarna kunde tiden i placering i olika barnomsorgsformer fram till skolstarten mätas.

Vidare inhämtades information om barnens prestationer i skolan från deras lärare vid 8 respektive 13 års ålder. Data över både kognitiva färdigheter och social anpassning i skolan samlades in. Andersson använde sedan regressionsanalys där han kontrollerade för bakgrundsvariabler såsom moderns utbildning (4 grupper), familjens sociala klass (3 grupper), familjetyp (en eller två föräldrar i hushållet) och förändring i familjetyp. De centrala barnomsorgsvariablerna specificerades med fyra variabler: ålder vid placering i omsorg utanför hemmet respektive tre former av barnomsorg (daghem, familjedaghem och blandad form). Överlag finner han att ju tidigare barnet börjar i omsorg utanför hemmet – och därmed i praktiken ju längre tid de vistats på daghem – desto bättre går det i skolan.⁶⁸ Däremot finner han inga signifikanta skillnader mellan de olika formerna för barnomsorg. Resultaten är i huvudsak desamma vid 8 och 13 års ålder.

En annan långsiktig uppföljningsstudie har genomförts av Broberg m.fl. (1997). De följde upp 146 barn som under 1982 och 1983 identifierades i barnomsorgskön hos Göteborgs kommun vid 12–24 månaders ålder. Uppföljningar ägde rum vid fyra tillfällen fram till ca. 8 års ålder. Mätningar av både kvaliteten i den omsorg barnen fick och barnens verbala förmåga genomfördes av utbildade psykologer. Även matematisk förmåga uppmättes vid skolstarten.

De empiriska skattningarna är något svåra att tolka då de inte presenteras i form av regressionskoefficienter som anger förväntad skillnad i utfallsvariablerna mellan barn som deltagit i daghemsverksamhet (av olika längd) och de barn som inte deltagit. Det framgår heller inte hur ransoneringen av platser bland dem som står i barnomsorgskön har gått till. Helhetsintrycket från studien är dock att längre vistelse på daghem överlag är förknippad med bättre resultat i de kognitiva utfallen.

Jonsson (1994) studerade de personer som ingick i Levnadsnivåundersökningen (LNU) 1991 och var födda mellan 1965 och 1973, sammanlagt knappt 600 personer. Hans utfallsvariabel var val av teoretisk linje på gymnasiet, dvs. ett ufall längre fram i utbildningskarriären än de som undersökts tidigare. Valet av linje på gymnasiet beror rimligen inte bara på skolresultat utan också på barnets och föräldrarnas preferenser för högre utbildning. Detta understryker vikten av att beakta faktorer som fångar upp sådana preferenser, exempelvis föräldrarnas egen utbildning samt sociala och ekonomiska ställning. När Jonsson beaktade faderns och moderns utbildning, familjens socialklass och

⁶⁸ Alla indikatorer på kognitiva färdigheter och anpassning i skolan uppvisar inte koefficienter som innebär signifikant positiva effekter, men inga punktskattningar är negativa.

ortstyp fann han för pojkar ett svagt signifikant positivt utfall för ”år på daghem”, dvs. ju fler år på daghem under förskoleåldern desto högre sannolikhet för teoretisk linje på gymnasiet. Däremot var effekten inte signifikant för flickor. Jonsson undersökte också om det finns belägg för att daghem verkar utjämnande, i meningen att effekten av daghem är större för barn från familjer med låg utbildning och samhällsklass än för andra barn. Han fann dock inget stöd för sådana utjämnande effekter.

I Jonsson (2004) utvecklades analysen med hjälp av LNU 2000 till att innehålla även barn födda till och med 1981; urvalet steg då till sammanlagt 1470 personer. Materialet tillåter då en intressant beskrivning av selektionen till daghem.⁶⁹ Det visade sig att denna hade förändrats över tiden. För åldersgrupperna födda 1966–69 fanns inga påtagliga skillnader i benägenheten att delta i daghemsverksamhet beroende på föräldrarnas socio-ekonomiska status eller utbildning. För åldersgrupperna födda 1970 och framåt är däremot barn till högre tjänstemän och universitetsutbildade klart överrepresenterade. I dessa avseenden har således selektionen förändrats över tiden. Ett stabilt mönster i selektionen till daghem är dock att barn till separerade föräldrar varit överrepresenterade för alla ålderskohorterna 1966–81.

I denna studie använde Jonsson också en regressionsmodell med fler kontrollvariabler än den första studien. Utöver föräldrarnas utbildning, socio-ekonomiska status och ortstyp under uppväxten tillkommer indikatorer på familjens ekonomi under uppväxten och om föräldrarna har separerat liksom en variabel för antalet syskon. Båda könen studeras i samma ekvation med en nivåskillnad för kön. Däremot är utfallsvariabeln densamma, nämligen val av teoretisk linje på gymnasiet. Med dessa kontrollvariabler fick Jonsson en positiv effekt av ”år på daghem”. Punktskattningen tolkas av Jonsson som relativt svag. Även i denna studie undersökte Jonsson om det fanns tecken på att effekterna verkar utjämnande. De interaktionseffekter som skulle kunna tolkas som utjämnande effekter var dock inte signifikanta.

En liknande studie som Jonssons men med ett annat datamaterial och ett annat utfallsmått har genomförts av Söderström m.fl. (1999). De använde UGU-data från den pedagogiska institutionen vid Göteborgs universitet. Mer specifikt använde de ett urval av cirka 6000 elever födda 1967 som genomgick tester i spatial, verbal och induktiv i slutet av årskurs sex; intervjuer med föräldrarna ger information om deras utbildning, yrke, bostadsort och familjetyp samt om barnens deltagande i daghems- och lekskoleverksamhet. Med

⁶⁹ Det inledande avsnittet i detta kapitel bygger på samma data.

detta material genomförde Söderström m.fl. regressionsanalyser med den sammanlagda poängen på de tre testerna som beroende variabel och erfarenhet av daghems- respektive lekskoleverksamhet som oberoende variabel av främsta intresse. Dessutom ingick föräldrarnas yrke/utbildning, familjetyp (separerade respektive sammanboende föräldrar) och bostadsort som kontrollvariabler.

Skattningarna gav positiva koefficienter för variabler som indikerade deltagande i daghems- respektive lekskoleverksamhet i förhållande till att inte ha deltagit i sådan verksamhet. Dessutom var koefficienten för lång (över två år) större än koefficienten för kort daghemserfarenhet (kortare än två år). Koefficienterna är dock relativt små (som högst cirka 1,8 extra poäng för deltagande i lekskola att jämföra med en medelpoäng kring 70) och bara denna skattning är signifikant skild från noll på fem procents signifikansnivå. Sammantaget är resultaten således i linje med dem som rapporterades av Jonsson (1994, 2004).

Nordström Skans och Liu (2010) genomförde en annan typ av studie av daghem. De utnyttjade en reform som genomfördes 1988 då den betalda föräldraledigheten förlängdes från 12 till 15 månader. Förlängningen genomfördes under en övergångstid på tre månader så att föräldrar till barn födda i augusti fick en extra betald månad, de med barn födda i september fick två extra månader och de med barn födda från och med oktober fick tre extra betalda månader. Studien baserades på svenska registerdata och utnyttjade därför skolbetyg och nationella prov i årskurs nio som utfallsvariabler.

Reformens upplägg (i kombination med svenska registerdata) medgav en studie som dels har en tydlig tolkning av effekterna, dels bygger på en trovärdig metod. För det första visar författarna (med olika data) att den förlängda föräldraledigheten i allt väsentligt innebar en senareläggning av daghemsstart och senare återgång till arbetet för mödrarna. Detta innebär att det är en studie där förlängd föräldraledighet jämförs med alternativet daghem och arbetande. För det andra kan författarna utnyttja den variation som skapades av reformen. De måste dock också beakta att variationen sammanfaller med födelsetid under året, något som görs genom att utnyttja data från närliggande år för att ”säsongkorrigera” skolresultaten.

Resultaten visar att reformen inte hade några genomsnittseffekter. Skattningarna är små och har hög precision. Däremot fann de svagt positiva effekter av förlängd föräldraledighet för barn till högutbildade mödrar. De använde registerdata för att försöka förklara via vilka mekanismer som effekten för dessa barn uppstår, men fann ingen entydig förklaring. Sammantaget framstår denna studie som den kanske mest övertygande rent metodmässigt av svensk

daghemsvistelse. Den övertygande metodansatsen erhålls dock på bekostnad av en mer avgränsad tolkning, nämligen som en förändring på marginalen av daghemsvistelse (eller omvänt av förlängd föräldaledighet) i åldern 1 till 1,5 år.

6.4 Studier från andra länder

I övriga nordiska länder har en motsvarande utbyggnad av förskoleverksamhet som den i Sverige genomförts, om än något senare än i vårt land. Resultaten från tillförlitliga effektstudier från dessa länder skulle möjligen kunna överföras till Sverige. Datta Gupta och Simonsen (2010) redovisar en intressant dansk studie där barn födda 1995 följs upp till åtta års ålder.⁷⁰ Vid denna ålder intervjuas (vanligen) modern, som besvarar frågor om förekomsten av olika problem hos barnen; dessa problem avser sådant som uppförande, hyperaktivitet och kamratrelationer. Svaren på dessa problem bildar ett index och är studiens utfallsvariabel. Detta utfall relateras sedan till vistelse i kommunal barnomsorg, uppdelad i daghem respektive familjedaghem, samt omsorg hos någon av föräldrarna. Den första analysen liknar de svenska studierna med kontroller i en regressionsansats för variabler såsom föräldrarnas utbildning och ålder. Därtill kommer kontroll för exempelvis födelsevikt som de svenska studierna inte hade tillgång till. I denna analys jämförs de båda kommunala omsorgsformerna med alternativet omsorg hos föräldrarna. Resultatet är då att familjedaghem är förknippat med signifikant fler beteendeproblem än omsorg i hemmet, medan en sådan skillnad inte finns för daghem. Effekten uppgår till ca. 0,3 *d*, men är trots över 4 000 observationer skattad med relativt låg precision. Vidare görs en jämförelse mellan olika lång vistelsetid per vecka inom den kommunala barnomsorgen givet att sådan har valts. Resultatet blir då att det är vistelsetider överstigande 30 timmar per vecka som skapar problem.

Den andra analysen utnyttjar att vissa kommuner har kö till den kommunala barnomsorgen medan andra har garanterad tillgång. Med en instrumentalvariabelteknik jämförs därför effekten av daghem i förhållande till familjedaghem. Detta innebär en jämförelse mellan dem som lockas till valet av daghem till följd av garanterad tillgång till sådan omsorg och dem som till följd av ransonering använder familjedaghem. Även denna analys utfaller till fördel för daghem; den skattade koefficienten är stor (nästan 0,5 *d*) men skattningen har låg precision.

⁷⁰ Datta Gupta och Simonsen noterar dock att vid denna tid hade alla barn rätt till kommunal barnomsorg oavsett föräldrarnas arbetsmarknadsstatus, dvs. en skillnad mot i Sverige.

Havnes och Mogstad (2009) är en färsk studie av de långsiktiga effekterna för de deltagande barnen av expansionen av daghemsplatser i Norge i mitten av 1970-talet. Ambitionen är att dels genomföra en studie av effekterna av utbyggnad av universell barnomsorg till skillnad från den mer selektiva typ av barnomsorg för ensamstående mödrar som dominerat i USA, dels att studera långsiktiga effekter. För att skatta effekterna av fler daghemsplatser utnyttjade Havnes och Mogstad att den snabba expansionen av daghemsplatser från 1975 och framåt var mycket ojämnt fördelad mellan landets kommuner. De kommuner med snabbare expansion än mediankommunen behandlas som experimentkommuner medan de med långsammare expansion är jämförelsekommuner. Skillnaden i expansion mellan dessa båda grupper av kommuner är betydande; täckningsgraden för de förra ökade från cirka 15 till 41 % från 1975 till 1979 mot bara 9 till 18 % för de senare.

Den påtagliga skillnaden mellan dessa båda typer av kommuner används sedan i en numera frekvent använd ansats, kallad ”differences-in-differences”. Havnes och Mogstad skattar en regressionsmodell som identifierar effekten av den utbyggda barnomsorgen som skillnaden i förändringarna i de långsiktiga utfallen mellan de båda kommungrupperna. Effekten skattas genom att jämföra experiment- och jämförelsekommunerna vad avser förändringen i utfall mellan, å ena sidan, kohorterna födda 1973–76, och, å andra sidan, tidigare kohorter. Med denna ansats behövs inte information om huruvida de enskilda barnen gick på daghem eller ej, utan det räcker att veta om de bodde i en experiment- eller jämförelsekommun. Man kan också säga att denna ansats fångar upp den samlade effekten av fler daghemsplatser på alla barn i de berörda kommunerna. Det är möjligt att det finns kamrateffekter så att de barn som gått på daghem i experimentkommunerna påverkar övriga barn, som ej gått på daghem, i samma kommuner. Sådana kamrateffekter är i princip också att beteckna som effekter av daghemmen och är således inbakade i de rapporterade effekterna. Havnes och Mogstad redovisade dock resultat och bedömningar som tyder på att sådana indirekta effekter av daghemsutbyggnaden är mycket små.

Havnes och Mogstad mätte utfallen år 2006 då de direkt berörda barnen var 30–33 år gamla. Utfallen är främst slutlig utbildning, arbetsinkomst och sannolikheten att erhålla socialbidrag. De finner en klart signifikant effekt på antal utbildningsår i form av 0,35 år per daghemsplats, eller 0,06 år per barn i hela kommunen där daghemsplatserna ökade. Utbildningseffekten återfanns både i den undre delen av fördelningen i form av minskad sannolikhet att avbryta gymnasiet och i den högre i form av högre sannolikhet att genomgå en hög-

skoleutbildning. De fann också en utjämnande effekt på arbetsinkomsterna, och sannolikheten att erhålla socialbidrag föll till följd av daghemsexpansionen. Ytterligare analyser visade att de positiva utbildningseffekterna var större för barn med lågutbildade mödrar.

En viktig fråga är via vilka mekanismer som sådana positiva effekter kan tänkas uppstå. För att bedöma detta redovisar Havnes och Mogstad också resultat för effekten av samma utbyggnad av daghemmen på mödrars sysselsättning och arbetsinkomster. De fann, kanske något förvånande, inga sådana effekter. Därför tolkade de resultaten som att daghemmen har trängt undan andra former av barnomsorg.

Studien ger således en mycket positiv bild av den norska utbyggnaden av daghemmen, som i sig påminner mycket om den svenska. Frågan är då om resultaten i alla avseenden är trovärdiga. Styrkan med den använda metoden är att den beaktar permanenta skillnader mellan de olika typerna av kommuner; dessa skillnader elimineras när det är skillnader mellan förändringar som analyseras. Det är dock möjligt att experiment- och jämförelsekommunerna har olika utvecklingsförlopp i sådana avseenden som är viktiga för dessa barngenerationers långsiktiga utveckling. Sådana skillnader beaktar inte metoden och de skulle därför kunna leda till felaktiga resultat. Havnes och Mogstad gör ett antal känslighetsanalyser för att undersöka om så är fallet, men finner inga tecken på detta. Studien, som är färsk och troligen kommer att göras föremål för ingående granskning av olika forskare under den närmaste tiden, får därför tills vidare betecknas som ett klart stöd för hypotesen att det finns gynnsamma långsiktiga effekter av utbyggd generell barnomsorg i form av daghem.

Det finns också några färska studier av förlängd föräldraledighet som påminner om den svenska av Nordström Skans och Liu (2010). Würtz Rasmussen (2010) utnyttjade en dansk reform från 1984 som förlängde föräldraledigheten från 14 till 20 veckor. Hon jämför utfallen för de barn som föddes just före och just efter tidpunkten då reformen trädde i kraft. Stora urval från register gör det möjligt att få god precision i skattningarna trots den tämligen lilla "behandling" som 6 veckor innebär. Utfallen var ett lästest vid 15 års ålder, en indikator på om individen vid 21 års ålder hade fullföljt gymnasieutbildning, samt genomsnittsbetyget från gymnasiet. Hon fann inga signifikanta effekter av reformen. I stort sett samma resultat, dvs inga effekter av förlängd föräldraledighet, fann Dustmann och Schönberg (2009) när de utnyttjade tyska reformer som förlängde föräldraledigheten.

Däremot har en färsk norsk studie uppvisat effekter av sådana reformer. Carneiro m.fl. (2009) genomförde en liknande studie av en förlängning av föräldraledigheten från 12 till 18 veckor som genomfördes i Norge 1977. Eftersom reformen genomfördes så pass tidigt kunde de (med hjälp av registerdata) studera utfallen för barnen ända fram till tidig vuxen ålder.⁷¹ De fann att reformen, dvs. längre betald ledighet, ledde till färre avbrott från gymnasiet, högre IQ (enbart för män då IQ-mått härstammade från militärtester) och något högre sannolikhet att påbörja universitetsutbildning. Däremot fann de inga signifikanta effekter på den totala längden på utbildningen eller på inkomsterna just efter arbetsmarknadsinträdet.

För att få indikationer på tänkbara mekanismer bakom dessa effekter undersökte Carneiro m.fl. också om effekterna varierade beroende på familjeförhållanden. De fann mer gynnsamma effekter för barn till lågutbildade mödrar liksom för dem som bodde nära mor- eller farföräldrar och där tillgången till daghem var dålig. En del ytterligare analyser gav resultat som var förenliga med hypotesen att reformen underlättade för mödrarna att amma barnen under dess första halva levnadsår. Det faktum att denna reform i praktiken avsåg en så pass tidig period som 12 till 18 veckor skulle således kunna förklara de annorlunda resultaten i denna studie. Men eftersom den danska reformen avsåg en förlängning från 14 till 20 veckor är skillnaden i resultat ändå något svår att förstå.

Frågan om avkastningen på förskoleliknande satsningar har också varit föremål för en intensiv diskussion i USA och Kanada, denna diskussion har pågått i både forskarsamhället och inom politiken. Det finns också en relativt stor empirisk litteratur med utvärderingar av olika satsningar på föreskoleliknande verksamhet. Två av dessa är närmast klassiska och kan sägas ha fått läroboksstatus, nämligen HighScope Perry respektive Abecedarian programmen från 1960- respektive 1970-talet. Detta var förskoleprogram som riktades mot barn från fattiga förhållanden på två olika platser i USA, i Michigan respektive Carolina. Var och en av studierna omfattade bara några hundra barn, men dessa fördelades slumpmässigt i förskoleprogram respektive i en kontrollgrupp. Resultaten från långsiktiga uppföljningar av dessa program är övervägande mycket positiva.⁷² Den experimentella designen gör dessutom trovär-

⁷¹ Denna studie liksom några av de övriga undersökte även de långsiktiga effekterna för mödrarna men vi går här inte igenom dessa resultat.

⁷² Se Barnett och Masse (2007) för en färsk reanalys av Abecedarian programmet och Heckman m.fl. (2010) för en reanalys av HighScope Perry programmet. Båda artiklarna ger rikliga referenser till andra analyser av programmen.

digheten hög vad gäller effekten för berörda grupper vid den aktuella tidpunkten i de aktuella platserna; den interna validiteten är hög. Med tanke på att det är frågan om riktade satsningar mot utpräglade problemgrupper i det amerikanska samhället är det svårt att argumentera för att resultaten kan överföras till svenska förhållanden.

Ett annat centralt program i USA har varit Head Start, ett förskoleliknande verksamhet för ensamstående föräldrar i USA. Det finns en omfattande litteratur även kring detta program; Currie (2001), Blau och Currie (2006) och Ludwig och Philips (2008) för olika översikter. Med tanke på att även detta är selektiva program i en helt annan miljö än den svenska avstår vi från att försöka sammanfatta de inte helt entydiga resultaten från denna litteratur, utan hänvisar den intresserade läsaren till översikterna.

I den nordamerikanska diskussionen har det också efterlysts studier av effekterna av mer universellt erbjuden barnomsorg av nordisk daghemskaraktär, eftersom det funnits förslag att genomföra mer generella satsningar i USA och Kanada. Baker m.fl. (2008) är just en sådan färsk studie av kanadensiska förhållanden. De utnyttjar det faktum att delstaten Quebec expanderade sin barnomsorg radikalt i förhållande till övriga landet. Expansionen skedde under några år i slutet av 1990-talet genom kraftigt sänkta avgifter för först 5-åringar under 1997 och sedan (år 2000) successivt ned till barn som var två år och yngre. Det faktum att det skedde en snabb utbyggnad i en del av Kanada skapade en analysmöjlighet som författarna utnyttjade. Den analytiska designen blev därför att studera skillnaderna i utvecklingstakten i centrala utfallsvariabler mellan Quebec och övriga landet.

Resultaten visar för det första, kanske föga förvånande, att mödrars förvärvsfrekvens steg kraftigare i Quebec än i övriga landet. De fann samma mönster i utnyttjandet av någon form av barnomsorg utanför det egna hemmet. För våra syften är det dock av störst relevans att notera resultaten som avser barnen. För detta syfte utnyttjar de informationen i en stor surveyundersökning av barns situation i Kanada, National Longitudinal Survey of Children and Youth (NLSCY). Denna innehåller uppgifter hämtade från intervjuer med både barn, föräldrar och lärare. Frågorna är utprovade med hjälp av experter inom utvecklingspsykologi.

För barnen finner författarna signifikanta försämringar – i Quebec i förhållande till övriga landet – i utfall som ”känslomässig stabilitet” ”fysisk aggression och motvilja” och ”motorisk och social utveckling”. Även indikatorer på hälsa visar en försämring. Den största försämringen avser ”motorisk

och social utveckling” som uppgår till $d = 0,1$. Däremot visar punktskattningen för ett test av verbal förmåga på en positiv effekt av daghemsutbyggnaden, men precisionen är mycket låg och skattningen är inte signifikant skild från noll.

Baker m.fl. använder även variabler som mäter föräldrars beteende. Dessa variabler är index som konstruerats från direkta frågor om deras beteende och upplevelser. Resultaten visar en klar försämring i utfall som ”aggressiv och ineffektiv uppfostran”, ”konsekvent uppfostran” och ”aversiv (bestraffande) uppfostran” i Quebec i förhållande till övriga landet. De finner även en viss relativ försämring av föräldrarnas självrapporterade hälsa i Quebec.

Sammantaget ger denna ambitiösa studie en negativ bild av effekterna för barn, effekter som delvis kan ha uppkommit via effekter på föräldrarna av deras förändrade situation. Författarna betonar dock att det är frågan om kortsiktiga effekter. De långsiktiga effekter på slutlig utbildning och framgång på arbetsmarknaden, som den norska studien som diskuterades ovan undersöker, skulle kunna vara mer positiva. Detta gäller om övergången till förskola eller omsorg utanför hemmet är förknippad med omställningsproblem av olika slag.⁷³ De barn som tidigt börjar på förskola eller liknande får då denna omställning tidigare än andra barn som kanske får den i samband med den regelrätta skolstarten. En reservation mot studien är också motiverad av att den bygger på en jämförelse mellan utvecklingen i den fransktalande delen av landet i förhållande till den engelsktalande. Det kan finnas andra förändringar än skillnaden i utvecklingen av barnomsorgen som skiljer de båda delarna av landet åt. Det är också möjligt att den mycket snabba utbyggnaden i Quebec innebar att barnomsorgens kvalitet i vissa avseenden blev lidande. Likafullt visar denna studie, den danska och den norska att det är mycket viktigt att skilja mellan kort- och långsiktiga effekter av förskola och barnomsorg. Det vore önskvärt med en framtida studie som kan kombinera dessa tidsperspektiv.

6.5 Studier av tidigare skolstart

Frågan om effekterna av tidigare skolstart är nära besläktad med frågan om effekterna av förskola eller daghemsvistelse. Med tidigare skolstart menar vi tidigare start av den obligatoriska skolan och att pedagogiken får ”skolliknande former”. Om man ställer den frågan i dagens svenska situation blir självfallet jämförelsealternativet för de allra flesta vistelse på daghem med den pedagogik som används där. För de barn som stannar hemma med en förälder blir dock

⁷³ Denna invändning kan även riktas mot den danska daghemsstudien som presenterades tidigare i detta kapitel.

alternativet detta i stället. En annan fråga är vilka effekterna skulle vara av att förändra pedagogiken på daghemmen mot mer ”skolliknande former”. Så länge förskolan är frivillig vore detta dock en annan fråga än den om tidigare obligatorisk skolstart.

En annan viktig distinktion när det handlar om att precisera vad man menar med skolstart är mellan, å den ena sidan, en tidigareläggning av en obligatorisk utbildning av en given längd och, å den andra sidan, en förlängning av den obligatoriska utbildningen (med exempelvis ett år) i dess början.

Oavsett hur man formulerar frågan som man önskar besvara stöter man på betydande metodproblem. Ett genomgående sådant problem är att man måste skilja mellann allmän ålderseffekt och effekten av att gå längre i skolan (se kapitel 5). Några relativt färska studier har dock utnyttjat data på ett innovativt sätt och kommit fram till intressanta resultat.

Fredriksson och Öckert (2008) har undersökt effekterna av en tidigareläggning av den obligatoriska utbildningen med hjälp av stora urval från svenska registerdata. De utnyttjar det faktum att barn som är födda tidigt på året i normalfallet börjar skolan vid en högre ålder än barn som är födda i slutet av året. Det är rimligt att tro att det är slumpmässigt med avseende på faktorer som bestämmer skolprestationer och senare arbetsmarknadskarriär om man föds just före eller efter ett årsskifte. Därför är det rimligt att utnyttja de skillnader i skolstartsålder som orsakas av att man är född före eller efter årsskiftet för en analys av skolstartsålderns betydelse. Fredriksson och Öckert mäter utfall i form av slutlig längd på utbildning respektive arbetsmarknadsinkomster vid samma tidpunkt i vuxen ålder. Det betyder att de jämför effekten av att den obligatoriska skolperioden var förlagd tidigare respektive senare, dvs de undersöker den renodlade effekten av tidigare skolstart och inte effekten av att förlänga skolperioden i början av densamma.

Resultaten visar att en högre skolstartsålder leder till längre sammanlagd utbildning. För grupperna födda 1935–44 och 1945–54 innebär skattningarna att ett års senare skolstart förlänger utbildningen med cirka 0,18 år, medan effekten är något svagare för grupperna födda 1955–64 och 1965–74. Utfallet är här uppmätt år 2000 då de flesta hunnit avsluta sin utbildning. Att effekten är något större för de äldre åldersgrupperna menar författarna kan bero på att dessa genomgick ett kortare och mer selektivt skolsystem.

Fredriksson och Öckert redovisar också resultat för effekterna på arbetsinkomster, som också uppmättes år 2000. Dessa är negativa för åldersgrupperna födda 1955 och senare, men positiva för dem som är födda mellan

1935 och 1954. Detta mönster i resultaten torde förklaras av att de som startar senare i skolan till en början halkar efter karriärmässigt, men sedan ges en chans att ta igen detta då effekterna av högre utbildning slår igenom. Mönstret innebär också att det är oklart huruvida effekterna av högre skolstartsålder är positiva på livsinkomsterna.

Black m.fl. (2010) utnyttjade i grunden samma metodansats i en studie av norska data. De fann att effekterna på utbildning och inkomster i vuxen ålder var obetydliga.

Leuven m.fl. (2010) undersökte situationen i Nederländerna utifrån en annan frågeställning, nämligen om en förlängning av skoltiden i form av tidigare start har gynnsamma effekter på barns kunskaper längre fram under skoltiden. För att erhålla variation mellan barn i skoltid som inte är perfekt korrelerad med ålder använder de gällande regler för skolstart i Nederländerna. Där erbjuds barnen skolstart den dag de fyller fyra år. I kombination med i förläggningen av (främst) sommarlovet skapar denna regel en variation i skoltid som är användbar för en empirisk studie. De finner att en extra månad i skolan höjer resultaten i språktester med 0,06 *d* och i matematiktester med 0,05 *d* för barn från studiesvaga hemförhållanden, medan de inte fann några effekter för övriga elever. Utfallen mättes dock redan i slutet av första klass varför det är oklart om dessa effekter består på längre sikt. Vidare skall resultaten tolkas som effekterna av att *erbjuda* tidigare skolstart. För dem som de facto utnyttjade denna möjlighet är effekterna större.

6.6 Slutsatser

Utbyggnaden av den svenska förskolan innebär en stor politisk satsning liksom en stor samhällsförändring som ägde rum under en relativt kort period. Underlaget för en övergripande utvärdering av denna satsning måste dock betecknas som tämligen magert. Det finns några studier med detaljerad mätning av barnens färdigheter tidigt i skolan och som i några fall gjort urval av deltagande och icke-deltagande barn ur kommunernas kösystem, men dessa studier bygger på små och lokala urval och effekterna är inte skattade med någon större precision. De studier som bygger på större och riksrepresentativa urval har tvingats använda utfall – såsom val av teoretisk linje på gymnasiet – som kommer senare i skolan vilket kan dölja en del av de verkliga effekterna. Dessa studier utgår heller inte från kommunernas kösystem varför skattningarna av effekter innebär att man (åtminstone delvis) jämför barn vars föräldrar

velat ha sina barn på daghem med barn vars föräldrar inte velat detta. Detta faktum gör studierna mer svårtolkade än vad de annars skulle vara.

Ett genomgående drag i de svenska studierna är att effekterna på kunskaper och färdigheter och på vidare prestationer i skolan är tämligen små. En rimlig bedömning utifrån befintliga studier är därför att det i varje fall inte finns några dramatiskt stora eller små effekter av att delta i förskoleverksamhet. Även den svenska studien av tidigare skolstart förstärker bilden av att effekterna är tämligen små, särskilt om man har livsinkomsten som måttstock.

Ett annat genomgående drag är att de skattade effekterna av daghem är osäkra, dvs. har höga standardfel. Detta hänger samman med att det inte finns någon studie som kunnat använda riktigt stora urval baserade på registerdata. Man skulle kunna tänka sig att använda s.k. metaanalys för att väga samman skattningarna så att precisionen blir högre, vilket här inte varit möjligt. Men man kan ändå argumentera för att förekomsten av ett flertal studier med skattade positiva effekter kan tolkas med lite större säkerhet än de enskilda studierna. Helhetsintrycket av de befintliga studierna blir därmed snarast att de tyder på svagt positiva effekter av daghem.

Om underlaget är osäkert vad avser svenska förhållanden skulle man kunna förlita sig på studier från övriga nordiska länder istället, i synnerhet om dessa framstår som mer tillförlitliga. En dansk och en norsk studie, som båda publicerats nyligen, ger dock något motstridiga resultat. Den norska tyder på mycket gynnsamma långsiktiga effekter på slutlig utbildning och arbetsmarknads-etablering, medan den danska tyder på obetydliga effekter och snarast negativa sådana för omsorg i familjedaghem i förhållande till omsorg hos någon av föräldrarna. Den danska studien undersöker dock utfall kring sju års ålder och dessa avser inte kognitiv förmåga. Om vistelse på daghem eller familjedaghem innebär en omställningsprocess som innebär en påfrestning på kort sikt, men ändå en gynnsam utveckling i skolan på lång sikt, vore dock resultaten förenliga med varandra.

Vår bedömning att kunskapsläget är torftigt när det gäller effekterna av utbyggd förskola och barnomsorg måste dock kvalificeras på en punkt. En studie för vardera Sverige, Danmark och Norge har undersökt effekterna av förlängd föräldraledighet där alternativet till föräldraledigheten är just vistelse i kommunal barnomsorg. Här handlar det således om vistelse i offentlig barnomsorg vid mycket unga år i jämförelse med föräldraomsorg. Dessa studier har kunnat utnyttja skillnader som uppstår mellan jämförbara barn till följd av reformer som ger längre föräldraledighet vid speciella datum. Vår bedömning

är att detta är trovärdiga studier. För Sverige och Danmark tyder de på att förlängd föräldraledighet och kommunal barnomsorg var likvärdiga som alternativ när man jämför skolprestationer i tonåren. Den norska studien tyder dock på positiva effekter av förlängd föräldraledighet, men här var det frågan om en förlängning från mycket låg ålder så faktorer som amning kan ha varit viktiga för resultaten.

En viktig fråga är om det hade varit möjligt att genomföra bättre studier på svenska data. En brist hade åtgärdats om det hade funnits tidigare utfallsvariabler än de som främst har använts i svenska studier. Tester genomförda tidigt i skolan som hade kunnat påföras befintliga datamaterial hade gjort studierna ovan klart bättre. Likaså hade en motsvarighet till den kanadensiska longitudinella databasen med information om barns utveckling upp till 11 års ålder underlättat genomförandet av svenska studier.

Det är också intressant att notera att särskilda satsningar på förskola, och kanske särskilt tidigare skolstart, hade kunnat genomföras som storskaliga kvasi-experiment, i samma anda som grundskolereformen på 1950-talet. Då genomfördes den längre och gemensamma grundskolan vid olika tidpunkter i olika delar av landet just för att kunna dra lärdomar och genomföra bra utvärderingar. Satsningar på olika former av förskola och tidigare skolstart hade med fördel kunnat läggas upp på samma sätt. Detta hade dock krävt politiska beslut.

Slutligen är det möjligt att förutsättningarna för bra studier av förskolans effekter kommer att förbättras något under de kommande åren då forskningen av allt att döma får tillgång till resultat från nationella prov i årskurserna tre och fem. Det blir då möjligt att relatera dessa skolprestationer till deltagande i olika former av förskoleverksamhet. Men samtidigt saknas ett heltäckande nationellt register över vilka som deltar i förskolan. Ett sådant register, i kombination med tidigare testresultat, skulle bidra till bättre framtida studier av förskolans verksamhet.

Referenser

- Andersson, B-E. (1989), Effects of public day-care: A longitudinal study, *Child Development*, 60, 857–866
- Andersson, B-E. (1992), Effects of day-care on cognitive and socioemotional competence of thirteen-year-old Swedish school children, *Child Development*, 63, 20–36.
- Baker M., J. Gruber och K. Milligan (2008), Universal child care, maternal labor supply and family well-being, *Journal of Political Economy*, 116, 709–745.
- Barnett W.S. och L.N. Masse (2007), Comparative benefit-cost analysis of the Abecedarian program and its policy implications, *Economics of Education Review*, 26, 113–125.
- Black S., P. Devereux och K. Salvanes (2010), Too young to leave the nest? The effects of school starting age, *The Review of Economics and Statistics*, under publicering.
- Blau, D. och J. Currie (2006), Preschool day care, and after school care: Who's minding the kids?, *Handbook of Economics of Education*.
- Broberg, A., H. Wessels, M.E. Lamb och P. Hwang (1997), Effects of day care on the development of cognitive abilities in 8-year-olds: A longitudinal study, *Developmental Psychology*, 33, 62–69.
- Carneiro, P., K. Løken och K. Salvanes (2009), A flying start? Maternity leave and long-term outcomes for mother and child, manuskript.
- Currie, J. (2001), Early childhood education programs, *Journal of Economic Perspectives*, 15, 213–238.
- Datta Gupta, N. och M. Simonsen (2010), Non-cognitive child outcomes and universal high quality child care, *Journal of Public Economics*, 30–43.
- Dustmann, C. och U. Schönberg (2009), The effect of expansions in maternity leave coverage on children's long-term outcomes, manuskript.
- Fredriksson P. och B. Öckert (2008), Är det bättre att börja skolan tidigare?, *Ekonomisk Debatt*, 36, 17–32.

- Havnes T. och M. Mogstad (2009), No child left behind: Universal child care and children's long-run outcomes, Discussion Papers No. 582, Statistics Norway, Research Department.
- Heckman, J. S. Urzua och E. Vytlacil (2006), Understanding instrumental variables in models with essential heterogeneity, *Review of Economics and Statistics*, 88, 389–432.
- Heckman, J., S. Moon, R. Pinto, P. Savelyev och A. Yavitz (2010), The rate of return to the high scope Perry preschool program, *Journal of Public Economics*, 94, 114–128.
- Jonsson J.O. (1994), Förskola – en strategi för jämlikhet?”, i Erikson R. och Jonsson J.O. (red) *Sorteringen i skolan*. Stockholm: Carlssons bokförlag.
- Jonsson J.O. (2004), Förskola för förfördelade, i Bygren, M., Gähler, M. och Nermo, M. (red.) 2004. *Familj och arbete – vardagsliv i förändring*. Stockholm: SNS förlag.
- Kågesson Per (2006), *Tid för barn?* Stockholm: SNS förlag.
- Leuven E., M. Lindahl, H. Oosterbeck och D. Webbink (2009), Expanding schooling opportunities for 4-year-olds, *Economics of Education Review* under publicering.
- Ludwig J. och D. Philips (2008), The long-term effects of Head Start on low-income children, *Annals of the New York Academy of Sciences*, 40, 1–12.
- Mörk, E., A. Sjögren och H. Svaleryd, (2009), Effekter av barnomsorgsavgifter på barnafödandet, *Ekonomisk Debatt*, 37, 59–69.
- Nordström Skans O. och Q. Liu (2010), The duration of parental leave and children's scholastic performance, *The B.E. Journal of Economic Analysis & Policy*, 10(1) (Contributions), Article 3.
- Stukát K-G. (1966), *Lekskolans inverkan på barns utveckling*. Stockholm: Almqvist & Wicksell.
- Söderström L., A. Björklund, P.G. Edebalk och A. Kruse (1999), *Från dagis till servicehus: Valfärdsolitik i livets olika skeenden*, Stockholm: SNS Förlag.

Würtz Rasmussen, A. (2009), Increasing the length of parents birth-related leave: The effect on children's long-term educational outcomes, *Labour Economics*, 17, 91–100.

7 Lärartäthet och lärarkompetens

Detta kapitel inventerar forskning kring resursfaktorer som påverkar utbildningsresultat. Översikten är selektiv: vi fokuserar i första hand faktorer som i någon mån är manipulerbara och därför intressanta i den utbildningspolitiska diskussionen. Eftersom det i ett sammanhang som detta är omöjligt att täcka av all forskning tar vi i första hand upp svenska studier, och sådana internationella studier som vi uppfattar som banbrytande. Vi presenterar och diskuterar också resultat från en del andra forskningsöversikter och metaanalyser.

I en nyligen publicerad studie sammanfattade Hattie (2009) resultat från över 800 olika metaanalyser av undersökningar av vilka faktorer som har betydelse för elevernas resultat. Denna meta-metaanalys baseras på drygt 50 000 enskilda undersökningar som omfattat mer än 80 miljoner personer. Genomgången omfattar flera olika kategorier av faktorer som eleven, hemmet, skolan, läraren, läroplanen och undervisningen. Bland skolfaktorerna återfinns flera av de faktorer som varit i fokus i forskningen om resursers betydelse för elevprestationer. Detta gäller bland annat klasstorlek, vilken dock enligt Hatties resultat har en relativt blygsam inverkan på elevresultaten ($d = 0,20$). Det mest framträdande resultatet i Hatties översikt är dock att de faktorer som avser olika aspekter av lärarnas undervisning är de som har de starkaste effekterna. Denna resultatbild överensstämmer i sina huvuddrag med den slutsats som drogs av Gustafsson och Myrberg (2002) (se också Gustafsson 2003) att det är lärarens kompetens som är den viktigaste faktorn, men därutöver identifierar Hattie mer specifika faktorer som feedback och direkt undervisning som några av de medel genom vilka de goda undervisningsresultaten uppnås.

En meta-metaanalys av det slag som Hattie (2009) genomfört ger givetvis en mycket grov och översiktlig bild av forskningsresultaten. Endast i begränsad utsträckning har Hattie (2009) tagit hänsyn till faktorer som påverkar en resursfaktors betydelse, men det är väl känt att vissa faktorer (s.k. moderatorer) påverkar hur stark effekt en viss faktor har. Exempelvis har elevernas ålder och hembakgrund, liksom kulturella faktorer, visat sig ha modererande påverkan på effekten av klasstorlek (Gustafsson och Myrberg 2002). Inte minst är det viktigt att vara medveten om denna begränsning då syftet är att applicera de internationella forskningsresultaten på svenska förhållanden, och det är nödvändigt att ta hänsyn till sådana interaktionseffekter då betydelsen av en resursfaktor värderas.

Detta är en anledning att mer utförligt behandla forskningen kring lärartäthet och klasstorlek. Vi behandlar även lärarkompetensen, eftersom den visat

sig ha stor betydelse för elevernas resultat. Fokus ligger alltså på lärartäthet och lärarkompetens som de huvudsakliga resursfaktorerna.

Vi vill här också hänvisa till en annan nyligen publicerad översikt av Gustafsson och Myrberg (2009), som också inventerar forskning kring resurser betydelse för elevernas resultat. Denna översikt ingår i en antologi som har som huvudsyfte att identifiera olika faktorer som påverkar elevernas resultat och att hitta förklaringar till den observerade nedgången i de svenska grundskoleelevernas resultat. Materialet som presenteras i detta kapitel kompletterar Gustafsson och Myrberg (2009); den senare översikten fokuserar i större utsträckning svensk och nordisk forskning, medan denna översikt i större utsträckning diskuterar metodfrågor och resultat i de enskilda undersökningarna.

7.1 Lärartäthet

Det är en allmän utbredd uppfattning bland föräldrar, lärare och elever att resultaten blir bättre då undervisningen bedrivs i små klasser än i stora klasser. Ur kostnadssynpunkt är också klasstorleken en betydelsefull faktor, eftersom den bestämmer hur många lärare som behövs. Den är därför den enskilda resursfaktor som har störst implikationer för skolans kostnader.

Ett stort antal studier har undersökt effekter av klasstorlek/lärartäthet på elevernas resultat. Fram till 1990-talets början var huvudslutsatsen att vare sig resurser i allmänhet eller klasstorlek i synnerhet hade någon större betydelse för de uppnådda resultaten (t.ex. Hanushek 1989). Därefter har dock undersökningar genomförts som pekar på att klasstorlek/lärartäthet kan vara betydelsefulla faktorer för elevernas resultat.

7.1.1 Några grundläggande metodproblem

En anledning till den komplexa resultatbilden är att många studier använt metoder med begränsad förmåga att ge underlag för slutsatser om kausala relationer. Huvudproblemet är att det finns mekanismer som skapar skillnader mellan elevernas förutsättningar i klasser av olika storlek. I första hand leder kompensatorisk resursallokering till att svagare elever tenderar att placeras i mindre klasser. Att sådana mekanismer kan vara verksamma inom svensk skola har noterats av Skolverket (1999); Lazear (2001) har utvecklat en teoretisk modell som förklarar varför det tenderar att föreligga ett positivt samband mellan elevernas studieförutsättningar och klasstorlek. Den kompensatoriska resursallokeringen medför att elever i mindre klasser i medeltal har sämre förutsättningar. Detta kan lätt leda till den felaktiga slutsatsen att större klasser

medför bättre elevresultat, i synnerhet i tvärsnittsdata. Korrekta kausala slutsatser förutsätter att detta problem kan lösas, exempelvis med hjälp av en experimentell ansats eller med statistiska metoder som har sådana selektions-effekter under kontroll. Det finns studier som gjort ambitiösa försök i denna riktning och vi kommer att fästa särskild uppmärksamhet vid dessa.

Ett annat problem inom detta forskningsfält är av begreppslig art, och avser distinktionen mellan klasstorlek å ena sidan och lärartäthet å den andra. Dessa begrepp har olika innebörd och olika forskare har definierat dem på olika sätt. Definitionen av lärartäthet varierar också mellan olika länder, vilket gör jämförelser svårtolkade.

Ett sätt att förstå begreppet *klasstorlek* är att definiera det som det antal elever som finns inskrivna i en och samma undervisningsgrupp under en längre tid. Ett problem med detta sätt att definiera klasstorlek är att det inte tar hänsyn till att undervisning ibland bedrivs i halvklasser eller i olika former av flexibla elevgrupperingar. Ett annat problem är att skolklassen åtminstone i Sverige är en mycket föränderlig företeelse, där vissa elever kommer till, och andra elever försvinner. Skolklassen som officiellt begrepp avskaffades också inom det svenska skolväsendet redan under 1970-talet.

Ett annat sätt att beskriva elevers tillgång till undervisning är att bestämma det genomsnittliga antalet heltidsläsande elever per heltidsundervisande lärare ("pupil-teacher ratio", PTR, i engelskspråkig litteratur). I den svenska statistiken presenteras istället i allmänhet *lärartätheten*, definierat som antal lärare per 100 elever (dvs. som "lärarprocent"). Även om lärartätheten i princip är enkel att definiera möter det ändå en hel del praktiska problem att bestämma den med exakthet. Ett problem är att bestämma antal heltidsläsande elever; ett annat är att bestämma antalet lärare, där det inte är entydigt hur denna kategori ska avgränsas från andra yrkeskategorier som är verksamma i skolan, t.ex. skolläding, elevvårdspersonal, förskollärare, fritidspedagoger och fritidsledare.

Måtten på lärartäthet och klasstorlek samvarierar perfekt då endast en lärare undervisar varje klass som hela tiden är sammanhållen, och då lärarens undervisningsskyldighet dessutom sammanfaller med antalet elevtimmar. Om lärarens undervisningsskyldighet är lägre än elevtimmarna blir lärartäthetsmättet stort i förhållande till klasstorleksmättet. Begreppen lärartäthet och klasstorlek står sålunda inte i något enkelt och entydigt förhållande till varandra. Det är inte heller troligt att effekterna av förändringar i klasstorlek respektive i lärartäthet är desamma, men trots detta har de i diskussionen ofta använts som synonyma. De mekanismer som påverkar utfallet kan också antas vara olika

beroende på om det är antalet elever i klassen som förändras eller om det är mängden lärartid som förändras. Som Lazear (2001) visat teoretiskt är sannolikheten för störningar i klassens arbete en funktion av antalet elever i klassen, men denna sannolikhet behöver inte hänga samman med mängden lärartid per elev. Den enskilda elevens möjlighet att få tillgång till lärarens stöd kan dock antas vara beroende av mängden lärartid per elev.

7.1.2 Metaanalyser

Som redan nämnts finns det flera metaanalytiska studier som explicit behandlar effekter av klasstorlek på inlärningsresultat. Metaanalys är en statistisk metod för att sammanställa resultat från flera undersökningar, som utvecklades av Glass (1978). En av de första tillämpningarna av metoden rörde effekten av klasstorlek. Smith och Glass (1980) drog slutsatsen att klasstorlek har betydelse för studieprestationer. Ett huvudresultat var att effekten av klasstorlek på såväl inlärningsresultat som attityder kunde beskrivas med en logaritmisk funktion. Detta innebär att den negativa effekten av en ökning av klasstorleken med en ytterligare elev blir större i en klass med få elever än i en klass med många elever. Effekten av klasstorlek var också större för yngre elever (12 år och yngre) än för äldre elever.

Robinson (1990) har också genomfört en mycket omfattande metaanalys av ett 100-tal undersökningar. I denna undersöktes om effekterna skiljde sig åt mellan grupper. Materialet delades därför upp med avseende på olika åldersgrupper, etniska grupper, ämnesområden, osv. Ett tydligt resultat var att effekten av klasstorlek varierar med elevernas ålder. För årskurserna 0–3 visade ett stort antal undersökningar på positiva effekter av små klasser; för årskurserna 4–8 erhöles en svagt positiv effekt. Undersökningarna som avsåg årskurserna 9–12 gav inget stöd för att mindre klasser skulle ge bättre inlärningsresultat. Ett annat tydligt resultat i Robinsons metaanalys är att elever som hör till utsatta grupper når bättre resultat i mindre klasser.

De ovan redovisade metaanalyserna av effekter av lärartäthet och klasstorlek ger inte stöd för slutsatsen att dessa faktorer har starka och entydiga effekter på elevresultat. Inte heller stödjer de den slutsats om total avsaknad av samband som förfäktats av bland andra Hanushek (1989, 2003). Resultaten pekar snarare på att effekterna är heterogena: klasstorlek har större betydelse för yngre elever än för äldre elever; de positiva effekterna av mindre klasser är större för elever från mindre gynnad socioekonomisk bakgrund.

7.1.3 Studier från andra länder

Metaanalyserna baseras på undersökningar som i många fall haft en svag empirisk design (se kapitel 5). En stor mängd studier har sålunda genomförts där man försökt hantera selektion och utelämnade variabler genom att hålla observerade egenskaper konstanta. En tanke med metaanalysen är visserligen att resultat från en stor mängd studier med svag design skall kunna ge grund för starka slutsatser, men det är givetvis en fördel om slutsatserna kan baseras på studier där det inte råder tvekan om korrektheten i slutsatserna om kausalitet. Någon uppsjö av sådana studier finns inte, men i den internationella forskningen har under de senaste 20 åren några studier med hög ambitionsnivå avseende kausalinferens genomförts. Dessa studier diskuteras kortfattat nedan, och tolkas mot bakgrund av svenska förhållanden.

En av de mest välaktade studierna av klasstorlekens betydelse är det s.k. STAR-projektet (Tennessee Student/Teacher Achievement Ratio) i USA. Denna studie har diskuterats utförligt av Gustafsson och Myrberg (2002), men pga. undersökningens stora betydelse och de många nya analyser som genomförts tar vi upp den även i detta sammanhang. Undersökningen genomfördes i Tennessee med början 1985 och inkluderade ett stort antal elever, lärare och skolor. Den första undersökningsperioden pågick fram till 1990, men uppföljningsundersökningar har genomförts vid senare tidpunkter. Undersökningen hade experimentell design och både elever och lärare tilldelades slumpvis endera av tre klasstyper: små klasser (mellan 13 och 17 elever), normalstora klasser (mellan 22 och 25 elever) och normalstora klasser med en assisterande lärare. Valet av antalet elever i små klasser grundades på resultaten från Glass metaanalyser, enligt vilka det krävs en minskning av klasstorleken till ca. 15 elever för att nå en tydlig effekt. Den slumpmässiga fördelningen skedde i allmänhet då eleverna var fem år och eleverna förväntades sedan tillhöra samma klasstyp under fyra år. Varje år fördelades lärarna slumpmässigt över de olika klasserna. Eleverna testades i läsning och matematik vid slutet av varje läsår under de fyra åren. Vidare undersöktes elevernas motivation och självkänsla och lärarna intervjuades om sin undervisning. I undersökningen deltog 11 600 elever från ett 80-tal skolor. Varje skola hade åtminstone en vardera av de tre klasstyperna, och den slumpvisa fördelningen skedde inom varje skola. Totalt representerades varje klasstyp av ca. 100 klasser.

Resultaten visade att elever i små klasser hade bättre inlärningsresultat än i normalstora klasser. Effekten av klasstyp var starkast för åk 1, där en tydlig fördel för små klasser kunde iaktas, med en effektstorlek om ca. $d = 0,25$. Man

fann också att klasstyp interagerade med socioekonomisk och etnisk bakgrund. Effekten var drygt dubbelt så stor för elever med minoritetsbakgrund ($d = 0,35$ mot $d = 0,15$), och svarta elever vann särskilt mycket på att undervisas i små klasser. Som visats av Finn och Achilles (1990) medförde detta att prestationsskillnaden mellan minoritetsgrupper och vita med avseende på resultat i läsproven i åk 1 var betydligt mindre i de små klasserna, jämfört med de normalstora klasserna.

En analys av uppföljningsdata mellan årskurserna 4 och 8 har presenterats av Nye m.fl. (1999). Denna undersökning tog hänsyn till det bortfall som alltid drabbar longitudinella undersökningar; genom att använda tvånivåanalyser tog man också hänsyn till slumpmässiga skoleffekter. Huvudresultaten var desamma som de som har erhållits i andra analyser av uppföljningsdata: de positiva effekterna av att ha gått i en liten klass kvarstod i stor utsträckning. Vid slutet av åk 8 var den uppmätta effekten minst 70 % av den initiala effekten. Matematikresultaten kvarstod i större utsträckning än resultat inom andra områden. Man kunde också konstatera att de elever som deltog i mindre klasser under hela fyraårsperioden uppvisade högre resultat än de som endast deltog i delar av experimentet.

Krueger och Whitmore (2001) fann i en uppföljningsundersökning att benägenheten att göra urvalsprov för högre studier, liksom resultaten på dessa prov, var högre för de elever som gått i en liten klass. Även här fann man att de positiva effekterna var betydligt mer markerade för svarta elever och för elever från socioekonomiskt mindre gynnade hem.

STAR projektet har fördelar jämfört med de flesta andra undersökningar; fördelarna ligger dels i det experimentella upplägget, och dels i att undersökningen är longitudinell. Storskaliga experimentundersökningar kan dock aldrig genomföras exakt som planerat, och alla longitudinella undersökningar drabbas av bortfall. Båda dessa problem kunde noteras i STAR. Hanushek (1998) har bl.a. hänvisat till bortfallet som en grund för att tvivla på resultatens pålitlighet. Krueger (1999) visar emellertid att det omfattande bortfallet inte kan anföras som en förklaring till de goda resultaten för små klasser.

Som understryks av Blatchford m.fl. (1998) kan det faktum att lärare och elever visste vilket slags klass de tillhörde ha påverkat resultaten, exempelvis genom att de gjort extra ansträngningar för att visa på goda effekter av små klasser, eller genom att lärarna i denna grupp kan ha kommit att reflektera över de möjligheter som en mindre klass ger och därför ha förändrat sin undervisning.

Krueger (1999) (se också Krueger 2003) drog slutsatsen att effektstorleken av att minska klasstorleken med 7 elever uppgår till $d = 0,20$. Precis som i metaanalyserna visade resultaten från STAR på effekterna varierade över grupper: de var starkare för elever med minoritetsbakgrund ($d = 0,35$) än för vita elever ($d = 0,15$) elever. Man fann också att effekterna var starkare för elever från mindre gynnad socioekonomisk och utbildningsmässig bakgrund. Det bör då noteras att Tennessee hör till den grupp av delstater som uppvisar de sämsta nationella skolresultaten i USA, och där kostnaderna per elev är betydligt lägre än medeltalet. I STAR-studien ingick vidare en större andel innerstadsskolor än i Tennessee som helhet, och andelen svarta elever som ingick i undersökningen var ungefär dubbelt så stor som andelen i landet. Efter amerikanska mått mätt är sålunda de elever som ingick i undersökningen en ur utbildningssynpunkt underprivilegierad grupp, vilket innebär att effektskattningarna för urvalet till STAR antagligen är högre än dem som kan förväntas gälla för USA som helhet. Mot bakgrund av att de sociala skillnaderna i Sverige är mindre än de i USA finns det därför skäl att tro att effekterna av minskad klasstorlek blir mindre i Sverige, även om det är svårt att precisera hur stor skillnaden kan förväntas vara. Det faktum att undersökningen omfattade de första skolåren, där effekten av mindre klasser tenderar att vara störst, kan också medföra att effektskattningar baserade på STAR utgör en övre gräns för klasstorlekens betydelse.

Andra studier har utnyttjat naturligt förekommande variation i klasstorlek. De har därmed kunnat separera den kausalt betydelsefulla variationen från den variation i klasstorlek som är förknippad med andra variabler och med selektionseffekter.

Angrist och Lavy (1999) genomförde en studie i Israel, där det finns en regel om delningstal; när klasstorleken passerar ett tak delas klassen i två. Enligt den s.k. Maimonides regel, som formulerades på 1100-talet, är delningstalet 40 elever/klass. Eftersom den variation i klasstorlek som orsakas av denna regel kan tänkas vara oberoende av faktorer som elevernas socioekonomiska bakgrund kan den utnyttjas för att bestämma den kausala effekten av klasstorlek på elevernas prestation. Angrist och Lavy (1999) använde en instrumentvariabelansats på data från ungefär 2 000 klasser i årskurs 4 och 5. De fann att klasstorlek hade en signifikant effekt på prestation i läsning och matematik i åk 5, och på läsning i åk 4. Elever i små klasser hade bättre resultat: skattningarna tyder på att minskning av klasstorleken med 7 elever hade något lägre effekter än de som erhöles i STAR-undersökningen. För läsning i åk 5 var

effekten $d = 0,18$, och för åk 4 var effekten något lägre. Även denna studie visade att effekterna var starkare för de klasser som innehöll en större andel elever från socioekonomiskt missgynnad bakgrund.

I en studie med likartad ansats undersökte Hoxby (2000) variationer i klassstorlek inom skola som orsakades av variationer i populationsstorlek. Även i det amerikanska skolsystem tillämpas klasstorleksregler. Hoxby använde de abrupta förändringar i klasstorlek som dessa medförde för att studera den kausala effekten av klasstorlek. Hon använde ett material som omfattade 649 skolor för vilka data om antal klasser, klasstorlek och elevresultat fanns tillgängliga under en 12-årsperiod. Trots att Hoxby visar att den statistiska styrkan i hennes analyser var tillräcklig för att kunna upptäcka effekter av den storleksordning som STAR-studien visat på, kunde hon inte belägga någon signifikant effekt av klasstorlek i någon av analyserna.

Hoxbys (2000) resultat skiljer sig sålunda från dem som erhöles i STAR. Hennes tolkning av detta utfall är att STAR har gett missvisande resultat på grund av att lärarna kände till vilken experimentgrupp de hörde till. Detta kan enligt Hoxby ha medfört att lärarna som hade de små klasserna ansträngde sig extra för att visa på värdet av små klasser. I sin reanalys av data från STAR-experimentet visar dock Krueger (1999) att en sådan effekt inte är trolig, eftersom effekten av klasstorlek inom de olika betingelserna var av samma storleksordning som effekten av klasstorleksvariationen mellan betingelserna.

Det finns emellertid en annan förklaring till varför inte Hoxby fann någon positiv effekt av mindre klasser. Hennes analys bygger på antagandet att de lärare som anställdes för att undervisa i de tillkommande klasserna har samma kompetens som den befintliga gruppen av lärare. I synnerhet om det är brist på kvalificerade lärare, vilket under lång tid varit fallet i USA (se Hanushek 2003), är detta antagande inte rimligt, utan den nyrekryterade gruppen kan antas ha haft lägre kompetens. I den mån lärarkvalitet har betydelse för elevernas resultat kan vi sålunda förvänta oss att ökningen av antalet lärare medför en sänkning av elevernas resultat, vilket kan motverka en positiv effekt av den lägre klasstorleken. I princip gäller naturligtvis samma kritik för Angrist och Lavys (1999) studie. Denna studie genomfördes dock som en tvärsnittsstudie, medan Hoxbys (2000) studie huvudsakligen baserades på förändringar i populationsstorlek över tid. Populationsökningar medför givetvis ett generellt ökat behov av lärare, vilket kan antas leda till lärarbrist.

Hanushek och Luque (2003) använde data från TIMSS 1995 för att undersöka effekter av resurser på pedagogiska resultat, med särskilt fokus på skill-

nader i effekter för olika länder. De analyserade data för såväl åk 4 som åk 8, och modellerade relationer på skolnivå inom varje land mellan resurs- och bakgrundsfaktorer å ena sidan och resultat å den andra. För klasstorlek erhöles det förväntade negativa sambandet för 14 av 17 skattade relationer för åk 4. För åk 8 däremot erhöles ett positivt samband i 23 av 33 skattade relationer. Dessa resultat pekar på att mindre klasser i större utsträckning är gynnsamma för yngre elever än för äldre elever, även om det bör påpekas att denna studie är metodmässigt svagare än övriga studier som vi diskuterat här.

Resultatbilden i de undersökningar som presenterats här är något splittrad. Flera undersökningar baserade på omfattande datamaterial och kvalificerad analys pekar på att minskning av klasstorleken leder till bättre elevresultat. Hit hör STAR-experimentet, Angrist och Lavys (1999) studie, och resultaten avseende åk 4 i Hanushek och Luque (2003). Mot detta står de negativa resultaten i Hoxbys (2000) undersökning och resultaten för åk 8 Hanushek och Luques (2003) studie.

Som redan noterats kan Hoxbys (2000) negativa resultat möjligen förklaras av att minskningen av klasstorleken kan ha varit förknippad med en sänkning av lärarkompetensen. De studier som visar på positiva effekter av minskad klasstorlek tycks dock i första hand avse yngre elever, vilket också är i linje med resultaten från metaanalyserna av studier av klasstorlekens betydelse, som visar på en avtagande effekt som en funktion av elevernas ålder.

Den huvudslutsats som dessa undersökningar tillåter är att klasstorlek har betydelse för elevernas resultat, men att de är större för yngre elever. Resultaten pekar också på att mindre klasser har större betydelse för elever som är missgynnade med avseende på socioekonomisk bakgrund.

Vi återkommer till en mer utförlig diskussion om dessa resultat. Först skall vi dock redovisa resultat från några svenska studier.

7.1.4 Resultat från svenska undersökningar

Endast ett fåtal undersökningar av betydelsen av lärartäthet/klasstorlek har genomförts i Sverige, och eftersom vi här fokuserar på undersökningar som ger rimligt god grund för kausala slutsatser är antalet än mer begränsat.

Lindahl (2005) har presenterat en ny empirisk strategi för att från longitudinella data dra slutsatser om den kausala effekten av klasstorlek. Lindahl samlade vid tre tillfällen in resultat från ett och samma matematikprov för 556 elever i 16 skolor i Stockholm. Det första mättillfället var på vårterminen i skolår 5, det andra på höstterminen i skolår 6, och det tredje på vårterminen i

skolår 6. Datainsamlingen omfattade också bakgrundsdata för eleverna, information om klasstorlek, och ett frågeformulär till läraren.

I analysen studerade Lindahl förändringen i resultat mellan det andra och tredje provtillfället, med kontroll för förändringar mellan det första och det andra provtillfället. Idén med denna analys är att den ger information om förändringar under ett läsår, när klasstorlek kan förväntas ha effekt, med kontroll för förändringar under sommaren, då klasstorlek inte kan förväntas ha någon effekt medan däremot faktorer utanför skolan har det. Denna analysansats ger bättre kontroll över effekten av icke-observerade skillnader mellan elever än statistiska tekniker som kontrollerar för skillnader i observerade variabler. Denna mer sofistikerade analys visade att små klasser gav bättre resultat än större klasser.

Lindahl (2005) genomförde också separata analyser för elever med utländsk bakgrund och svenska elever. Dessa analyser gav det intressanta resultatet att det inte förelåg någon effekt för elever som inte hade utländsk bakgrund, medan en stark effekt kunde påvisas för elever med utländsk bakgrund.

Lindahl (2005) drog slutsatsen att i medeltal ökar elevprestationen med 2,8 percentilpoäng ($d = 0,10$) då klasstorleken minskas med 7 elever. Den starka interaktionen mellan klasstorlek och migrationsbakgrund gör det dock mindre meningsfullt att bara beakta medeleffekten. För elever med utländsk bakgrund förelåg en relativt stark effekt (ca. $d = 0,30$ vid minskning med 7 elever), medan effekten för elever med svensk bakgrund var nära 0 ($d = 0,04$). Det faktum att Lindahls undersökningsgrupp omfattade en icke-representativt stor andel elever med utländsk bakgrund (23 %) gör också att medelresultaten inte är direkt generaliserbara till skolor/kommuner med annan sammansättning av elevernas migrationsbakgrund. Huvudslutsatsen i Lindahls (2005) studie är sålunda att klasstorlek inte har någon betydelse för elever i åk 6 för elever med svensk bakgrund, men att den har en betydande effekt på resultaten för elever med utländsk bakgrund.

Fredriksson och Öckert (2008) har presenterat en studie som bygger på en likartad "difference-in-differences" teknik som den som användes av Lindahl (2005). De genomförde analyser på kommunnivå och relaterade skillnader i lärartäthet mellan åren 1992 och 2001 till skillnader i elevernas resultat, mätt med slutbetyg i åk 9 och med testresultat vid 18 års ålder. Utgångspunkten för denna analys var den avreglering och decentralisering av skolans styrning, inklusive resursfördelning, som genomfördes under början av 1990-talet. Som Fredriksson och Öckert visar ledde denna till en variation mellan kommunerna

i resurstilldelningen till skolan, och tanken var att utnyttja denna variation för att undersöka den kausala effekten av lärartäthet.

Huvudresultatet var att nedgången i lärartäthet från 8,8 år 1992 till 7,5 år 2001 ledde till en nedgång i elevprestation med 1,1 percentilpoäng ($d = 0,04$). Fredriksson och Öckert (2008) visar att om denna effektskattning översätts till en förändring som motsvarar STAR-experimentets minskning av klasstorleken med 7 elever skulle den motsvara 2,6 percentilpoäng, eller $d = 0,09$.

Fredriksson och Öckert (2008) gör också beräkningar där de använder det militära inskrivningsprovet som beroendevariabel, vilket givetvis medför att endast män ingår i analysen. Denna analys visade att effekten på testresultat var ungefär dubbelt så stor som effekten på betygen, vilket Fredriksson och Öckert (2008) tolkar som uttryck för att de mer objektiva testen bättre förmår fånga upp effekterna än betygen. De drar därför slutsatsen att effekten av en minskning av klasstorleken med 7 elever uppgår till $d = 0,17$, och konstaterar att detta är i linje med den effektstorlek som Krueger (1999) anger för STAR.

Fredriksson och Öckert (2008) gjorde också analyser där de kontrollerade för effekten av förändring i andelen behöriga lärare, utan att detta påverkade resultaten. De undersökte också förekomst av interaktionseffekter med kön, föräldrarnas utbildningsnivå, och utländsk bakgrund, utan att finna några signifikanta effekter. Det fanns dock en tendens att lärartätheten betydde mer för flickornas än för pojkarnas resultat, vilket var ett oväntat resultat. Det fanns också svaga tendenser att lärartätheten betydde mer för elever som inte hade föräldrar med postgymnasial utbildning, liksom för elever med utländsk bakgrund. Jämfört med den mycket starka interaktion med migrationsbakgrund som framkom i Lindahls (2005) studie var dock den interaktionseffekt som Fredriksson och Öckert (2008) fann av blygsam storlek.

Detta är en mycket intressant studie, som baseras på ett omfattande datamaterial. Det finns dock skäl att tolka resultaten med viss försiktighet. Difference-in-differences metoden kontrollerar för inflytande av sådana kommunegenskaper som är konstanta över tid. Därtill inkluderade Fredriksson och Öckert (2008) i sina regressionsmodeller både individkaraktistika (kön, ålder, migrationsbakgrund, och föräldrarnas utbildning) och variabler på kommunnivå (inkomst, skolstorlek, andel behöriga lärare, och medeltal av individegenskaper). Resultaten visade att när de olika kontrollvariablerna inkluderades i modellen reducerades effekten av lärartäthet, men effekten var fortfarande signifikant. Även om Fredriksson och Öckert (2008) visade att resultaten inte

påverkas av kontrollera för trender i elevresultat på kommunnivå, är det möjligt att ytterligare kontroller skulle ha reducerat effekten än mer.

Det finns också anledning att tolka den slutsats som dras på grundval av testresultaten med försiktighet. Den skattade effekten på testprestation är högre än med betygsutfall vilket är förvånande då det finns en inverkan av ytterligare två år, i allmänhet i gymnasiet, med en effekt som är delvis oberoende av effekten av lärartäthet i grundskolan.

Huvudslutsatsen från denna studie är sålunda att lärartätheten har en positiv effekt på elevernas resultat. Effekten, som uppgår till 0,09 standardavvikelseenheter är lägre än skattningar för USA. Eftersom detta är den enda storskaliga svenska studien som med avancerade metoder undersökt effekten av lärartäthet framstår det som mycket angeläget att forskningen intensifieras inom detta område, och att i synnerhet frågan om klasstorleken olika betydelse för elever med olika ålder görs till föremål för analys. En sådan studie förutsätter dock att data samlas in för tidigare årskurser än årskurs 9 vilket hittills varit fallet.

Även om studien avser gymnasienivån finns det skäl att kort nämna en finsk studie (Häkkinen m.fl. 2003), eftersom den bygger ett likartat angreppssätt som det som användes i Fredriksson och Öckerts (2008) studie. Även i Finland skedde under 1990-talets början kraftiga neddragningar av skolresurser pga. den ekonomiska krisen, vilka framförallt yttrade sig i en minskad lärartäthet. Häkkinen m.fl. (2003) undersökte effekterna av resursförändringar på resultatförändringar på både kommun- och skolnivå med en likartad ansats som Fredriksson och Öckert (2008). Trots god precision i skattningarna fann de dock inga effekter av minskningar i lärartäthet på gymnasisternas examensresultat. Resultaten från denna studie faller väl in i mönstret att lärartäthet/klasstorlek inte tenderar att ha effekt på studieresultat för äldre elever.

7.1.5 Diskussion och slutsatser

Vi har vid upprepade tillfällen kunnat konstatera att klasstorlek och lärartäthet har betydelse för elevernas resultat, men vi har också dragit slutsatsen att effekternas storlek i allmänhet varierar som en funktion av elevernas ålder och bakgrund. Effekter av klasstorlek är särskilt starka för årskurserna 1–6, och i STAR kunde uppkomsten av den positiva effekten av att gå i en mindre klass i huvudsak lokaliseras till den första årskursen. Det är emellertid viktigt att understryka att denna positiva effekt var relativt stabil över tid, och att den visade sig även i form av intresse för högre studier. Vidare är effekten av

klasstorlek betydligt starkare för elever med utländsk bakgrund och elever som är missgynnade vad gäller socioekonomisk bakgrund.

Det har framförts flera olika förklaringar till varför små klasser tenderar att ge bättre elevresultat än stora klasser (Gustafsson och Myrberg 2002; Lazear 2001), men vi har inte möjlighet att föra en detaljerad diskussion kring dessa i detta sammanhang. En av de mer intressanta och plausibla förklaringarna som har framförts till STAR-resultaten är dock att de små klasserna var mer effektiva när det gällde att socialisera eleverna till skolmiljön, där det bland annat ingår att lära sig reglerna för när och hur man skall arbeta med olika typer av skoluppgifter; Babcock och Betts (2009) har funnit visst stöd för denna hypotes. Det är rimligt anta att denna typ av färdigheter och vanor, när de väl etablerats, har stor generell betydelse, och då också över längre tid. Det är också rimligt att anta att inläring av denna specifika skolkultur och dessa skolspecifika färdigheter är av störst betydelse för de barn som inte haft tillfälle att förvärva dessa färdigheter i hemmiljön. En närliggande förklaring fokuserar på betydelsen av tidig läsinläring, och det är rimligt att tro att även denna gynnas av mindre klasser.

7.2 Lärarkompetens

I diskussioner om vad som faktiskt betyder något för resultaten i skolan finns konsensus om att en faktor, bortsett från elevernas familjbakgrund, är viktigare än någon annan faktor, och det är lärarkompetensen. När detta konstaterande är gjort upphör dock enigheten. Ett viktigt skäl till detta är oklarheter i definitionen av lärarkompetens. Det är därför nödvändigt att först diskutera vilken innebörd som skall läggas i denna term.

En grundläggande betydelse, som lyfts fram av bland andra Eric Hanushek, är att lärarkompetens avser lärarens förmåga att förbättra elevresultat: "High quality teachers are ones who consistently obtain higher than expected gains in student performance, while low quality teachers are ones who consistently obtain lower than expected gains." (Hanushek 2003, s. F90). Denna form av lärarkompetens kan betecknas som "prestationsrelaterad" kompetens eller som lärareffektivitet. Även om Hanushek beskriver lärareffektivitet i termer av faktiska elevprestationer är det naturligtvis möjligt att definiera denna i mer abstrakta och generella termer som lärarens förmåga att bidra till utveckling av elevernas kunskaper och färdigheter.

En annan innebörd refererar till den formella kompetensen, som att ha genomgått lärarutbildning och uppfylla andra krav för anställning som behörig

lärare. I den allmänna diskussionen är det vanligtvis denna innebörd som avses. Ytterligare en innebörd är "bedömd" eller "upplevd" lärarkompetens, som är den grad av kompetens som tillmäts en lärare av exempelvis elever, föräldrar, kollegor och skolledare.

Prestationsrelaterad, upplevd och formell lärarkompetens är tre mycket olika kompetensaspekter, men ändå används de ofta som om de vore utbytbara; denna sammanblandning har i sin tur skapat förvirring.

Det är sålunda vanligt att resultat som visar på avsaknad av samband mellan elevresultat och läraregenskaper, exempelvis i form av lärar- eller ämnesutbildning, tolkas som att lärare inte har någon betydelse för de resultat som elever uppnår i skolan (se t.ex. Coleman m.fl. 1966). Det kan givetvis finnas stora skillnader mellan lärare i prestationsrelaterad kompetens, samtidigt som dessa inte har samband med den formella lärarkompetensen. En anledning till detta kan vara att det faktiskt inte finns något samband, men det kan också vara så att sambandet inte visar sig på grund av mätproblem, eller andra metodproblem.

Det finns många metodproblem i anslutning till studiet av effekter av olika läraregenskaper på elevprestationer. Ett problem är att de mått som funnits tillgängliga på lärarakaraktistika ofta varit både grova och opålitliga, och dessutom svåra att jämföra över tid och mellan länder. De svenska studier som utförts har ofta tagit utgångspunkt i distinktionen mellan formellt behöriga/obehöriga lärare, där denna distinktion varit av administrativ snarare än vetenskaplig karaktär. Gränsdragningarna mellan de två kategorierna är därför mer eller mindre godtyckliga. Ett annat problem är att de flesta undersökningarna har varit av tvärsnittstyp. I denna typ av undersökningar är kausalitetsriktningen i observerade samband i allmänhet oklar, och det finns därför en stor mängd alternativa tolkningar av såväl observerade samband, som avsaknad av samband. Ett vanligt problem i utbildningssammanhang är omvänd kausalitet. Detta innebär att elevernas prestationer (om det är vår beroende variabel) påverkar den undervisning de får. Ofta används kompensatorisk resursallokering, vilket t.ex. innebär att elever med stora behov placeras i små klasser tillsammans med lärare med speciallärarutbildning; detta kan skapa det felaktiga intrycket att små klasser och speciallärarutbildning leder till låga resultat. Ett annat, mycket likartat, problem är att den variabel som i själva verket orsakat sambandet, och som är korrelerad med undersökningsvariabeln, inte finns med i undersökningen, vilket gör det omöjligt att nå fram till en korrekt tolkning av resultaten.

I dessa sammanhang är det svårt av såväl praktiska som etiska skäl, att använda sig av randomiserade experiment, vilka annars ger god grund för kausalinferens. Men det finns ändå några få undersökningar, där forskarna antingen kunnat genomföra randomisering, eller där de på grundval av omfattande datamaterial och uppfinningsrika analyser lyckats komma fram till vad som framstår som övertygande slutsatser. Vi diskuterar först dessa studier, för att sedan göra en genomgång av de svenska undersökningarna kring lärarkompetens.

7.2.1 Prestationsrelaterad lärarkompetens

Det är naturligt att ta utgångspunkt i frågan om det finns skillnader mellan lärare i prestationsrelaterad kompetens, och hur stora dessa skillnader i så fall är. Om det kan visas att det finns skillnader i prestationsrelaterad lärarkompetens blir det meningsfullt att ställa frågan om hur denna kan utvecklas, och vilka olika läraregenskaper som har samband med den. Om det däremot inte går att påvisa skillnader i elevernas uppnådda resultat som kan relateras till lärarna, är inte heller frågan om betydelsen av formell eller bedömd lärarkompetens av någon relevans.

Ett sätt att få kunskap om lärarfaktorns betydelse är att använda sig av en longitudinell ansats, där samma elever, lärare och skolor följs över tid. Rivkin m.fl. (2005) har redovisat inträngande analyser av ett omfattande undersökningsmaterial med kohortsekventiell design, insamlat i Texas. Studien omfattar samtliga elever i Texas (ca. 1,5 miljoner) som gick i åk 4 åren 1993, 1994 eller 1995. Dessa tre kohorter av elever följdes upp till åk 7 med en årlig insamling av reguljära prov i läsning och matematik. Till skillnad från svenska elever byter eleverna i Texas lärare varje år. Detta gjorde det möjligt att studera effekten av lärarkompetens på elevernas resultat, trots att det inte fanns information som knöt samman enskilda elever och enskilda lärare. Däremot fanns det uppgift om vilken skola eleverna och lärarna tillhörde, liksom uppgifter om flytt från en skola till en annan.

Den grundläggande analysansatsen och huvudresultaten kan illustreras av några enkla korrelationer mellan genomsnittliga resultatförändringar på skolnivå från ett år till ett annat. Om man följer en given kohort över årskurser är korrelationen nära 0, vilket visar att det inte fanns några systematiska skillnader i hur förändringarna såg ut på skolnivå. Men korrelationen över kohorter i en given årskurs var positiva och relativt höga. Grunden till detta var att dessa grupper av elever mötte samma lärare, så korrelationerna återspeglar det fak-

tum att det finns en systematisk variation i de resultatförändringar som olika lärare åstadkommer.

Analysmodellen som Rivkin m.fl. (2005) faktiskt använde för att skatta effekterna av variation i lärarkompetens var dock betydligt mer komplicerad än så. Den utnyttjade den kohortsekventiella designen och undersökte förändring på elevnivå, skolnivå och lärarnivå, där den senare typen av förändring undersöktes genom att studera effekten av andelen av lärarna som fanns kvar på skolan från ett år till ett annat. Genom att analysen fokuserade på förändring för fixa enheter kontrollerade den för utelämnade variabler. I analysen inkluderades också skolegenskaper som klasstorlek, och lärarvariabler, som utbildning på mastersnivå och undervisningserfarenhet.

Resultaten visade att den i särklass mest betydelsefulla faktorn för förändringarna i elevresultat var lärarfaktorn, i den bemärkelsen att olika lärare nådde olika goda resultat. Någon effekt av hög ämneskunskap i form av mastersexamen fanns inte. Däremot visade resultaten på en effekt av erfarenhet. I synnerhet var denna märkbar under år 1 och 2, men planade sedan ut.

Resultaten visade på något starkare lärareffekt på matematik än på läsning. Enligt Rivkin m.fl. (2005) var denna effekt så stark att en serie goda lärare förmår kompensera för de effekter på skolresultat av en mindre gynnsam hemmiljö. De visar också att effekten av att öka lärarkvaliteten en standardavvikelse är starkare än effekten av en minskning av klasstorleken med 10 elever.

Även om det inte finns anledning att ifrågasätta huvuddragen i resultatbilden från denna studie bör det påpekas att storleken på lärareffektskattningen kan vara svår att generalisera till andra länder, och att den förmodligen är underskattad. Anledningen till detta är att den avser variation i en grupp som är selekterad in i läraryrket, och där huvuddelen av lärarna har mångårig erfarenhet. Detta leder till en beskuren variation i denna population, och den beskurna variationen medför sänkta effektskattningar. Det bör också betonas att Rivkin m.fl. avsiktligt konstruerade en skattningsmetod som skulle ge skattning av ett lägre gränsvärde för lärarfaktorn, vilket var nödvändigt därför att det inte var möjligt att koppla resultaten för enskilda elever till enskilda lärare.

Rockoff (2004) har publicerat resultat från en longitudinell studie av lärareffektivitet som inte lider av denna begränsning. Han hade möjlighet att utnyttja data från två skoldistrikt, i en undersökning som sammantaget omfattade ca. 10 000 elever och 300 lärare. Eleverna följdes upp under de första 5–6 skolorn. Mätningar av elevernas resultat gjordes flera gånger per år i läsning och matematik. Rockoff (2004) kunde utnyttja dessa longitudinella data

för att skatta betydelsen av fixa lärareffekter på elevernas resultat. Detta gjordes med en så kallad "random coefficients" modell, där standardavvikelsen i en antagen normalfördelning av sanna lärareffekter inom skolorna skattades.

Med denna metod fann Rockoff (2004) att standardavvikelsen i fördelningen av fixa lärareffekter var 0,10 i såväl matematik som läsning. Detta betyder att om lärarkvaliteten ökar med 1 standardavvikelse ökar elevernas resultat med 0,1 standardavvikelse, vilket sammanfaller med resultatet i Rivkin m.fl. (2005). Men även Rockoffs (2004) resultat måste betraktas som en skattning av ett minimivärde av lärares faktiska betydelse, eftersom estimatet endast avser inomskolsvariationen. I USA finns en stor variation mellan skolor i lärarnas anställningsförhållanden och lönevillkor, liksom i elevernas resultat; Rockoff (2004) understryker att den totala variationen i lärarkvalitet därför kan antas vara betydligt större än vad som visas i variationen inom skolor.

Det finns även andra studier som använt angreppssättet att studera lärarvariation med hjälp av "fixa effekter". Nye m.fl. (2004) redovisar en sammanställning av 18 estimat av lärareffektens storlek baserade på 7 undersökningar. Denna sammanställning visar att mellan 7 % och 21 % av variationen i förändring i elevresultat kan hänföras till variation mellan lärare.

I sin artikel presenterar Nye m.fl. en undersökning med bättre metodegenskaper än någon av de tidigare presenterade undersökningarna, eftersom denna bygger på slumpmässig fördelning av eleverna till lärare. Studien baseras på data från den ovan beskrivna STAR-undersökningen. I reanalysen av data fokuserade dock Nye m.fl. på att undersöka variationen i lärareffektivitet inom skolor och experimentgrupp, vilket var möjligt därför att det i de större skolorna fanns mer än en klass inom varje grupp. Vidare gjorde man analyser som dels fokuserade på prestationsnivå i slutet av varje årskurs, dels förändring från föregående årskurs. Eftersom STAR var ett randomiserat experiment borde det egentligen vara tillfyllest att enbart analysera prestationsnivå, men eftersom det fanns en del svårigheter att upprätthålla den randomiserade designen på grund av in- och utflyttningar av elever valde man att även analysera förändringen.

Resultaten visade att de fixa lärareffekterna förklarade ungefär 8 % av variansen i läsning, medan den förklarade ungefär 11 % av variansen i matematik. När man fokuserade på förändringen var motsvarande siffror ungefär 7 % för läsning och 13 % för matematik.

Detta är en enda storskaliga experimentella undersökning av den prestationsrelaterade lärarkompetensen som vi känner till och det är svårt att se att estimaten i denna studie är påverkade av snedvridande faktorer. Det är också

intressant att notera att estimaten i denna studie sammanfaller med medelresultatet från de tidigare undersökningar som Nye m.fl. har med i sin meta-analys. I sammanfattning kan man säga att dessa resultat visar att lärarfaktorn förklarar 10 % av variansen i elevresultat, vilket är en betydligt högre skattning jämfört med vad Rivkin m.fl. (2005) och Rockoff (2004) erhöill. Deras metoder var dock konstruerade att ge en uppskattning av den lägre gränsen för lärarfaktorns betydelse, så det är rimligt att dessa studier gav lägre estimat.

Rivkin m.fl. noterade också att deras studie visade på en stor lärareffekt i jämförelse med exempelvis effekten av klasstorlek. Den starka lärareffekten framstår som mycket viktig, samt av stort teoretiskt och praktiskt intresse. Som en annan intressant jämförelse kan nämnas att lärarfaktorn enligt dessa resultat är lika betydelsefull för elevernas resultat som deras sociala och utbildningsmässiga bakgrund. En stor mängd forskning visar nämligen att ca. 10 % av elevernas prestationsskillnader kan hänföras till deras familjebakgrund.

Det finns anledning att nämna ytterligare ett intressant resultat från studien av Nye m.fl. (2004). De redovisar också analyser av storleken på lärareffekten inom skolor med olika socioekonomisk status (SES), varvid de jämförde de 25 % av skolorna som hade lägst SES med de 25 % av skolorna som hade högst SES. Resultaten visade på en betydligt starkare lärareffekt för både läsning och matematik inom gruppen av skolor med låg SES än inom gruppen av skolor med hög SES. Den förklaring som författarna lägger fram till detta resultat är att det sker en selektion av lärare med hög prestationsrelaterad kompetens från områden med låg SES till områden med hög SES. Detta innebär en homogenisering av lärarkvaliteten inom skolor i områden med hög SES, och den beskurna variationen medför att estimaten av lärarfaktorns betydelse blir lägre. Om denna förklaring är korrekt pekar den på hur beroende skattningarnas storlek är av omfattningen av variationen i den prestationsrelaterade lärarkompetensen.

Även om det är svårt att komma fram till en helt entydig bestämning av omfattningen av lärarfaktorns betydelse för elevernas resultat kan vi sammanfatta resultaten med att skillnader i prestationsrelaterad kompetens är mycket stora. Detta är av fundamental betydelse: om det inte går att visa att det finns en sådan variation, är fortsatt forskning om lärarkompetens av begränsat intresse. Det skulle i så fall inte finnas anledning att ställa frågor om hur kompetensen influeras av faktorer som utbildning, erfarenhet, och kunskaper, för att bara nämna några.

7.2.2 Vad påverkar den prestationsrelaterade lärarkompetensen?

Mot bakgrund av att den prestationsrelaterade lärarkompetensen uppvisar så stor variation är det av stort intresse att få reda på vilka faktorer som påverkar denna kompetens.

Det är då intressant att notera att vissa forskare intar den bestämda hållningen att observerade faktorer inte förklarar lärareffektiviteten. Hit hör Rivkin m.fl. (2005) som i diskussionen kring sina resultat säger:

Though it is tempting to tighten standards for teachers in an effort to raise quality, the results in this paper and elsewhere raise serious doubts about that more restrictive certification standards, education levels, etc. will succeed in raising the quality of instruction. Rather the substantial differences in quality among those with similar observable backgrounds highlight the importance of effective hiring, firing, mentoring and promotion practices. (Rivkin m.fl., 2005, s. 450).

Denna slutsats grundar sig på det faktum att man i denna studie, liksom i vissa andra studier, inte funnit några tydliga samband mellan olika läraregenskaper, t.ex. utbildning, och förändring i elevernas resultat. Ett problem är dock att man i undersökningen inte hade någon annan information om lärarnas utbildning än den högsta avlagda akademiska examen, vilken visade sig sakna betydelse. På denna grund kan Rivkin m.fl. (2005) inte uttala sig om betydelsen av ämnesmässig kontra pedagogisk utbildning, eller om betydelsen av ämnesfördjupning i undervisningsämnet. Resultatet att lärarnas utbildningsnivå inte har betydelse är i överensstämmelse med vad som framkommit i en lång rad andra studier. Dock ger det inte någon grund för generella slutsatser om lärarutbildningens betydelse. Härtill kommer det generella problemet att Rivkin m.fl. (2005) försöker sig på konststycket att bevisa icke-existensen av ett samband. Att detta är en omöjlighet följer av att det finns ett antal omständigheter som kan förhindra att ett existerande samband framträder som statistiskt signifikant, t.ex. problem med mätinstrumentens reliabilitet eller validitet och låg statistisk styrka.

Det finns en omfattande forskning kring samband mellan olika läraregenskaper och elevers prestationer. Lärarens ämnesmässiga kunskaper och färdigheter, pedagogiska utbildning, undervisningserfarenhet, och lärarbehörighet är exempel på faktorer som undersökts. Denna forskning är inte lika metodologiskt stringent som de studier som vi redovisat ovan, och resultatbilden framstår som mycket heterogen. Enligt vissa forskare (t.ex. Greenwald m.fl. 1996) pekar resultaten på att vissa observerbara lärarkaraktistika har starka

samband med prestationsrelaterad lärarkompetens, medan andra forskare (t.ex. Hanushek 2003) ställer sig avvisande till tanken att det skulle finnas några observerade faktorer som påverkar eller predicerar lärareffektivitet.

7.2.3 Lärares kunskaper och färdigheter

En hypotes är att lärares egna kunskaper och färdigheter inom ämnesområdet är en viktig faktor bakom den prestationsrelaterade lärarkompetensen. I flera amerikanska studier har man haft tillgång till olika typer av mått på lärares kunskaper och färdigheter, och har kunnat relatera dessa till elevresultat. Bland annat har man kunnat utnyttja resultat på prov som använts i samband med lärarcertifiering.

Ett metodproblem i dessa studier är dock att lärare och elever praktiskt taget aldrig kombineras på ett slumpmässigt sätt, och det är lätt att hitta exempel på att duktiga elever väljer duktiga lärare, liksom att duktiga lärare väljer duktiga elever. Det finns olika sätt att hantera denna typ av selektionseffekter, och det vanligaste är att utnyttja en longitudinell ansats. I longitudinell design kan man använda ett eller flera mått på elevernas prestationsnivå innan de fått undervisning av de lärare som ingår i undersökningen och sedan upprepa mätningen efter undervisningen. Med en sådan design är det åtminstone i princip möjligt att bestämma "value added" för eleverna, och relatera detta till läraregenskaper. Rothstein (2008) visade dock med ett omfattande empiriskt material att även "value added" modeller som innehåller rik information om elevernas tidigare prestationer leder till systematiska fel vid skattning av individuella lärareffekter. En anledning till detta är att fördelningen av elever över lärare inte är slumpmässig även efter kontroll för tidigare elevresultat. Rothstein (2008) visar också på betydande skillnader i skattningar av kort- och långsiktiga lärareffekter, vilket bl.a. kan förklaras av att kortsiktiga effekter kan uppstå genom att läraren förbereder eleverna genom testträning. Det måste dock understrykas att dessa metodproblem framförallt leder till svårigheter då syftet är att använda individuella skattningar av lärareffekter för beslut angående enskilda lärare avseende t.ex. lön eller fortsatt anställning, medan de lättare kan hanteras då syftet är att studera variation i lärareffekter på gruppnivå, som i de studier som presenteras i avsnitt 7.2.1.

En likartad ansats är att använda regressionsmodeller med s.k. fixa elev-effekter för att kontrollera för selektion. Ett exempel där vi kan använda sådana modeller är om vi mäter elevens resultat inom flera olika ämnen men där bara ett ämne har undervisats av en viss lärare. Avvikelsen mellan medelresultatet

över alla ämnen och det ämne som den i undersökningen ingående läraren undervisat i innebär då en god kontroll över selektionseffekterna. Nackdelen med denna ansats är att den i allmänhet medför underskattning av läraregenskapernas effekter.

En översiktsartikel av Wayne och Youngs (2003) sammanställer forskningsresultat kring effekter av lärarkaraktistika på elevers resultat. Genom litteratursökningar identifierade de ett mycket stort material av amerikanska studier, men efter att de applicerat selektionskriterier som avsåg möjligheten att på grundval av studiens design dra korrekta slutsatser om kausalitet återstod endast 21 undersökningar. Dessa undersökningar var i allmänhet longitudinella 'value added' studier. Resultaten från dessa redovisas kortfattat nedan, tillsammans med resultat från europeiska, och då i första hand svenska undersökningar.

Wayne och Youngs (2003) översikt innehåller sju undersökningar där man använt sig av olika typer av test på lärares kunskaper och färdigheter, som läsförmåga och ordförråd, liksom kunskaper inom olika ämnesområden. Av de sju undersökningarna visade fem på positiva relationer med förändring av elevernas resultat. I de två andra studierna som inte visade på ett positivt samband hade man dock i analysen inkluderat ytterligare en prestationsvariabel som avsåg college-resultat, som i mycket mätte samma sak som testen, vilket medförde att de partiella regressionskoefficienterna för både testen och college-resultaten underskattades. Wayne och Youngs (2003, s. 100) drar slutsatsen att eleverna lär sig mer då deras lärare har goda testresultat, men det är inte möjligt att dra mer specifika slutsatser om vilka typer av förmågor som är viktiga komponenter i lärarkompetensen på grundval av dessa studier.

Frågan om vilka kunskaper som en lärare bör besitta är mycket omdiskuterad, där vissa menar att ämneskunskaper är viktigast, medan andra menar att pedagogiska kunskaper är de viktigaste. Shulman (1986) gjorde en diskntktion mellan "content knowledge" (ämneskunskaper), "pedagogical knowledge" (pedagogiska kunskaper) och "pedagogical content knowledge", vilket ungefär motsvarar ämnesdidaktiska kunskaper. Han menade att alla dessa kategorier av kunskap är viktiga, men att i synnerhet de ämnesdidaktiska kunskaperna är centrala. Baumert m.fl. (2009) har noterat att mycket av forskningen kring betydelsen av lärares kunskaper fokuserat på allmänna kunskaps- och färdighetsmått, och att man endast i begränsad utsträckning försökt skilja mellan de tre kategorier av lärarprofessionell kunskap som Shulman (1986) introducerade.

Baumert m.fl. (2009) redovisar dock en undersökning där man studerat olika aspekter av lärares matematikkunskaper, och deras effekter på elevers lärande. Den tyska delen av PISA 2003 vidareutvecklades till en longitudinell design genom att man följde upp eleverna efter ett år, då de gick i åk 10. Elevernas kunskaper, färdigheter och attityder i matematik mättes med PISA-instrumenten i åk 9 och åk 10. För lärarna mätte man både ämneskunskaper och ämnesdidaktiska kunskaper i matematik med egenutvecklade mätinstrument (se Krauss, m.fl., 2008). Provet på ämneskunskaper i matematik omfattade uppgifter av relativt hög komplexitetsgrad, som krävde matematiska bevis och argumentation. Det ämnesdidaktiska provet var nyutvecklat och bestod av tre kategorier av uppgifter: den första kategorin innehöll uppgifter som skulle besvaras med så många olika lösningar som möjligt; i den andra presenterades hypotetiska klassrumssituationer och uppgiften var att upptäcka, analysera eller förutse typiska förståelseproblem eller missuppfattningar bland eleverna; i den tredje kategorin skulle lärarna ange olika sätt att presentera och förklara vanliga matematikproblem, som *Hur skulle Du förklara varför $(-1) * (-1) = 1$?* Som visats av Krauss m.fl. (2008) mätte matematikprovet och det ämnesdidaktiska provet två olika, men relativt högt korrelerade, dimensioner.

I analysen användes strukturell ekvationsmodellering på två nivåer (klasser och elever) för att analysera vilka faktorer som påverkade elevernas resultatförändring mellan åk 9 och åk 10. Både lärarnas matematikkunskaper och deras ämnesdidaktiska kunskaper hade effekt på elevernas kunskapsutveckling, men sambandet var starkare med de ämnesdidaktiska kunskaperna. I studien samlades också data in avseende olika kvalitetsaspekter på undervisningen, som uppgifternas kognitiva nivå och lärarens bemötande av eleverna i undervisningen. Effekten av de ämnesdidaktiska kunskaperna visade sig medieras av dessa kvalitetsaspekter, vilket inte var fallet med lärarnas matematikkunskaper.

Denna studie visar att det är möjligt att separera ämneskunskaper och ämnesdidaktiska kunskaper, och att det är de senare som är mest betydelsefulla för elevernas resultatutveckling. Av detta följer givetvis inte att ämneskunskaper är oviktiga för matematiklärare; slutsatsen är snarare att ämneskunskaper är nödvändiga men inte tillräckliga för att utveckla de ämnesdidaktiska kunskaperna. Denna studie ger också grund för viss optimism: de kunskaper och färdigheter som är betydelsefulla för den prestationsrelaterade lärarkompetensen verkar vara avgränsningsbara och mätbara, men detta måste göras med en högre grad av teoretisk och metodmässig precision än vad som kännetecknat stora delar av den tidigare forskningen inom fältet.

Grönqvist och Vlachos (2008) har presenterat en mycket intressant studie av effekten av kognitiva och icke-kognitiva läraregenskaper på elevers resultat i åk 9 i den svenska grundskolan. En källa till information om läraregenskaper var registerdata från inskrivningsprövningarna för militärtjänst, som var obligatorisk för 18-åriga män. På detta sätt fick man uppgifter om allmänbegåvning, mätt med olika test, liksom psykologbedömningar av ledarskapsförmågor, som psykisk stabilitet, initiativkraft, ansvarskännande, och social interaktionsförmåga. En annan källa till information var registerinformation om lärarnas avgångsbetyg från gymnasiet. Medan data från inskrivningsprövningarna bara omfattade män, fanns betygsdata tillgängliga för både män och kvinnor. Dessa uppgifter kombinerades med information från Lärarregistret om befattningar. Lärarinformationen kombinerades även med resultat på nationella prov i åk 9 för lärarnas elever under åren 2003–07. Elevdata hämtades dels från centrala register över resultat på nationella prov i svenska, engelska och matematik, dels från nio stora kommuner i landet som byggt egna register över lärare och elever. Två datamängder skapades, den ena med information från inskrivningsprovet, och den andra med betygsdata för lärarna. Den första datamängden omfattade 704 ämneslärare födda under perioden 1951 till 1981 och ca. 25 000 elever, medan den andra omfattade 1589 ämneslärare som avslutat gymnasiestudierna år 1985 eller senare och ca. 45 000 elever.

Grönqvist och Vlachos (2008) analyserade data med regressionsmodeller där de inkluderade läraregenskaper tillsammans med ett antal kontrollvariabler dels för lärarna (bl.a. födelsekohort, ämne, gymnasieprogram), dels för eleverna (kön, föräldrarnas utbildningsnivå, och migrationsbakgrund). I syfte att helt utesluta risken för att samband mellan läraregenskaper och elevresultat skall orsakas av selektionseffekter tog de i sin specifikation också in fixa elev-effekter. Detta innebär att den beroendevariabel som studerats är avvikelsen mellan resultatet på det nationella provet i ett visst ämne och medeltalet över de tre nationella proven.

Grönqvist och Vlachos (2008) huvudslutsats var att det inte finns någon generell effekt av lärarnas egenskaper i kognitiva och icke-kognitiva hänseenden. En annan slutsats var att det finns betydelsefulla interaktionseffekter som gör det nödvändigt att ta hänsyn till både lärar- och elevgenskaper.

För manliga lärare fanns det en positiv effekt av gymnasiebetyg på elevresultat, men inte för kvinnliga lärare. Olika elevgenskaper interagerade också med läraregenskaper. Bland annat hade elever med hög studieförmåga, mätt med betyg i åk 9, bättre resultat då läraren hade hög kognitiv förmåga, medan

icke-kognitiva läraregenskaper hade större betydelse för elever med låg studieförmåga. För pojkar var det negativt att ha lärare med höga gymnasiebetyg, medan det inte fanns någon effekt för flickor.

Det finns dock skäl att betrakta huvudslutsatsen om avsaknad av generella effekter av läraregenskaper med viss försiktighet. Modellspecifikationer med denna typ av fixa eleveffekter riskerar att ta bort en del av de "sanna" lärareffekterna, ett problem som kan accentueras att resultaten på de nationella proven är betygssatta i en fyrgradig skala (IG, G, VG, och MVG); denna skala är alltför grov för att på ett precist sätt kunna fånga upp prestationsvariation över ämnen.

7.2.4 Den formella kompetensens betydelse

Stort intresse har också riktats mot frågan om vilken betydelse som examina på olika nivåer har; i den amerikanska forskningen har fokus framförallt riktats mot betydelsen av en mastersexamen. Som redan påpekats visade studien av Rivkin m.fl. (2005) inte på någon effekt av mastersexamen, och många andra studier har fått fram samma resultat. Wayne och Youngs (2003) kommer dock fram till att dessa negativa slutsatser har sin grund i att man i forskningen i allmänhet inte tagit hänsyn till om lärarens undervisningsämne är det som mastersexamen avser, liksom att man inte heller tagit hänsyn till elevernas åldersnivå. Deras forskningsöversikt visar att samtliga studier där man undersökt elevernas framsteg i matematik, visar på en positiv effekt av mastersexamen i matematik. Dessa studier avser elever på vad som motsvarar svensk gymnasienivå, och författarna understryker att det inte finns underlag att dra några slutsatser om detta är giltigt för grundskolenivån eller ej.

Wayne och Youngs (2003) noterar också att många undersökningar inte gjort någon distinktion mellan fördjupade ämnesstudier och ämnesdidaktisk fördjupning, och de påpekar att "Studies that use subject-specific measures of teacher preparation and that distinguish between subject preparation and preparation in the methods of teaching a subject would provide valuable new information" (s. 104). Resultaten från Baumert m.fl. (2009) understryker vikten av detta påpekande.

En mycket omfattande forskning har fokuserat på lärarbehörighetens ("teacher certification") betydelse för elevernas resultat. Denna fråga är dock, om möjligt, än svårare att angripa än frågan om betydelsen av andra läraregenskaper. En anledning till detta är att regelverket för lärarbehörighet varierar mellan olika länder, liksom inom olika länder. I t.ex. USA bestämmer varje

delstat behörighetsreglerna. Möjligheterna att dra direkta slutsatser om lärarbehörighetens betydelse från denna forskning för svenska förhållanden är därför mycket begränsad, och det är därför nödvändigt att här särskilt fokusera på den svenska forskningen.

Det finns dock anledning att nämna de slutsatser som Wayne och Youngs (2003) kom fram till på grundval av sin översikt. Detta är snabbt gjort då de endast kunde identifiera två studier som uppfyllde de uppställda kriterierna. Liksom för mastersexamen är huvudslutsatsen att certifiering som avser adekvat utbildning för undervisningsämnet har betydelse för elevernas resultat. Studierna visar sålunda att allmän behörighet inte har betydelse medan lärare med behörighet för undervisning i matematik åstadkom bättre resultat i matematikundervisning än lärare utan behörighet för matematikundervisning.

De svenska studierna av lärarkompetensens betydelse har tagit utgångspunkt i uppgifter i Lärarregistret, som anger om läraren är behörig eller ej. För behörighet krävs att ha genomgått lärarutbildning i form av pedagogisk högskoleutbildning. Bland dem med pedagogisk högskoleutbildning görs också en distinktion om utbildningen är adekvat för den aktuella befattningen eller ej. Med dessa uppgifter som grund går det dock inte att identifiera den kategori av lärare som har en för tjänsten adekvat ämnesmässig utbildning men som saknar pedagogisk utbildning. Ett annat problem är att kategoriseringen som bestämmer om lärare med pedagogisk högskoleutbildning har en befattning som motsvarar utbildningen eller ej är relativt grov. På grundval av uppgifterna i Lärarregistret är det därför svårt att undersöka frågor om den relativa betydelsen av ämnesutbildning och pedagogisk utbildning.

Ett annat metodproblem är att jämförelser mellan behöriga och obehöriga lärare, avser grupper som utsatta för olika selektionsmekanismer till anställning som lärare. Personer som genomgår lärarutbildning måste exempelvis satsa en stor mängd tid, kraft och pengar för att fullfölja utbildningen. Det finns anledning att tro att denna selektion medför ett positivt urval av personer till läraryrket. Härtill kommer effekten den påverkan på professionsrelevanta kunskaper och färdigheter som utbildningen i sig medför, och som vi vill bestämma effekten av.

Personer utan lärarutbildning som fått tillsvidareanställning har även de genomgått en selektion. Enligt Skollagen får en lärare som saknar utbildning inte anställas om det finns utbildade lärare att tillgå. En person utan lärarutbildning kan dock ges tillsvidareanställning under följande förutsättningar: (1) om det saknas behöriga sökande till tjänsten; (2) om det finns särskilda skäl att

anställa den obehörige; (3) om den obehörige har motsvarande kompetens för den undervisning som anställningen avser; samt (4) om det finns skäl att anta att den obehörige är lämpad att sköta undervisningen. Dessa krav innebär att en person utan lärarutbildning för att få anställning tills vidare måste prövas avseende både ämnesmässig utbildning och förmåga att undervisa. Detta kan medföra att lärare utan utbildning är mer noggrant utvalda än lärare med utbildning. Det finns skäl att anta att en person utan utbildning först prövar läraryrket genom en eller flera visstidsanställningar och att erfarenheter från dessa ger underlag både för personens beslut om fortsatt verksamhet som lärare och kommunens beslut om eventuell tillsvidareanställning. En sådan selektionsprocess kan antas vara effektiv i att identifiera lärare med god kompetens för läraryrket. Det är också rimligt att tro att denna selektion är mer effektiv än selektionen till och inom lärarutbildningen, eftersom den sker direkt med avseende på yrkeskompetensen. Om denna hypotes är korrekt kan detta innebära att en direkt jämförelse mellan personer med och utan lärarutbildning leder till underskattning av lärarutbildningens effekt, eftersom en sådan jämförelse inte tar hänsyn till att två grupper som selekterats på olika grunder jämförs med varandra.

Andersson och Waldenström (2007) har genomfört en analys där de använt sig av en tydlig empirisk strategi för att skatta effekten av behöriga lärare på elevprestationer. De använde sig av avgångsbetygen för samtliga elever som slutat grundskolan under åren 1997/98–2003/04 och kombinerade denna information med information från bl.a. Lärarregistret. Genom att utnyttja resultaten från de olika åren för samma skolor kunde de utnyttja sig av en ”fixed effects” ansats och undersöka förändring på skolenivå. För att skapa ett instrument i en IV-ansats kombinerade de information om omfattningen av arbetslösheten bland behöriga lärare inom kommunerna med information om medel som tilldelats kommunerna för att öka personaltätheten inom ramen för de s.k. Wärnerssonpengarna.

I analysen användes såväl vanlig regressionsanalys, som regressionsanalys med fixa effekter, och IV-estimering med fixa effekter. Dessa analyser gav olika resultat, men Andersson och Waldenström (2007) fäste avseende enbart vid IV-resultaten. Dessa visade på en starkt positiv effekt av andelen behöriga lärare på skolan på elevernas resultat. I syfte att konkretisera innebörden i resultaten räknade Andersson och Waldenström (2007) den förväntade resultat-skillnaden mellan skolor med en andel behöriga lärare vid 5:e respektive 95:e percentilen. De fann att denna skillnad var lika stor som könsskillnaden i betyg, vilken kan betraktas som betydande.

De starkaste positiva effekterna av behöriga lärare uppträdde för elever med högutbildade föräldrar. Andersson och Waldenström (2007) tolkar tentativt detta resultat som orsakat av att behöriga lärare kan vara mer betydelsefullt för de högpresterande eleverna än för lågpresterande elever.

Skolverket (2006) har redovisat resultat från fördjupade analyser av den Nationella Utvärderingen 2003 där bland annat elev- och lärarenkäter analyserats tillsammans med mått på elevresultat. Studien omfattade även elevers bedömningar av lärarens kompetens liksom lärares självbedömda kompetens. Undersökningen bygger på ett omfattande material som består av 120 skolor, 1 688 lärare och 6 788 elever.

I denna undersökning utgjorde lärarnas utbildning en av de variabler som fokuserades; dessa relaterades bl.a. till elevernas bedömningar av lärarnas undervisningsskicklighet och till lärarnas självbedömningar. I studien gjordes en uppdelningen av lärarna i två ungefär lika stora grupper, där den ena hade både lärarutbildning och adekvat ämnesutbildning för undervisningsämnet, och den andra inte uppfyllde något av dessa villkor. Denna dikotoma lärarutbildningsvariabel hade samband med elevernas bedömningar av lärarskickligheten – ett samband som var starkare bland högpresterande elever, liksom i Andersson och Waldenströms (2007) studie. Det fanns även ett samband med utbildningsvariabeln och lärarnas självbedömda kompetens i metodik och didaktik. Det måste dock betonas att denna undersökning endast undersöker samvariation mellan lärarutbildning och de olika elevvariablerna, vilket ger ett stort utrymme för inflytande från andra faktorer.

7.2.5 Rekrytering till läraryrket

Mot bakgrund av att lärares kunskaper och färdigheter visat sig vara betydelsefulla för elevernas lärande är det rimligt att förvänta sig att graden av selektion på grundval av studieprestationer har betydelse för den prestationsrelaterade lärarkompetensen.

Fredriksson och Öckert (2007) visar att andelen lärarexamina bland det totala antalet högskoleexamina har minskat kraftigt sedan slutet av 1960-talet. Med hjälp av information från ett begåvnings-test som givits till urval om ca. 10 % av eleverna i åk 6 för fem kohorter födda mellan 1948 och 1977 inom det s.k. UGU-projektet undersökte de också den relativa placeringen i resultaten på dessa test för dem som blev lärare, jämfört med dem som gick till andra yrken. För gruppen födda 1948 presterade lärare 2 percentilpoäng sämre än andra,

medan motsvarande skillnad för dem födda 1977 uppgick till nästan 8 procentpoäng. Dessa resultat pekar på en försämring av rekryteringen till läraryrket.

Fredriksson och Öckert (2007) visar också att den ekonomiska avkastningen av lärarutbildning successivt försämrats under samma tidsperiod, att detta skett i större utsträckning än för andra högskoleutbildningar, och att lönefördelningen uppvisar mindre spridning än fördelningen inom andra yrken. Deras resultat pekar också på att det är personer med goda testresultat som i störst utsträckning lämnat läraryrket. Slutsatsen är att de minskade monetära incitamenten att bli lärare förklarar den försämrade rekryteringen till läraryrket.

Grönqvist och Vlachos (2008) har i den ovan beskrivna undersökningen också redovisat förändringar i resultat på begåvningstest och gymnasiebetyg bland nyblivna ämneslärare på grundskolans högstadium under de senaste decennierna. För de män som ingick i studien noterar de en uppgång från ett medelvärde vid 66 vid 1980-talets början med ca. 6 steg vid 1990-talets början. Därefter har det skett en kraftig nedgång med ca. 20 steg till år 2006. En analys av förändringar i testresultat för bröderna till manliga och kvinnliga lärare visade på samma mönster för båda könen, även om nedgången var kraftigare för de kvinnliga lärarna. Även för gymnasiebetygen (som de kunde studera under perioden 1993 till 2006) fann de samma mönster av nedgång, dock utan några tydliga könsskillnader. Denna undersökning visar på samma huvudresultat som det som erhöles av Fredriksson och Öckert (2007), men med bättre precision på grund av det större materialet och flera mätpunkter.

Det bör i detta sammanhang också påpekas att den främsta anledningen till att andelen lärarexamina bland det totala antalet examina minskat är att det skett en kraftig utbyggnad av antalet högskoleplatser sedan 1960-talet (se kapitel 10), och att ökningen av platser avsett andra utbildningsområden än lärarutbildning. Ökningen av platsantalet medför med nödvändighet att medelprestationen på betyg och prov bland dem som antagits till högskolestudier sjunker, och i den mån lärarutbildningen inte förmår att upprätthålla sin konkurrenskraft om de sökande kommer medelprestationen att minska mer bland de lärarstuderande än bland studerande på andra utbildningar. Detta gör det också av intresse att i den fortsatta forskningen göra jämförande studier av hur utvecklingen av rekryteringen till olika utbildningar gestaltat sig.

7.3 Diskussion och slutsatser

De studier som redovisats kring betydelsen av lärarvariationen som sådan ger en stabil grund för slutsatsen att ungefär 10 % av variationen i elevernas resul-

tat kan hänföras till lärarfaktorn. Detta är en mycket betydelsefull variationskälla som är ungefär lika betydelsefull som elevernas sociala och utbildningsmässiga bakgrund. En implikation av detta är att en serie goda lärare skulle kunna kompensera för det utbildningsmässiga handikapp som elever som kommer från hem utan utbildningstradition har. Medan en elev i allmänhet endast har två föräldrar möter dock varje elev i allmänhet ett stort antal lärare under sin utbildningstid, vilket tenderar att medföra en utjämnande effekt. I den mån det finns en systematisk variation i lärareffektivitet mellan skolor bidrar denna till prestationsvariation mellan skolorna. Det måste dock understrykas att skattningarna av lärarvariationens betydelse helt och hållet bygger på amerikansk forskning och det är inte självklart att dessa resultat är direkt överförbara till svenska förhållanden. Det finns sålunda ett stort behov att genomföra sådana studier även i Sverige.

Även om de faktorer som är associerade med variationen i lärareffektivitet endast är delvis kända framstår det som en alltför extrem ståndpunkt att hävda att denna variation inte är relaterad till några observerbara variabler. Kunskapsöversikter visar att lärares kunskaper och färdigheter delvis förklarar lärareffektivitet – inte minst framstår mått på ämnesdidaktisk kompetens som viktiga förklaringsvariabler. Denna forskning är dock fortfarande i sin linda och behöver utvecklas för att den skall kunna ge en solid kunskapsbas. Det finns också visst stöd för att både lärarutbildning och ämnesutbildning har betydelse för elevernas resultat, där det dock också framstår som viktigt att undersöka interaktionseffekter med elevernas prestationsnivå.

Några slutsatser från en nyligen utgiven dansk kunskapsöversikt (Nordenbo m.fl. 2008) är också värda att lyfta fram i detta sammanhang. I en systematisk litteraturinventering identifierade de 55 studier med acceptabel forskningskvalitet, vilka utgjorde grund för en narrativ forskningsöversikt. Huvudslutsatsen var att man identifierade tre kompetensdimensioner som betydelsefulla för framgångsrikt lärararbete. Den första dimensionen betecknade man ”relationskompetens” och den tar sig form av en positiv social kompetens där läraren visar respekt, tolerans och intresse för eleverna. Lärare med hög nivå i denna kompetensdimension kombinerar ledning av elevernas arbete med elevaktiverande och elevmotiverande insatser.

Den andra dimensionen betecknade man som ”regelledningskompetens” och avser förmågan att etablera regler för klassens uppförande och arbete. Detta innebär att läraren involverar eleverna i strukturering och val av aktiviteter. Det innebär också att lektionerna startas enligt schema, och att huvud-

delen av tiden använts till undervisning. Denna dimension innebär vidare en förmåga att fokusera klassens uppmärksamhet på de centrala delarna av innehållet, repetition av viktiga principer, och snabb och relevant feedback.

Den tredje dimensionen betecknades ”didaktikkompetens”. Detta innebär att läraren har goda ämneskunskaper, på vilken grund undervisningen kan genomföras på flexibla sätt, och anpassas efter elevernas behov och lokala förhållanden. Denna dimension innebär också att läraren ställer upp tydliga mål, såväl på längre som kortare sikt, och planerar och genomför undervisningen så att undervisningstiden maximeras.

Den ovan presenterade inventeringen av vilka resursfaktorer som är betydelsefulla för elevernas resultat har fokuserat dels på de kvantitativa lärarresurserna, dels på kvalitetsmässiga aspekter av lärarresurserna. Från den samlade resultatbilden framstår några huvudslutsatser, vilka också kan vara intressanta i en diskussion om den utbildningspolitiska inriktningen.

En huvudslutsats är att kvantitativa resursfaktorer – lärartäthet och klasstorlek – har betydelse för elevernas resultat. Positiva effekter av en minskning av klasstorleken verkar i första hand förekomma i lägre årskurser. Och effekterna är starkast för minoritets elever och elever med svagare socio-ekonomisk bakgrund. En viktig uppgift för den fortsatta forskningen är att klargöra de mekanismer som förklarar dessa effekter, och att tydliggöra betydelsen av lärartäthet respektive klasstorlek för elevernas resultat.

En annan huvudslutsats är att det finns stora skillnader i utbildningsresultat mellan olika lärare. I viss utsträckning går det att koppla dessa skillnader i resultat för olika lärare till observerbara och påverkbara egenskaper, som ämneskunskaper, pedagogisk utbildning, undervisningserfarenhet och fortbildning, liksom till vissa personliga egenskaper. En stor del av förklaringsfaktorerna till variationen i lärareffektivitet återstår dock ännu att upptäcka.

Det är intressant att notera att en enda undersökning, nämligen STAR-experimentet, bidragit med underlag för båda dessa huvudslutsatser. Experimentet genomfördes med syfte att undersöka effekten av klasstorlek, men som visats av Nye m.fl. (2004) ger denna studie grund för slutsatser också avseende lärarnas betydelse. Initialt kom dock endast resultaten kring klasstorlek att uppmärksammas i den utbildningspolitiska diskussionen i USA, liksom även i Sverige (Krueger och Lindahl 2002). På flera håll i USA genomfördes också reformprogram i syfte att minska klasstorleken, i första hand i de lägre årskurserna. I Kalifornien startades ett omfattande sådant program vid mitten av 1990-talet, vilket följts upp och utvärderats (Bohrnstedt och Becher 2002).

Programmet innebar en minskning av klasstorleken i åk 0–3, och var mycket populärt bland elever och föräldrar. Några positiva effekter på elevresultaten kunde dock inte påvisas (Bohrnstedt och Becher 2002). Förklaringen till detta synes vara att det stora behov av nyanställning av lärare som programmet medförde ledde till att det var nödvändigt att rekrytera mindre kvalificerade lärare; sänkningen av lärarkompetensen motverkade den positiva effekten av den minskade klasstorleken. Erfarenheterna från Kalifornien visar således på farorna med ett alltför ensidigt fokus på lärartätheten utan att beakta lärarkompetensen.

Referenser

- Andersson, C. och N. Waldenström (2007), Teacher certification and student achievement in Swedish compulsory schools. IFAU Working paper 2007:6.
- Angrist, J., och Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics*, 114, 533–575.
- Babcock, P., och J. Betts (2009), Reduced-class distinctions: Effort, ability, and the education production function. *Journal of Urban Economics*, 65, 314–322.
- Baumert, J., M. Kunter, W. Blum, M. Brunner, T. Dubberke, A. Jordan, U. Klusman, S. Krauss, M. Neubrand, och Y-M. Tsai (2009), Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*. doi:10.3102/0002831209345157
- Blatchford, P., H. Goldstein, och P. Mortimore (1998). Research on class size effects: a critique of methods and a way forward, *International Journal of Educational Research*, 29, 691–710.
- Bohrnstedt, G.W. och B.M Stecher (red.) (2002), *What we have learned about class size reduction*. Sacramento, CA: California Department of Education.
- Coleman, J.S., E.Q. Campbell, C.J. Hobson, J. McPartlant, A.M. Mood, F.D. Weinfeld, och R.L. York (1966). *Equality of educational opportunity*. Washington, DC: U. S. Government Printing Office.
- Finn, J. D., och C.M. Achilles (1990). Answers and questions about class size: A statewide experiment, *American Educational Research Journal*, 28, 557–577.
- Fredriksson, P. och B. Öckert (2007), The supply of skills to the teacher profession, manuskript, IFAU
- Fredriksson, P. och B. Öckert (2008), Resources and student achievement - evidence from a Swedish policy reform, *Scandinavian Journal of Economics*, 110, 277–296
- Glass, G.V. (1978), Integrating findings: The meta-analysis of research. In L. Shulman (red.), *Review of Research in Education* (vol. 5). Itasca, Illinois: Peacock.

- Greenwald, R., L.V. Hedges och R.D. Laine (1996), The effect of school resources on student achievement, *Review of Educational Research*, 66(3), 361–396.
- Grönqvist, E., och Vlachos, J. (2008), One size fits all? The effects of teacher cognitive and non-cognitive abilities on student achievement. IFAU Working paper 2008:25.
- Gustafsson, J.-E. (2003). What do we know about effects of school resources on educational results? *Swedish Economic Policy Review*, 10, 77–110.
- Gustafsson, J.-E., och E. Myrberg (2002). *Ekonomiska resursers betydelse för pedagogiska resultat – en kunskapsöversikt*. Stockholm: Skolverket.
- Gustafsson, J.-E., och E. Myrberg (2009). Resursers betydelse för elevers resultat. Kap 6 i Skolverket (2009). *Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer*. Stockholm: Skolverket.
- Hattie, J.A. (2009), *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hanushek, E. (1989). The impact of differential expenditures on school performance, *Educational Researcher*, 18, 45–65.
- Hanushek, E. (1998), The evidence on class size. Occasional Paper Number 98-1. W. Allen Wallis Institute of Political Economy. University of Rochester, Rochester, NY, February 1998.
- Hanushek, E. (2003), The failure of input-based schooling policies. *Economic Journal*, 113, F64–F98.
- Hanushek, E. och J.A. Luque (2003), Efficiency and equity in schools around the world. *Economics of Education Review*, 22, 481–502.
- Hoxby, C. (2000), The effects of class size on student achievement: New evidence from population variation, *Quarterly Journal of Economics*, 115, 1239–1285.
- Häkkinen, I., T. Kirjavainen och R. Uusitalo (2003), School resources and student achievement revisited: new evidence from panel data, *Economics of Education Review*, 22, 329–335.

- Krauss, S., M. Brunner, M. Kunter, J. Baumert, W. Blum, M. Neubrand och A. Jordan (2008), Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, 100, 716–725.
- Krueger, A. (1999), Experimental estimates of educational production functions, *Quarterly Journal of Economics*, 114, 497–532.
- Krueger, A. (2003), Economic considerations and class size. *Economic Journal*, 113, 34–63.
- Krueger, A., och M. Lindahl (2002), *Klassfrågan – En ESO-rapport om lärartätheten i skolan*. Finansdepartementet, Expertgruppen för studier av offentlig ekonomi. Ds 2002:11
- Krueger, A., och D. Whitmore (2001), The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. *Economic Journal*, 111, 1–28.
- Lazear, E. (2001), Educational production, *Quarterly Journal of Economics*, 116, 777–803.
- Lindahl, M. (2005), Home versus school learning: A new approach to estimating the effect of class size on achievement, *Scandinavian Journal of Economics*, 107, 375–394.
- Nordenbo, S.E., M. Søgaaard Larsen, N. Tiftikçi, R.E. Wendt och S. Østergaard (2008), Teacher competences and pupil achievement in pre-school and school. A systematic review carried out for The Ministry of Education and Research, Oslo. Technical report, Danish Clearinghouse for Educational Research, Copenhagen.
- Nye, B., L.V. Hedges och S. Konstantopoulos (1999), The long-term effects of small classes: A five-year follow-up of the Tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 21, 127–142.
- Nye, B., S. Konstantopoulos och L.V. Hedges (2004), How large are teacher effects? *Educational Evaluation and Policy Analysis*, 26, 237–257.
- Rivkin, S., E. Hanushek och J.F. Kain (2005), Teachers, schools, and academic achievement, *Econometrica*, 73, 417–458.
- Robinson, G.E. (1990), Synthesis of research on the effects of class size. *Educational Leadership*, 47, 80–90

- Rockoff, J. (2004), The impact of individual teachers on student achievement: Evidence from panel data, *AEA Papers and Proceedings*, 94, 247–252.
- Rothstein, J. (2008), Teacher quality in education production: Tracking, decay and student achievement. NBER Working Paper 14442.
- Shulman, L.S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15, 4–14.
- Skolverket (1999), Samband mellan resurser och resultat. En studie av landets grundskolor med elever i årskurs 9. Skolverkets rapport nr 170.
- Skolverket (2006), Lusten och möjligheten. Om lärarens betydelse, arbetssituation och förutsättningar. Skolverkets rapport 282, 2006.
- Smith, M. L. och G.V. Glass (1980), Meta-analysis of research on class size and its relationship to attitudes and instruction, *American Educational Research Journal*, 17, 419–433.
- Wayne, A.J. och P. Youngs (2003), Teacher characteristics and student achievement gains: A review, *Review of Educational Research*, 73, 89–122.

8 Organisatoriska frågor

Detta kapitel behandlar frågor kring skolsystemets organisation, styrning och utvärdering.

En huvudfråga avser differentiering av elever. I Sverige har frågan om hur skolan skall utformas för att både passa elevers olika behov och tillgodose samhällets behov ständigt varit föremål för diskussioner. Diskussionen har ofta varit präglad av ideologiska ståndpunkter som i olika grad prioriterar jämlikhets- och effektivitetsmål.

Under 1940- och 1950-talen var en politisk stridsfråga om den organisatoriska differentieringen i form av parallellskolsystemet med folkskola och realskola skulle ersättas med en mer sammanhållen enhetsskola. Förespråkarna för enhetsskolan argumenterade huvudsakligen utifrån jämlikhetsargument, medan motståndarna i första hand anförde effektivitetsargument. Efter att beslutet fattades om att införa en 9-årig grundskola vid 1960-talets början kom frågor om linjesystem, nivågruppering, individualisering av undervisning, utformning av specialpedagogiskt stöd och andra mer flexibla former av differentiering att fokuseras istället för den organisatoriska differentieringen av den grundläggande utbildningen. Frågor om organisatorisk differentiering kom istället att diskuteras i anslutning till utformningen av gymnasieskolan. I avsnitt 8.1 presenterar vi forskning kring differentieringsfrågor med särskilt fokus på grundskolan, medan forskning kring organisatorisk differentiering på gymnasienivå behandlas i kapitel 9.

Varje utbildningssystem behöver former för utvärdering, tillsyn och kontroll. Dels är det centralt att kunna bedöma hur väl de övergripande politiska målen har uppfyllts, dels är det viktigt att verksamheten har bedrivits på ett kostnadseffektivt sätt. Sådan tillsyn och kontroll kan ske i många former.

En central avvägning finns mellan centralisering och decentralisering. Den nationella utbildningspolitiken kan föreskriva, inte bara vilka mål som skall uppnås, utan även hur verksamheten skall bedrivas i form av krav på lärare, lokaler och undervisningsmaterial. Även finansieringen av utbildningen kan ligga på central nivå eller decentraliseras till exempelvis kommuner eller andra skoldistrikt. Valet mellan centralisering och decentralisering kan även formuleras i termer av mål- eller regelstyrning; den nationella politiken kan formulera både övergripande mål för verksamheten och regler som föreskriver hur utbildningen ska bedrivas, eller nöja sig med att formulera mål och decentralisera besluten över verksamheten.

En annan dimension är inslagen av skolval och konkurrens i skolsystemet. Genom att tillåta elever (och föräldrar) att välja mellan befintliga skolor är det möjligt att skapa en konkurrens mellan skolorna. Sådan konkurrens skulle kunna befrämja effektiviteten i verksamheten men också leda till skillnader i utbildningskvalitet som står i strid mot fördelningsmål för politiken. Skolval kan begränsas till en uppsättning av offentliga skolor, men det är även möjligt att tillåta att fristående skolor etablerar sig och erbjuder ekonomiska villkor som liknar de offentliga skolorna. Hotet om friskoleetablering skulle kunna förstärka det konkurrenstryck som skolorna upplever i sin verksamhet och därmed vara effektivitetsfrämjande.

Ytterligare en dimension i utbildningssystemet gäller i vilken utsträckning det skall förlita sig på nationella eller internationella kunskapstester av olika slag, respektive olika typer av skolinspektioner, för att få en bild av hur målen uppfylls och hur verksamheten bedrivs.

Alla dessa dimensioner hänger rimligen ihop. En långt gången decentralisering torde exempelvis underlättas av att det finns tillförlitlig information om hur väl målen uppfylls. Vidare kan vissa nackdelar med fristående skolor förhindras av en nationell översyn av dessa. Kunskapsmätningar kan ersätta direkta inspektioner av skolor, men det kan också krävas bra inspektioner för att tester och betyg skall fungera som det är tänkt.

Detta är självfallet stora och komplicerade frågor och vi har ingen möjlighet att ge dem full rättvisa i denna rapport. Det finns dock tre viktiga skäl till att aktualisera dem. För det första har det svenska utbildningssystemet sedan början av 1990-talet genomgått stora förändringar när det gäller flera av dessa dimensioner på styrning: decentralisering från central statlig nivå till kommunerna, skolval och etableringsmöjlighet för friskolor, (ny)satsning på skolinspektioner. För det andra finns en del kvalificerad aktuell forskning som ger intressant belysning av frågorna. För det tredje menar vi att det finns möjlighet till ytterligare studier för att bättre utvärdera det svenska utbildningssystemets egna utvärderingsinstrument.

Utöver differentieringsfrågan behandlar vi i detta kapitel betydelsen av fristående skolor och den potentiella konkurrens mellan skolor som uppkom i samband med införandet av friare skolval i början av 1990-talet; se avsnitt 8.2. Avsnitt 8.3 behandlar decentralisering och avsnitt 8.4 tar upp skolinspektioner. Kapitlet sammanfattas i avsnitt 8.5 där vi också diskuterar möjligheten att förbättra det svenska dataunderlaget för vidare studier på området.

8.1 Differentiering

Forskningen kring effekten av olika former av differentiering är alltför omfattande och resultatmässigt komplex för att det skall vara möjligt att gå igenom denna i detalj. Här sammanfattar vi huvudresultaten, varvid vi börjar med frågan om organisatorisk differentiering och därefter tar upp mer flexibla former av differentiering.

8.1.1 Organisatorisk differentiering

Under 1950-talet genomfördes en försöksverksamhet med sammanhållna, odifferentierade, klasser som alternativ till parallellskolsystemet; i anslutning till denna genomfördes även forskning. En klassisk studie är den s.k. Stockholmsundersökningen av Svensson (1962), som utnyttjade det faktum att vissa delar av Stockholm deltog i försöksverksamheten, medan andra inte gjorde det. Det huvudresultat som framkom var att det inte förelåg några resultatskillnader mellan elever i de differentierade och odifferentierade systemen. Detta pekade på att det var möjligt att åstadkomma en högre grad av jämlikhet utan att behöva betala ett pris i form av lägre effektivitet. I en mycket uppmärksamrad reanalys av data från Stockholmsundersökningen ifrågasatte emellertid Dahllöf (1967) denna slutsats. Han visade att de kunskapsprov som användes i undersökningen framförallt fokuserade på enkla frågor och innehåll, och att de därför inte så väl representerade det mer avancerade innehåll som eleverna i realskolan nådde fram till med samma tidsinsats.

Även om differentieringsfrågan var i starkt fokus i både den utbildningspolitiska diskussionen och den pedagogiska forskningen fram till början av 1960-talet, har endast ett mycket begränsat antal svenska studier genomförts sedan en politisk kompromiss ledde fram till beslutet om införande av grundskolan år 1962. Meghir och Palme (2005) har dock i en studie med elegant design studerat effekter av införandet av grundskolan (se kapitel 5 för en närmare beskrivning). Resultaten visade att grundskola i genomsnitt ledde till 0,3 års längre utbildning i vuxen ålder samt att det framförallt var elever med goda begåvningsmässiga förutsättningar och fäder med låg utbildning som gynnades. Effekten var också starkare för kvinnor än för män. Någon signifikant medeleffekt på inkomst erhöles inte, men däremot fanns en positiv effekt på inkomst för de grupper för vilka positiva utbildningseffekter observerats. Det fanns också en negativ effekt på inkomst för personer med fäder med högre utbildning, vilket Meghir och Palme (2005) förklarar med att det var denna grupp som tidigare framförallt gått i realskola, och som med den nya grundsko-

lan fått en utbildning med lägre kvalitet. En annan möjlig förklaring är att real-examen i sig hade ett signalvärde som inte kom grundskoleeleverna till del.

Pekkarinen m.fl. (2009) har genomfört en studie som till sin design är mycket lik Meghir och Palme (2005). De utnyttjade det faktum att den finska grundskolereformen genomfördes successivt i olika delar av landet 1972–77 för att skatta effekter av reformen på elevernas kognitiva förmågor. Dessa mättes i 20-årsåldern under de första veckorna av militärtjänsten med hjälp av tre olika prov som avsåg mäta verbal förmåga, matematisk förmåga och logisk resonemangsförmåga. De kunde även utnyttja familjeinformation om förekomst av bröder som eventuellt gått i det gamla parallellskolsystemet för att hantera eventuella selektionsproblem som har att göra med att olika typer av familjer väljer att bosätta sig där skolreformen genomförts.

Analysresultaten visade på en svag positiv huvudeffekt på verbal förmåga av grundskolereformen. Framförallt erhöles dock starka interaktionseffekter mellan reform och föräldrarnas utbildningsnivå med avseende på samtliga utfallsvariabler: för elever vars föräldrar hade postgymnasial utbildning förelåg ingen effekt, medan det fanns en positiv effekt av grundskolereformen för övriga elever. Även med föräldrarnas inkomst erhöles interaktioner, men dessa var svagare än för utbildningsvariabeln. Effekterna var sålunda i linje med de förhoppningar som knöts till grundskolereformen, nämligen att denna skulle minska ojämlikheten i utbildningsresultat, utan negativa effekter för grupper som traditionellt varit framgångsrika i utbildningssystemet. Effekterna var dock relativt svaga, och uppgick för elever vars föräldrar hade lägre utbildningsnivå till $d = 0,02$ för de icke-verbala förmågorna, och $d = 0,04$ för verbal förmåga. Det finns dock anledning att understryka att det föreligger många omständigheter som hindrar starka effekter från att framträda i resultaten. En sådan är att beroendevariabeln observerades fyra år efter att eleverna slutat grundskolan; en annan att grundskolereformen ännu inte var fullt utvecklad vid införandet. I matematik användes sålunda fortfarande nivågruppering i stor utsträckning.

En studie av Hanushek och Wössman (2006) har också gett intressanta resultat kring frågan om effekter av organisatorisk differentiering på nivån och spridningen av elevresultaten. De utgick från de internationella undersökningar av elevers kunskaper och färdigheter som genomförts efter 1995, och använde resultat både från studier av kunskapsresultaten i de tidiga skolåren (PIRLS i åk 4 och TIMSS i åk 4) och i slutet av sekundärskolan (TIMSS åk 8 och PISA för 15-åringar).⁷⁴ Genom att använda en ”difference-in-differences” metod under-

⁷⁴ Dessa internationella undersökningar beskrivs närmare i Kapitel 2.

sökte de sedan förändringar i nivå och spridning mellan de tidiga åren och de sena åren. Dessa omfattade mellan 18 och 26 länder, där ungefär hälften hade odifferentierade utbildningssystem och hälften hade differentiering i teoretiskt eller praktiskt inriktade studier efter årskurserna 4–6.

Resultaten visade att spridningen i elevernas resultat mellan tidiga och sena år ökade i länder med differentierade system. Man fann också en tendens att de differentierade systemen gav lägre medelresultat. En analys av utfallet för elever på olika prestationsnivåer visade att differentierade system gav negativa effekter för alla nivåer, men i avtagande grad som en funktion av prestationsnivå. Hanushek och Wössman (2006, s. C75) avslutar med följande slutsats. *”From a policy perspective, it seems incumbent on those advocating early tracking in schools to identify the potential gains from this. These preliminary results suggest that countries lose in terms of the distribution of outcomes, and possibly also in levels of outcomes, by pursuing such policies.”*

Även om resultatbilden var mycket konsistent över olika länder fanns det dock ett land som avvek, nämligen Sverige. I analysen var Sverige klassificerat som ett odifferentierat land men resultaten sammanföll snarare med den resultatbild som observerades för länderna med differentierade skolsystem. En möjlig tolkning av detta är att det svenska utbildningssystemet inte är odifferentierat pga. möjligheter att välja skola (t.ex. Söderström och Uusitalo 2010); dessutom bedrivs undervisningen i vissa ämnen fortfarande i nivågrupperade former.

Jakubowski (2007) har reanalyserat material som användes av Hanushek och Wössman (2006) i syfte att undersöka resultatens robusthet för olika antaganden och analysmetoder. Ett av de problem som Jakubowski uppmärksammar är att de jämförelser som involverar PIRLS och PISA innebär att man också för in en skillnad i urvalsmodell, eftersom PIRLS använder årskurs (åk 4) som grund för populationsdefinitionen, medan PISA använder ålder (15 år) som grund för populationsdefinitionen. En följd av detta är att det uppkommer en ålderskillnad mellan olika länder när man beräknar skillnaden i ålder mellan PISA- och PIRLS-urvalen. Denna ålderskillnad var olika stor för den grupp av länder som har differentierade skolsystem och den grupp av länder som har odifferentierade skolsystem; den var också korrelerad med resultatskillnaden mellan PISA och PIRLS. När åldersskillnaden inkluderades i regressionerna minskade den skattade negativa effekten av differentierade skolsystem och blev icke-signifikant. Jakubowski (2007) konstaterade därför att slutsatsen om negativa effekter på medelprestation av organisatoriskt differentierade

skolsystem inte är hållbar. Inte heller effekterna på spridning i elevresultat kunde Jakubowski (2007) replikera i sin reanalys.

Ett annat problem som Jakubowski (2007) uppmärksammade var att effekterna av organisatorisk differentiering kan skilja sig åt för elever med olika social och utbildningsmässig bakgrund (SES). Denna fråga undersöktes med en lång rad olika modellspecifikationer, både på aggregerad landsnivå och på individnivå. De olika analyserna gav delvis olika resultat, men huvudslutsatsen var att det förväntade mönstret, med starkare negativ effekt av organisatorisk differentiering för elever med låg SES, inte fick stöd.

På grundval av dessa resultat drog Jakubowski (2007) slutsatsen att ”... *it is doubtful whether a simple difference-in-differences approach is able to estimate unbiased tracking effects without further adjustments and methodological developments*” (s. 24). Detta framstår som en rimlig slutsats. Samtidigt bör det understrykas att den kritik Jakubowski (2007) riktar mot Hanushek och Wössman (2006) också kan kritiseras. Idén att ta hänsyn till det varierande åldersspannet mellan de två mätpunkterna i olika länder är intressant, men eftersom antal årskurser är en viktigare faktor än ålder bör hänsyn tas också till denna faktor. Vad gäller det oväntade resultatet att det inte förelåg någon interaktion mellan social bakgrund och organisatorisk differentiering i de flesta specifikationer kan det noteras att en kraftfull interaktion uppträdde i analysen som också inkluderade fixa landeffekter. Eftersom fixa landeffekter är en grundläggande egenskap hos de ”difference-in-differences” analyser som Jakubowski baserar reanalysen på framstår det som mycket förvånande att sådana inte användes i alla specifikationer.

Den av de här presenterade studierna som framstår som mest övertygande är Pekkarinen m.fl. (2009). Även om resultaten från Hanushek och Wössmans (2006) studie uppenbarligen kan ifrågasättas ger inte Jakubowskis (2007) reanalys några resultat som stödjer användning av organisatorisk differentiering i tidig ålder. Den motsatsställning mellan jämlikhet och effektivitet som varit i fokus i de utbildningspolitiska diskussionerna får sålunda inte stöd i de empiriska resultaten från studier av organisatorisk differentiering. Inom detta fält behövs dock ytterligare forskning.

8.1.2 Nivågruppering

En vanlig form av pedagogisk differentiering är nivågruppering (”ability-grouping”) på grundval av prestationer, begåvning eller val. Nivågruppering kan utformas på en lång rad olika sätt, beroende på stadium och skolsystemets

utformning i övrigt. På primärnivån (ungefär åk 1–6) är nivågruppering ovanlig i de flesta länder, och i den mån den förekommer är det i allmänhet i form av nivågruppering inom klassens ram i matematik och läsning. På sekundärnivån (åk 7–12) förekommer en rad olika former av nivågruppering. En variant är att använda nivågruppering inom några centrala ämnen, medan övriga ämnen läses i heterogena grupper. Fram till 1994 fanns i den svenska grundskolan möjlighet för skolorna att på frivillig bas använda ett sådant system med alternativkurser på två svårighetsnivåer (allmän och särskild kurs) i matematik och engelska, vilket utnyttjades av en stor andel av skolorna. En annan möjlighet är att inrätta klasser med olika prestationsnivåer, där eleverna läser alla ämnen tillsammans. Denna längre drivna form av nivågruppering betecknas med olika termer som ”between-class grouping”, ”tracking” och ”streaming” och har i vissa länder varit vanligt förekommande i åk 7–9. På gymnasienivån är också olika linjer eller ”tracks” med fokus på olika ämnen och yrkesförberedande inslag en vanlig utformning.

Nivågruppering som pedagogisk modell har varit mycket omdiskuterad. Förespråkare för nivågruppering, ofta lärare och skolledare, har hänvisat till effektivitetsargument, och då framförallt att nivågruppering gör det möjligt att anpassa undervisning efter de olika elevgruppernas behov, vilket inte är möjligt i heterogena elevgrupper. Förespråkare för undervisning i heterogena grupper har i allmänhet hänvisat till jämlikhetsargument, dels med hänvisning till att en skola med demokratiska ideal förutsätter att alla elever, oberoende av klass, ras, kön och begåvning undervisas tillsammans, dels med hänvisning till att de elever som placeras på de lägre nivåerna drabbas resultatmässigt av att inte ha tillgång till de duktigare eleverna i klassen (se t.ex. Slavin 1990).

En betydande mängd forskning har behandlat effekter av nivågruppering. I huvudsak har två ansatser använts i denna forskning. I den ena har man med experimentell ansats jämfört resultaten för elever som undervisats i homogena grupper med resultaten för elever som undervisats i heterogena grupper, och där man genom randomisering eller matchning försäkrat sig om att eleverna i de två betingelserna initialt var jämförbara. I den andra ansatsen har man med longitudinell design studerat resultatutvecklingen för elever som placerats på olika nivåer i ett nivågrupperat system, och med hjälp av statistiska metoder för kontroll av initiala skillnader mellan elever försökt bestämma behandlings-effekterna av de olika nivåerna.

Slavin (1987, 1990) har publicerat två metaanalyser av den internationella forskningen med experimentell ansats, den ena avseende primärnivån (åk 1–6,

huvudsakligen åk 4–6), och den andra avseende sekundärnivån (åk 7–12, huvudsakligen åk 7–9). Trots att dessa metaanalyser publicerades för ungefär två decennier sedan är de fortfarande aktuella. En anledning till detta är att, som Slavin (1987, 1990) konstaterar, den experimentella ansatsen i huvudsak upphörde att användas inom detta område under 1970-talet, inte heller efter publiceringen av Slavins metaanalyser har sådana studier genomförts i någon egentlig omfattning. En annan anledning att fokusera på Slavins översikter är att dessa innehåller utförliga metoddiskussioner och diskussioner av de enskilda studierna.

Slavin (1987) fann 14 studier som undersökte ”between-class grouping” på primärnivån med experimentell ansats. För de 17 jämförelser som var möjliga på grundval av dessa studier var medeleffekten exakt 0. Ungefär lika många studier undersökte effekten av flexibel, årskurslös, nivågruppering inom läsning. Med denna form av nivågruppering, som Slavin (1987) benämner Joplinplan, placerades alla elever i närliggande årskurser i olika grupper på grundval av deras aktuella läsprestationer. Detta innebar att det i samma grupp återfanns högpresterande elever från lägre årskurser och lågpresterande elever från högre årskurser. Undervisningen var noga anpassad med avseende på nivå och hastighet efter de olika nivåerna. I jämförelse med kontrollgruppen, som undervisades i heterogena grupper, hade de nivågrupperade eleverna ett högre medeltal ($d = 0,45$). Slavin (1987) identifierade också åtta studier av effekter av nivågruppering inom klassen i matematik. Även dessa studier visade på en fördel för de nivågrupperade eleverna ($d = 0,36$).

Dessa effektiva former av nivågruppering förekommer dock sällan i praktiken, utan där dominerar ”between-class grouping”. En huvudslutsats som Slavin drar från studierna av mellanklassformerna av nivågruppering är att dessa inte ger de förväntade effektivitetsvinster: *”it is surprising to see how unequivocally the research evidence refutes the assertion that ability-grouped class assignment can increase student achievement in elementary schools.”* (Slavin 1987, s. 307).

En andra huvudslutsats är att under vissa omständigheter kan nivågruppering ge effektivitetsvinster på den primära nivån. Men grupperingen *“... must create true homogeneity on the specific skill being taught and instruction must be closely tailored to students' levels of performance”* (Slavin 1987, s. 323). En tredje huvudslutsats som Slavin (1987) drar är att eftersom det finns mycket forskning som visar på negativa effekter på de lågpresterande elevernas självkänsla av nivågruppering (t.ex. Oakes 1985) bör undervisningen i huvudsak

bedrivs i heterogena grupper. Nivågruppering bör endast användas då den faktiskt innebär en anpassning av undervisningen efter elevernas behov; det är då viktigt att eleverna inte upplever grupperingen som stigmatiserande.

Slavins (1990) metaanalys av studier på sekundärnivån omfattade 29 studier, som använde sig av tre typer av designer: (1) experiment med randomisering (6 studier); (2) experiment med matchning (9 studier); och (3) studier av befintliga skolor med eller utan nivågruppering med kontroll för initiala skillnader (14 studier). I samtliga de randomiserade experimenten var resultaten bättre för heterogena klasser än för nivågrupperade klasser, men skillnaderna var överlag mycket små. För de experimentella studierna sammantaget förelåg en liten skillnad till förmån för heterogena klasser ($d = -0.06$). Effekterna var i huvudsak desamma över ämnen, förutom att det fanns en tendens till bättre resultat i heterogena klasser i samhällsvetenskap. För 21 av de 29 studierna var det också möjligt att bestämma effekter för elever på tre olika prestationsnivåer. För alla tre nivåerna var dock de skattade effekterna nära 0. Slutsatsen från metaanalysen var sålunda att varken genomsnittsprestationer eller spridning påverkades av nivågruppering.

Som redan nämnts var Slavins översikter begränsade till experimentella och kvasiexperimentella studier som jämför homogena och heterogena klasser. Efter att de experimentella studierna upphört under 1970-talet har istället studier med longitudinell design kommit att dominera forskningen om effekter av nivågruppering. Även om sådana studier inte omfattas av Slavins översikter diskuterar Slavin (1990) skillnader i utfall mellan de två typerna av studier. Enligt de översikter som Slavin (1990) hänvisar till framträder ett annorlunda resultatmönster i de longitudinella studierna: elever på högre linjer gör större framsteg än elever på lägre linjer även efter kontroll för individ- och bakgrundsvariabler. Detta står i kontrast till de experimentella studierna som i huvudsak inte hittar några effekter av nivågruppering. Slavin framför flera möjliga förklaringar till denna skillnad.

En förklaring tar fasta på att de experimentella och longitudinella ansatserna undersöker delvis olika fenomen. I de experimentella studierna är det alltid samma kunskapsstoff (t.ex. kursen Algebra 1) som behandlas i de olika betingelserna, men i de longitudinella studierna bestäms kunskapsstoffet inom olika linjer delvis av elevernas val av kurser med olika inriktning och svårighetsgrad, i synnerhet på gymnasienivån. Enligt Slavin är det rimligt att anta att nivågruppering som erbjuder elever olika möjligheter till fördjupning och studietakt

kan ge upphov till stora skillnader i kunskapsstillägnet, men detta är i så fall inte orsakat av skillnader i undervisningseffektivitet mellan de olika linjerna.

En andra förklaring till utfallen i de longitudinella studierna är att kontrollen för initiala skillnader inte är tillräcklig för att hantera effekten av de mycket stora initiala prestationskillnader mellan elever som går på olika linjer. Dessa skillnader kan uppgå till flera standardavvikelser, och endast begränsade mängder mätfel eller måttliga avvikelser från andra antaganden som regressionsjusteringen bygger på, leder under dessa omständigheter till stora felskattningar av effekten av att gå på olika linjer; Betts och Shkolnik (2000) har kritiserat studier som visar på heterogena effekter av nivågruppering av likartade skäl.

Den slutsats som Slavin (1990) drar av metaanalysen är att nivågruppering inte ger de resultatvinster som utlovats. Han menar därför att det inte finns några argument för nivågruppering, men starka argument mot:

Given the antidemocratic, antiegalitarian nature of ability grouping, the burden of proof should be on those who would group rather than those who favor heterogeneous grouping, and in the absence of evidence that grouping is beneficial, it is hard to justify continuation of the practice. The possibility that students in the low groups are at risk for delinquency, dropout, and other social problems ... should also weigh against the use of ability grouping. (Slavin, 1990, s. 494).

Avsaknaden av effekter av nivågruppering framstår som överraskande, men den omfattande forskningen, som i många fall är av osedvanligt hög kvalitet, ger litet utrymme för att betvivla hållbarheten i de empiriska resultaten. Det är dock möjligt att det bakom detta mönster av icke-skillnader döljer sig effekter som går i olika riktningar, och som balanserar varandra. Det är sålunda möjligt att det för gruppen som sorteras till den lägre nivån finns en positiv undervisningseffekt som motverkas av en negativ kamrateffekt (se avsnitt 8.1.4). Mot detta argument talar dock att det i så fall borde föreligga en stark positiv effekt för gruppen som sorteras till den högre nivån, eftersom det för denna grupp skulle finnas både en positiv undervisningseffekt och en positiv kamrateffekt. En annan möjlighet är att resultaten av de experiment som genomförts inte är generaliserbara till de normala skolsituationerna, eftersom de genomförts med strikt kontroll av faktorer som lärarkompetens och kursinnehåll. Det måste dock betonas att det är denna höga grad av kontroll som gjort det möjligt att pröva och inte förkasta hypotesen att undervisning i heterogena elevgrupper är lika effektivt som undervisning i homogena grupper. Samtidigt måste det understrykas att det finns aspekter av nivågrupperingsfenomenet som inte är

åtkomliga med den experimentella ansatsen. Resultaten från de longitudinella undersökningarna som visar på samband mellan nivågruppering och kursval är ett exempel på detta, liksom studier som visar att undervisningsgrupper på de låga nivåerna tenderar att få sämre lärare (Oakes 1985, 1990).

Det finns endast ett begränsat antal svenska studier av nivågruppering. Sund (2007) har i en mycket intressant svensk studie undersökt effekten av nivågruppering i form av alternativkurser med olika svårighetsgrad (allmän och särskild kurs) i matematik och engelska på grundskolans högstadium. Dessa, och några ytterligare alternativkurser, kvarstod som den enda resten av det omfattande system av alternativkurser som avskaffades med den läroplan som introducerades 1969. Med 1980 års läroplan avskaffades alla alternativkurser utom dem i engelska och matematik och det blev valfritt för skolorna att använda dessa eller ej. De användes dock av en stor majoritet av skolorna fram till införande av Lpo 1994 då de helt avskaffades. Denna förändring utnyttjar Sund (2007) för att undersöka effekterna av den differentiering av eleverna i homogena undervisningsgrupper som alternativkurserna utgjorde. Det faktum att vissa skolor gick från ett differentierat till ett odifferentierat system medan andra behöll sitt odifferentierade system kunde Sund utnyttja för att studera resultatförändringen inom skolor. Den beroendevariabel Sund (2007) fokuserade på var betyget i Matematik A från gymnasiet. Först undersökte dock Sund (2007) om alternativkurserna påverkat sannolikheten att ha avslutat gymnasiet och fann att så inte var fallet. Inte heller fann Sund (2007) någon medelvärdeskillnad mellan elever som gått i de differentierade och odifferentierade matematikkurserna. I nästa steg undersökte Sund (2007) om effekterna varierade med föräldrarnas utbildningsnivå. Analysresultaten visade att så var fallet: elever från familjer med låg utbildningsnivå hade sämre resultat i det differentierade systemet, medan elever från familjer med hög utbildningsnivå hade bättre resultat i det differentierade systemet. En analys med materialet uppdelat efter kön visade dock att de positiva effekterna var begränsade till att gälla för flickor, medan utfallet för pojkar inte berodde på utbildningsbakgrund.

Denna studie visar resultat som stödjer de som erhållits i flera andra studier (förutom de experimentella), nämligen att differentiering tenderar att ge en negativ effekt för elevgrupper på lägre prestationsnivåer. Av Sunds (2007) studie framgår dock inte om det är låg utbildningsnivå i hemmet eller låg initial prestationsnivå som är den betydelsefulla faktorn, eftersom det inte var möjligt att separera dessa faktorer. Det kan också noteras att de skattade effekterna är små, och trots det stora datamaterialet var dessa inte mer än svagt signifikanta. En

förklaring till de svaga effekterna skulle kunna vara att skolorna i viss utsträckning fortsatt använda nivågruppering i matematik även efter det att alternativkurserna togs bort. Enligt 2006 års PISA-undersökning (OECD 2008) fanns sålunda inte mindre än 70 % av de svenska eleverna i skolor där skolledaren angivit att man regelbundet använde nivågruppering i vissa ämnen. Denna andel var betydligt högre än i exempelvis Norge och Finland.

8.1.3 Specialpedagogiskt stöd

Specialpedagogiskt stöd kan ges i många olika former. Ett förhållande som kännetecknar många former av specialpedagogiska insatser är att eleven tas ut från sin ordinarie undervisningsgrupp, och under längre eller kortare tid placeras i en annan undervisningsgrupp. I extrema fall handlar det om överflyttning till särskola eller annan organisatoriskt differentierad lösning. Det kan också röra sig om mer eller mindre permanenta former av nivågruppering inom ordinarie klass eller skola, såväl som mer flexibla former av specialundervisningsinsatser. Sådana lösningar motiveras i allmänhet med att de ger möjlighet att erbjuda eleven en undervisning som är anpassad efter hans eller hennes behov.

Särskolan är en egen skolform, som finns för de elever som på grund av en utvecklingsstörning, autism eller autismliknande tillstånd inte kan nå målen i grundskolan. Särskolan är också öppen för barn och ungdomar som fått en betydande och bestående begåvningsmässig funktionsnedsättning på grund av hjärnskada. Utbildningen ska i så stor utsträckning som möjligt motsvara den som finns i grundskolan och gymnasieskolan. Särskolan har samma läroplaner som grundskolan och gymnasieskolan men har egna kursplaner, mål och betygskriterier. Eleverna i särskolan går antingen i särskoleklasser eller integrerade i grundskoleklasser.

Under 1990-talet mer än fördubblades antalet elever i grundsärskolan, från ca. 4 000 elever läsåret 1991/92 till ca. 8 600 elever läsåret 1999/00. Ökningen fortsatte sedan fram till läsåret 2003/04 då grundsärskolan omfattade ca. 10 500 elever. Därefter har antalet minskat marginellt till ca. 10 000 elever 2007/08. En förklaring till det ökade antalet elever i grundsärskolan är de ökade kraven på skolprestation som följde av det mål- och kunskapsrelaterade betygssystem som introducerades med Lpo 94 tillsammans med de höjda behörighetskraven för tillträde till gymnasieskolans program.

Särskiljande former av specialpedagogiskt stöd som särskola kan ses som en form av nivågruppering, men normalt brukar inte forskning kring effekter av

sådana former behandlas tillsammans med andra studier av effekter av nivå-gruppering, och så var inte heller fallet i Slavins (1987, 1990) metaanalyser. Forness (2001) har sammanfattat resultaten från 24 metaanalyser av undersökningar som studerat effekter av olika typer av specialpedagogiska interventioner. En av dessa (Carlberg och Kavale 1980), som i sin tur byggde på 50 empiriska studier, avsåg effekter av placering i specialklass; detta var den enda typ av intervention som gav en svag negativ effekt ($d = -0,12$). De flesta studierna i Forness gällde elever med lätt utvecklingsstörning. De få studier som fanns att tillgå när det gällde elever med inlärningssvårigheter, eller beteendestörningar av annat slag, indikerade en positiv effekt av placering i specialklass ($d = 0,29$).

Liksom var fallet med de effektivitetsargument som framförts för nivå-gruppering, tycks inte heller effektivitetsargument för undervisning i specialklass få något entydigt stöd av forskningsresultaten. Som påpekades av Slavin (1987, 1990) finns det många observationer av negativa upplevelser i anslutning till särskiljande i form av placering i specialklass. En betydande mängd internationella forskningsresultat visar också på förekomsten av stigmatiseringseffekter av särskilt stöd och särskilt om det organiseras under segregande former. Självvärdering och motivation påverkas negativt, och det har i många studier observerats att de akademiska kraven tenderar att sänkas och att tolkningen av läroplanens intentioner tenderar att bli förenklad (Darling-Hammond och Bransford 2005).

Enligt Forness (2001) meta-metaanalys erhöles de största positiva effekterna för riktad undervisning i speciella färdigheter som systematisk minnesträning och för strategier för läsförståelse. Även interventioner med inriktning på beteendeförändringar uppvisade stora effektstorlekar. Dessa resultat är i överensstämmelse med den metaanalys över generella undervisningseffekter som nyligen publicerats av Seidel och Shavelson (2007), där ett huvudresultat var att de starkaste effekterna erhöles för domänspecifika färdigheter. Dessa resultat stöds också av resultaten i en metaanalys av Kroesbergen och Van Luit (2003) som omfattade 58 interventionsstudier med fokus på matematik. De högsta effektstorlekarna erhöles för direkt undervisning av grundläggande färdigheter inom matematiken.

Dessa resultat kan kort sammanfattas så att det inte finns något stöd för att särskiljande lösningar som innebär mer permanent nivågruppering skulle vara en effektiv specialpedagogisk metod. Specialpedagogiska interventioner som syftar till träning av domänspecifika färdigheter och inlärningsstrategier har däremot visat sig ha positiv effekt.

8.1.4 Kamrateffekter

Effekter av differentiering diskuteras i allmänhet i termer av skillnader i undervisningskvalitet mellan homogena och heterogena elevgrupper. En annan skillnad är dock att elevsammansättningen skiljer sig; i den mån det finns kamrateffekter är det nödvändigt att ta hänsyn även till dessa vid tolkningen av resultaten.

Coleman m.fl. (1966) visade tidigt att elevernas hembakgrund har stor betydelse för skolresultaten; man fann också att klasskamraterna hade stor inverkan. Man drog slutsatsen att *“a pupil's achievement is strongly related to the educational backgrounds and aspirations of the other students in the school”* (Coleman m.fl. 1966, s. 302). Man drog också slutsatsen att *“children from a given family background, when put in schools of different social composition, will achieve at quite different levels”*.

Denna studie pekar på att de skol- och klasskamrater en elev har kan vara en viktigare faktor än undervisningens resurser och kvalitet. Efter denna tidiga studie har en lång rad undersökningar genomförts kring kamrateffekters existens och betydelse. Resultatbilden har dock varit mycket splittrad, vilket framförallt har sin grund i de stora metodologiska utmaningar som studiet av kamrateffekter innebär. Eftersom det sällan är möjligt att använda en experimentell ansats i dessa studier har forskarna varit hänvisade till olika statistiska metoder som hjälpmedel för att särskilja den kausala effekten av elevsammansättningen från sådan likhet mellan medlemmarna i en grupp (t.ex. skolklass) som beror av selektion eller gemensam miljö. Svårigheterna att på ett korrekt sätt åstadkomma detta är utomordentligt stora, och resultaten från forskningen kring kamrateffekter har betraktats med skepsis.

Det finns dock studier som väckt optimism om möjligheterna att fastställa betydelsen av kamrateffekter; resultaten från dessa studier pekar på att effekterna kan vara stora. Zimmerman (2003) studerade den effekt som rumskamratens prestationsnivå utövar på prestationer på ett amerikanskt college, varvid data från slumpmässigt hopparade rumskamrater utnyttjades. En medelpresterande student påverkades negativt av att bli tilldelad en rumskamrat med låga akademiska prestationer, men positivt av att få en rumskamrat med höga akademiska prestationer. Sacerdote (2001) har i en motsvarande undersökning fått liknande resultat. Dessa undersökningar kommer nära idealet att använda randomiserad design för att studera kamrateffekter.

Hoxby (2000) har använt en annan teknik för att särskilja inverkan från kamratfaktorer från andra faktorer. Med utgångspunkt i ett mycket stort admi-

nistrativt material från samtliga skolor i Texas har hon utnyttjat den slumpmässiga variationen år från år i könsmässig och etnisk sammansättning av elevgruppen vid olika skolor. Resultaten visade både på kraftiga kamrateffekter, och på att sammansättningen av elevgruppen har olika effekt som funktion av kön och etnicitet. Hoxby (2000) uppskattade att för läsning i åk 3–6 leder en prestationsnivå bland klasskamraterna som är i medeltal en poäng högre till en höjning av en viss elevs resultat med 0,3–0,4 poäng. För matematik var effekterna större, och i synnerhet för de lägre årskurserna. Hoxby (2000) visade också att en större andel flickor i klassen medförde bättre resultat i matematik. Denna studie visar sålunda på betydande kamrateffekter på elevernas resultat.

Hoxby (2000) pekar på att det finns anledning att göra en distinktion mellan additiva och icke-additiva kamrateffekter. Enligt den additiva modellen är den kamrateffekt som en viss elev utsätts för en linjär funktion av medelvärdet för övriga elever i gruppen. Därför blir totalmedelvärdet detsamma oberoende av hur eleverna grupperas, medan variationen i resultat mellan eleverna påverkas av olika grupperingar. Den största variationen uppstår då lågpresterande elever grupperas tillsammans med andra lågpresterande elever, och högpresterande elever grupperas tillsammans med andra högpresterande elever.

Kamrateffekter kan dock också vara icke-additiva, vilket innebär att effekterna är olika för olika undergrupper av elever. Det är exempelvis tänkbart att kamrateffekten är starkare för lågpresterande elever än för högpresterande elever, eller att den är starkare för pojkar än för flickor. I den mån kamrateffekterna är icke-additiva påverkas inte bara spridningen utan även totalmedelvärdet av olika sätt att gruppera eleverna. Hoxby (2000) fann att kamrateffekterna delvis var icke-additiva, men för närvarande är kunskapen kring de icke-additiva kamrateffekterna alltför begränsad för att det skall vara möjligt att föra någon mer konkret diskussion kring konsekvenser av dessa.

I en svensk studie undersökte Sund (2009) kamrateffekter med hjälp av ett mycket omfattande longitudinellt material över elever och lärare i Stockholmskolor. Han konstaterade att det finns två huvudsakliga metodproblem vid studiet av kamrateffekter, nämligen selektionsproblem och simultanitetsproblem. Detta förra problemet är välkänt från andra sammanhang, men simultanitetsproblemet är specifikt för bestämning av kamrateffekter. Innebörden är att om det finns kamrateffekter kommer en person som påverkas av dessa också att påverka sina klasskamrater, osv. För att hantera selektionsproblemet använde sig Sund (2009) av ett longitudinella data som innehöll information om exakt vilka elever och lärare som fanns i olika klasser i både gymnasiet och åk

9 i grundskolan. Med hjälp av detta material skattade han en modell som inkluderade fixa effekter för skola, lärare och elev. Beroendevariabeln var betyg på gymnasiets A-kurser i matematik, svenska, svenska som andraspråk, och idrott. Kamraternas prestationer mäts med betyget i åk 9 på motsvarande kurs för elever som tar kursen tillsammans med en viss elev. Härigenom behandlades simultanitetsproblemet genom att använda ett ”laggat” prestationsmått.

Resultaten visade på signifikanta kamrateffekter. En förändring av kamraternas prestationsnivå med en standardavvikelse gav en effekt på $d = 0,08$. Även spridningen i kamraternas prestationer hade en positiv effekt som uppgick till $d = 0,02$. Sund (2009) visade också att kamrateffekterna var icke-linjära: de positiva effekterna var starkast för de elever som hade låga grundskoleresultat; för elever med höga grundskoleresultat var kamrateffekterna insignifikanta. Även denna studie visar sålunda att kamrateffekter är betydelsefulla, i synnerhet för de svaga eleverna.

Ammermüller och Pischke (2009) använde data från PIRLS 2001 för flera europeiska länder, däribland Sverige, för att skatta kamrateffekter med avseende på läsförståelse i åk 4. De utnyttjade de skolor där mer än en klass ingick i urvalet och använde variationen mellan dessa klasser, som de antog vara slumpmässig, för att skatta kamrateffekter. De undersökte flera olika variabler, som kön och migrationsbakgrund, men fokuserade särskilt på hembakgrund, representerad av antalet böcker hemma. Ett skäl för att fokusera på denna variabel var att den har ett starkt samband med läsprestation; ett annat var att det fanns uppgift om antal böcker hemma från både eleven och föräldrarna. Denna dubbla information gjorde att Ammermüller och Pischke (2009) kunde hantera mätfelsproblem med en IV-ansats (se kapitel 5). Innan de genomförde skattningarna av kamrateffekter undersökte de om antagandet om slumpmässig fördelning av elever över klasser inom skolor var korrekt. För de flesta länder fann de ingen anledning av förkasta detta antagandes korrekthet, men för Sverige fanns indikationer på att skillnaderna mellan klasserna med avseende på medelantal böcker hemma var större än vad som kunde förväntas av slumpen. Skattningarna av kamrateffekternas storlek skiljde sig dock inte systematiskt mellan Sverige och de övriga länderna.

Med korrigering för mätfel var den skattade kamrateffekten $d = 0,17$ vid förändring av klassmedelvärdet med en standardavvikelse; utan denna korrigering var den skattade effekten endast hälften så stor. Detta resultat illustrerar vikten av att ta hänsyn till mätproblem, vilket sällan görs i denna typ av studier. Skattningen pekar också på kamrateffekter av en ganska betydande storlek: den

motsvarar en sänkning av klasstorleken med 5 elever enligt STAR-experimentets resultat (se kapitel 7).

De studier som redovisats i detta avsnitt visar alla på att kamrateffekter påverkar elevernas resultat. Det är dock slående i hur hög grad estimaten varierar över olika studier, från mycket höga i Hoxby (2000) till relativt blygsamma i Sund (2009). Denna variation kan ha sin grund i att de olika studierna i varierande grad lyckats bemästra de metodmässiga utmaningarna, men den kan också ha sin grund i att kamrateffekternas storlek varierar med elevernas ålder och de kunskaper och färdigheter som undersökts. Det framstår som angeläget att forskningen inom detta fält intensifieras, bland annat för att det skall vara möjligt att systematiskt undersöka olika typer av interaktionseffekter.

8.1.5 Diskussion och slutsatser

Detta avsnitt redovisar studier som undersökt olika aspekter av differentiering och gruppering av elever. En vanlig form av differentiering är att undervisningen bedrivs i grupper som homogeniserats med avseende på prestation eller begåvning, så kallad nivågruppering. Intressant nog visar den experimentella forskningen inte på några effekter av nivågruppering på medelprestation, eller på prestation för olika grupper av elever. Ett undantag utgörs dock av vissa former av nivågruppering i läsning och matematik i åk 4–6, under förutsättning att elevens initiala prestationsnivå inom ämnet används som utgångspunkt för en noggrann anpassning av undervisningen till elevens behov. Normalt kan man inte ta avsaknad av effekter till intäkt att en intervention inte har någon effekt, eftersom olika former av metodproblem kan hindra att en existerande effekt blir synlig i experimentet. Men i detta fall är de rapporterade experimentella studierna enligt Slavin (1987, 1990) av en sådan omfattning och kvalitet att de skulle förmått att identifiera även svaga effekter av nivågruppering.

Avsaknaden av effekt av nivågruppering på grundval av prestation är anmärkningsvärd då företrädare för nivågruppering enkom hänvisar till effektivitetsargument, som att undervisning i homogena elevgrupper gör det möjligt för läraren att anpassa undervisningen efter de behov som elever på olika prestationsnivåer har. Eftersom nivågruppering har andra negativa konsekvenser (t.ex. stigmatisering) finns det starka skäl som talar mot nivågruppering.

En speciell form av nivågruppering utgörs av specialundervisning i särskilda klasser till vilka eleverna fördelas för en längre tid. Några positiva effekter av denna typ av specialundervisning har inte kunnat påvisas. Positiva effek-

ter av specialundervisningsinsatser som avser träning av olika specifika färdigheter och inlärningsstrategier har dock konstaterats i en lång rad studier.

Det finns också stöd för slutsatsen att organisatorisk differentiering på grundskolenivån orsakar ökad variation i elevresultaten, utan att vara förknippad med en ökad medelprestation. Snarare pekar resultaten på att organisatorisk differentiering leder till sänkning av medelprestationen, även om dessa resultat är osäkra. Effekterna av organisatorisk differentiering kan uppkomma på flera sätt – det är ännu inte helt klart hur de resultatmönstren skall tolkas.

En tänkbar förklaring är att kamrateffekter är orsaken till den organisatoriska differentieringens effekter, eftersom differentieringen sorterar eleverna efter prestationsnivå. Litteraturen pekar på att den enskilde elevens prestation förbättras om kamraterna presterar. Om dessa effekter är additiva kommer vi i huvudsak se ökad spridning som resultat av organisatorisk differentiering, vilket ligger i linje med studierna på området. Frågan är dock om kamrateffekter är en tillräcklig förklaringsfaktor för att förklara de observerade effekterna av organisatorisk differentiering.

Studier av bland andra Oakes (1985, 1990) har visat att de högpresterande grupperna vid nivågruppering i större utsträckning får de mer erfarna och välutbildade lärarna, vilket kan leda till kvalitetsskillnader i undervisningen för de olika elevgrupperna. Resultaten i Baumert (2008) pekar på att detta kan vara en betydelsefull faktor i de stora skillnader i prestationer mellan studie- och yrkesförberedande inriktningar i Tyskland, till vilka elever i de flesta delstater differentieras efter åk 4 eller åk 6. Eleverna i dessa studieinriktningar undervisas av lärare med olika lång lärarutbildning, och Baumert m.fl. (2009) har visat på betydande prestationsskillnader mellan dessa kategorier av lärare på ett prov som mäter matematikdidaktiska kunskaper (se kapitel 7). Dessa kunskaper har i sin tur en kausal effekt på utvecklingen av elevernas matematikkunskaper, vilket skulle kunna förklara förändringarna i prestationsskillnader över tid mellan elever i studie- och yrkesförberedande studieinriktningar.

8.2 Skolval och konkurrens

Från svenskt perspektiv är det naturligt att fokusera på de reformer som genomfördes 1 juli 1992. Dessa innebar dels ökade möjligheter att välja skola, dels möjlighet att etablera fristående skolor. Före denna reform rådde närhetsprincipen och eleverna placerades i närmaste skola. Reformen innebar en möjlighet för föräldrar att välja skola för sina barn, men i de fall det fanns fler sökande än platser till en skola skulle närhetsprincipen fortfarande gälla. Sam-

tidigt blev det möjligt att etablera fristående skolor och erhålla ersättning för sin verksamhet. Den som önskar etablera en friskola måste ansöka om detta hos Skolverket som ska ta del av synpunkter på ansökan från den berörda kommunen. Om ansökan bifalls, vilket normalt varit fallet, är den fristående skolan berättigad till finansiell ersättning från kommunen. Denna ersättning har legat nära den genomsnittliga kostnaden i de kommunala skolorna. De fristående skolorna har däremot inte rätt att ta ut avgifter av eleverna. De kan inte heller välja vilka elever de vill ta emot. Om det finns fler sökande än platser till en fristående skola är det kötiden som avgör.

Fram till 1992 var fristående skolor mycket ovanliga i Sverige: de omfattade bara en ca. halv procent av landets grundskoleelever. De få fristående skolorna hade fram till 1992 antingen en speciell pedagogisk inriktning (som Waldorf) eller en särskild etnisk eller religiös inriktning. Reformen följdes dock av en snabb tillväxt av fristående skolor. Fram till 2003 hade andelen ökat till cirka 5 % och läsåret 2008/09 ligger siffran kring hela 10 %. Expansionen har varit mycket ojämnt fördelad över landet: den har varit särskilt koncentrerad till storstadsregionerna. År 2003 fanns fristående skolor (med elever i åk 9) i 93 av landets 284 kommuner. I kommunen med högsta andelen elever i sådana skolor var andelen 39 %. Även med avseende på föräldrars utbildnings- och invandrarbakgrund har det varit en ojämn fördelning av elever på fristående skolor; elever med högutbildade föräldrar och med föräldrar födda utomlands har varit påtagligt överrepresenterade (se Björklund m.fl. 2005).

Den svenska friskolereformen reser ett antal viktiga frågor. En första fråga är vilka effekterna har varit på elevernas prestationer i skolan. Med tanke på reformens utformning bör man här skilja mellan två olika effekter: dels effekten av att gå i en fristående skola (i stället för i en kommunal), dels effekten av att de fristående skolorna kan ha påverkat de kommunala skolorna. Den senare effekten kan uppkomma via konkurrens om eleverna; risken att förlora elever till de fristående skolorna skulle kunna påverka hur de kommunala skolorna arbetar. Den senare effekten skulle också kunna uppkomma utan att några fristående skolor etableras. Hotet om konkurrens kan påverka de kommunala skolornas verksamhet.

Man kan även föreställa sig att det finns av effekter av att införa skolval även i kommuner och landsdelar där fristående skolor inte alls förväntas etableras. Det fria skolvalet och skolpengen innebär nämligen en slags konkurrens om eleverna mellan de kommunala skolorna. Skolpengen medför att budgeten minskar i skolor som får få sökande elever. Som vi kommer att se

finns inga studier som undersökt denna fråga. De svenska studier som genomförts fokuserar på effekterna av att fristående skolor de facto har införts.

Konkurrens behöver dock inte bara påverka kvaliteten på den vara som produceras (i detta fall elevernas prestationer). Den kan också påverka kostnaderna för att framställa varan i fråga. Konkurrensen skulle kunna ha haft en kostnadsreducerande effekt. I debatten har dock också påpekats att effekten på skolkostnaderna kan vara den motsatta genom att försvåra kommunernas planering av skolverksamheten när elevernas kan välja fristående skolor i stället för de kommunala. Av båda dessa skäl framstår det som angeläget att undersöka friskolornas effekter på skolkostnaderna.

Det är också möjligt att den svenska friskolereformen haft ofördelaktiga fördelningseffekter, antingen genom att positiva effekter enbart kommit starka elever till del, eller genom att direkt försämra resultaten för svagare elever. Av dessa skäl finns anledning undersöka om eventuella effekter på elevernas prestationer varit heterogena i meningen att de varierat med exempelvis elevernas familjebakgrund. Det kan också vara av intresse att undersöka om reformen ökat segregationen i skolorna. Ökad segregation så att elever med samma familjebakgrund i ökad omfattning går i samma skolor uppfattas ofta negativt. Man bör dock hålla i minnet att ökad segregation i alla dimensioner inte behöver vara negativt för elevernas prestationer. Exempelvis tyder resultat från Åslund m.fl. (2009) på att invandrabarn tenderar att prestera bättre när de, allt annat lika, har närhet till fler vuxna med samma etniska bakgrund (mätt med invandringsland) medan de tenderar att presentera sämre med närhet till fler invandrare generellt sett.

Det finns nu ett antal svenska studier av effekterna av friskolereformen från 1992. I Tabell 8.1 presenterar vi en översikt av hur dessa studier varit upplagda, vilka frågor de försökt besvara, och vilka resultat de gett. Vi presenterar studierna i den ordning de offentliggjorts i den svenska debatten och inte efter året för den slutliga vetenskapliga publiceringen.

Tabell 8.1. Översikt över svenska studier av effekter av fristående skolor

Studie	Data	Metod	Frågeställning	Resultat
Sandström & Bergström (2005, 2007)	Register-data över elever som gick ut grundskolan 1998.	Tvärsnittsanalys (OLS), instrumentvariabel-ansats med individdata och paneldata på aggregerade kommundata	Effekten av andelen friskoleelever i kommunen på meritvärdet för eleverna i de kommunala skolorna.	Genomgående positiva effekter på prestationer i de kommunala skolorna. Den, enligt författarna, mest tillförlitliga metoden tyder på att 10 procentenheter högre friskoleandel höjer meritvärdet med 14,6 enheter.
Ahlin (2003)	UGU-data över 8000 elever födda 1982.	Value-added modell med kontroller för testresultat i åk 6	Effekten av andelen friskoleelever på nationella pro-ven i åk 9 för samtliga elever.	10 procentenheter högre friskoleandel höjer resultaten i matematik med 6 percentiler. Effekterna är homogena med avseende på elever med olika bakgrund.
Björklund, m.fl. (2003, 2005)	Register-data över elever som gick ut grundskolan 1998–2001.	Difference-in-differences, dvs. förändring i friskoleandel relaterades till förändring i elevprestation.	Effekten av andelen friskoleelever på betyg (och testresultat) i åk 9 för samtliga elever.	10 procentenheter högre friskoleandel höjer betygen med 4 percentiler. Lägre effekter för elever med svagare socioekonomisk bakgrund.
Böhlmark & Lindahl (2007, 2008)	Register-data över elever som gick ut grundskolan 1993–2003	Difference-in-differences, dvs. förändring i friskoleandel relaterades till förändring i elevprestation.	Effekten av andelen friskoleelever på: (1) betyg (och testresultat) i åk 9 för samtliga elever; (2) skolkostnader; (3) övergångar till teoretiskt gymnasium och högskola; (4) skolsegregation.	10 procentenheter högre friskoleandel: (1) höjer betygen med 1 percentil. (inga påtagliga skillnader mellan elever med olika socioekonomisk bakgrund); (2) höjer kostnaderna med 2 %; (3) ökar övergångar till teoretiskt gymnasium, men har ingen effekt på övergång till högskola; (4) ger tendens till ökad segregation, särskilt baserad på invandrarstatus.
Söderström & Uusitalo (2010)	Register-data över gymnasieelever i Stockholm 1998–2001	Difference-in-difference, dvs. förändring i segregation i Stockholms kommun jämfört med förändring i angränsande kommuner	Förändring i skolsegregation till följd av en antagningsreform år 2000 med större valmöjligheter	Klart ökad segregation och rörlighet (bl.a. längre resavstånd till skolan) till följd av antagningsreform som ökade de geografiska valmöjligheterna.

Först ut att studera effekterna av reformen var Fredrik Bergström och Mikael Sandström. De använde SCBs registerdata och fokuserade på effekterna av konkurrens från friskolor på elevprestationer i de kommunala skolorna. De studerade (främst) elever som gick ut grundskolan 1998, vilket innebär att de studerade den allra första ökningen av friskolor. I sina analyser har de använt olika utfallsvariabler som mått på elevprestationer, såväl betygsgenomsnitt (meritvärde), som betyg i enskilda ämnen och resultat från de nationella proven. De använder också en uppsättning olika metoder. Några av dessa kräver tämligen starka antaganden för att resultaten skall kunna tolkas som kausala effekter, och vi menar därför att det finns anledning till en hel del försiktighet med tolkningen av resultaten. Även författarna själva framhåller att vissa av antagandena kan ifrågasättas, men de fäster vikt vid att resultaten från olika modellansatser pekar åt samma håll, nämligen i riktning mot positiva effekter för barnen i de kommunala skolorna. Ett resultat från den av författarna föredragna modellen pekar på att en höjd friskoleandel med tio procentenheter höjer meritvärdet med 14,6 enheter, eller med 0,23 standardavvikelser.

Ahlin (2003) studerade effekterna på resultaten på de nationella proven i matematik, engelska och svenska i årskurs 9, men hade en något annorlunda metodansats. Hennes utgångspunkt är att de skattningar som utnyttjar tvärsnittssamband kan vara snedvridna då friskolorna kan ha etablerats i kommuner som redan tidigare hade bättre presterande elever och detta i sådana dimensioner att de inte beaktades av de kontrollvariabler som användes av Bergström och Sandström. Ahlin använde därför UGU-data över drygt 8 000 elever födda 1982, som gick i tredje klass 1991/92 i 34 slumpmässigt utvalda kommuner. I detta material finns även tester gjorda i åk 6. Hon tillämpade en s.k. value-added modell där testresultaten i åk 6 är kontrollvariabler. Vidare inkluderas elever både från friskolor och kommunala skolor i analysen. Det senare betyder att effekten av andelen elever i friskolor skall tolkas som summan av två olika effekter, nämligen effekten av att gå i friskola samt den indirekta ”konkurrens-effekten” av friskoleandelen på eleverna i de kommunala skolorna. Resultaten visade att andelen elever i friskolor hade en signifikant positiv effekt på testresultaten i matematik. Skattningen innebär att 10 procentenheters ökning av andelen friskoleelever leder till bättre testresultat motsvarande 6 percentiler i fördelningen av testresultat. Ytterligare resultat visade dessutom att effekterna var tämligen homogena: låg- och högpresterande elever påverkas ungefär lika. Däremot var de skattade effekterna på resultaten i engelska och svenska små och inte signifikanta.

Nästa forskningsprojekt som studerade effekterna av friskolor finns avrapporterat i Björklund m.fl. (2003, 2005). Även här studerades huruvida ökad friskoleandel påverkade studieresultaten för samtliga grundskoleelever i kommunen. De kunde använda data som låg något längre fram i tiden (1998–2001) då friskoleandelen på grundskolenivån hade ökat ytterligare. För att hantera problemet med att friskoleandelen torde vara korrelerad med skolprestationer användes metoden ”difference-in-differences” så att *förändringen* i skolresultat relaterades till *förändringen* i friskoleandelen. De fann då positiva effekter på betygen i matematik, engelska och svenska som motsvarar ungefär 0,4 percentiler i betygsfördelningen till följd av en procentenhet högre friskoleandel. Denna effekt var dessutom lägre för elever med invandrarbakgrund respektive lågutbildade föräldrar, varför den positiva effekten i huvudsak är att hänföra till socioekonomiskt mer gynnade elever.

Böhlmark och Lindahl (2007, 2008) har i två studier utvecklat analyserna vidare. De utnyttjar den kommunala variationen i ökningstakten i friskoleandelen (för grundskolan) under hela perioden 1993–2003. I deras första studie är ett huvudresultat att effekterna på grundskolebetyget är lägre än i Björklund m.fl.: Böhlmark och Lindahl fann att ökningen till följd av ökning av friskoleandelen med den procenhet bara var 0,1 percentiler (dock fortfarande statistiskt signifikant skild från noll). En förklaring till denna skillnad kan vara att Böhlmark och Lindahl undersökte en längre tidsperiod. Den främsta skillnaden torde dock vara att de mätte friskoleandelen på ett annat sätt, nämligen som friskoleandelen i årskurs 7–9 medan både Bergström och Sandström och Björklund m.fl. utgick från andelen friskoleelever i årskurserna 1–9.

Med lägre totala effekter skulle man kunna tänka sig att det finns elevgrupper som förlorar på ökad friskoleandel. Böhlmark och Lindahls (2007) resultat tyder dock inte på detta. Deras analyser för olika delgrupper tyder inte på att det finns några förlorande grupper och skillnaderna mellan gruppern (definierade i termer av familjebakgrund) är inte särskilt stora.

Böhlmark och Lindahl diskuterade och analyserade också huruvida användandet av betyg som utfallsvariabel kan vara missvisande. Detta är en naturlig kritisk fråga för dessa studier då det är möjligt att konkurrensen från friskolorna om eleverna har effekter i sig via s.k. betygsinflation. Deras analyser med hjälp

av befintliga data från nationella prov pekade dock inte på att resultaten som bygger på betyg är missvisande.⁷⁵

I sin studie från 2008 utvecklar Böhlmark och Lindahl analysen ytterligare i en mycket intressant riktning. De studerar nämligen mer långsiktiga effekter av friskolor på grundskolenivå, nämligen effekter på vidare gymnasiestudier och påbörjade universitetsstudier. Det visade sig att de positiva effekterna (för de elever som gick ut grundskolan 1993–2003) kvarstod när det gällde val av teoretisk linje på gymnasiet medan det inte fanns några signifikanta effekter på gymnasiebetyg och benägenheten att påbörja universitetsstudier. För framtida utvärderingar blir det även relevant att studera arbetsmarknadseffekterna i form av arbetslöshet och inkomster.

Några av studierna ovan har också undersökt friskolornas effekter på andra utfall såsom kostnader och segregation. Med tanke på den längre undersökningsperioden (fram till 2003) framstår här Böhlmark och Lindahls (2007) resultat som mest trovärdiga.⁷⁶ De fann med samma metodansats som för elevprestationer att friskolor har en viss kostnadshöjande effekt. Skattningarna gav som resultat att medan en procentenhet högre friskoleandel höjer elevprestationen med 0,1 % leder en sådan förändring också till 0,2 % högre kostnader. Jämfört med andra kostnadshöjande reformer menade dock Böhlmark och Lindahl att detta var en kostnadseffektiv reform.

Böhlmark och Lindahl (2007) undersökte också en ökad friskoleandel ledde till ökad segregation. De fann att en ökad friskoleandel ledde till att de kommunala skolorna tenderade att förlora elever med högutbildade föräldrar men inte elever med föräldrar som har höga inkomster. Även för invandrarbakgrund var effekterna splittrade med ökad andel första generationsinvandrare men lägre andel andra generationens invandrare i de kommunala skolorna till följd av ökad fler friskolor.

En intressant studie av effekterna av skolval på segregation är Söderström och Uusitalo (2010). De studerade konsekvenserna av en reform år 2000 av antagningssystemet till gymnasiet i Stockholms kommun. Reformen innebar en övergång till renodlad betygsantagning, medan tidigare system använde en kombination av betyg och närhetsprincip då antalet sökande till en skola över-

⁷⁵ Även Bergström och Sandström (2005) och Björklund m.fl. (2003, 2005) undersökte effekterna på testresultat. Vid den tidpunkt då dessa studier gjordes var dock informationsläget vad gäller testresultat klart sämre än för betyg så dessa analyser är osäkra av detta skäl.

⁷⁶ Se även Björklund m.fl. (2003, 2005), Antelius (2007) för analyser av kostnadseffekter. Dessa studier ger också referenser till debattinlägg i frågan. Von Greiff (2009) diskuterar också vissa kostnadsaspekter av friskolereformen.

steg antalet platser. Söderström och Uusitalo undersökte effekterna genom att jämföra förändringen i segregation på skolnivå i Stockholms kommun med förändringen i angränsande kommuner. De fann tämligen kraftiga effekter av reformen. För det första visar de att benägenheten att välja en skola längre bort från bostaden ökade påtagligt; avståndet mellan bostad och skola ökade relativt kraftigt. För det andra visade olika mått på segregation att denna ökade klart mer i Stockholms kommun än i angränsande kommuner. Segregationen ökade i flera dimensioner: efter elevernas betyg i grundskolan, efter föräldrarnas utbildning och efter invandrarsstatus. De relativt starka effekterna kan bero på att det handlade om skolval vid relativt hög ålder (gymnasiet) och det var frågan om en storstad med goda kommunikationsmöjligheter.

Dessa svenska studier har kunnat genomföras tack vare tillgången till data baserade på administrativa register. Studierna är i flera avseenden avancerade och tillförlitligheten och precisionen i studierna har ökat när förändringar i friskoleandelen över längre tidsperioder har kunnat användas i analyserna. Om de senare studierna med längre tidshorisont ges en högre vikt framstår det som rimligt att säga att effekterna av friskolereformen är tämligen begränsade. Ett undantag här är dock studien av ändrad antagning till gymnasier i Stockholms kommun som fann att effekterna på segregation var relativt stora. Samtidigt är det en brist i det svenska kunskapsunderlaget att studierna kring effekterna på elevernas prestationer är begränsade till betyg och tester i årskurs nio som enda utfall. Under de kommande åren torde det vara möjligt att undersöka även effekterna av friskolor på gymnasienivå.

Allmänt sett framstår det som rimligt att både positiva och negativa effekter av skolval och konkurrens beror mycket på institutionella förhållanden. Mot denna bakgrund är det intressant att notera att den svenska forskningen bara är en liten del av en pågående forskningsvåg i många länder. Det är möjligt att denna forskning inom några år kan ge en tydligare bild av under vilka omständigheter olika inslag av skolval och konkurrens kan ge effekter av olika slag. Studier från andra länder kommer också under flera år framåt vara enda möjligheten att studera effekterna på elevprestationer i tidig ålder.

8.3 Decentralisering

En central avvägningsfråga för varje nationellt utbildningssystem är hur långt beslut om verksamheten skall decentraliseras, till den enskilda kommunen eller skolan. Även finansieringen av verksamheten kan ske på nationell eller kommunal nivå. För Sveriges del innebär 1990-talet en stor förändring härvidlag då

både beslut över verksamheten och finansieringen kom att förflyttas från central nivå till kommunerna: skolan ”kommunaliserades”. Somliga bedömare menar att denna utveckling förändrade det svenska skolsystemet från ett av västvärldens mest centraliserade system till ett av de mest decentraliserade (OECD 1998).

Det finns troligen för- och nackdelar med ett decentraliserat system. En fördel kan vara att beslut om verksamheten tas av dem som också skall genomföra denna och därför har den bästa kunskapen om hur undervisningen bör anpassas till elevunderlaget. Det kan också vara en fördel att skolan avspeglar den lokala befolkningens preferenser vad avser skolan. Mot dessa fördelar står ett fördelningspolitiskt argument. Om den enskilda kommunen får möjlighet att själv välja hur mycket resurser den vill lägga på skolan kan det uppstå stora skillnader i undervisningens kvalitet mellan barn som växer upp i rika och fattiga kommuner. Detta kan undvikas med central (statlig) finansiering av skolan. Även en central styrning av undervisningens utformning kan bidra till större likhet i undervisningens kvalitet mellan olika delar av landet.

Att belysa dessa för- och nackdelar är en viktig uppgift för forskningen. Två svenska studier har studerat effekter av den kommunaliseringen under 1990-talet och båda fokuserar på fördelningsfrågan, om fördelningen av utbildningsresurser mellan kommuner, förändrades till följd av decentraliseringen.

Björklund m.fl. (2003, 2005) visar att spridningen i lärartätheten mellan elever ökade mellan 1990/91 och 2002/03. Spridningen ökade framför allt i den undre delen av fördelningen så att kvoten mellan medianen och den 10:e percentilen av fördelningen ökade. Samtidigt visar de att fördelningen av lärarresurserna över landet varit kompensatorisk – dvs. till förmån för kommuner med lägre inkomster – under hela perioden och att denna kompensation inte var lägre i slutet av perioden än i början.

Ahlin och Mörk (2008) undersökte också om decentraliseringen påverkat fördelningen av resurser men på ett något annorlunda sätt. För det första beaktar de att decentraliseringen genomfördes i tre etapper: (1) 1991 blev kommunerna ansvariga för grundskolorna och de tidigare öronmärkta statsbidragen till kommunerna blev ett samlat bidrag till skolan; (2) 1993 försvann statsbidraget till skolan och ersattes med ett generellt statsbidrag som kommunerna fritt disponerade över; (3) 1996 blev lönesättningen för lärare decentraliserad. För det andra undersöker de om effekterna av vissa egenskaper hos kommunerna på: (1) skolutgifterna per elev och (2) lärartätheten per elev förändrades vid de tre tidpunkterna.

De fann att sambandet mellan kommunernas skattebas och de båda utgiftsmåtten inte blivit starkare (mer positivt) till följd av decentraliseringen. Vidare fann de att statsbidragen fick en mindre effekt på utgifterna – och särskilt på lärartätheten – när de blev mer generella och mindre öronmärkta.

Sammantaget tyder dessa studier på att resursfördelningen har påverkats mellan kommuner till följd av decentraliseringen, men inte på något fördelningspolitiskt ofördelaktigt sätt. Negativa fördelningseffekter kan dock ta lång tid att uppkomma varför fler studier med längre uppföljningsperioder är angelägna. Det är också möjligt att skillnaderna i resurser inom kommunerna har påverkats under perioden och åtminstone delvis som en följd av reformerna.

8.4 Inspektioner

8.4.1 Skolinspektioner i Sverige

Skolinspektioner har en lång bakgrund i Sverige.⁷⁷ En statlig skolinspektion med betydande kyrkligt inflytande infördes 1863. Det kyrkliga inflytandet försvann dock helt 1955. I slutet av 1950-talet infördes Länskolnämnderna som bl.a. utövade tillsyn över skolan. Reguljär inspektion hade dock en underordnad roll fram till 1980-talet då den intensifierades och länskolnämnderna ålades att utöva ”aktiv tillsyn” med syfte att stödja skolan i sitt utvecklingsarbete. I samband med den omfattande kommunaliseringen av skolan i början av 1990-talet avskaffades både skolöverstyrelsen och länskolnämnderna. Under denna tid hade vissa kommuner egen inspektionsverksamhet. Det kan således noteras att den kraftiga decentraliseringen av den svenska skolan inte följdes upp med en förstärkt central inspektionsverksamhet, utan även denna decentraliserades.

År 2004 skedde dock återigen en förstärkning av inspektionsverksamheten på central nivå då Skolverket fick en särskild inspektionsnämnd och anställde nya inspektörer. Ytterligare en förstärkning av denna verksamhet kom 2008 då den nya myndigheten Statens skolinspektion bildades. Denna inspektion ska granska kvaliteten i skolväsendet, förskolan och skolbarnsomsorgen. En första uppgift var att slutföra den första cykeln av ett uppdrag som Skolverket fick år 2003, nämligen att inspektera samtliga skolor vart sjätte år tom. år 2009. Till denna tillsynsuppgift kommer att utföra tillsyn med anledning av anmälningar och signaler om problem i olika skolor. Ytterligare en uppgift för Statens skolinspektion är tillståndsprövning vad avser fristående skolor. Inom myndigheten

⁷⁷ Se Lander och Granström (2000) för en jämförelse mellan svensk och engelsk skolinspektion och historiska referenser.

finns även ett särskilt ombud för barn och elever, barn- och elevombudet. Skolinspektionen har också verkställt en central omrättning av ett urval av nationella prov. Sammantaget har denna nya myndighet således fått omfattande uppgifter. Budgeten för 2010 uppgick till ca. 340 miljoner kronor.

Skolinspektionens verksamhet avser hela skolsystemet till och med gymnasiet. Högskoleverket har sedan länge genomfört utvärderingar av olika högskolor och av olika ämnen inom högskolan. Inspektioner har varit viktiga inslag i dessa utvärderingar.

8.4.2 Erfarenheter av inspektioner från andra länder

Skolinspektioner är och har varit en viktig del av utbildningspolitiken även i en del andra europeiska länder. Däremot har de aldrig varit centrala i USA.

I Storbritannien finns, liksom i Sverige, en gammal tradition av skolinspektioner som också först utövades av prästerna. Av särskild relevans för den svenska debatten är en nysatsning som 1992 genomfördes i England och Wales genom inrättandet av Office of Standards in Education (OFSTED). Denna myndighets verksamhet har blivit föremål för ett par intressanta kvantitativa utvärderingar.⁷⁸ Den bakomliggande lagen krävde att OFSTED skulle inspektera varje offentligt finansierad skola minst en gång under en period av fyra år. OFSTEDs övergripande uppdrag var att förbättra undervisningens kvalitet genom inspektioner, offentliga rapporter och rådgivning. Skolorna kunde informeras om inspektionerna upp till ett år i förväg. De påbörjade då normalt ett tämligen omfattande förberedelsearbete som underlag för inspektionen, t.ex. genom att tillhandahålla information om elevernas bakgrund och olika testresultat. Inspektörerna observerade sedan eleverna i klassrummen och genomförde intervjuer med lärare och elever. Slutligen publicerades en detaljerad rapport som ofta fått stor uppmärksamhet i lokala media.

Två likartade men oberoende utvärderingsstudier av inspektionerna under 1990-talet har genomförts av Shaw m.fl. (2003) och Rosenthal (2004). Båda studierna undersökte om inspektionerna hade några effekter på resultaten från ett nationellt prov för 15-åringar (General Certificate of Secondary Education, GCSE), som tillmätts stor betydelse i det engelska skolsystemet. Studierna är baserade på resultat på skolnivå med observationer för ett antal år, både före och efter inspektionen. Analysen är således uppbyggd så att effekter av inspektionen identifieras som en förändring för skolan under inspektionsåret, eller året efter detta, i förhållande till skolans egen långsiktiga nivå och till utveck-

⁷⁸ Se också Lander och Granström (2000) för ingående diskussion och jämförelse med Sverige.

lingen i riket. Båda studierna visar inledningsvis att inspektionernas förläggning i tiden var helt oberoende av skolornas tidigare prestationer.

Resultaten är tämligen enstämmiga. Rosenthal höll sig till den dominerande gruppen av skolor, s.k. comprehensive schools, som inte hade något selektivt urval av elever baserat på studieprestationer. Han fann svagt negativa effekter av inspektioner. Shaw m.fl. genomförde separata analyser av comprehensive schools och skolor med selektivt intag. De fann också negativa effekter av inspektioner för comprehensive schools, men positiva för de selektiva skolorna; effekterna var dock inte så stora: andelen elever med någon av de tre högsta betygsnivåerna föll från 35,7 till 35,1 % för comprehensive schools men ökade från 92,9 till 94,0 % i skolorna med selektiv intagning.

Även Nederländerna har ett liknande system för skolinspektioner. Den ansvariga myndigheten har i uppdrag att inspektera skolorna för att se till att de följer gällande regler men också med syfte att förbättra undervisningen. Sedan 1998 inspekteras alla grundskolor med standardiserade instrument som involverar enkäter till elever, lärare och föräldrar, observationer av undervisningen och strukturerade intervjuer. Inspektionen utmynnar i rekommendationer och ett antal utvärderingsresultat vilka publiceras på internet och får ofta medial uppmärksamhet. Inspektionerna är av två slag, antingen korta som varar en dag eller längre som varar 2–3 dagar.

Luginbuhl m.fl. (2009) har utvärderat dessa inspektioner med två ansatser varav en är snarlik den som använts i de brittiska studierna. Utfallet är standardiserade testresultat för 12-åringar. Dessa tester är viktiga för både elever och skolor. Studien använder data för åren 1999–2003 för över 6 000 skolor. Den ena ansatsen använder fixa effekter för varje skola, varför analysen bygger på förändringar på skolnivå under åren just efter en inspektion. Dessutom kontrolleras för bl.a. elevernas socioekonomiska bakgrund vilken kan variera från år till år. Med denna ansats fann forskarna signifikant positiva effekter under åren efter inspektionerna på 2–3 % av en standardavvikelse. Dessa effekter var något kraftigare för ett aritmetiskt delprov än för två språkliga. De var också kraftigare för de längre inspektionerna än för de korta. Forskarna kunde också använda ett mindre urval av slumpmässigt utvalda skolor som blev föremål för inspektion. De fick också då positiva effekter av inspektioner men med lägre precision; enbart effekten på det aritmetiska delprovet var signifikant.

Sammantaget visar dessa studier att det är möjligt att genomföra meningsfulla utvärderingar av skolinspektioner. När man begrundar resultaten bör man dock beakta att studier som dessa inte fångar upp alla effekter av ett inspek-

tionssystem. Det är möjligt att enbart förekomsten av ett sådant har effekter på alla eller de flesta skolor oavsett om dessa blivit föremål för inspektion eller ej. Den metod som använts i studierna fångar bara upp effekterna av själva inspektionerna. Det är fullt möjligt att dessa i vissa system är positiva och negativa i andra. Rosenthal, som genomförde en av de engelska studierna, nämner som en möjlig förklaring till de negativa effekterna av själva inspektionerna att dessa kan vara krävande i sig och ta bort både uppmärksamhet och resurser från själva undervisningen. För att genomföra bra utvärderingar av inspektioner är det också centralt att ha tillgång till meningsfulla tester snarare än betyg. Inspektioner, som tar sikte på att granska skolornas betygssättning, skulle kunna bidra till att reducera tendenser till så kallad betygsinflation.

8.5 Diskussion och slutsatser

De studier kring effekter av differentiering med särskilt fokus på grundskolenivån som diskuteras i avsnitt 8.1 visar att de förväntade effektivitetsvinsterna av nivågruppering inte kunnat påvisas i den omfattande internationella forskningen. Däremot visar många studier på negativa effekter av nivågruppering på elevers självkänsla och skolmotivation.

Den internationella forskningen ger också visst stöd för slutsatsen att organisatorisk differentiering på grundskolenivån orsakar ökad variation i elevresultaten, utan förbättrad medelprestation. Snarare pekar resultaten på att organisatorisk differentiering leder till sänkning av medelprestationen, även om dessa resultat är osäkra. En tänkbar förklaring är att kamrateffekter, åtminstone delvis, medverkar till uppkomsten av den organisatoriska differentieringens effekter; flera studier pekar på att kamratsammansättningen har betydelse för elevprestationer. Det är dock tveksamt om kamrateffekter är tillräckliga för att förklara de observerade effekterna av organisatorisk differentiering. En annan möjlig förklaring är att den organisatoriska differentieringens effekter har sin grund i kvalitetsskillnader i den utbildning som ges i olika studieinriktningar.

Ett antal svenska studier kring effekterna av fristående skolor och decentralisering som genomförts under senare år har gett värdefulla insikter. Det övergripande intrycket är att effekterna av de stora förändringar i styrningen av svensk utbildningspolitik som genomfördes under första halvan av 1990-talet är tämligen begränsade. Studierna tyder på att både positiva och negativa effekter är tämligen små. Man kan dock invända att studierna de undersökt en aspekt på det stora reformpaketet under 1990-talet i taget, och att effekten av hela paketet kan vara relativt stort. Exempelvis har ingen studie specifikt undersökt

effekterna av att kunna välja kommunal skola. Likafullt är det frågan om kvalificerade svenska studier som bidragit till en växande internationell litteratur kring olika styrningsfrågor inom utbildningssystemet.

Det är den svenska infrastrukturen med för hela landet rikstäckande registerdata som har gjort dessa studier möjliga. Det finns anledning att förvänta sig fler studier av detta slag. Framför allt kommer det att bli möjligt att studera utfallen på längre sikt när de elever som berördes av de stora förändringarna under 1990-talet kommer ut på arbetsmarknaden så att deras sysselsättning och löner kan användas som utfallsvariabler.

Men man kan också peka på avsaknaden av flera önskvärda svenska studier där det är brister i infrastrukturen som förklarar att studierna inte kunnat göras. Studierna har främst använt SCB:s Årskurs 9 register med avgångsbetygen som utfallsvariabel. Här finns information om huruvida eleverna examinerades från fristående skola eller kommunal skola i årskurs 9.

Däremot saknas information om hur länge eleverna gått i respektive typ av skola. Detta är en brist. Men framför allt innebär det faktum att det saknats heltäckande registerdata på betyg och tester före årskurs 9 att de viktiga första skolåren inte kunnat bli föremål för liknande analyser.

Likaså har studierna av friskolor i första hand använt betyg som utfallsvariabel. Eftersom det finns anledning befara s.k. betygsinflation som en följd av införandet av fristående skolor hade det varit bättre om studierna genomgående också hade kunnat använda resultat från nationella prov eller andra standardiserade tester. De nationella prov som genomförs i årskurs 5 har ännu inte använts i utvärderingsstudier av det slag som diskuterats ovan.

Vi har också redovisat några resultat från studier av skolinspektioner i andra länder. Dessa har gett blandade resultat, negativa effekter i England och positiva i Nederländerna. Såvitt vi vet saknas motsvarande studier för Sverige. Den svenska registerstatistiken skulle vara en bra grund för liknande studier även i vårt land. Här är det dock också helt centralt med data på testresultat. Eftersom inspektionerna avser även skolornas betygssättning kan försämrade betyg vara en önskad effekt av skolinspektioner.

Vi vill avsluta med att understryka att skolinspektioner och nationella standardiserade tester av olika slag är komplement snarare än substitut när det gäller att övervaka och styra ett skolsystem. Med hjälp av skolinspektioner är det möjligt att se till att tester genomförts på ett betryggande sätt och att betygen är rimliga i förhållande till testresultaten. Med bra testdata är det också möjligt att utvärdera själva inspektionerna och successivt förbättra dessa.

Referenser

- Ahlin, A. (2003), Does school competition matter? Effects of a large-scale school choice reform on student performance, Working Paper 2003:2, Nationalekonomiska institutionen, Uppsala Universitet.
- Ahlin A och E Mörk (2008), Effects of decentralization on school resources, *Economics of Education Review*, 27, 276–284.
- Ammermüller, A. Och S. Pischke (2009), Peer effects in European primary schools: Evidence from PIRLS, *Journal of Labor Economics*, 27(3), 315–348.
- Antelius, J. (2007). Fristående skolor och kommunernas kostnader, i A Lindbom (red), *Friskolorna och framtiden – segregation, kostnader och effektivitet*. Stockholm: Institutet för Framtidsstudier.
- Baumert, J. (2008), Lecture at the seminar “School tracking, achievement and inequality”. Royal Swedish Academy of Sciences, Stockholm, Nov. 12.
- Baumert, J., M. Kunter, W. Blum, M. Brunner, T. Dubberke, A. Jordan, U. Klusman, S. Krauss, M. Neubrand och Y-M. Tsai (2010), Teachers’ mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, 47(1), 133–180.
- Betts, J.R. och J.L. Shkolnik, (2000). Key difficulties in identifying the effects of ability grouping on student achievement. *Economics of Education Review*, 19(1), 21–26.
- Björklund A, P-A Edin, P Fredriksson och A Krueger (2003), *Den svenska skolan – effektiv och jämlik?*, Välfärdspolitiska rådets rapport. Stockholm: SNS förlag.
- Björklund A, M Clark, P-A Edin, P Fredriksson och A Krueger (2005), *The Market Comes to Education in Sweden: an evaluation of Sweden’s surprising School Reforms*. New York: Russel Sage Foundation.
- Böhlmark A. och M. Lindahl (2007), The impact of school choice on pupil achievement, segregation, and costs: Swedish evidence, IZA DP 2786.
- Böhlmark A och M Lindahl (2008), Does school privatization improve educational achievement? Evidence from Sweden’s voucher reform, IZA DP 3691.

- Carlberg, C. och K.A. Kavale, (1980), The efficacy of special versus regular class placement for exceptional children: A meta-analysis, *Journal of Special Education*, 14, 296–309.
- Coleman, J.S., E.Q. Campbell, C.J. Hobson, J. McPartlant, A.M. Mood, F.D. Weinfeld och R.L. York (1966). *Equality of educational opportunity*. Washington, DC: U. S. Government Printing Office.
- Dahllöf, U. (1967). *Skoldifferentiering och undervisningsförlopp*. Stockholm: Almqvist & Wiksell.
- Darling-Hammond, L. och J. Bransford (2005), *Preparing teachers for a changing world: What teachers should learn and be able to do*. San Francisco, CA: Jossey-Bass.
- Forness, S.R. (2001), Special education and related services: What have we learned from meta-analysis? *Exceptionality*, 9(4), 185–197.
- von Greiff, C. (2009). Lika skola med olika resurser? En ESO-rapport om likvärdighet och resursfördelning, ESO 2009:5. Stockholm: Fritzes.
- Hanushek, E.A. och L. Wössmann (2006). Does educational tracking affect performance and inequality. Differences-in-differences evidence across countries. *Economic Journal*, 116(510), C63–C76.
- Hattie, J.A. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Hoxby, C. (2000). Peer effects in the classroom: Learning from gender and race variation. NBER WP 7867.
- Jakubowski, M. (2007). Effects of tracking on achievement growth. Exploring difference-in-differences approach to PIRLS, TIMSS and PISA data. Discussion Paper 2006-7, EUF-S.
- Kroesbergen, E.H. och J.E.H. Van Luit (2003), Mathematics interventions for children with special educational needs, *Remedial and Special Education*, 24(2), 97–114.
- Lander R och K Granström (2000), Skolinspektionen i England och Sverige- hjälp till självhjälp eller självstyre med blets i munnen?, *Pedagogisk Forskning i Sverige*, 5(3), 215–234.

- Luginbuhl R., D. Webbink och I. de Wolf (2009), Do school inspections improve primary school performance?, *Educational Evaluation and Policy Analysis*, under publicering.
- Meghir, C. och M. Palme (2005), Educational reform, ability, and family background, *American Economic Review*, 95(1), 414–424.
- Oakes, J. (1985), *Keeping track: How schools structure inequality*. New Haven: Yale University Press.
- Oakes, J. (1990). Multiplying inequalities: the effects of race, social class and tracking on opportunities to learn mathematics and science. Santa Monica; CA: The Rand Corporation.
- OECD (1998), *Education at a glance. OECD indicators 1998*. Paris: Centre for Educational Research and Innovation, OECD.
- Pekkarinen, T., R. Uusitalo och S. Kerr (2009), School tracking and development of cognitive skills, IFAU Working Paper 2009:6.
- Rosenthal L. (2004), Do school inspections improve school quality? Ofsted inspections and school examinations in the UK, *Economics of Education Review*, 23, 143–151.
- Sacerdote, B. (2001), Peer effects with random assignment: results for Dartmouth roommates, *Quarterly Journal of Economics*, 116, 681–704.
- Sandström M. och F. Bergström (2005), School vouchers in practice: competition will not hurt you, *Journal of Public Economics*, 89, 351–380.
- Sandström M. och F. Bergström (2007), Konkurrens mellan skolor – för kvalitetens skull!, i A Lindbom (red.) *Friskolorna och framtiden – segregation, kostnader och effektivitet*. Stockholm: Institutet för Framtidsstudier.
- Seidel, T. och R.J. Shavelson (2007), Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results, *Review of Educational Research*, 77(4), 454–499.
- Shaw, I., D.P. Newton, M. Aitkin och R. Darnell (2003), Do Ofsted inspections of secondary schools make a difference to GCSE results?, *British Educational Research Journal*, 29(1), 63–75.

- Slavin, R.E. (1987), Ability grouping and student achievement in elementary schools: A best evidence synthesis, *Review of Educational Research*, 57(3), 293–336.
- Slavin, R.E. (1990), Achievement effects of ability grouping in secondary schools: A best evidence synthesis, *Review of Educational Research*, 60(3), 471–499.
- Sund, K. (2007), *Teachers, family and friends: Essays in economics of education*, nationalekonomiska institutionen, Stockholms universitet.
- Sund, K. (2009), Estimating peer effects in Swedish high school using school, teacher, and student fixed effects', *Economics of Education Review*, 28, 329–336.
- Svensson, N-E. (1962). *Ability grouping and scholastic achievement. Report on a five-year follow-up study in Stockholm*. Stockholm: Almqvist & Wiksell.
- Söderström, M., och R. Uusitalo (2010), School choice and segregation: evidence from an admission reform, *Scandinavian Journal of Economics* 112(1), 55–76
- Zimmerman, D. (2003). Peer effects on academic outcomes. Evidence from a natural experiment. *Review of Economics and Statistics*, 85, 9–23.
- Åslund O., P-A. Edin, P. Fredriksson och H. Grönqvist, H. (2009). Peers, neighborhoods and immigrant student achievement – evidence from a placement policy, IFAU Working Paper 2009:20.

9 Gymnasieskolan och vuxenutbildningen

Många relevanta skolpolitiska frågor är gemensamma för gymnasiet och grundskolan. Alla frågor som vi behandlat i kapitel 7 och 8 är studieöverskridande och analysen har därför relevans på gymnasie- såväl som på grundskolenivå.

Men det som definierar gymnasiet i nuvarande utbildningsstruktur är att det är första gången elever gör ett val av inriktning som har betydelse för deras framtida utbildningsnivå. Som vi noterade i inledningskapitlet innebar införandet av 1994 års läroplan att den formella uppdelningen i allmän och särskild kurs på högstadiet försvann. Därmed försköts den differentieringen, åtminstone formellt, till gymnasievalet.

Sedan 1960-talet har gymnasiestudier delats upp i studieförberedande och yrkesförberedande utbildningar. 1991 års gymnasiereform minskade skillnaderna mellan studie- och yrkesförberedande inriktningar. Den organisatoriska differentieringen kvarstod, men yrkesprogrammen förlängdes med ett år och förlängningen bestod i huvudsak av teoretiska studier; efter reformen gav även yrkesprogrammen allmän behörighet för högskolestudier. Ambitionen med denna reform var att skapa en väsentligen mer sammanhållen utbildningsstruktur än tidigare. I och med reformen kom den svenska ungdomsutbildningen att närma sig den amerikanska modellen, där nästan alla elever fortsätter till High School vars huvudsakliga syfte är att förbereda för vidare studier.

Denna utveckling reser två fundamentala och närliggande frågor: Vilka effekter har specialisering vid olika åldrar? Och gynnas alla elever av generell utbildning? Dessa frågor hänger nära samman med differentieringsfrågan som vi behandlade i kapitel 8. Forskningen tyder på att sammanhållna system minskar spridningen i utbildningsutfall men att genomsnittliga studieprestationer inte påverkas särskilt mycket. Detta är resultatet från både jämförande länderstudier (se Hanushek och Wössmann 2006) och studier av enskilda länder (t.ex. Pekkarinen m.fl. 2009). Men det är viktigt att komma ihåg att denna forskning antingen jämför länder där selektionen sker vid olika åldrar (någonstans mellan 11 och 16 års ålder) eller förändringar inom land där ålder vid uppdelningen förskjutits från 11 till 16 års ålder. Studierna belyser alltså inte frågan om att minska inslaget av specialisering för ungdomar i gymnasieåldern.

Det finns dock en liten litteratur som berör värdet av yrkesutbildning för ungdomar i gymnasieåldern. Det är framförallt denna fråga vi berör här. Vi fokuserar på yrkesutbildningen och hur 1990-talets förändringar påverkade utbildnings- och arbetsmarknadsutfall för de elever som berördes.

1991 års gymnasiereform ökade inslaget av generell utbildning i yrkesutbildningarna. Argumentationen som låg till grund för denna förändring var rimligen att en generell utbildning är en bättre försäkring då ekonomin är i snabb omvandling. Men ett ökat inslag av generell utbildning innebär definitionsmässigt att den direkta kopplingen till arbetsmarknaden försvagas. Och det är också möjligt att ett ökat inslag av teoretiska studier skapar särskilda problem för elever på yrkesutbildningar, som tenderar att komma från studiesvagare miljöer och som kanske också är mindre studiemotiverade. I detta kapitel undersöker vi bland annat vilket argument som verkar ha störst bäring.

Frågan om graden av specialisering på gymnasienivå är synnerligen aktuell. En statlig utredning (SOU 2008:27) har nyligen föreslagit att gymnasieutbildningen ska reformeras igen. Avsikten är att minska teoriinslaget och öka praktikinnehållet i de yrkesförberedande programmen. I detta reformpaket ingår också förslaget att införa en lärlingsutbildning. Inslaget av yrkesutbildning kommer kort sagt att öka om dessa förslag blir verklighet.

Hur man ser på värdet av generell utbildning kontra värdet av en specialiserad utbildning beror på möjligheterna till omskolning. I det svenska systemet finns stora möjligheter till omskolning via kommunal vuxenutbildning (komvux); existensen av en omfattande andrachansutbildning mildrar eventuella negativa konsekvenser av en för tidig specialisering. Eftersom ett genomgående tema i detta kapitel är konsekvenser av specialisering är det naturligt att också ta upp frågan om effekterna av komvux. Dessutom är det huvudsakligen gymnasial utbildning som bedrivs på komvux: t.ex. deltog 78 % av eleverna i komvux i gymnasial vuxenutbildning 2007 (Utbildningsstatistisk årsbok 2009).

Ett slående inslag i 1990-talets utveckling i svensk skola är att andelen elever som inte har godkända slutbetyg har ökat, ett fenomen som gäller såväl grundskolan som gymnasiet. Uppgången sker under andra halvan av 1990-talet och är samtida med övergången från det relativa till det målrelaterade betygssystemet. Andelen med icke-godkända slutbetyg är särskilt hög bland grupper från studieovana miljöer, grupper som är överrepresenterade på yrkesutbildningar, vilket gör att vi särskilt uppmärksammar fenomenet i detta kapitel.

Kapitlet är disponerat på följande sätt. I avsnitt 9.1 beskriver vi gymnasieutbildningen med särskilt fokus på yrkesförberedande utbildningar. Vi beskriver både utvecklingen i det längre perspektivet och presenterar inkomstutfall för olika studieinriktningar. Avsnitt 9.2 behandlar konsekvenserna av gymnasiereformen från 1991. Den grundläggande frågan är: Hur påverkades de elever som gick in på yrkesutbildningarna av reformen? Vi studerar effekterna på

examensfrekvenser, övergångar till högre utbildning och framtida inkomster. I avsnitt 9.3 undersöker vi om det går att hänföra nedgången i andelen med godkända betyg till införandet av målrelaterade betyg. Vi undersöker också om individers arbetsmarknadsutsikter påverkades av betygsreformen. Avsnitt 9.4 behandlar vuxenutbildningen och vad forskningen kan säga om arbetsmarknadseffekterna av denna utbildningsform. En sammanfattning och diskussion av resultaten återfinns i avsnitt 9.5.

9.1 Gymnasiet – en beskrivning

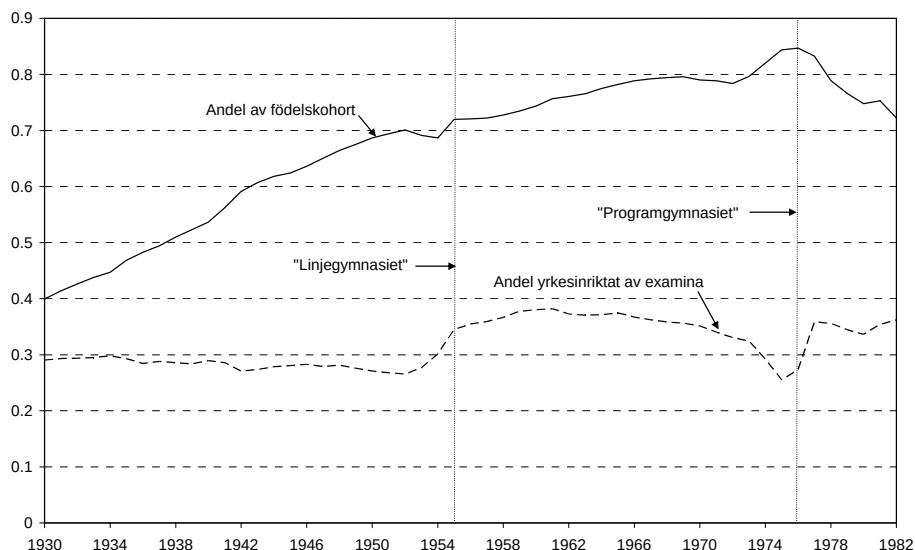
I likhet med de flesta frivilliga utbildningsformer har gymnasieutbildningen expanderat kraftigt under de senast 50 åren. Den heldragna kurvan i Figur 9.1 illustrerar denna expansion genom att redovisa andelen av individer med ett visst födelseår som har en gymnasial utbildning på minst 2 år. Beräkningarna är baserade på 2004 års version av SCB:s utbildningsregister. Notera att vi genomgående har definierat alla utbildningar som kom att föras till gymnasiet i samband med 1971 års gymnasiereform som ”gymnasiala”.⁷⁹ I figuren har vi också indikerat vilka årgångar som berörs av de två stora gymnasiereformerna som genomförts under de senaste 50 åren. Individer som är födda 1955 är de som framförallt berörs av 1971 års gymnasiereform och införandet av ”linjegymnasiet”.⁸⁰ Individer som är födda 1976 är de som framförallt berörs av 1991 års gymnasiereform och införandet av ”programgymnasiet”.

Det är tydligt att den stora expansionen av den gymnasiala utbildningen sker mellan födelsekohorterna 1930 och 1960, då andelen med en gymnasial utbildning stiger från 40 % (födda 1930) till knappt 75 % (födda 1960). Därefter ökar förvisso andelen med gymnasieutbildning trendmässigt under en tid, men i samband med införandet av programgymnasiet och det målrelaterade betygssystemet, sjunker andelen med gymnasieexamen avsevärt: andelen med gymnasieexamen bland individer födda 1980 uppgår återigen till knappt 75 %.

⁷⁹ Vi inkluderar alltså genomgående individer som har en 2-årig utbildning från t.ex. fackskolor och yrkesskolor trots att dessa utbildningar organisatoriskt inte fördes till gymnasiet förrän 1971. Detta höjer andelen med en gymnasial utbildning bland individer som är födda före 1955.

⁸⁰ Beslutet om att 1971 års gymnasieskola togs 1968. Notera även att gymnasieskolan reformerades 1964.

Figur 9.1. Andelen med minst 2 års gymnasial utbildning och andelen med en yrkesinriktad examen av totala antalet examina, efter födelsekohort



Not: Egna beräkningar med hjälp av 2004 års årgång av utbildningsregistret. Utbildningsklassificeringen följer SUN 2000. När vi beräknat andelen yrkesinriktade examina har vi antagit att alla individer med högre utbildning än gymnasium har gått en studieförberedande utbildning. Detta antagande kommer att leda till en underskattning av andelen som har gått en yrkesförberedande utbildning.

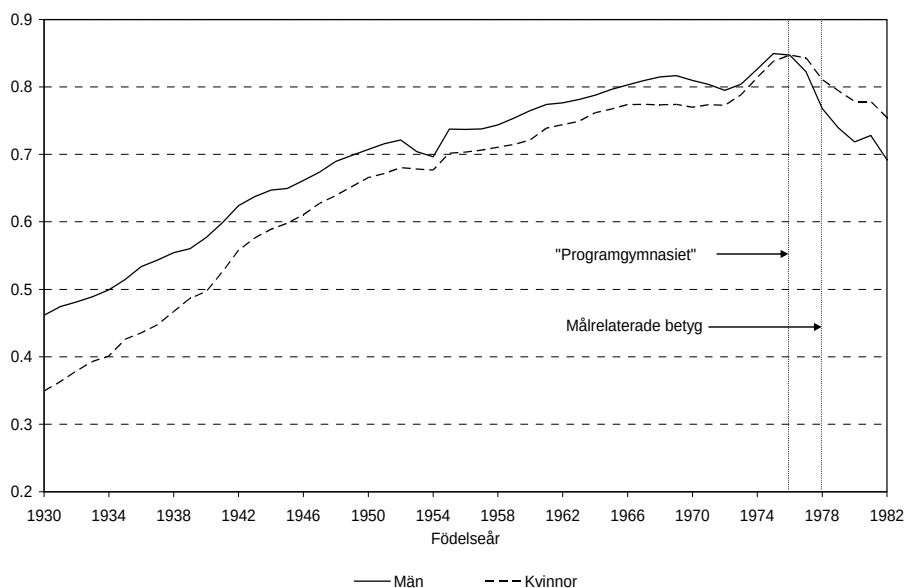
Den streckade linjen i Figur 9.1 visar antalet yrkesexamina som andel av det totala antalet examina. Klassificeringen är baserad på Svensk utbildningsnomenklatur (SUN) och bygger på antagandet att alla individer som har en högre utbildning än en gymnasial utbildning har gått en studieförberedande utbildning. Detta antagande kommer att underskatta andelen yrkesexamina något, men detta är rimligen ett försumbart problem om vi är intresserade av utvecklingen över tid.

Andelen yrkesexamina av det totala antalet examina framstår som förhållandevis konstant. Förvisso förändras andelen yrkesexamina i samband med de två stora gymnasiereformerna – införandet av linjegymnasiet 1971 och programgymnasiet 1991 – med detta verkar till stor del vara följderna av att utbildningarna klassificeras om i samband med reformerna.

Figur 9.2 redovisar gymnasieexpansionen efter kön. Om vi går långt tillbaka i tiden har männen ett försteg: bland män födda 1930 har 46 % åtminstone en examen från en 2-årig gymnasial utbildning; motsvarande andel bland kvinnor

är 35 %. Men precis på samma sätt som i övriga delar av utbildningssystemet minskar, och försvinner, männens försteg över tid.

Figur 9.2. Andelen med minst 2 års gymnasial utbildning, efter kön och födelsekohort



Not: Egna beräkningar med hjälp av 2004 års årgång av utbildningsregistret.

I den födelsekohort som framförallt påverkas av 1991 års gymnasiereform – individer födda 1976 – är examensfrekvensen för både män och kvinnor 85 %. Därefter faller examensfrekvensen snabbare för män än för kvinnor, särskilt när det målrelaterade betygssystemet införs som huvudsakligen påverkar individer födda 1978 eller senare. Bland individer födda 1982 uppgår andelen män med gymnasieexamen till 69 % och andelen kvinnor till 75 %.

Eftersom mäns studieprestationer generellt sett är sämre än kvinnors, och män i större utsträckning väljer yrkesutbildningar, finns det skäl att tro att både införandet av programgymnasiet och de målrelaterade betygen haft negativa konsekvenser för examensfrekvenserna hos studiesvagare elever. Avsnitt 9.2 och 9.3 innehåller en rigorös undersökning av dessa frågor.

Tabell 9.1 innehåller delvis samma information som Figur 9.1 och Figur 9.2. Den visar bland annat examensfrekvenserna för olika kohorter som stude-

rats inom ramen för UGU-projektet, dvs. kohorterna födda 1948, 1953, 1967, 1972, 1977, och 1982.⁸¹

Tabell 9.1. Andel med minst 2 års gymnasium (%) och förmåga vid 13 års ålder (rangordning)

Kohort	Minst 2 års gymnasieexamen					
	Totalt		Yrkes		Studie	
	Andel	Rang	Andel	Rang	Andel	Rang
1948	68,6	55,5	18,9	44,5	49,7	59,7
1953	71,1	55,3	18,9	43,5	52,2	59,5
1967	83,0	52,6	32,1	40,3	50,8	60,4
1972	81,4	52,9	28,6	39,1	52,8	60,4
1977	88,0	51,8	35,3	39,9	52,6	59,9
1982	75,5	53,3	27,7	43,1	47,8	59,3

Källa: Egna beräkningar från UGU och 2004 års utbildningsregister.

Anm.: 1948 och 1953 är individernas födelseår. Från 1967 och framåt är urvalet baserat på årskurs (3:e eller 6:e klass). "Födelseåret" indikerar därför när individer som går i 3:e eller 6:e klass normalt är födda. Rangorden är baserad på förmågetest som eleverna genomför vid 13 års ålder. Rangordningen går från 0 (sämst) till 100 (bäst).

Mervärdet med tabellen är att den också illustrerar hur olika utbildningskategorier positionerar sig i fördelningen av förmågetester vid 13 års ålder. Tabellen visar t.ex. att när gymnasieutbildningen expanderar så tenderar genomsnittliga studieprestationer att försämrats något (jämför t.ex. kohorterna födda 1948 och 1977). Det är även tydligt att elever som väljer yrkesinriktade program i genomsnitt presterar påtagligt sämre på förmågetest vid 13 års ålder än de som väljer studieförberedande program. Återigen visar data att examensfrekvensen faller avsevärt efter att de målrelaterade betygen införs – jämför kohorterna födda 1977 och 1982. Notera att den genomsnittliga rangen bland elever med en yrkesexamen stiger relativt mycket i samband med detta, vilket betyder att nedgången i examensfrekvensen är koncentrerad bland elever som presterande sämre vid 13 års ålder.

Ett möjligt problem med den beskrivning som hittills presenterats är att individernas utbildningsnivå observeras vid olika åldrar. Erikson m.fl. (2007) redovisar istället individernas utbildning vid givna åldrar, 20 och 23 års ålder. De grundläggande mönstren kvarstår dock vid en sådan analys. Examensfrekvensen sjunker dramatiskt bland individer som omfattas av betygsreformen.

⁸¹ Eftersom UGU-undersökningen är en urvalsundersökning kan examensfrekvenserna skilja sig från utvecklingen som visas i Figur 9.1, som i sin tur är baserad på populationsdata.

Vilken arbetsmarknadskoppling har då yrkeslinjerna? I Tabell 9.2 underöker vi denna fråga genom att redovisa inkomstpremier för olika yrkesinriktningar.

Tabell 9.2 Sambandet mellan yrkesprogram och årsarbetsinkomst, procent

Program	Ålder					
	20 år		23 år		26 år	
(Andel elever, %)	Inkomst	Inkomst-premie	Inkomst	Inkomst-premie	Inkomst	Inkomst-premie
Bygg (6,1)	4,0 (0,5)	35,7 (1,9)	2,3 (0,4)	26,3 (1,5)	0,8 (0,4)	20,5 (1,5)
El (11,3)	-2,1 (0,4)	6,6 (1,6)	0,7 (0,3)	16,0 (1,3)	1,3 (0,4)	15,9 (1,3)
Energi (2,1)	-0,6 (0,7)	15,0 (2,9)	-0,1 (0,6)	19,6 (2,2)	0,9 (0,7)	15,5 (2,4)
Fordon (9,2)	1,4 (0,4)	47,3 (1,7)	2,7 (0,3)	37,7 (1,3)	3,2 (0,4)	29,7 (1,3)
Handel och administration (14,6)	Referens	Referens	Referens	Referens	Referens	Referens
Hotell och restaurang (8,5)	2,1 (0,4)	9,0 (1,7)	0,4 (0,3)	1,8 (1,4)	1,0 (0,4)	-4,6 (1,4)
Hantverk (2,4)	-5,0 (0,7)	-3,4 (2,7)	-8,9 (0,7)	-8,1 (2,6)	-10,0 (0,8)	-19,6 (2,9)
Naturbruk (5,8)	-3,7 (0,5)	-0,3 (1,9)	-1,4 (0,4)	-4,2 (1,7)	-1,3 (0,5)	-5,2 (1,7)
Livsmedel (1,8)	1,0 (0,7)	12,4 (3,0)	-1,6 (0,7)	5,2 (2,6)	-0,6 (0,7)	-0,2 (2,6)
Media (6,3)	-4,6 (0,5)	-26,2 (1,9)	-4,1 (0,4)	-23,8 (1,8)	-3,3 (0,5)	-19,9 (1,8)
Barn och fritid samt omvårdnad (30,8)	2,2 (0,3)	3,7 (1,3)	1,5 (0,3)	-3,4 (1,1)	2,2 (0,3)	-5,5 (1,1)
Industri (1,0)	-4,0 (1,0)	24,8 (4,5)	0,0 (0,9)	20,0 (3,6)	0,6 (0,9)	18,2 (3,2)
Antal individer	117902	103709	117048	108122	116499	105411

Anm.: Skattningarna är baserade på individer födda mellan 1976 och 1979 vars inkomst observeras vid en given ålder (20, 23, eller 26) mellan 1996–2005. Bara individer som startat en yrkesutbildning på gymnasiet vid 16 års ålder ingår i skattningarna. Vi konstanthåller för kohort (vilket fångar både kohort- och tidseffekter), kön, samt grundskolebetyg i regressionerna. Skattningarnas (robusta) standardfel anges inom parentes.

Skattningarna baseras på individer födda 1976 till 1979 som slutade grundkolan på normalt tid (dvs. vid 16 års ålder) och påbörjade en gymnasial yrkesutbildning samma år. Inriktningskoden anger det program de började på. Vi betingtar inte på att de ska ha tagit examen, eftersom det är ett möjligt utfall av inriktningsvalet. Individernas årsinkomster observeras mellan 1996–2005.

Även om vi konstanthåller för individernas grundskolebetyg ska dessa skattningar rimligen tolkas som de beskrivningar de faktiskt är. Vi har här inte ambitionen att skatta kausala inriktningseffekter.

Tabell 9.2 visar att det finns förhållandevis stora skillnader mellan inriktningar och att dessa existerar både ett år efter att de i normalfallet ska ha slutat gymnasiet och sju år efter det normala examenstillfället. De inriktningar som kommer bäst ut i dessa jämförelser är traditionella yrkesprogram, som bygg- och fordonsprogrammet. Även om inkomstpremierna avtar något med ålder är det tydligt att de ger en snabb koppling till arbetsmarknaden. Dessa program domineras av pojkar med förhållandevis låga grundskolebetyg.

Individer som har gått media- och hantverksprogrammet har de lägsta inkomsterna vid 26 års ålder. Båda dessa program attraherar elever med relativt höga grundskolebetyg; individer som började dessa program har dock låga inkomster, även sju år efter den normala examenstidpunkten.

Den enkla jämförelsen i Tabell 9.2 faller alltså ut till fördel för yrkesprogram som bygg- och fordonsprogrammet. Individer som går dessa program kommer snabbt ut på arbetsmarknaden och har högre inkomster längre fram i livet.

9.2 Effekter av ett tredje år i yrkesutbildningen

Gymnasiereformen 1991 förlängde de praktiska utbildningarna med ett tredje (huvudsakligen) teoriinriktat år. En intressant fråga är vilka effekter denna reform hade på utbildnings- och arbetsmarknadsutfall. Vi undersöker här hur reformen påverkade sannolikheten att ta en gymnasieexamen, sannolikheten att gå vidare till högskolan, samt framtida arbetsinkomster, för individer på yrkesutbildningarna.

Reformen av de praktiska utbildningarna introducerades på försöksbasis (det s.k. ÖGY-försöket), vilket gör att den faktiskt kan utvärderas. Det egentliga försöket pågick under läsåren 1988/89–1990/91, men de 3-åriga programmen introducerades i liten skala redan 1987/88 (då det omfattade ca. 500 studieplatser). Läsåret 1988/89 allokerades drygt 6 000 studieplatser till 3-åriga yrkesutbildningar; läsåret 1989/90 knappt 10 000 platser och läsåret 1990/91 drygt 11 000 platser. Valet av försökskommuner var inte slumpmässigt. Kommunerna ansökte om att få delta i försöket; dock betonade regeringen att det var viktigt att försöket skulle omfatta kommuner med olika industri- och befolkningsstruktur.

Inom försöksverksamheten ökade undervisningen i kärnämnen. I tillägg till svenska (som var det enda obligatoriska ämnet på alla 2-åriga utbildningar) så

innehöll de 3-åriga utbildningarna även engelska, samhällskunskap och en valbar kurs (vilket var matematik i de allra flesta fallen). Försöksverksamheten hade också en större arbetsplatsförlagd del i jämförelse med de tvååriga utbildningarna. Den yrkesutbildning som infördes i och med 1991 års gymnasiereform hade en mindre andel arbetsplatsförlagd utbildning och ett större fokus på teoretiska kärnämnen.

Ekström (2002) har redovisat ett första försök att utvärdera försöksverksamheten. Hon studerade hur individers utbildningsnivå och framtida arbetsmarknadsanknytning påverkades. Det finns dock två potentiella problem med utvärderingen. För det första saknade hon information om gymnasierregistreringar vilket innebar att hon var tvungen att anta att det 3:e året inte påverkade sannolikheten att få fullständigt slutbetyg. För det andra var det inte slumpmässigt vilka kommuner som blev försökskommuner, vilket inte hanterades på ett helt tillfredställande sätt. Det första problemet är ett reellt problem vilket illustreras i en färsk studie av Hall (2009), se nedan. Det andra problemet har att göra med att Ekström (2002) jämförde utfallen för elever bosatta i kommuner med försöksverksamheten med utfallen för elever i kommuner utan försöksverksamheten. Om försökskommuner skiljer sig från icke-försökskommuner leder denna ansats till snedvridna skattningar. Hall (2009) jämförde istället förändringar i utfall för elever som påverkades i olika utsträckning av försöket i en s.k. Differences-in-differences analys.

Halls studiepopulation utgörs av individer som avslutade en grundskoleutbildning åren 1986–90 och direkt påbörjade en 2- eller 3-årig yrkesutbildning. Individer som började 1986 omfattas inte alls av försöket. Individer som började övriga år påverkas i varierande grad av försöket. Omfattningen på försöket i varje kommun används som ett instrument för sannolikheten att påbörja en treårig yrkeslinje. Tabell 9.3 visar huvudresultaten från Halls studie.

Tabellen visar att de treåriga yrkesutbildningarna ökade sannolikheten att inte ta examen i jämförelse med de tvååriga utbildningarna. Effekten uppgår till 3,8 procentenheter vilket i relativa termer motsvarar en ökning med 35 %. Effekterna är koncentrerade till elever med låga grundskolebetyg – det finns inga effekter för elever vars grundskolebetyg överstiger snittet. För elever med låga grundskolebetyg uppgår effekten till 8,3 procentenheter, vilket motsvarar en ökning med 50 % i förhållande till medelvärdet under hela tidsperioden. Hall visar också att sannolikheten att inte ta examen är betydligt större för elever som har lågutbildade föräldrar.

Tabell 9.3. Effekter av att påbörja en treårig yrkesutbildning på utbildningsutfall

	Sannolikheten att ta examen			Sannolikheten att påbörja högskoleutbildning		
	Alla	Efter grundskolebetyg		Alla	Efter grundskolebetyg	
		Under snittet	Över snittet		Under snittet	Över snittet
Påbörjat 3-årig yrkeslinje	−0,038 (0,019)	−0,083 (0,023)	0,017 (0,023)	0,006 (0,016)	0,002 (0,016)	0,010 (0,030)
Antal individer	184 101	101 543	82 558	184 101	101 543	82 558
Medelvärde, beroende variabel	0,109	0,158	0,049	0,184	0,100	0,288

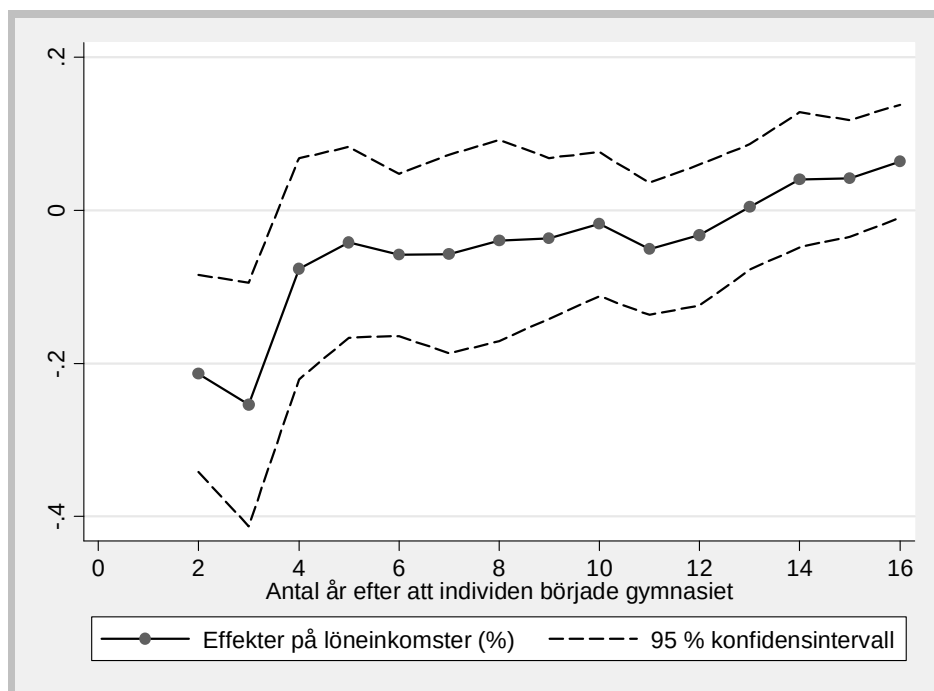
Anm.: En individ klassas som att ha tagit examen om denne har fullständiga gymnasiebetyg inom 6 år efter påbörjad yrkesutbildning. En individ klassas som att ha påbörjat en högskoleutbildning om detta skett senast vid 31 års ålder. Studiepopulationen utgörs av individer som slutade grundskolan 1986–1990 och direkt började en 2- eller 3-årig yrkesutbildning. Alla regressioner har skattats med en instrumentvariabelmetod. I regressionerna hålls kohort och kommun konstant genom s.k. fixa effekter. I regressionerna hålls även individens grundskolebetyg, kön, och invandrarbakgrund, samt föräldrarnas utbildningsnivå och invandrarbakgrund konstant. Standardfelen, som visas inom parentes, är klusterkorrigerade på kommunnivå.

Källa: Hall (2009)

Det tredje året på yrkesutbildningarna hade ingen effekt på övergången till högre studier eller sannolikheten att ta en högskoleexamen. Tabell 9.3 ger ett exempel på dessa resultat för övergångsfrekvensen till högskoleutbildning. Effekterna varierar inte heller med elevernas grundskolebetyg. En tolkning av dessa resultat är att studiemotiverade yrkeselever innan reformen använde Komvux för att läsa in allmän behörighet. Att de treåriga yrkesutbildningarna gav allmän behörighet hade därför ingen effekt på övergångar till högskolestudier. Detta är ett anmärkningsvärt resultat eftersom ett viktigt motiv till reformen var att underlätta övergången till högre studier.

Figur 9.3 visar inkomsteffekterna av att påbörja en 3-årig yrkesutbildning. Initialt är effekterna negativa, vilket naturligen följer av att de 3-åriga yrkesutbildningarna är ett år längre. Ur ett samhällsekonomiskt perspektiv måste de långsiktiga inkomsteffekterna vara positiva och tillräckligt stora för att uppväga denna initiala kostnad. Det finns inga starka tecken på att så är fallet, även om de skattade inkomsteffekterna åtminstone är positiva (dock inte statistiskt signifikanta) 14–16 år efter påbörjad gymnasieutbildning.

Figur 9.3. Inkomsteffekter av att påbörja en treårig yrkesutbildning



Anm. Figuren visar resultaten från regressioner som skattats separat för varje år efter påbörjad yrkesutbildning.
Källa: Hall (2009).

Vad förklarar resultaten i Hall (2009)? Inom försöksverksamheten ökade både det akademiska innehållet i utbildningen och den arbetsplatsförlagda delen i yrkesutbildningen. Det är sannolikt inte ökningen av den arbetsplatsförlagda delen som gör att examensfrekvensen går ner bland elever med sämre förutsättningar. Den rimligaste tolkningen är att nedgången har att göra med kombinationen av ett tredje utbildningsår och en ökad betoning på kärnämnen.

I vilken utsträckning finns det stöd för Halls slutsatser i den internationella litteraturen? Det finns ett antal studier som är baserade på amerikanska data.⁸² Det amerikanska high-school systemet är generellt till sin natur, men det är möjligt att välja yrkeskurser inom ramen för high-school. I några studier har man undersökt arbetsmarknadsutfallen för dem som valt yrkeskurser (se t.ex.

⁸² Det finns även en färsk studie av Malamud och Pop-Eleches (2008) som har studerat en rumänsk reform som liknar den svenska reformen. Även i denna studie fann man att övergångar till högre studier inte påverkades.

Kang och Bishop 1989; Light 1999). Det finns ingen egentlig strategi för att hantera den systematiska selektionen av elever som väljer yrkeskurser, varför dessa studier antagligen bör ges mindre vikt. Likväl är det typiska resultatet att yrkeskurser har positiva effekter på inkomster både på kort sikt (upp till två år efter examen) och medellång sikt (6–7 år efter examen).

En experimentell utvärdering av "Career Academies" levererar dock trovärdigare evidens. En serie rapporter kopplar till denna utvärdering; t.ex. Kemple och Scott-Clayton (2004) och Kemple och Willner (2008). Career Academies existerar i hela USA och är ett särskilt high-school program. Det består av en kombination av akademiska ämnen och yrkesämnen; inom ramen för programmet har man etablerat samarbeten med lokala arbetsgivare. Sedan 1993 har man följt en kohort av drygt 1400 individer som sökte till Career Academies på 9 high-schools över hela USA. Skolorna låg i ekonomiskt utsatta områden och 85 % av individerna i studien hade latinamerikanskt eller afroamerikanskt ursprung. Eftersom antalet sökande översteg antalet platser på utbildningen tillämpade man ett lottningsförfarande. På så sätt skapades två jämförbara grupper: en behandlad grupp (deltagarna i Career Academies) och en kontrollgrupp (som gick på vanlig high-school).

Utvärderingen visar att denna form av yrkesutbildning minskade andelen avhopp, framförallt bland grupper med stor risk för avhopp. Inkomsteffekterna var överlag positiva, men koncentrerade bland män. Åtta år efter den normala examenstidpunkten var arbetsinkomsterna i genomsnitt 11 % högre för dem som antogs till Career Academies. Det finns inga effekter på sannolikheten att fortsätta till högre studier.

Om man tolkar införandet av de treåriga yrkesutbildningarna i Sverige som ett steg mot traditionell high-school utbildning (dvs. minskat fokus på yrkesutbildning) så är dessa resultat anmärkningsvärt konsistenta med resultaten i Hall (2009). Skillnaden som finns rör hur inkomsterna påverkades, där Hall (2009) inte fann långsiktiga inkomsteffekter. Man ska dock ha i åtanke att den svenska reformen också förlängde utbildningarna med ett år vilket i sig torde ha en positiv effekt på arbetsinkomsterna. Med detta i åtanke så kan inkomstresultaten mycket väl vara förenliga med varandra.

9.3 Konsekvenser av det målrelaterade betygssystemet

Här är syftet att undersöka konsekvenserna av det målrelaterade betygssystemet. Införandet av ett nytt betygssystem påverkar naturligtvis både grundskole- och gymnasieelever. Skälet till att fokusera på effekterna på gymnasienivå är

att det är möjligt att jämföra elever med samma (relativa) grundskolebetyg som har betygsatts på olika sätt på gymnasienivå eftersom de är födda olika år.

De målrelaterade betygen på grundskolenivån innebär att andelen som inte har fullständiga betyg från grundskolan ökat, vilket således påverkar inflödet till gymnasieskolan. Elever utan fullständiga grundskolebetyg påbörjar det individuella programmet: andelen som påbörjar det individuella programmet låg på 6 % under perioden 1993–97. Mellan 1998 och 2007 har den varierat mellan 8 och 9 %.

De målrelaterade betygen på gymnasienivå har antagligen bidragit till att andelen som inte fullföljer gymnasieskolan ökat. I årskullen födda 1985 får 23 % av dem som påbörjat en gymnasieutbildning inte slutbetyg; ytterligare 8 % får slutbetyg men har inte grundläggande behörighet (Skolverket 2008).

Här fokuserar vi på hur mycket de målrelaterade betygen påverkat elever som påbörjat gymnasiestudier. Det grundläggande skälet till detta fokus är att vi kan kontrollera för elevernas tidigare studieprestationer, via grundskolebetyget. Den grundläggande ansatsen är som följer. Vi analyserar de individer i årskullarna födda 1976–79 som avslutat grundskolan på normaltids (dvs. vid 16 års ålder) och påbörjat en treårig gymnasieutbildning samma år. Fördelen med detta är att alla dessa individer är betygssatta enligt det relativa betygssystemet på grundskolenivå. Därmed påverkar de målrelaterade betygen inte vilka som är behöriga för gymnasiestudier. Kohorterna födda 1976 och 1977 betygsattes enligt det relativa betygssystemet på gymnasienivå, medan kohorterna födda 1978 och 1979 betygsattes enligt meritvärdessystemet på gymnasienivå.

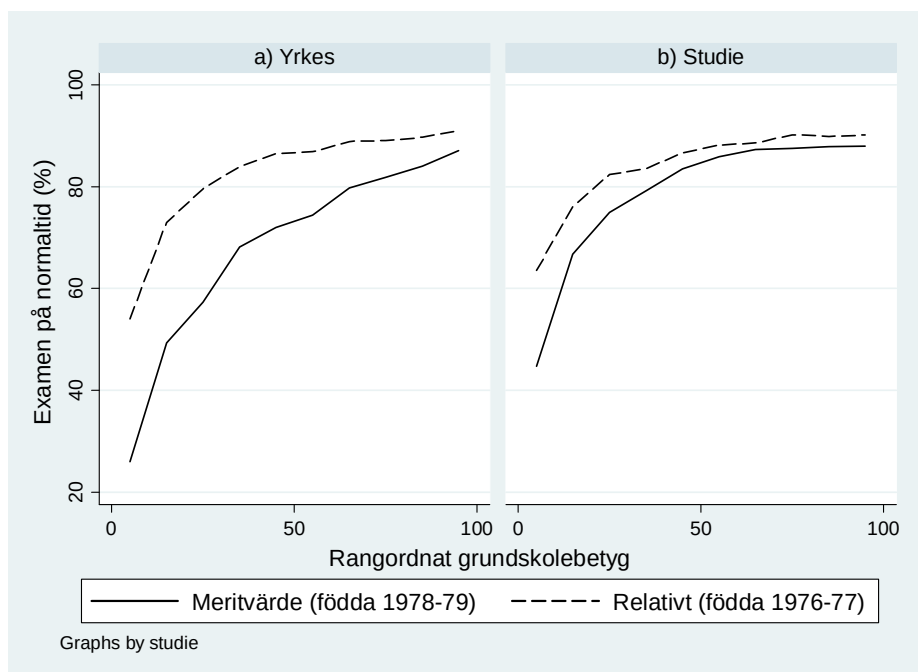
Figur 9.4 visar hur sannolikheten att ta examen på normaltids (dvs. examen vid 19 års ålder) från en gymnasieutbildning varierar över årskullar (och därmed betygssystem) och position i fördelningen av grundskolebetyg. Panel a) visar examensfrekvensen för yrkesutbildningar medan panel b) visar examensfrekvensen för studieförberedande utbildningar.

Det är tydligt att examensfrekvensen är betydligt lägre för de årskullar som betygsattes enligt meritvärdessystemet och att skillnaden är särskilt stor för individer med låga grundskolebetyg. Det gäller både för yrkeselever och elever på studieförberedande utbildningar. Likväl är nedgången i examensfrekvensen mycket kraftigare för elever som påbörjat yrkesutbildningar.

Nedgången i sannolikheten att ta examen på normaltids är en direkt följd av att elever inte godkänns i olika ämnen. I det tidigare relativa betygssystemet fick elever ett slutbetyg så länge som de deltagit i undervisningen. Men meritvärdessystemet har också ett antal indirekta effekter. En del elever väljer

sannolikt att hoppa av eftersom det har blivit svårare att få ett slutbetyg; andra kanske väljer att anstränga sig hårdare för att passera ”ribban”. En ytterligare grupp av elever väljer att fortsätta studera och får ett godkänt slutbetyg, om än vid en senare tidpunkt.

Figur 9.4. Examen på normaltids efter årskull och position i fördelningen av grundskolebetyg



Not: ”Examensåret” mäts med avgångsåret från gymnasiet. En elev har avgått på normaltids om avgångsåret är senast tre år efter påbörjade gymnasiestudier. Elevernas grundskolebetyg har rangordnats inom respektive utbildningstyp (yrkes eller studie).

Källor: Egna beräkningar baserade på olika årgångar av sökanderegistret och avgångsregistret på gymnasiet.

Alldeles oavsett vad som händer är det rimligt att tänka sig att betygsreformen har arbetsmarknadskonsekvenser på individnivå. Om arbetsgivare använder ett godkänt slutbetyg som en signal – och inte bryr sig så mycket om vad eleven har för betyg i t.ex. svenska – kommer de elever som inte har ett slutbetyg att vara i en sämre situation än vad som annars skulle vara fallet. De elever som tar examen senare är också i en sämre position än annars eftersom de förlorar värdefull arbetslivserfarenhet. Via dessa två mekanismer kan alltså

individernas arbetsinkomster påverkas. Huvudsyftet med detta avsnitt är att undersöka om det går att spåra arbetsmarknadseffekter av betygsreformen.

För att svara på denna fråga genomför vi regressionsanalyser. Regressionerna håller genomgående kön och grundskolebetyg konstant. Vi inkluderar dessutom linjära ”trender” i skattningarna.⁸³ Vi undersöker också om det finns särskilt starka effekter i olika delar av fördelningen av grundskolebetyg, eftersom Figur 9.4 antyder att det borde finnas ett sådant mönster.

Tabell 9.4 redovisar regressionsresultat för två utbildningsutfall: sannolikheten att ta examen på normaltids och sannolikheten att ha åtminstone 3 års gymnasieutbildning. Utbildningsnivån observeras 2005 då individerna är mellan 26 och 29 år gamla. I tabellen visas genomsnittseffekterna, dvs. de genomsnittliga effekterna över hela fördelningen av grundskolebetyg.

Tabellen visar att andelen som tar examen på normaltids minskade med drygt 13 procentenheter för elever som påbörjade 3-åriga yrkesutbildningar och betygsattes enligt meritvärdessystemet. Motsvarande effekt uppgår till en minskning med 7,6 procentenheter för elever som påbörjade 3-åriga studieförberedande utbildningar.

Nedgången i andelen individer som har åtminstone 12 års utbildning är inte lika stark. Effekten uppgår till -5,1 procentenheter för dem som påbörjade yrkesutbildning och -1,4 procentenheter för de studieförberedande utbildningarna. Att nedgången inte är lika stark för detta utfall betyder att de kohorter som är berörda av betygsreformen antingen tar längre tid till examen eller hittar andra vägar (t.ex. komvux) för att uppnå 12 års utbildning.

⁸³ Ett delikat problem som uppstår i detta sammanhang är att det är ett exakt linjärt samband mellan årskull, ålder och observationsår. Eftersom meritvärdeseffekterna skattas genom att jämföra olika årskullar kan vi inte både hålla ålder och observationsår konstant. Vi hanterar detta problem genom att fixera antingen ålder eller observationsår och inkludera linjära ”trender”. Trenderna fångar olika saker beroende på vad vi har fixerat: om vi studerar utfall vid en given ålder (som vi gör för examensfrekvensen och inkomster) fångar trenderna linjära tidseffekter; om vi studerar utfall vid ett givet observationsår (som vi gör för utbildning) fångar trenderna linjära ålderseffekter. En ännu svårare fråga som följer av detta är om man ska tillåta dessa trender/ålderseffekter att variera med individens position i betygsfördelningen. Argumentet för att göra detta skulle t.ex. vara att det även utan betygsreformen skulle ha gått sämre för ungdomar med lägre grundskolebetyg. När det gäller examensfrekvensen är detta argument långsökt, framförallt eftersom dessa kohorter ligger nära varandra i tid. När det gäller inkomster är det svårare att ta ställning eftersom konjunkturen kan påverka individer med olika betyg på skilda sätt. När det gäller utbildning kommer vi att tillåta att ålderseffekterna varierar med positionen i betygsfördelningen eftersom olika studiebegåvade individer fortsätter olika länge i utbildningssystemet.

Tabell 9.4. Effekter av meritvärdessystemet på utbildningsutfall, procent

Utfall	Meritvärdeseffekt
<u>Yrkesutbildningar</u> (Antal individer: 116 560)	
(1) Examen på normaltids (senast vid 19 års ålder)	-12,9 (0,5)
Medelvärde beroende variabel, kohorter födda 1976–77	82,4
(2) Åtminstone 12 utbildningsår (utbildning observerad 2005)	-5,1 (0,4)
Medelvärde beroende variabel, kohorter födda 1976–77	91,1
<u>Studieförberedande utbildningar</u> (Antal individer 176 539)	
(1) Examen på normaltids (senast vid 19 års ålder)	-7,6 (0,4)
Medelvärde beroende variabel, kohorter födda 1976–77	84,1
(2) Åtminstone 12 utbildningsår (utbildning observerad 2005)	-1,4 (0,2)
Medelvärde beroende variabel, kohorter födda 1976–77	96,6

Not: Populationen som studeras består av alla individer födda 1976–79 som avgick grundskolan på normaltids och påbörjade en 3-årig gymnasieutbildning samma år. Meritvärdeseffekten skattas genom en jämförelse av individer födda 1978–79 (som betygsattes enligt meritvärdessystemet på gymnasiet) med individer födda 1976–77 (som betygsattes enligt det relativa systemet). I regressionerna med examensfrekvensen som beroende variabel hålls kön och grundskolebetyg samt en gemensam trend konstant; regressionerna med utbildning som beroende variabel inkluderar även en linjär ålderseffekt som varierar med position i betygsfördelningen. Robusta standardfel visas inom parentes. Regressionerna utnyttjar data från 2005 års utbildningsregister samt olika årgångar av sökanderegistret och avgångsregistret på gymnasiet.

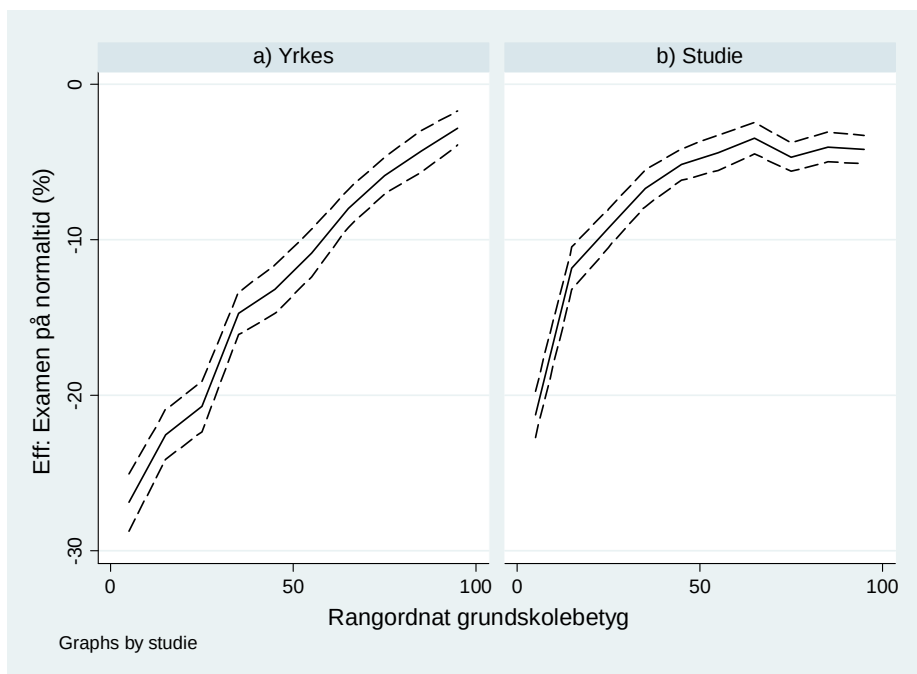
Figur 9.5 och Figur 9.6 redovisar separata skattningar för varje decil i fördelningen av grundskolebetyg; panel a) innehåller skattningar för elever som påbörjat yrkesutbildningar, medan panel b) gäller de studieförberedande utbildningar. Figur 9.5 visar effekterna på examensfrekvensen och Figur 9.6 visar effekterna på sannolikheten att ha åtminstone 12 års utbildning.⁸⁴

Det är tydligt att examensfrekvensen minskar särskilt mycket för elever med låga grundskolebetyg. Detta gäller oberoende av utbildningstyp (yrkes- eller studieförberedande) men det är uppenbart att mönstret är starkast bland yrkeseleverna. För en yrkeselev som tillhörde de tio procenten med lägst grundskolebetyg (inom gruppen av yrkeselever) minskade examensfrekvensen

⁸⁴ I klartext innebär detta att individer som har 3 års utbildning efter grundskolan tillhör kategorin som har åtminstone 12 års utbildning.

med hela 26,9 procentenheter; medan den minskade med 2,8 procentenheter för yrkeseleverna med högst grundskolebetyg.

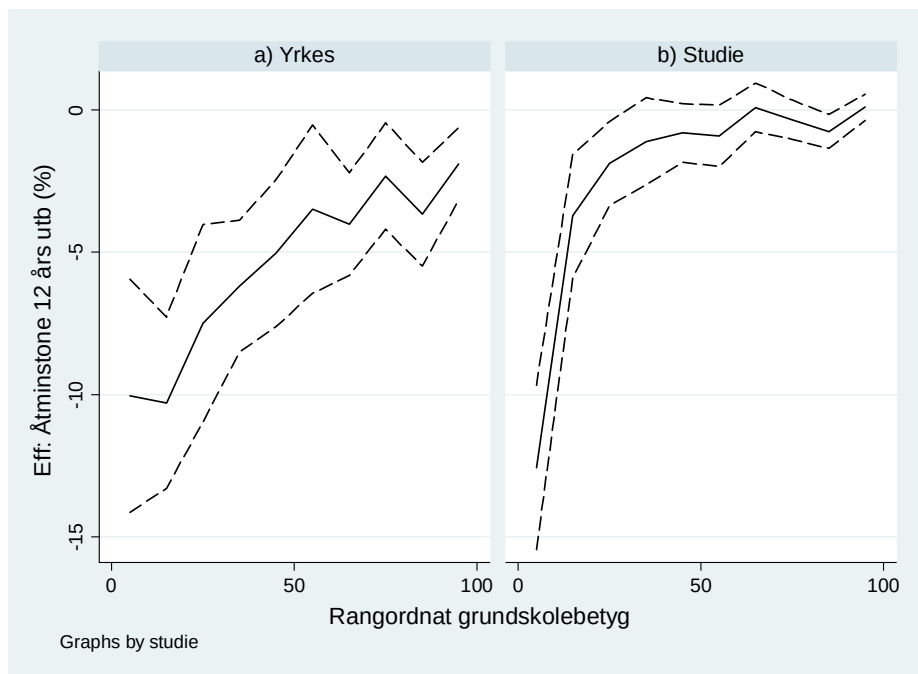
Figur 9.5. Effekter på sannolikheten att ta examen på normaltids efter position i fördelningen av grundskolebetyg



Not: De heldragna linjerna visar effektskattningarna och de streckade linjerna är skattningarnas konfidensintervall. Effekterna har skattats separat per decil. Regressionerna håller kön och en gemensam trend konstant. Elevernas grundskolebetyg har rangordnats inom respektive utbildningstyp (yrkes eller studie).

Figur 9.6 visar att mönstret från Figur 9.5 i mångt och mycket kvarstår. Storleken på effekterna är mindre eftersom en del av eleverna uppnår en 3-årig gymnasieutbildning vid en senare tidpunkt. Panel b) visar t.ex. att det bara finns en statistiskt säkerställd effekt för elever som börjat en teoretisk gymnasieutbildning och tillhörde den lägsta kvartilen i betygsfördelningen.

Figur 9.6. Effekter på sannolikheten att ha åtminstone 12 års utbildning efter position i fördelningen av grundskolebetyg



Not: De heldragna linjerna visar effektskattningarna och de streckade linjerna är skattningarnas konfidensintervall. Effekterna har skattats separat per decil. Regressionerna håller kön och en linjär ålderseffekt konstant. Elevernas grundskolebetyg har rangordnats inom respektive utbildningstyp (yrkes eller studie).

Den kanske intressantaste frågan är dock om det också går att spåra konsekvenser på regelrätta arbetsmarknadsutfall. För att besvara denna fråga har vi matchat på information om individers löneinkomster åren 2002 till 2005. Med hjälp av dessa data undersöker vi om individers årsinkomster för en given ålder (26 år) påverkas av betygsreformen. Individerna i våra data är ännu för unga för att det ska vara meningsfullt att studera elever som har valt studieförberedande utbildningar; huvudsyftet med dessa är ju att förbereda för högskolestudier och många av dessa individer är fortfarande i utbildning. Analysen av arbetsmarknadsutfall genomförs därför bara för yrkeselever.

Tabell 9.5 redovisar skattningar av meritvärdeseffekter på olika inkomstutfall för individer som påbörjat en yrkesutbildning. Vi undersöker två olika inkomstutfall: dels sannolikheten att ha en årsinkomst och dels årsinkomsten som andel av medianinkomsten i populationen som arbetar vid varje tidpunkt.

Tabell 9.5. Effekter av meritvärdessystemet på inkomstutfall, yrkesutbildningar, procent

Utfall	Meritvärdeseffekt
(1) Årslöneinkomst > 0 (vid 26 års ålder)	−0,4 (0,4)
Medelvärde beroende variabel, kohorter födda 1976–77	91,7
(2) Årslöneinkomst (vid 26 års ålder) i relation till medianlöneinkomst	−4,6 (0,6)
Medelvärde beroende variabel, kohorter födda 1976–77	76,6

Anm.: Studiepopulationen är individer födda 1976–79 som slutade grundskolan på normaltid och direkt påbörjade en yrkesutbildning. Löneinkomsterna mäts vid 26 års ålder och observeras alltså under perioden 2002–05. Robusta standardfel visas inom parentes. Det fåtal individer som inte finns i befolkningsregistret under observationsåren har exkluderats. Genomgående hålls en trend, grundskolebetyg och kön konstant. Skattningarna är baserade på 116 834 individer.

Källa: Egna beräkningar baserade på sökanderegistret till gymnasiet och inkomstregistret.

Den första panelen i Tabell 9.5 visar resultaten för sannolikheten att ha en löneinkomst. Vi finner inga effekter i genomsnitt. Eftersom vi har valt en låg nedre gräns (0 kronor) och det är årsinkomster som studeras är detta inte så förvånande. Det stora flertalet (91,7 %) har också en registrerad löneinkomst och det finns därför inte så mycket variation att exploatera i skattningarna.

Den andra panelen av tabellen visar hur individernas inkomster påverkas. Vi mäter inkomsterna som en andel av medianinkomsten vid varje observations-tidpunkt. Den sista raden i tabellen visar att yrkeselever som betygsatts enligt det relativa systemet tjänade knappt 77 % av medianinkomsten vid 26 års ålder. Skattningarna visar att inkomsterna i genomsnitt minskade med 4,6 procentenheter som en konsekvens av meritvärdesreformen.

Figur 9.7 visar skattningar där vi tillåter effekterna att variera över betygs-fördelningen på samma sätt som tidigare. Panel a) visar att, även om det inte finns några genomsnittseffekter, så påverkas individer längst ner i fördelningen av betygsreformen. För individer i den lägsta decilen uppgår minskningen till 3,5 procentenheter och effekterna är statistiskt säkerställda för de individer som tillhör den lägsta kvartilen.

Figur 9.7. Effekter på inkomster vid 26 års ålder efter position i fördelningen av grundskolebetyg, elever som påbörjat yrkesutbildningar.



Not: De heldragna linjerna visar effektskattningarna och de streckade linjerna är skattningarnas konfidsensintervall. Effekterna har skattats separat per decil. Regressionerna håller kön och en gemensam trend konstant.

Panel b) visar inkomsterna för alla individer i fördelningen påverkas. Att skattningarna är större (i en absolut mening) för individer längst ner i fördelningen har till största delen att göra med att sannolikheten att inte ha inkomst där är större. Detta betyder att inkomsteffekten är mer eller mindre konstant för individer som har positiva inkomster.

Varför uppstår då dessa inkomsteffekter? Vi har visat att betygsreformen både minskar utbildningsnivån och förlänger tiden till att uppnå en given utbildningsnivå. Detta medför att individer vid en given ålder både har lägre utbildning och mindre arbetslivserfarenhet. Eftersom både utbildning och erfarenhet har positiva effekter på inkomster har individers arbetsmarknadsutsikter försämrats av betygsreformen. Inkomsteffekterna förstärks om det finns ett positivt signalvärde av att ha ett godkänt slutbetyg.

9.4 Kommunal vuxenutbildning

I internationell jämförelse har Sverige en omfattande vuxenutbildning.⁸⁵ Kommunal vuxenutbildning handlar idag framförallt om studier på gymnasienivå: 78 % av eleverna bedrev gymnasial vuxenutbildning 2007 (Utbildningsstatistisk årsbok 2009). Kvinnor studerar på komvux i större utsträckning än män: 68 % av eleverna var kvinnor 2007. Den vanligaste åldersgruppen i 20–24 år: 30 % av eleverna tillhörde denna åldersgrupp 2007. Sannolikheten att studera på komvux avtar sedan med ålder. Cirka 80 % av alla komvuxelever 2007 var mellan 20 och 44 år.

Figur 9.8 visar hur omfattningen på komvux varierat över tid. Vi har relaterat antalet elever på komvux (oavsett ålder) till storleken på befolkningen i åldern 20–44 år.

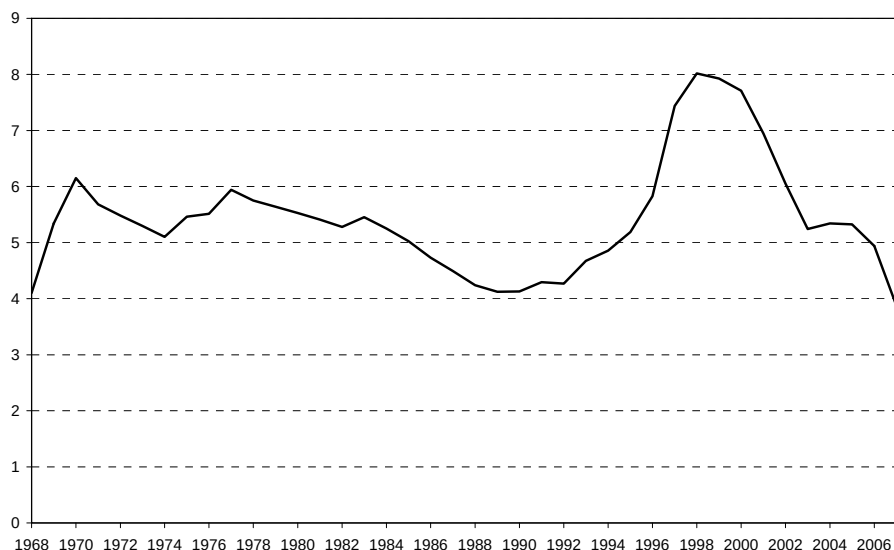
Figur 9.8 visar att en anmärkningsvärt konstant andel på 4–6 % deltar i kommunal vuxenutbildning, om vi bortser från kunskapslyftet (som pågick åren 1997–2001).

Den kommunala vuxenutbildningen startade 1968 då yrkesskolornas deltidskurser samt kvällsgymnasierna överfördes till komvux. Yrkesskolans deltidskurser var en mycket omfattande verksamhet: 1961 deltog t.ex. 110 000 elever; 1969 deltog 92 000 elever i yrkesskolans deltidskurser inom ramen för komvux (se Utbildningsstatistisk årsbok 1978).⁸⁶ Tillgängliga data visar alltså att svensk vuxenutbildning varit omfattande under de senaste 50 åren.

⁸⁵ Det är svårt att ge exakta siffror på hur omfattande vuxenutbildningen är i olika länder, framförallt pga. utbildningsstrukturen varierar över länder. OECD (2005) redovisar siffror på andelen 25–64 åringar som deltar i någon form av utbildning. Dessa siffror inkluderar alltså alla utbildningsnivåer och även utbildning inom ramen för arbetsmarknadspolitiken och företagsutbildning. De nordiska länderna tillsammans med England och Schweiz har de högsta deltagandeandelarna (bland de 18 länder som redovisas i undersökningen).

⁸⁶ Dessa 92 000 elever motsvarar runt 50 procent av eleverna vid komvux 1969.

Figur 9.8. Antalet elever i kommunal vuxenutbildning som andel av befolkningen i åldern 20–44 år, procent, 1968–2007



Källa: Olika årgångar av Utbildningsstatistisk årsbok

Vilka argument finns för en så omfattande vuxenutbildning? Komvux brukar i litteraturen gå under epitetet ”andrachansutbildning”. Och detta epitet stämmer relativt väl med vad det handlar om. Om individer gjort felaktiga utbildningsval i unga år, får de genom komvux en andra chans att korrigera dessa beslut. Argumentet för komvux förstärks i ett system med utpräglad och tidig specialisering.

Mot en omfattande vuxenutbildning kan man anföra att den samhällsekonomiska avkastningen på utbildningsinvesteringar (för en given utbildningsnivå) avtar med ålder. Detta förstärks av att individerna inte bekostar utbildningen själva. Eftersom kostnaderna inte varierar med ålder i någon större utsträckning, och utbildning vid en högre ålder innebär en kortare intjäningsperiod, blir utfallet att avkastningen för samhället som helhet avtar med ålder. OECD:s kalkyler som vi redovisade i kapitel 4 illustrerar detta.

Med perfekt information borde utbildningsinvesteringar göras vid ung ålder. Men utbildningsbeslut fattas under osäkerhet. Och därför kan det finnas goda argument att ha möjligheter till förkovran även vid vuxen ålder. I kapitel 3 illustrerade vi i ett exempel att marginalavkastningen på investeringar kan vara högre om man investerat för lite i unga år än om man gjort optimala val. Denna

utsaga finner visst stöd i de skattningar som Jacobson m.fl. (2005) redovisade. Deras resultat var att marginalavkastningen är hög för vissa utbildningsinriktningar (t.ex. vårdutbildningar) och obefintlig för andra (t.ex. historia).

Det finns några relativt färskta studier på svenska data som berör värdet av Komvux. Det finns goda skäl att skilja på studier som handlar om traditionell vuxenutbildning och studier som utvärderar Kunskapslyftet, då Kunskapslyftet mer har karaktären av en arbetsmarknadspolitisk åtgärd än en permanent utbildningspolitisk åtgärd. Kunskapslyftet var riktat till arbetslösa individer som kunde studera inom den kommunala vuxenutbildningen med bibehållen a-kassa. Det faktum att individerna kunde behålla sin a-kassa innebär att incitamenten för vuxenutbildningen förändrades. Detta kommer i sin tur att påverka vilka grupper som går in i utbildning och det är rimligt att tänka sig att inkomsten efter avslutad utbildning har lägre betydelse som argument för vuxenutbildning än om individerna inte kunnat behålla sin a-kassa. Vi kommer därför att diskutera effekterna av traditionell kommunal vuxenutbildning och Kunskapslyftet separat.

Vad säger utvärderingsforskningen om effekterna av traditionell vuxenutbildning? Ekström (2003) har genomfört den första svenska studien av de långsiktiga inkomsteffekterna av komvux. Hon studerade de individer som påbörjade komvux någon gång mellan våren 1988 och våren 1993. Frågan är hur komvuxstudier påverkade deras inkomster år 2000. För att besvara denna fråga antog Ekström att observerade egenskaper styr vilka som börjar studera på komvux. Ett enkelt sätt att tänka på den empiriska strategin är att hon jämförde två individer – en som började på komvux och en som inte gjorde det – som hade samma kön, utbildning, och inkomstutveckling åren innan den påbörjade vuxenutbildningen.⁸⁷

Ekström fann att effekterna av komvux varierade mellan män och kvinnor samt svenskfödda och invandrare. För svenskfödda män var arbetsinkomsten 3,5 % lägre bland dem som hade gått på komvux än bland män som inte gjort detta. För svenskfödda kvinnor fanns ingen effekt; detsamma gällde invandrade män, även om detta resultat delvis hade att göra med låg precision i skattningarna. För invandrade kvinnor var effekten positiv (och statistiskt signifikant på 10 procentsnivån): invandrade kvinnor som hade gått på komvux hade 9 % högre inkomst än dem som inte gjort detta.

Stenberg (2007b) är ett exempel på en serie uppsatser som involverar samme författare och handlar om komvux samt kunskapslyftet. Stenberg under-

⁸⁷ I tekniska termer använde hon en kombination av en regressions- och en matchningsansats.

sökte individer födda 1970 som antingen hade deltagit på komvux eller ej. Dessa individer kan ha deltagit inom ramen inom reguljär komvux eller inom ramen för kunskapslyftet; han genomförde dock separata analyser för dessa två grupper. Mervärdet i förhållande till Ekström (2003) är att Stenberg hade information antalet kurspoäng på komvux. Han kunde därmed mäta antalet studieår på komvux, till skillnad från Ekström som bara kunde mäta om individerna deltagit eller ej. I övrigt är ansatsen i grunden densamma som i Ekström. Det grundläggande antagandet är att det är individers observerade egenskaper som styr om individen studerar på komvux eller ej.

Ett problem med Stenberg (2007b) är dock att han i praktiken betingar på val eller utfall som observeras efter avslutade studier. Skälet till att detta problem uppstår är att han studerade individer som fortsatt till högskolestudier och individer som inte gjort detta (dvs. individer som trädde in på arbetsmarknaden) separat. Individer som fortsatt till högre studier kan därmed vara både positivt och negativt selekterade; en negativ selektion uppstår t.ex. om individen fortsätter till högre utbildning bara om denne inte hittat jobb efter komvux. Hur som helst är resultaten svårtolkade eftersom Stenberg betingade på utfall som uppkommer efter – och sannolikt som ett resultat av – komvux.⁸⁸

Med reservation för ovanstående problem verkar ett studieår generera 5 % högre inkomster för individer som hade 2 års gymnasieutbildning innan komvux och som inte fortsatte till högre studier. Denna siffra ska som sagt betraktas med ett sunt mått av skepsis. Men även om denna genomsnittseffekt är svårtolkad, kan det vara intressant att titta på den relativa avkastningen i olika dimensioner snarare än avkastningens nivå. Den relativa avkastningen kan ges en tolkning under förutsättning att det selektionsproblem som noterades ovan är lika stort i de grupper vars avkastning jämförs.

Det finns en tendens i skattningarna som tyder på att kvinnor har något högre avkastning än män, vilket ligger i linje med Ekström (2003). Resultaten tyder även på att yrkesinriktade studieval (särskilt vårdutbildningar) hade större avkastning än matematik, svenska och engelska, vilket är ett resultat som ligger i linje med resultaten i Jacobson m.fl. (2005).

Vad säger då studier av kunskapslyftet? En kort sammanfattning är att det beror på vilka jämförelser som görs och vilka grupper som studeras. I två uppsatser jämförde Stenberg deltagarna i kunskapslyftet med deltagare i arbetsmarknadsutbildning (AMU); se Stenberg (2005, 2007a). Argumentet för att

⁸⁸ Detta är alltså ett problem som uppkommer även om man är beredd att acceptera antagandet att selektionen in i Komvux styrs av observerade egenskaper.

göra denna jämförelse är att kunskapslyftet har stora likheter med en arbetsmarknadspolitisk åtgärd, vilket kan göra antagandet om att bara observerade egenskaper styr deltagande i kunskapslyftet (eller AMU) mer plausibelt. Även om det ligger något i denna argumentation är det tveksamt om jämförelsen med AMU mildrar selektionsproblemet. Eftersom AMU utgörs av förhållandevis korta omskolningskurser kan man lika gärna hävda att det är individer med förhållandevis goda förutsättningar som går till AMU eftersom de "bara" är i behov av en kort omskolning.

Resultaten – framförallt Stenberg (2007a) men även i viss mån Stenberg (2005) – tyder på Kunskapslyftet hade negativa effekter i förhållande till AMU. Deltagarna i Kunskapslyftet hade inkomster efter utbildning som var avgjort lägre än deltagarna i AMU: för män var de i genomsnitt 28 % lägre och för kvinnor i genomsnitt 11 % lägre. När man tolkar dessa skillnader bör man också hålla i åtanke att AMU består av yrkesinriktade kurser, vilka visat sig ha större avkastning än allmänteoretiska ämnen, även inom komvux.

En ambitiösare ansats är att jämföra deltagare i kunskapslyftet med likartade individer som var registrerade som arbetslösa vid motsvarande tidpunkt. Detta görs i Albrecht m.fl. (2004) och Stenberg och Westerlund (2008). Populationerna som studeras skiljer sig dock åt: Albrecht m.fl. (2004) studerar alla deltagare i kunskapslyftet, medan Stenberg och Westerlund särskilt fokuserar på individer som varit arbetslösa under väldigt lång tid (minst två år).

Den övergripande bilden i Albrecht m.fl. (2004) är att det inte finns effekter av kunskapslyftet på individers inkomster, även om denna slutsats delvis drivs av att precisionen i skattningarna är låg. Stenberg och Westerlund (2008) fann dock positiva och förhållandevis stora effekter av kunskapslyftet: effekten uppgick till en ökning av inkomsten med 14 % för män och 23 % för kvinnor. Stenberg och Westerlund tolkade skillnaderna i resultat jämfört med Albrecht m.fl. (2004) som att de positivare effekterna framförallt har att göra med att de fokuserar på riktigt långtidsarbetslösa.⁸⁹ Denna tolkning är inte orimlig och ligger i linje med litteraturen som rör arbetsmarknadspolitiska program (se Calmfors m.fl. 2002). Program behöver vara tydligt riktade till individer med störst behov för att de ska ha positiva effekter för deltagande individer. Det är avsevärt svårare att observera positiva effekter av breda satsningar.

⁸⁹ Det finns även vissa skillnader i metod som kan bidra till att förklara skillnaderna i resultat. Båda studier använde en s.k. matchningsmetod. Albrecht m.fl. (2004) tog i matchningen hänsyn till individers tidigare inkomster, vilket också Ekström (2003) gjorde. Stenberg och Westerlund (2008) betingade förvisso på att individerna ska ha varit arbetslösa under drygt 1,5 år innan men tar i övrigt inte hänsyn till tidigare inkomster.

Även om Stenberg och Westerlunds (2008) tolkning är rimlig bör man notera att deras empiriska ansats är behäftad med ett potentiellt problem. De betingade på att individerna i jämförelsegruppen ska ha varit fortsatt arbetslösa under den första terminen som komvuxgruppen deltog i utbildning. Detta leder antagligen till en överskattning av effekterna, eftersom jämförelsegruppen – i större utsträckning än behandlingsgruppen – blir negativt selekterad över tiden.

Vilka övergripande slutsatser kan man då dra utifrån de studier som gjorts av vuxenutbildningen? En uppenbar slutsats är att det saknas studier med en övertygande empirisk design. Alla studier på området är studier av observationskaraktär där antagandet är att det räcker med observerade egenskaper för att i sanning göra individerna jämförbara. De skattade nivåerna på avkastningen måste därför betraktas med viss försiktighet; se kapitel 5.

Några generella slutsatser tycker vi det ändå går att dra från studierna. Vuxenutbildning för individer med svagare förutsättningar verkar ha en högre avkastning än breda satsningar. Detta är resultatet i Stenberg och Westerlund (2008) och det är också ett resultat i Jacobson m.fl. (2005a). Resultaten i båda svenska och internationella studier (se även Gill och Leigh 2003) tyder på att avkastningen på yrkesinriktade kurser är högre än avkastning än på studier i teoretiska ämnen. Denna slutsats ligger också i linje med resultaten i Hall (2009) under det rimliga antagandet att individer som går till komvux har lite sämre studieförutsättningar. Till sist tycks avkastningen på komvux vara högre för kvinnor än för män.

9.5 Diskussion och slutsatser

Vilka allmänna slutsatser kan vi dra utifrån det material som presenterats i detta kapitel? En allmän slutsats är att ett större inslag av yrkesutbildning gynnar elever från studiesvagare miljöer. Både svenska och internationella studier på gymnasienivå (eller motsvarande) visar att så är fallet. Litteraturen tyder även på yrkesinriktade vuxenutbildningar har högst avkastning.

En annan allmän slutsats är att en specialisering vid 16 års ålder inte är orimlig. De svenska erfarenheterna av införandet av det programgymnasiet är relevanta här. Denna reform minskade graden av specialisering eftersom ett tredje år av huvudsakligen teoretiska studier lades till yrkesutbildningarna. Som en konsekvens av reformen gick examensfrekvensen ner, särskilt bland dem med sämre studieförutsättningar, och det går inte att spåra några effekter på vare sig övergångsfrekvensen till högre studier och arbetsinkomster (trots ett ytterligare skolår).

En aspekt av införandet av programgymnasiet som vi inte analyserat är att reformen även hade konsekvenser för de studieförebereðande inriktningarna. Ett led i reformen var att de grundläggande ämneskurserna (t.ex. Matematik A) blev gemensamma för samtliga program. Detta kan ha lett till att undervisningen för elever på de teoretiska programmen är sämre anpassade till deras förutsättningar (se kapitel 8), vilket i sin tur kan vara en bidragande orsak till den dramatiska resultatnedgång som kan observeras i matematik och naturvetenskap i de internationella mätningarna (se Skolverket 2009). Det finns dock ingen forskning som direkt studerat denna fråga och vi har därför valt att lämna den därhän.

En uppfattning om effekterna av specialisering vid olika åldrar påverkas också av möjligheterna till utbildning vid vuxen ålder. I Sverige finns förhållandevis goda möjligheter till s.k. andrachansutbildning. När möjligheterna till en andra chans är goda, är kostnaderna som är förenande med felaktiga utbildningsval relativt små. Här ska man också komma ihåg att denna kostnad ska ställas mot den avsevärt större kostnaden som följer av att ställa större studiekrav på det stora flertalet yrkeselever som inte är intresserade av högskoleutbildning.

Den mest slående utvecklingen i gymnasieskolan under 1990-talet är nedgången i examensfrekvensen. Bland elever födda 1972–74 som påbörjade en 2-årig yrkeslinje var det 80–82 % som tog sin examen på normalt tid. Motsvarande siffra för elever födda 1982–84 som påbörjade ett 3-årigt yrkesprogram är 62–64 %. I runda tal minskade examensfrekvensen med 18 procentenheter.

En del av förklaringen ligger i övergången från 2-åriga till 3-åriga yrkesutbildningar. Men den stora förklaringen är övergången från ett relativt till ett absolut betygssystem. Till viss del är nedgången per konstruktion: medan närvaro var tillräcklig för ett fullständigt betyg i det gamla systemet, krävs en viss studieprestation för ett slutbetyg i meritvärdessystemet. Men den stora frågan är vilka signaler ett icke-godkänt betyg skickar till presumtiva arbetsgivare och vilka kostnaderna är av tiden till examen förlängs. För att svara på dessa frågor har vi jämfört elever som är födda olika år men har samma grundskolebetyg. Vår analys visar att meritvärdessystemet försämrade arbetsmarknadsutsikterna för yrkeseleverna. Även denna reform slog särskilt hårt mot elever med sämre studieförutsättningar.

Tanken är att meritvärdessystemet ska fylla tre olika funktioner. För det första ska det rangordna elever för antagning till vidare studier. För det andra ska det mäta kunskapsutvecklingen över tid. För det tredje ska det mäta skolans

måluppfyllelse. Sunt förnuft (och elementär analys) säger att ett instrument aldrig kan fylla alla tre olika funktioner på ett bra sätt. I själva verket skulle vi vilja påstå att meritvärdessystemet inte fyller någon av dessa funktioner på ett bra sätt. Vi kommer att återkomma till en diskussion om konstruktionen av betygssystemet i det avslutande kapitel 11.

Referenser

- Albrecht, J., G. van den Berg, och S. Vroman (2004), The Knowledge Lift: The Swedish adult education program that aimed to eliminate low worker skill levels, IFAU Working Paper 2004:17.
- Calmfors, L., A. Forslund, och M. Hemström (2002), Vad vet vi om den svenska arbetsmarknadspolitikens sysselsättningseffekter? IFAU rapport 2002:8.
- Ekström, E. (2002), Värdet av en treårig yrkesinriktad gymnasieutbildning, IFAU rapport 2002:12.
- Ekström, E. (2003), Inkomsteffekter av kommunal vuxenutbildning, IFAU rapport 2003:13.
- Erikson, R., O. Nordström Skans, A. Sjögren och O. Åslund (2007), Ungdomars och invandrades inträde på arbetsmarknaden 1985–2003, IFAU-rapport 2007:18.
- Hall, C. (2009), Förlängningen av yrkesutbildningarna på gymnasiet: effekter på utbildningsavhopp, utbildningsnivå och inkomster, IFAU rapport 2009:7.
- Hanushek, E. och L. Wössmann (2006), Does educational tracking affect performance and inequality? Differences-in-differences evidence across countries, *Economic Journal*, 116, C63–C76.
- Jacobson, L., R. Lalonde och D. Sullivan (2005), The impact of community college retraining on older displaced workers: Should we teach old dogs new tricks? *Industrial and Labor Relations Review*, 58, 398–415.
- Kang, S. och J. Bishop, (1989), Vocational and academic education in high school: Complements or substitutes?, *Economics of Education Review*, 8, 133–148.
- Leigh, D. och A. Gill (1997), Labor market returns to community colleges: Evidence for returning adults, *Journal of Human Resources*, 22, 334–353.
- Gill, A. och D. Leigh (2003), Do the returns to community college differ between academic and vocational programs?, *Journal of Human Resources*, 28, 134–155.
- Light, A. (1999), High school employment, high school curriculum, and post-school wages, *Economics of Education Review*, 18, 291–309.

- Malamud, O. och C. Pop-Eleches (2008), School tracking and access to higher education among disadvantaged groups, Harris School Working Paper #08.10, University of Chicago.
- OECD (2005), *Promoting Adult Learning*. Paris: Organisation for Economic Co-operation and Development.
- Pekkarinen, T., R. Uusitalo och S. Kerr (2009), School tracking and development of cognitive skills, IFAU Working Paper 2009:6.
- Skolverket (2008), Studieresultat i gymnasieskolan – en statistisk beskrivning av ofullständiga gymnasiestudier, *Skolverkets aktuella analyser 2008*, Skolverket.
- Skolverket (2009) TIMSS Advanced 2008, Rapport 336. Stockholm: Fritzes.
- SOU 2008:27, *Framtidsvägen – en reformerad gymnasieskola*, Betänkande av Gymnasieutredningen. Stockholm: Fritzes.
- Stenberg, A. (2005) Adult education for the unemployed: The experience of the Swedish adult education initiative, *Labour*, 19, 123–146
- Stenberg, A. (2007a), Comprehensive education or vocational training for the unemployed? *International Journal of Manpower*, 28, 42–61.
- Stenberg, A. (2007b), Hur påverkar gymnasialt komvux löneinkomster och vidare studier? IFAU rapport 2007:7.
- Stenberg, A. och O. Westerlund (2008), Does comprehensive education work for the long term unemployed? *Labour Economics*, 15, 54–67.

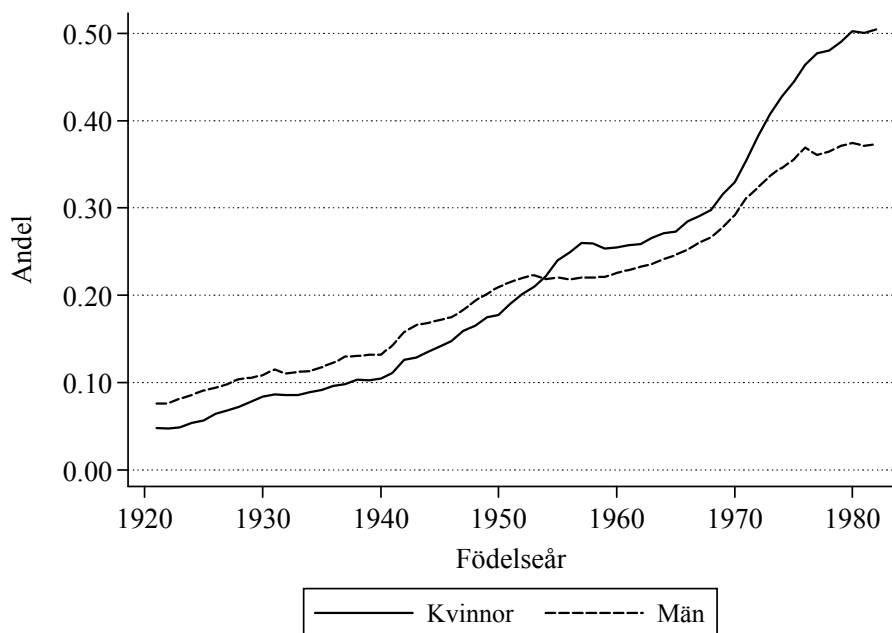
10 Högre studier

Högskoleutbildningen har byggts ut kraftigt under de senaste 50 åren, både vad gäller antalet högskolor och antalet studieplatser. Samtidigt har utbyggnaden haft en tydlig regional prägel. Expansionen av den högre utbildningen har drivits på av arbetsmarknadens ökande behov av välutbildad arbetskraft såväl som av ett tilltagande studieintresse bland ungdomar i takt med att alltfler uppnått grundläggande behörighet. Det har också funnits andra politiska motiv till utbyggnaden, såsom breddning av rekryteringsbasen till högskolan liksom stimulering av den regionala tillväxten.

Utbyggnaden av den högre utbildningen har skett dels genom etablering av nya lärosäten, dels genom en utökning av antalet platser på existerande högskolor. I mitten på 1940-talet bedrevs högre utbildning vid universiteten i Uppsala och Lund, vid högskolorna i Stockholm och Göteborg samt vid några fackhögskolor inom medicin, ekonomi och teknik. Idag finns 16 universitet och 16 högskolor som erbjuder högre utbildning inom olika ämnesområden. Därtill kommer 29 andra utbildningsanordnare som har rätt att utfärda examina på grundutbildningsnivå. Den regionala spridningen av högre utbildning är stor och idag finns minst en högskola i varje län.

Expansionen av högskoleutbildningen under det senaste halvsekle illustreras kanske bäst med förändringen av antalet studieplatser. Figur 10.1 visar andelen män och kvinnor födda olika år som påbörjade högskolestudier före de fyllt 26 år. Omkring 7–8 % av personer födda på 1920-talet påbörjade en universitetsutbildning. Flertalet högskolestuderande var också män. Under de efterföljande årtiondena skedde en kontinuerlig utbyggnad av högskoleutbildningen, med ökande studenttillströmning som följd. Drygt 15 % av personer födda på 1940-talet gick på högskolan. Högskolestudier var dock fortfarande vanligast bland män. Detta kom att ändras från och med personer födda på 1950-talet, vilket delvis kan förklaras av att flera kvinnodominerade eftergymnasiala utbildningar kom att klassas som högskoleutbildningar i och med högskolereformen 1977. Den kraftiga utbyggnaden av högskolan under 1990-talet kom framförallt personer födda 1970-talet till del. Drygt hälften av alla kvinnor födda i början på 1980-talet har påbörjat högre studier före 26 års ålder. Motsvarande siffra för män är knappt 40 %. Notera också att den positiva trenden i andelen högskolenybjörjare planar ut i slutet av perioden, framförallt för män.

Figur 10.1. Andel högskolenybörjare före 26 år, födda 1920–82

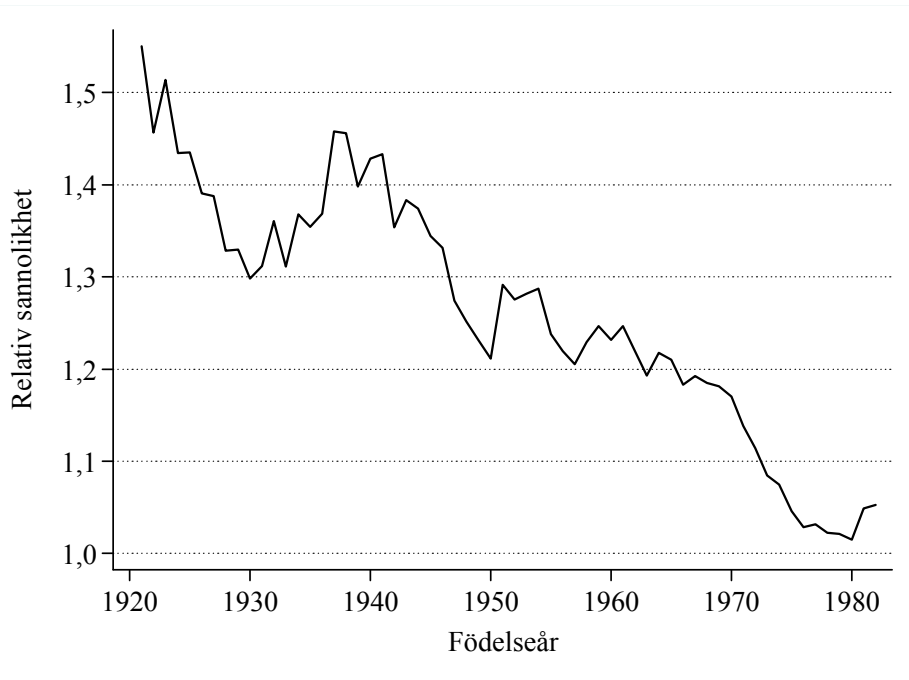


Not: Figuren baseras på uppgifter om tidpunkt för första registrering på högskolan 1977–2007, indikator på om personen varit registrerad på högskolan någon gång under perioden 1956–76 samt tidpunkt för högskoleexamen (om sådan finns). Uppgifterna har i allmänhet högre kvalitet för personer födda 1951 eller senare. För att verifiera uppgifterna för tidigare kohorter har jämförelser gjorts med uppgifter om examen enligt FoB 1970. Skattningen av andelen som påbörjat en högskoleutbildning före 26 års ålder för personer födda på 1930-talet, stämmer då väl överens med andelen av dem som har en högskoleexamen enligt FoB 1970.

Som beskrivits ovan har utbyggnaden av högskolan har haft en tydlig regional prägel sedan 1960-talet. Tanken har bland annat varit att stimulera regional tillväxt samt att bredda urvalet till högskolan. Den ökade regionaliseringen av högskoleutbildningen tycks också ha minskat skillnaden i sannolikheten att läsa vidare för personer födda i olika delar av landet. Figur 10.2 visar den relativa sannolikheten att börja läsa på högskolan före 26 års ålder för personer födda i något av de gamla universitetslänerna (Stockholms län, Uppsala län, Malmöhus län samt Göteborgs och Bohus län) i förhållande till personer födda i övriga län. Personer födda i ett universitetslän i början av 1920-talet hade mer än 50 procents större sannolikhet att påbörja högskolestudier före 26 års ålder än personer födda i andra län. I och med utbyggnaden av högskolan, och med tilltagande grad av regionalisering, kom dessa skillnader att minska

kontinuerligt för personer födda efter 1940. För personer födda i början av 1980-talet finns nästan inga skillnader i övergångssannolikheter till högskolan för personer födda i de olika länstyperna. Den tilltagande regionaliseringen tycks därmed ha verkat utjämnande i möjligheterna att läsa vidare för personer födda i olika delar av landet.

Figur 10.2. Relativ sannolikhet att börja på högskolan före 26 år för personer födda i universitetslän respektive i andra län



Not: Figuren visar relativ andel högskolenybjörjare för 26 års ålder födda i gamla universitetslän (Stockholms län, Uppsala län, Malmöhus län samt Göteborgs och Bohus län) respektive andra län i förhållande till befolkningen i sin helhet.

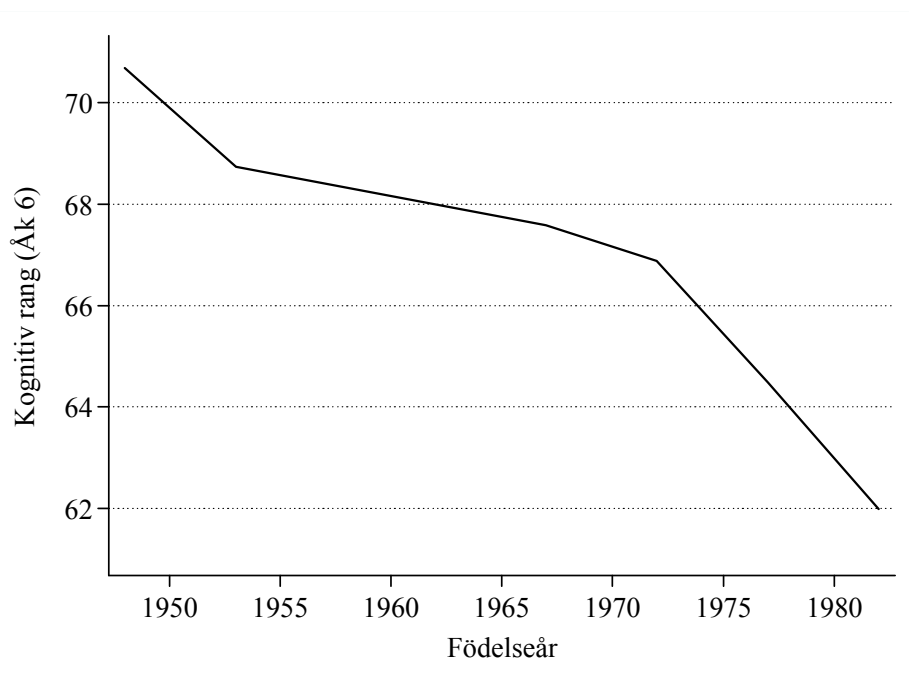
Ett syfte med den kraftiga expansionen av den högre utbildningen som skett under de senaste 30–40 åren, har varit att öppna upp högskolan för grupper som traditionellt sett varit underrepresenterade. Tanken var att underrepresenterade grupper – framförallt barn från hem utan studievana – skulle kunna utgöra en underutnyttjad ”begåvningsreserv” (Husén 1951). En utökning av antalet platser på högskolan skulle därmed vara ett sätt att locka dessa grupper att läsa vidare. Mot denna tanke står en mer mekanisk effekt, där en utökning av

antalet platser bara lockar personer med relativt sämre kognitiva förutsättningar till högskolan.

Den kraftiga utbyggnaden av högskolan har mycket riktigt medfört sämre ”studentkvalitet”. Figur 10.3 visar genomsnittlig förmåga i åk 6 för högskolestuderenter födda 1948–82.

Högskolestuderenter födda i slutet av 1940-talet och i början av 1950-talet var en starkt selekterad grupp. I genomsnitt kom de från den 70:e percentilen i förmågefördelningen i åk 6. Under den kontinuerliga utbyggnaden av högskolan under 1970- och 1980-talen skedde en långsam försämring av den genomsnittlige studenten. I samband med den kraftiga utbyggnaden under 1990-talet kom den genomsnittlige studentens förkunskaper försämrats betydligt. Högskolestuderenter födda i början av 1980-talet kommer i genomsnitt 10 procentenheter längre ned i fördelningen av förmågor än personer födda i slutet av 1940-talet. Den försämrade studentkvaliteten skapar förstås nya utmaningar för högskoleutbildningen.

Figur 10.3. Genomsnittlig kognitiv rang för högskolenyborjare före 26 år



Som illustrerats inledningsvis har den kraftiga utbyggnaden av högskolan under framförallt senare delen av 1900-talet i många avseenden förändrat förutsättningarna för högskolornas verksamhet. I det här kapitlet försöker vi att lyfta fram några av de viktigaste utbildningspolitiska frågorna – och den därtill relaterade forskningen – för den högre utbildningen. Avsnitt 10.1 behandlar högskolornas utbud av studieplatser; principiella utgångspunkter för högskolans dimensionering, hur antalet platser bestäms i praktiken samt vilka effekter eventuella avvikelser från den optimala planen har för individen och samhället (s.k. överutbildning). Avsnitt 10.2 tar upp studenternas efterfrågan på de studieplatser som högskolorna erbjuder. Där behandlas betydelsen av ekonomiska incitament för individers utbildningsval samt varför vissa grupper (t.ex. barn till lågutbildade föräldrar samt män) är underrepresenterade på högskolan. Betydelsen av studiemedelssystemet för att överbrygga eventuella ekonomiska begränsningar för högskolestudenter belyses särskilt.

Avsnitt 10.3 belyser fördelningen av högskolornas utbildningsplatser bland de sökande studenterna, dvs. antagningssystemets utformning; särskild vikt läggs vid hur väl olika urvalsinstrument (gymnasiebetyg, högskoleprov samt arbetslivserfarenhet) förutsäger studieframgång och olika arbetsmarknadsutfall. Avsnitt 10.4 behandlar utbildningskvalitet, dvs. om löneskillnader mellan personer som läst vid olika lärosäten kan förklaras av skillnader i utbildningarnas kvalitet. Skillnader i avkastning mellan olika utbildningsinriktningar analyseras i avsnitt 10.5. Det sista avsnittet sammanfattar kapitlet.

10.1 Högskolornas utbud av studieplatser

Den ökande efterfrågan på högutbildad personal på arbetsmarknaden i kombination med ett högt söktryck till högskolan aktualiserar frågan om hur den högre utbildningen ska dimensioneras och inriktas. I det här avsnittet diskuterar vi några principiella utgångspunkter för den högre utbildningens dimensionering. Därtill gör vi en kort beskrivning av hur dimensioneringen sker i praktiken. Slutligen diskuterar vi vilka konsekvenser eventuella skillnader mellan den ”optimala” och verkliga dimensioneringen får för individen och samhället (s.k. överutbildning).

10.1.1 Några principiella utgångspunkter

En rimlig utgångspunkt för högskolans dimensionering och inriktning är att försöka maximera samhällets samlade produktion av varor och tjänster. Detta förutsätter dock att man skapar sig en uppfattning om det framtida behovet av

olika typer av arbetskraft, något som självfallet är förknippat med en viss osäkerhet. Den samhällsekonomiska kalkylen måste sedan – precis som vid individers utbildningsbeslut – väga intäkterna av en given dimensionering med de därtill kopplade kostnaderna. De samhällsekonomiska kostnaderna av att tillhandahålla platser på högskolan består dels av de direkta resurser som undervisningen tar i anspråk (lärare, lokaler, etc.), dels av utebliven produktion för studenterna under studietiden. Intäkterna kommer sedan i form av ökad produktivitet på arbetsmarknaden efter utbildningen. Vid en optimal fördelning av platser med olika inriktningar kommer kostnaden för att inrätta ytterligare en högskoleplats att vara högre än intäkterna, liksom att intäkterna av att ta bort en plats kommer att överstiga kostnaderna.⁹⁰

10.1.2 Dimensioneringen i praktiken

I praktiken bestäms inte antalet högskoleplatser med olika inriktningen av en samhällsplanerare på det sätt som beskrivs ovan. I officiella dokument sägs högskolans dimensionering och inriktning istället fastställs utifrån två olika principer: studenternas önskemål eller arbetsmarknadens behov. Vilken av dessa som varit vägledande har varierat över tiden. Många gånger har antalet platser med olika inriktning bestämts av ytterligare en faktor: högskolepersonalens kompetens (Högskoleverket 1997).

Frågan är dock i vilken utsträckning studenternas efterfrågan eller prognoser om arbetsmarknadens behov kan tänkas motsvara den samhällsekonomiskt optimala dimensioneringen av högskolan? Studenternas önskemål är oftast relativt lätta att observera, om inte annat i antagningsprocessen. På grund av statliga ingripanden (frånvaro av studieavgifter, generösa studiemedel och progressivt skattesystem), samt möjlig förekomst av externa effekter, finns det emellertid anledning att misstänka att studenternas utbildningsval inte överensstämmer med en samhällsekonomiskt optimal platsfördelning. Det kan därför vara problematiskt att dimensionera utbildningen enbart med utgångspunkt från studenternas önskemål.

Arbetsmarknadens behov på kort sikt är också relativt enkelt att observera, exempelvis genom att studera arbetslöshet och vakanser inom olika yrken. Det långsiktiga behovet av välutbildad arbetskraft är förstås betydligt svårare att förutsäga. Här finns dock en del olika prognoser att tillgå, exempelvis från SCB

⁹⁰ Det samhällsekonomiskt optimala antalet högskoleplatser med olika inriktningar kommer sannolikt inte motsvara studenternas efterfrågan, eftersom individen vanligen inte bär hela kostnaden för sina utbildningsval (pga. avsaknaden av studieavgifter samt generösa studiemedel) och inte heller får ta del av alla intäkterna (pga. skattesystemet).

och Arbetsförmedlingen. Problemet är att förutse snabba förändringar av efterfrågan på personal inom och mellan olika yrken. Prognosproblemet ska dock inte överdrivas. Erikson m.fl. (2006) visar på en förbluffande stabilitet över tid i sannolikheten att snabbt få ett välbetalt jobb för olika studieinriktningar. De visar också att utbildningsinriktningar med goda arbetsmarknadsutsikter inte attraherar studenter med högre gymnasiebetyg än andra utbildningar, vilket skulle kunna tolkas som om varken högskolornas utbud eller studenternas efterfrågan på utbildningsplatser svarar mot arbetsmarknadens behov. Kunskapen om högskolornas dimensionering och dess effekter är dock bristfällig, och mer forskning på detta område vore önskvärd.

10.1.3 Konsekvenser av överutbildning

Det hävdats emellanåt att antalet utbildningsplatser generellt sett är för stort, eller att utbildningsinriktningarna har varit felaktigt dimensionerade (se t.ex. Korpi och Tåhlin 2009). Detta brukar i litteraturen benämnas överutbildning (eller underutbildning). Överutbildning är ett ganska otydligt begrepp. Vad som avses borde rimligtvis vara att det faktiska antalet högskoleplatser med olika inriktningar inte motsvarar den samhällsekonomiskt optimala dimensioneringen. En central fråga är dock hur överutbildning kan uppstå.

I ett utbildningssystem där antalet utbildningsplatser med olika inriktningar bestäms genom högskolornas utbud och studenternas efterfrågan (marknadsanpassat system), skulle över/underutbildning kunna uppstå om individer av olika anledningar gör val som avviker från den samhällsekonomiska planen. Kanske kan andra personer dra nytta av att en person utbildar sig (externa effekter), vilket inte beaktas i individens utbildningsbeslut. Det kan också finnas olika kreditbegränsningar som gör att personer med lågutbildade föräldrar har sämre möjligheter att låna till sin utbildning. Det är också möjligt att personer baserar sina beslut på felaktiga förväntningar om framtiden eller att de värderar välbefinnandet idag betydligt mer än välbefinnandet i morgon. Det är dock inte denna typ av "marknadsmislyckanden" som brukar avses när man talar om överutbildning.

Den kanske vanligaste förklaringen till överutbildning utgår från tanken att arbetsgivare har svårt att avgöra produktiviteten hos potentiella arbetstagare. Under vissa förutsättningar kan då högproduktiva arbetare finna det lönsamt att vidareutbilda sig bara för att signalera sin produktivitet, även om utbildningen i sig inte är produktiv. Under förutsättning att utbildningssignalen inte förbättrar matchningen mellan arbetstagare och arbetsgivare, utan bara påverkar vilka

som får högre lön, kommer en sådan utbildningsdimensionering att överstiga samhällets ”optimala plan” (överutbildning). Notera dock att modellen förutsätter att det stora flertalet av dem som utbildar sig får en lön som motsvarar deras högre produktivitet. Om många högskoleutbildade får en låg lön, blir arbetslösa eller hamnar i låglöneyrken – vilket är några av de utfall som litteraturen om överutbildning studerar – skulle det inte längre vara privat-ekonomiskt försvarbart att utbilda sig bara för att signalera sin produktivitet. En signalmodell skulle således kunna ge upphov till överutbildning, men däremot inte medföra de negativa konsekvenser som ofta förknippas med begreppet.⁹¹

I Sverige bestäms antalet platser på högskolan inte genom marknadsmekanismer utan av statsmakterna. Som diskuteras ovan är det inte helt klart vad som styr högskolornas dimensionering. Det verkar dock inte orimligt att beslutsfattare har andra målsättningar med dimensioneringspolitiken än att uppnå största effektivitet i utbildningssystemet. En kraftig utbyggnad av högskolan skulle exempelvis kunna ses som ett medel för att ge människor med olika bakgrund samma möjligheter att läsa på högskolan. Ett alltför stort utbud av högskoleplatser är i allmänhet inte tillräckligt för att skapa överutbildning, eftersom individer inte förväntas utbilda sig längre än vad som är privatekonomiskt motiverat. I Sverige finns dock flera interventioner som stärker individernas incitament att läsa vidare, såsom frånvaro av studieavgifter och ett generöst studiemedelssystem. Detta ska dock vägas mot ett progressivt skattesystem som gör högskolestudier mindre lönsamma. Sammantaget är det emellertid möjligt att ett högt utbud av högskoleplatser i kombination med låga studiekostnader skulle kunna leda till överutbildning. Notera dock att de negativa konsekvenserna av överutbildning – såsom låg lön, hög arbetslöshet eller risk att hamna i låglöneyrke – inte kan vara alltför stora. Då skulle människors intresse för att utbilda sig minska och problemet med överutbildning avta.⁹²

Det finns en relativt stor internationell litteratur om överutbildning (se översikten av Hartog 2000). Ett problem med detta forskningsområde är att teorierna om överutbildning inte har några enkla testbara implikationer. Om överutbildning orsakas av att arbetstagare använder utbildning för att signalera sin produktivitet skulle vi förvänta oss att högutbildade har högre löner, lägre arbetslöshet och mindre risk att hamna i låglöneyrken än lågutbildade personer. Detta är också precis vad humankapitalteorin förutspår. På teoretiska grunder

⁹¹ Se också Card (2002) för en kritisk diskussion om överutbildning.

⁹² Vad vi förstår avser begreppet överutbildning vanligen en mer långvarig obalans mellan utbud och efterfrågan på arbetskraft och inte tillfälliga konjunktursvängningar. I praktiken är det dock svårt att skilja mellan lång- och kortsiktiga obalanser.

skulle vi således inte förvänta oss att överutbildning ska medföra lägre löner, ökad arbetslöshet och risk att hamna i ”fel” yrke. Tanken är istället att individer till viss del utbildar sig ”i onödan” för att signalera sin produktivitet till arbetsgivarna, något som är mycket svårt att testa empiriskt.

Om överutbildning istället orsakas av en kraftig utbyggnad av högskolan i kombination med frånvaro av studieavgifter och ett generöst studiemedelsystem, skulle högutbildade kunna ha relativt låga löner, en viss arbetslöshetsrisk och risk att hamna i låglöneyrken. Det är dock viktigt att poängtera att sådana utfall inte behöver vara tecken på överutbildning, utan kan bero på att utbyggnaden av högskoleutbildningen attraherar sämre studenter och medför lägre utbildningskvalitet (se nedan). Även om avkastningen väntas bli lägre när högskolan byggs ut, skulle det fortfarande kunna vara samhällsekonomiskt motiverat att utbilda dessa personer.

Empiriska studier av över/underutbildning analyserar vanligen dels sannolikheten att hamna i ”rätt” eller ”fel” yrke, dels avkastningen på utbildning för de som har hamnat rätt eller fel. Valet av utfall antyder att studierna utgår från att över/underutbildning orsakats av en alltför kraftig utbyggnad av högskolan i kombination med stimulans av individernas efterfrågan, eftersom en signalmodell inte förutsäger att andelen som hamnar i fel yrke skulle öka. Avsikten är att försöka mäta utbildning som över/understiger ”utbildningskraven” för ett visst yrke, antingen vad gäller utbildningsnivå eller inriktning. Problemen är dels att utbildningskraven för olika yrken oftast inte är väldefinierade, dels att många utbildningar är mer eller mindre generella till sin natur. Det är dessutom troligt att utbildningskraven för olika yrken skiftar med tiden eller över yrkeskarriären, i samband med att arbetsuppgifterna blir mer (eller mindre) specialiserade eller krävande.

Empiriska studier om överutbildning använder tre olika angreppssätt för att mäta överutbildning. Den första metoden baseras på analyser av arbetsbeskrivningar för olika yrken för att därigenom avgöra vilka förkunskaper som krävs. Den andra ansatsen utgår från frågor till arbetstagarna, där de får ange om de har en utbildning som motsvarar deras yrke eller inte. Det tredje angreppssättet baseras på analyser av den vanligaste utbildningen inom ett yrke, och låter dem som har en avvikande utbildning definieras som över/underutbildade. Dessa ansatser är problematiska eftersom de dels rent mekaniskt kommer att definiera vissa individer som över/underutbildade, dels bygger på antagande om stabila utbildningskrav över tiden/yrkeslivet. Groot och Maassen van der Brink (2000)

visar dessutom att resultaten är känsliga för val av metod, och att individer oftare definieras som överutbildade om man använder enkätfrågor.

Le Grand m.fl. (2001, 2004) och Åberg (2003) visade att överutbildningen i Sverige har ökat under perioden 1974–2000. I mitten av 1970-talet bedömdes omkring 14 % av arbetskraften vara överkvalificerade för sitt yrke. Motsvarande siffra för 2000 var nästan 33 %. Andelen som upplever att de är underkvalificerade har legat relativt konstant kring 17 %. De fann också att avkastningen för överutbildade personer var lägre än för de som var rätt matchade.⁹³ Korpi och Tåhlin (2009) visade att antalet utbildningsår som överstiger yrkeskraven har stigit från omkring 0,5 år 1974 till 1,3 år 2000. Antalet utbildningsår som understiger yrkeskraven har dock legat konstant på omkring 0,5 år. De fann också att avkastningen på utbildning utöver yrkeskraven (överutbildning) var lägre än för utbildning som motsvarade kraven, medan avkastningen på krävd utbildning som inte uppnåts (underutbildning) var negativ. De fann dock inte några större skillnader i löneutveckling mellan rätt matchad utbildning och överutbildning.

En förklaring till de negativa konsekvenserna av överutbildning skulle kunna vara att det fångar andra skillnader mellan individer med olika bra matchning. Nordin m.fl. (2008) studerade överutbildning bland högskoleutbildade personer födda 1964–75. De definierade överutbildning genom att studera de vanligaste utbildningarna inom olika yrken, och fann att år 2003 var omkring 60–70 % av alla arbetstagare matchade till rätt yrke, 20 % hade en ”svag” match och 10–15 % var felmatchade. Personer utan examen löpte en större risk att klassificeras som felmatchade.

De fann också att personer som klassificerats som felmatchade i genomsnitt hade ungefär 30 % lägre lön än rättmatchade personer. Detta motsvarar en löneeffekt på 3–5 års högskolestudier. För att undersöka om effekterna skulle kunna förklaras av andra skillnader mellan personer med olika matchningsstatus, kontrollerade de för resultatet på inskrivningsprovet vid den militära mönstringen (män). Detta hade dock bara en mindre inverkan på resultatet. De lägre lönerna för ”felmatchade” personer kan således inte förklaras av skillnader i kognitiv förmåga före högskolestudierna. Det utesluter dock inte att det kan bero på skillnader i exempelvis icke-kognitiva förmågor eller studieresultat på högskolan.

⁹³ Johansson och Katz (2007) analyserar överutbildningens betydelse för lönegapet mellan män och kvinnor. De fann att kvinnor har större sannolikhet att klassificeras som överutbildade, vilket också delvis förklarar en del av löneskillnaderna mellan män och kvinnor.

Vi ser problem med litteraturen om överutbildning. Det saknas ofta en klar teoretisk motivering till varför överutbildning skulle uppstå. Vi har diskuterat två olika tänkbara mekanismer: signalering av produktivitet eller en politiskt medveten strategi för att bredda urvalet till högskolan. Dessa teorier ger dock helt olika implikationer för empiriska studier om överutbildning. Om arbetstagare använder utbildning för att signalera sin produktivitet, borde överutbildning inte medföra en ökning av ogynnsamma arbetsmarknadsutfall. En sådan förklaringsmodell kan därför inte motivera studier av andelen som hamnar i rätt eller fel yrke. Om beslutsfattare medvetet försöker öka utbudet på utbildning, skulle överutbildning kunna medföra lägre löner, högre arbetslöshet och större sannolikhet att hamna i fel yrke. Kunskaper om konsekvenserna av sådan utbildningspolitik skulle förstås vara mycket värdefull.

I det fall man skulle vilja studera förekomsten av överutbildning är det inte uppenbart att andelen som hamnar i rätt eller fel yrke är det utfall som är av primärt intresse. Om utbudet av högskoleutbildad arbetskraft är för högt eller har fel inriktning, borde detta först och främst resultera i högre arbetslöshet och lägre löner.⁹⁴ Detta fångas upp i den typ av avkastningsmått som diskuteras i kapitel 4. Det är värt att poängtera att förekomsten av överutbildning – såsom den uppmäts i litteraturen – har ökat sedan 1980-talet (Korpi och Tåhlin 2009; Le Grand och Tåhlin 2001, 2004; Åberg 2003). Detta har skett samtidigt som avkastningen på högskolestudier ökat.

Att avkastningen på högskoleutbildning stigit samtidigt som uppmätt överutbildning ökat, talar för att överutbildning inte är ett stort problem för samhällsekonomin. Detta utesluter dock inte att överutbildning är ett problem för de personer som ”drabbas”, dvs. för dem som har ett yrke vars utbildningskrav inte motsvarar deras formella kompetens. Empiriskt är det emellertid svårt att mäta effekterna av att vara över/underutbildad, eftersom dessa personer kan ha egenskaper som skiljer dem från andra arbetstagare. Studien av Nordin m.fl. (2008) visade att de som kategoriserades som över/underutbildade inte hade avsevärt sämre (eller bättre) kognitiva egenskaper än andra arbetstagare. Det är dock möjligt att de skiljer sig vad gäller andra egenskaper, såsom icke-kognitiv förmåga eller motivation. Vi ser de empiriska utmaningarna förenade med att skatta kausala effekter av att vara över/underutbildad som betydande.

⁹⁴ Risken att hamna i fel yrke skulle kunna vara en konsekvens av överutbildning. Det kan dock också fånga upp många andra fenomen, exempelvis att många utbildningar förmedlar generella kunskaper eller att arbetsinnehållet i olika yrken förändras över tiden.

10.2 Studenternas efterfrågan på studieplatser

Vi har tidigare liknat individers utbildningsval vid en investeringskalkyl, där intäkterna av ett visst alternativ vägs mot dess kostnader. I idealfallet kommer individernas utbildningsval också att motsvara samhällets önskade dimensionering. Det kan dock finnas olika typer av marknadsmisslyckande som gör att individer inte väljer den utbildning som ger störst avkastning, varken för dem själva eller för samhället. Ett problem som diskuteras i utbildningslitteraturen är förekomsten av olika typer av kreditbegränsningar.

Studenter är en ekonomiskt svag grupp. Dessutom uppkommer kostnaderna för högskolestudier tidigt i livet medan intäkterna kommer långt senare. I de flesta länder förekommer begränsningar på kreditmarknaden, vilket gör det svårt för enskilda individer att låna pengar till sina studier inom det vanliga banksystemet. Därmed är många studenter hänvisade till sina föräldrar för sin försörjning. Eftersom lågutbildade föräldrar i många fall saknar möjlighet att låna till sina barns studier, finns en uppenbar risk att selektionen till högskoleutbildning delvis kommer att baseras på ekonomiska förutsättningar snarare än på studiebegävning.⁹⁵

Idag har många länder mer eller mindre utvecklade studiemedelssystem som ska hjälpa studenterna med inkomstöverföring över livscykel. I Sverige introducerades ett heltäckande studiemedelssystem 1965. Därmed borde få studenter vara kreditbegränsade i ordets egentliga mening. Trots det finns stora skillnader i sannolikheten att läsa vidare på högskolan – även vid samma skolprestationer – för personer med olika familjebakgrund. Betydelsen av föräldrarnas utbildningsnivå för barnens utbildningsdeltagande brukar benämnas intergenerationellt samband i den ekonomiska litteraturen (se t.ex. Holmlund m.fl. 2008), medan den sociologiska litteraturen använder man begreppet social snedrekrytering (se t.ex. Erikson och Jonsson 2002).

Figur 10.4 visar den relativa andelen högskolenyborjare födda 1960–82 som har högutbildade respektive lågutbildade föräldrar, i förhållande deras andel i befolkningen. För personer födda på 1960-talet var det 2,5 gånger mer sannolikt att en person med välutbildade föräldrar började läsa på högskolan än en person med lågutbildade föräldrar. Skillnaderna avtar något för personer födda

⁹⁵ Numera har dock innebörden av begreppet kreditbegränsningar vidgats något, och avser också andra skillnader i utbildningsdeltagande mellan personer som växt upp i familjer med olika ekonomiska förutsättningar. Carneiro och Heckman (2003) skiljer på kortsiktiga kreditbegränsningar som avser problem för ekonomiskt svaga familjer att finansiera barnens studier, och långsiktiga begränsningar som speglar fattiga familjers brist på kunskapsfrämjande miljö och preferenser.

i början av 1970-talet, vilket sammanfaller med den kraftiga utbyggnaden och regionaliseringen av högskolan på 1990-talet. Sannolikheten att barn till högtbildade föräldrar födda i slutet av 1970-talet skulle börja läsa på högskolan var ungefär dubbelt så stor som för barn från studieovana hem. Den sociala selektionen till högskolan är således fortfarande betydande. Eftersom Sverige under hela denna period haft ett väl fungerande studiemedelsystem borde dessa skillnader i studiedeltagande med avseende på föräldrabakgrund inte kunna beror på skillnader i möjligheten att finansiera sina studier (kreditrestriktioner).

En annan grupp som historiskt sett varit underrepresenterade på högskolan är kvinnor. Det framgår av Figur 10.1 som visar den relativa andelen kvinnor och män födda 1925–80 som börjat på högskolan före 26 års ålder. För de som föddes i mitten på 1920-talet var det 50 % mer sannolikt att en pojke började på högskolan än att en flicka gjorde det. För personer födda på 1930- och 1940-talen skedde en långsam ökning av sannolikheten för kvinnor att läsa vidare. Den stora förändringen kom i och med utbyggnaden av högskolan på 1970-talet, vilket delvis kan förklaras av att eftergymnasiala utbildningar inom traditionella kvinnoyrken fick högskolestatus. Från och med personer födda i mitten på 1950-talet kom kvinnor att bli överrepresenterade på högskolan. Utvecklingen stannade upp något för 60-talisterna, men tog ny fart i samband med 1990-talets expansion av högskolan.

Figur 10.4. Relativ sannolikhet att börja på högskolan för barn till högutbildade respektive lågutbildade föräldrar



10.2.1 När uppträder selektion till högskolan?

Svensk högskolepolitik har sedan mitten av 1960-talet försökt stävja den sociala selektionen till högre studier. Politiken har bland annat innefattat införandet av ett heltäckande studiemedelsystem, utbyggnad och regionalisering av högskolan samt utformning av behörighets- och antagningsregler. Dessa förändringar har huvudsakligen varit avsedda att påverka gymnasieungdomars utbildningsval. Vi går nedan igenom litteraturen som behandlar effekter av dessa förändringar. För att veta vilken möjlighet olika reformer har att påverka den sociala (och könsmässiga) selektionen till högskolan, är det emellertid viktigt att först skapa sig en uppfattning om vid vilken ålder selektionen uppstår.

Studier av bl.a. Boalt (1947), Fägerlind (1975) samt Erikson och Jonsson (1993) visade att familjebakgrunden var viktig för att förklara både skolprestationer och studieval under hela utbildningskarriären för personer födda olika år mellan 1925 och 1967. I deras anda genomför vi därför en analys av i vilken utsträckning de samlade prestationerna och utbildningsvalen vid en viss ålder

kan förklara selektionen till högskolan. Tabell 10.1 utnyttjar UGU-materialet för att studera sannolikheten att börja läsa på högskolan före 26 års ålder för personer födda 1972. Genom att beakta studieprestationer och utbildningsval vid olika åldrar försöker vi illustrera när skillnader i övergångssannolikheter till högskolan för olika grupper tycks uppstå.⁹⁶

Den första kolumnen i Tabell 10.1 anger det beskrivande sambanden mellan föräldrarnas utbildningslängd, kön och sannolikheten att läsa vidare. För personer födda 1972 var sannolikheten att läsa vidare nästan 7 procentenheter högre för varje ytterligare skolår som föräldrarna hade. Skillnaden mellan pojkar och flickor var av samma storleksordning. Den andra kolumnen visar att omkring 30 % av den sociala selektionen till högskolan kan förklaras av skolprestationer vid 10 års ålder. Motsvarande andel för den könsmässiga selektionen är 15 %. Detta skulle kunna spegla allt från skillnader i ärftliga faktorer och uppväxtmiljö till olikheter i studiemotivation mellan grupperna. Bilden förändras inte nämnvärt när vi även beaktar studieprestationer vid 13 års ålder. När vi också tar hänsyn till skillnader i provresultat vid 16 års ålder kan vi dock förklara nästan hälften av den sociala selektionen till högskolan, men bara knappt en tredjedel av skillnaderna mellan kvinnor och män.

Den femte kolumnen i Tabell 10.1 visar att den resterande delen av de könsmässiga skillnaderna helt försvinner när vi också beaktar grundskolebetygen i årskurs 9. Den huvudsakliga orsaken till att kvinnor är överrepresenterade på högskolan är således att de i allmänhet har högre betyg i nian än män, givet provresultaten i årskurs 3, 6 och 9. Effekten på den sociala selektionen av att beakta grundskolebetyg är betydligt blygsammare, vilket tyder på att elever med olika familjebakgrund bedöms likartat vid betygsättningen. Kolumn sex beaktar skillnader i val av gymnasieprogram mellan olika grupper, liksom om eleverna avslutat gymnasiet överhuvudtaget. Detta förklarar ytterligare en del av den sociala selektionen till högskolan. I den sista kolumnen beaktas skillnader i gymnasiebetyg, vilket inte har någon större inverkan på andelen som läser vidare.

⁹⁶ Det är viktigt att poängtera att denna analys inte gör anspråk på att mäta ”kausala effekter” av kön eller familjebakgrund på sannolikheten att läsa på högskolan. Tanken är bara att beskriva hur övergångssannolikheter ser ut för personer som har samma skolprestationer och utbildningsval vid en viss ålder.

Tabell 10.1. Betydelsen av kön och familjebakgrund för sannolikheten att läsa på högskolan före 26 år

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Föräldrarnas skolår	0,067*** (0,003)	0,047*** (0,003)	0,044*** (0,002)	0,035*** (0,003)	0,029*** (0,003)	0,020*** (0,002)	0,018*** (0,002)
Kvinna	0,071*** (0,014)	0,060*** (0,013)	0,063*** (0,013)	0,048*** (0,012)	-0,017 (0,012)	-0,006 (0,013)	-0,006 (0,013)
<i>Kontrollvariabler:</i>							
- Provresultat i Åk 3		X	X	X	X	X	X
- Provresultat i Åk 6			X	X	X	X	X
- Provresultat i Åk 9				X	X	X	X
- Betyg i Åk 9					X	X	X
- Val av gymnasieprogram						X	X
- Betyg på gymnasiet							X
Antal personer	8 923	8 923	8 923	8 923	8 923	8 923	8 923

Noter: Tabellen visar regressionsskattningar av sambandet mellan olika egenskaper och sannolikheten att börja läsa på högskolan före 26 års ålder. Alla modeller kontrollerar för kommun, skola och klass i Åk 3. Resultaten baseras på ett urval av personer födda 1972. Föräldrarnas skolår avser utbildningslängden för den förälder som har uppnått den högsta utbildningsnivån. Provresultat i Åk 3 och Åk 6 utgörs av ämnesprov och begåvningsprov, medan provresultaten i Åk 9 bara utgörs av ämnesprov. Val av gymnasieprogram avser vilket gymnasieprogram eleven slutfört (om något). För att korrigera för eventuella mätfel i uppgiften om föräldrarnas utbildningslängd i utbildningsregistret, har uppgift om föräldrarnas utbildningslängd från en enkät till föräldrarna i UGU-undersökningen använts som instrument. *** anger att estimatet är signifikant skilt från noll på 1 % -nivån.

Sammantaget kan skillnader i skolprestationer, och i viss utsträckning även utbildningsval, från tredje klass till gymnasiet förklara nästan 75 % av den sociala och hela den könsmässiga snedrekryteringen. Omkring 30 % av skillnaderna i högskoleval för barn med olika uppväxtmiljö kan spåras tillbaks till skolresultaten i tredje klass. Ytterligare 25 % kan förklaras av prestationerna under resten av grundskoletiden. Val av gymnasieprogram manifesterar en del av den sociala selektionen, medan prestationerna på gymnasiet bara har marginell betydelse. Vad gäller skillnaden i övergångssannolikheter mellan pojkar och flickor, kan omkring en tredjedel hänföras till skillnader i provresultat under grundskoletiden medan två tredjedelar beror på skillnader i betygssättning givet provresultat.

Dessa resultat tyder på att politik som syftar till att bryta den sociala och könsmässiga selektionen antagligen har större chanser att lyckas om den inriktas mot elever som går i grundskolan än mot elever som avslutat gymnasiet.⁹⁷ Detta går också i linje med den teoretiska modell som diskuteras i Kapitel 3. Mycket av den politik som syftat till att motverka den sociala snedrekryteringen till högskolan har dock handlat om fler utbildningsplatser, etablering av regionala högskolor och universitet samt ändrade antagningsregler.

10.2.2 Studieavgifter och studiemedel

Historiskt sett har undervisningen vid svenska universitet och högskolor varit så gott som avgiftsfri.⁹⁸ Högskoleförordningen föreskriver att "[u]tbildningen vid högskolorna skall vara avgiftsfri för studenterna" (Högskoleförordningen 1993:100, 1 kap. 10 §). I praktiken utgör avgiftsfriheten för offentliga högskolor en prisrestriktion för offentligt drivna högskolor. Avgiftsfriheten påverkar utbildningsefterfrågan genom att minska studenternas utbildningskostnader. Därmed öppnas högskolan upp för grupper som saknar goda lånemöjligheter (i avsaknad av studiemedelssystem) eller som annars skulle ha funnit det privat-ekonomiskt olönsamt att läsa vidare. Prisrestriktionen påverkar dock inte bara omfattningen av studenternas efterfrågan utan även dess inriktning. Eftersom studenterna inte känner av utbildningarnas verkliga kostnad, finns risk att de söker sig till dyrare utbildningar än vad de annars skulle göra.

Studiemedelssystemet introducerades år 1965. Dess utformning har varierat över tiden, men det har alltid bestått av två delar: bidrag och lån. Tillsammans

⁹⁷ Notera dock att senare utbildningspolitiska insatser även kan påverka elevernas incitament att prestera väl redan i grundskolan.

⁹⁸ Vissa avgifter har förekommit, framför allt vid kommunal högskoleutbildning (SOU 1948:42).

bildar de studiemedlets totalbelopp. Studiebidraget är en generell subvention till studenterna, medan studielånet ska utjämna inkomster och kostnader över livscykeln. Därmed skapas förutsättningar för en utbildningsefterfrågan som inte är lika beroende av studenternas ekonomiska förutsättningar.

Studielånet är generellt men troligen är det främst personer med små ekonomiska resurser som väljer att låna. Det utgör därför ett potentiellt effektivt utbildningspolitiskt medel för att nå en minskad social skiktning till högre studier. En möjlig invändning mot lånesystemet är dock att viljan att skuldsätta sig kan skilja sig mellan olika grupper. Dohmen m.fl. (2006) visade exempelvis att barn till lågutbildade föräldrar var mindre villiga att ta risker både i allmänhet och när det gällde hypotetiska investeringsbeslut än barn till välutbildade föräldrar.

Till skillnad från studiebidraget måste studielånet betalas tillbaka. Före den 30 juni 2001 var återbetalningstakten relaterad till högskolestudenternas inkomst och skulden avskrevs vid pensionen. Tanken var att därmed minska individernas risktagande. Nackdelen med systemet var dock att studenterna blev mindre känsliga för utbildningarnas ekonomiska avkastning. Systemet innebar i praktiken en subventionering av utbildningar med låg ekonomisk avkastning. Från 1 juli 2001 är studielånet istället ett annuitetslån, vilket innebär att studenterna ska betala av en fast andel av lånet under en 25-årsperiod. Därmed skjuts också risktagandet för gjorda utbildningsval i större utsträckning över från samhället till individen.⁹⁹

Det finns en stor internationell litteratur om betydelsen av kreditbegränsningar för övergången till högre studier. En sammanställning av resultaten från studier med en tydlig strategi för att mäta orsakssamband tyder på att en ökning av studiestöden med omkring 10 000 kronor per år ökar övergången till högre studier med 4–14 procentenheter (McPherson and Shapiro 1991; Kane 1995; Dynarski 2000; Turner 2002; van der Klaauw 2002; Dynarski 2003). Även om dessa uppsatser inte uttryckligen studerar effekten av studiestöden på den sociala snedrekryteringen till högskolan, är det rimligtvis i första hand barn till föräldrar med begränsade ekonomiska resurser som påverkas av studiestöden. Man skulle därför kunna tolka resultaten som att de studiestöd som finns i USA tycks ha minskat den sociala selektionen till högre utbildning. Eftersom de institutionella förhållandena i USA skiljer sig mycket från de i Sverige är det svårt att överföra dessa resultat till Sverige.

⁹⁹ Det finns fortfarande möjlighet att få återbetalningen anpassad till inkomsternas storlek.

Nielsen m.fl. (2008) studerade effekterna av en studiemedelsreform i Danmark 1988. Före reformen var studiebidraget behovsprövat för studenter under 22 år; ju högre inkomst föräldrarna hade desto lägre blev bidraget. För äldre studenter var bidragets omfattning inte beroende av föräldrarnas inkomster. Efter reformen togs behovsprövningen bort för studenter som var 20–21 år. Dessutom höjdes nivån på bidraget. Förändringen av studiebidraget före respektive efter reformen var således störst för yngre studenter med rikare föräldrar, medan den var mindre för äldre studenter samt yngre studenter med fattigare föräldrar.

Nielsen m.fl. (2008) utnyttjade skillnaderna i bidragsförändringen för olika grupper för att mäta effekterna av studiebidraget på sannolikheten att läsa på högskolan (Difference-in-differences). De fann att en ökning av bidraget med omkring 10 000 kronor per år ökade inflödet till högskolan med omkring 2 %. En möjlig förklaring till de blygsamma effekterna av studiebidraget på sannolikheten att läsa vidare kan vara att reformen framförallt ökade bidragen för barn till rikare föräldrar. Det är troligt att de hade kunnat läsa på högskolan även med ett lägre bidrag. Dessutom var det danska studiemedelssystemet relativt generöst även före reformen.¹⁰⁰

Det finns några svenska studier som på olika sätt försökt mäta effekterna av studiemedelssystemet på studenternas utbildningsefterfrågan i allmänhet och den sociala selektionen i synnerhet. Fredriksson (1997) studerade hur olika faktorer påverkar efterfrågan på högre utbildning. Han använde aggregerade data över andelen högskolenyborjare i förhållande till gymnasieutbildade, nettolöner för gymnasie- och högskoleutbildade, arbetslöshet och studiemedel för perioden 1967–91. Analyserna begränsades till yngre män. Resultaten visade en stark samvariation mellan andelen högskolenyborjare och nettolöner. Högre löner för högskoleutbildade medförde ett ökat antal nyborjare medan det omvända gällde för högre löner för gymnasieutbildade. Ett mer gynnsamt studiemedelssystem tycktes också öka efterfrågan på utbildning. Ökad arbetslöshet för yngre personer liksom minskad arbetslöshet för högskoleutbildade ökade tillströmningen till högskolan. Resultaten är robusta för olika mått på andelen högskolenyborjare och för olika åldersgrupper.¹⁰¹

Erikson och Jonsson (1993) använde uppgifter från undersökningen om levnadsförhållanden (ULF) för att studera sambandet mellan andelen som tar

¹⁰⁰ Häkkinen och Uusitalo (2003) studerade effekten av en studiemedelsreform i Finland på tiden till examen, och fann mycket små effekter.

¹⁰¹ Bornmalm-Jardelöw (1988) finner också effekter av ekonomiska incitament på efterfrågan till högre utbildning. De är emellertid ganska små.

studiemedel och den sociala selektionen till högskolan för personer födda omkring 1900–70. Med hjälp av tidsserieanalys fann de att ju större andel som erhållit studiemedel vid 20 års ålder desto lägre tycks den sociala selektionen till högskolan vara. Även om de också tog hänsyn till variationer i arbetslöshet, så är det mycket möjligt att andelen med studiemedel fångar upp andra tidstrender under perioden. Dessa resultat måste därför tolkas med stor försiktighet.

Reuterberg och Svensson (1992) försökte uppskatta studiemedlens inverkan på den sociala selektionen till högskolan genom att ställa retrospektiva frågor till ett urval av högskolestudenter födda 1948, 1953 och 1963. Frågorna avsåg betydelsen av studiemedelsystemet för beslutet att satsa på en högskoleutbildning. Omkring 30 % av högskolestudenterna menade att de troligtvis inte skulle ha läst vidare om studiemedelsystemet inte funnits. Andelen som inte skulle ha läst på högskolan i frånvaro av studiemedel var dessutom högre bland personer från studieovana hem, vilket tolkades som att studiemedlen medfört minskad social snedrekrytering.

Reuterberg och Svensson (1992) visade att bland dem som ansåg att studiemedlen varit avgörande för valet att läsa vidare var andelen som lyckades avsluta sina studier lägre för dem med bättre familjebakgrund än för dem med sämre. Det skulle kunna tolkas som att de med högutbildade föräldrar som lockades att läsa på högskolan på grund av studiemedelsystemet befann sig längre ner i fördelningen av studiebegåvningar. Resultaten i Reuterberg och Svensson (1992) måste dock tolkas med stor försiktighet eftersom att det är mycket svårt för personer att bedöma hur de skulle ha agerat i olika hypotetiska situationer, i det här fallet om de skulle ha sökt till högskolan i frånvaro av studiemedel. Det är inte heller uppenbart vad som avses med en situation utan statliga studiemedel. Rimligen skulle en marknad för privata studielån då komma att etableras.

CSN (2006) genomförde en liknande enkätundersökning som Reuterberg och Svensson (1992). Studenter med studiemedel fick ta ställning till om de skulle börjat läsa på högskolan i det hypotetiska fall att studiemedel inte hade funnits. Upp emot 60 % angav att de inte skulle ha studerat utan statliga bidrag och lån. Denna siffra är högre än vad Reuterberg och Svensson fann, men avser också bara studenter som tagit studiemedel. När svaren delades upp med avseende på föräldrarnas utbildningsnivå, var rekryteringseffekten hela 71 % för personer med lågutbildade föräldrar, medan den var 42 % för barn till högutbildade föräldrar. Detta antyder således att studiemedlen kan ha motver-

kat den sociala selektionen. Vi har dock samma invändningar mot denna studie som mot Reuterberg och Svensson (1992).

Sammantaget är evidensen om studiemedelssystemets rekryteringseffekter tämligen svag. Ingen av de studier som baseras på svenska data har en trovärdig strategi för att mäta orsakssamband. Om vi ändå skulle våga oss på att tolka resultaten finns det vissa tecken som tyder på att studiemedelssystemet underlättat övergången till högre utbildning, framförallt för personer från studieovana hem. Studier från USA och Danmark, som i många fall har betydligt bättre strategier för att mäta orsakssamband, tyder också på att studiemedlen motverkar den sociala snedrekryteringen till högre studier. Eftersom studiemedelsystemet ofta framhävs som den kanske viktigaste åtgärden för att motverka den sociala selektionen till högskolan i Sverige, är det dock angeläget med mer forskning på detta område.

10.2.3 Regionalisering av högskolan

Högskolereformen år 1977 innebar en utlokalisering av högskoleutbildningen, då 11 nya högskolor etablerades runt om i landet. Samtidigt omvandlades universitetsfilialerna i Karlstad, Växjö och Örebro till högskolor. Därmed kom en betydligt större andel av befolkningen att bo nära en högskola. Även den kraftiga utbyggnaden av högskolan under 1990-talet hade en tydlig regional prägel. Under perioden 1992–2002 ökade antalet heltidsstudieplatser med 127 % på de regionala högskolorna. Motsvarande siffra för de traditionella universitetsorterna var 34 %. Under senare år har regionaliseringen av högskoleutbildningen befästs i och med att flera regionala högskolor har givits universitetsstatus.

Ett viktigt argument för utlokaliseringen av den högre utbildningen har varit att minska den sociala snedrekryteringen till högskolan. Genom att etablera av lärosäten på nya orter skulle högskolornas upptagningsområde vidgas. Eftersom befolkningsstrukturen i högskoleorterna såg väsentligen annorlunda ut än i universitetsorterna – med en relativt större andel lågutbildade personer – var förhoppningen att snedrekryteringen till högskolan därmed skulle motverkas. Studiemotiverade personer med lågutbildade föräldrar hade förstås haft möjlighet att läsa på högskolan även tidigare, men det har ibland hävdats att personer från studieovana hem ser avståndet till studieorten som ett större hinder än vad andra gör, se t.ex. SOU 1973:2.

Utlokaliseringen av högskoleutbildningen har också tänkbara regionalpolitiska konsekvenser. Högskoleutbildningen kan främja den regionala utvecklingen på olika sätt. För det första kan efterfrågan på högskoleutbildade vara

svår att tillgodose för vissa regioner. Ett större utbud av högutbildad personal skulle därmed rätta till lokala brister. För det andra kan närheten till forskning och utbildning skapa möjligheter för samverkan med det lokala arbetslivet och därmed öka regionens tillväxt. För det tredje har högskolorna en direkt inverkan på den lokala arbetsmarknaden i sin roll som arbetsgivare och kund. Efterfrågan på välutbildad personal i regionen borde därmed öka. I ett längre perspektiv skulle man också kunna tänka sig att ett stort utbud av högskoleutbildade skapar sin egen efterfrågan, dvs. att den lokala arbetsmarknaden med tiden absorberar en växande andel av de högskoleutbildade.

Holm och Häggström (1972) genomförde en tidig studie om sambandet mellan utlokaliseringen av den högre utbildningen och social snedrekrytering. De utnyttjade data från de s.k. studiekorten som de studerande fyllde i vid inskrivningstillfället till en högre utbildning under 1960- och 1970-talen. Förutom information om studieort och ämnen innehåller korten uppgifter om bl.a. familjebakgrund, ålder och födelseort. Författarna studerade personer som började en utbildning läsåren 1962/63, 1966/67, 1967/68 och 1969/70 för att undersöka olika demografiska förändringar vid de traditionella universiteten i samband med utökade utbildningsmöjligheter i Umeå, Linköping, Karlstad, Växjö och Örebro.

Holm och Häggström fann att antalet studenter minskade vid samtliga traditionella universitet i samband med regionaliseringen av den högre utbildningen. Minskningen var störst för personer som bor mer än 100 kilometer från de traditionella universitetsorterna. De fann också en större andel personer utan akademisk bakgrund vid de nya högskoleorterna än vid de traditionella universiteten. Det kan tolkas som en tendens till minskad snedrekrytering. Vi vill dock betona att resultaten bör tolkas med försiktighet då det finns vissa metodproblem.

Dryler (1994) använde aggregerade data från Folk- och bostadsräkningarna samt Universitets- och högskoleregistret. Uppgifterna avsåg personer i åldrarna 17–24 för perioden 1968–90. Hon undersökte om etableringen av högskolorna i Linköping, Luleå och Östersund minskade den sociala snedrekryteringen till högre utbildning. Dessa högskolor etablerades i slutet på 1960-talet och i början på 1970-talet. Effekterna skattades genom att jämföra andelen studenter från arbetarhem med de från akademikerhem före och efter etableringen av högskolorna. Som jämförelse användes personer med samma familjebakgrund men som inte växte upp nära en högskola eller ett universitet (Difference-in-

differences). Resultaten tydde inte på någon ökad rekrytering av personer från arbetarhem i samband med etableringen av de regionala högskolorna.

Kjellström och Regnér (1999a) undersökte sambandet mellan avstånd till en högskola och sannolikheten att påbörja en högre utbildning. De studerade regionaliseringens allmänna rekryteringseffekt och inte på den sociala snedrekryteringen. Författarna utnyttjade data från UGU-undersökningen för personer födda år 1967. Resultaten visade på ett signifikant negativt samband mellan avstånd till en högskola och sannolikheten att påbörja en högskoleutbildning. Det gäller när man kontrollerar för såväl föräldrarnas utbildning som resultat av förmågetest.¹⁰² Effekten är dock liten, vilket pekar på att avstånd inte är en speciellt viktig faktor för individens val att påbörja en högre utbildning.

Holzer (2009) studerade konsekvenserna av den kraftiga utbyggnaden av de regionala högskolorna under 1990-talet för utvecklingen av den sociala selektionen till högskolan. Hon studerade förändringen av andelen som började läsa på högskolan under 1980-talet och 1990-talet för dem som bodde i kommuner med en regional högskola i förhållande till dem som bodde i andra kommuner. Hon hävdar att andelen som började läsa på högskolan ökade med 42 procentenheter mer i kommuner med en regional högskola i förhållande till personer som bodde i andra kommuner. Detta måste dock vara fel då det motsvarar andelen som läste på högskolan i dessa kommuner under 1990-talet överhuvudtaget. Vi har därför replikerat Holzers analys och finner då att andelen som började läsa på högskolan som var födda i län med en regional högskola ökade med 2,4 procentenheter mer än personer födda i andra län mellan 1980-talet och 1990-talet. Om man delar upp analysen med avseende på föräldrarnas utbildningsnivå finner man dock ingen skillnad. Dessa resultat tyder således inte på att utbyggnaden av de regionala högskolorna skulle ha minskat den sociala snedrekryteringen.

Sammantaget tycks regionaliseringen av högskolan haft liten effekt på den sociala snedrekryteringen till högskolan. Evidensen på området är dock inte särdeles övertygande, och mer forskning behövs. Mycket tyder dock på att den sociala och könsmässiga snedrekryteringen till högskolan kan spåras längre tillbaka i elevernas utbildningskarriärer, i form av tidiga skolprestationer och utbildningsval. Det är därför troligt att tidiga insatser som riktas mot pojkar respektive barn till lågutbildade föräldrar kan hjälpa till att motverka den sociala och könsmässiga snedrekryteringen till högskolan på längre sikt.

¹⁰² Kjellström och Regnér (1999b) finner dock inga signifikanta samband mellan avstånd och valet av högre utbildning för personer födda åren 1948 och 1953.

10.3 Antagningssystemets utformning

Frånvaro av studieavgifter tillsammans med ett relativt generöst studiemedels-system skapar troligen en överefterfrågan på högre utbildning. Eftersom behovet av högutbildad arbetskraft är begränsat – och resursåtgången för högskoleutbildning stor – har antalet platser till högskolan begränsats. Frågan är hur man ska avgöra vilka studenter som skall ges tillträde till högskolan. En allmän utgångspunkt borde vara att ge företräde till personer med högst (samhälls-ekonomisk) avkastning på utbildningen.

I Sverige sker urvalet till högskolan mer eller mindre uteslutande utifrån olika förmågeindikatorer. Hösten 2008 fördelades ungefär 50 % av alla högskoleplatserna på utbildningsprogram på basis av gymnasiebetyg, 25 % med hjälp av högskoleprovet och resterande 25 % med särskilda urval.¹⁰³ Tre frågor infinner sig: Hur väl förutsäger urvalsindikatorerna studieframgång? Påverkar olika urvalsinstrument vilka grupper som antas till högskolan? Finns det ett samband mellan indikatorerna och privat- och samhällsekonomisk lönsamhet?

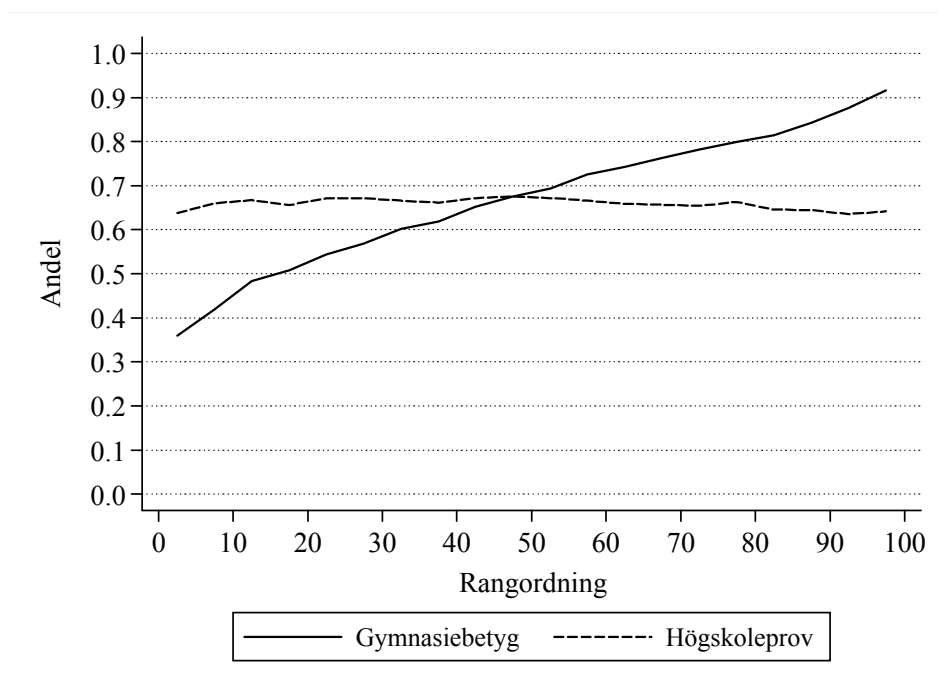
10.3.1 Prediktionsvärdet för olika urvalsinstrument

Gemensamt för behörighets- och urvalsreglerna är att de nästan uteslutande skall fånga förväntad studieprestation. Det finns en omfattande pedagogisk litteratur om kopplingen mellan olika urvalsinstrument och studieframgång, se t ex Willingham m.fl. (1990), Henriksson och Wolming (1998), Högskoleverket (2001), Svensson m.fl. (2001), Cliffordson (2008). Huvudbudskapet i dessa studier är att framförallt betyg har det högsta prediktionsvärdet. Däremot är värdet av högskoleprovet för att förutsäga studieframgång betydligt mer begränsat, framförallt givet studenternas gymnasiebetyg.

Figur 10.5 visar våra egna skattningar av sambandet mellan olika urvalsinstrument och sannolikheten att uppnå en kandidatexamen inom fem år. För att göra de två urvalsindikatorerna jämförbara har vi rangordnat provresultaten och betygen mellan 1 och 100. Analyserna är begränsade till personer som börjat på högskolan direkt efter gymnasiet under perioden 1994–2002, och som har uppgift om både avgångsbetyg från gymnasiet och resultat på högskoleprovet. För att mäta sambanden för alla som har skrivit högskoleprovet – och inte bara de som antagits – predicerar vi sannolikheten att ta examen. Det sker med hjälp av sambandet mellan urvalsinstrumenten och studieframgång för de som antagits (se Gustafsson och Reuterberg 2000). De skattade sambanden för det ena urvalsinstrumentet gäller givet värdet på det andra.

¹⁰³ Egna beräkningar baserade på HSV:s antagningsstatistik för utbildningsprogram hösten 2008.

Figur 10.5. Andel med minst kandidatexamen efter fem år på högskolan



Not: Figuren visar skattningar av andelen som uppnår minst kandidatexamen (examen eller poäng) efter fem år från högskolestarten, vid olika delar av fördelningen av gymnasiebetyg och högskoleprov. Sambandet för betyg (prov) gäller givet värdet på provet (betyget). Analysen omfattar både personer som antagits och de som inte antagits till högskolan. Sannolikheten att ta examen har predicerats för alla personer med hjälp av betyg och provresultat för de som har antagits. Analysen tar hänsyn till skillnader i sannolikheten att ta examen med avseende på interaktionen mellan kön, år, lärosäte och program/ämne.

Liksom tidigare studier finner vi ett mycket starkt positivt samband mellan gymnasiebetyg och studieframgång. Bara omkring 35 % av studenterna med de lägsta gymnasiebetygen har tagit en kandidatexamen efter fem år. Motsvarande siffra för studenter med de högsta betygen är mer än 90 %. Ett annat sätt att beskriva sambandet är att en ökning av gymnasiebetygen med en standardavvikelse är förknippad med en ökning av sannolikheten att ta examen med 14 procentenheter. Sambandet mellan högskoleprovresultat och studieframgång är mycket svagt.¹⁰⁴ I den mån det finns något samband är det negativt: Ju bättre

¹⁰⁴ Sambandet mellan högskoleprovresultat och studieframgång blir positivt om man inte biter på gymnasiebetyg. Personer med de högsta provresultaten har då omkring 20 procentenheter högre sannolikhet att ta examen inom fem år än de med de lägsta resultaten. Eftersom ca. 80 % av alla personer som skriver högskoleprovet också har slutbetyg från gymnasiet, är dock

högskoleprovsresultat desto lägre sannolikhet att slutföra studierna i tid. Den dåliga prognosförmågan för högskoleprovet är alarmerande med tanke på att 25 % av alla studenter antas via högskoleprovet. Användandet av högskoleprovet som urvalsinstrument tenderar således att minska examensfrekvensen på högskolan.

Frågan är dock hur stor betydelse valet av urvalsinstrument har i praktiken. Instrumentens prediktionsvärde är förstås viktigt, men effekterna av förändrade urvalsregler beror också på andra faktorer. För det första är frågan om val av urvalsinstrument relevant först om söktrycket till högskolan är tillräckligt högt. I Sverige antas omkring 60 % av alla som söker till högskolan (SCB 2009), vilket öppnar upp för en viss – men inte en extrem – selektion. För det andra beror betydelsen av valet av urvalsinstrument på hur mycket instrumenten samvarierar. Många av dem som har höga gymnasiebetyg skriver också bra på högskoleprovet och vice versa, men överensstämmelsen är förstås inte perfekt. Under början av 2000-talet låg korrelationen mellan gymnasiebetygen och högskoleprovet på nästan 0,4. Samvariationen mellan betyg och provresultat gör valet av urvalsinstrument ytterligare något mindre betydelsefullt.

För att undersöka hur genomströmningen på högskolan påverkas av valet av urvalsinstrument, har vi predicerat sannolikheten att ta kandidatexamen inom fem år för alla personer som slutade gymnasiet 1993–99 och som har uppgift om både gymnasiebetyg och högskoleprov.¹⁰⁵ Under antagandet att 60 % av dessa personer antas till högskolan, har vi beräknat förväntad examensfrekvens när vi låter andelen som antas via högskoleprovet variera mellan 0 och 100 procent.¹⁰⁶ Resultatet redovisas i Figur 10.6.¹⁰⁷

Om man helt slopade högskoleprovet som urvalsinstrument – och bara antog sökande med hjälp av gymnasiebetygen – skulle nästan 81 % av studenterna ta examen från högskolan inom 5 år. Om man successivt utökade andelen som antas med hjälp av högskoleprovet, skulle examensfrekvensen gradvis förväntas att sjunka. Försämringen av examenssannolikheten är dock inte så

den intressanta frågan vilken av de två urvalsindikatorerna som har det högsta prediktionsvärdet givet den andra indikatorn.

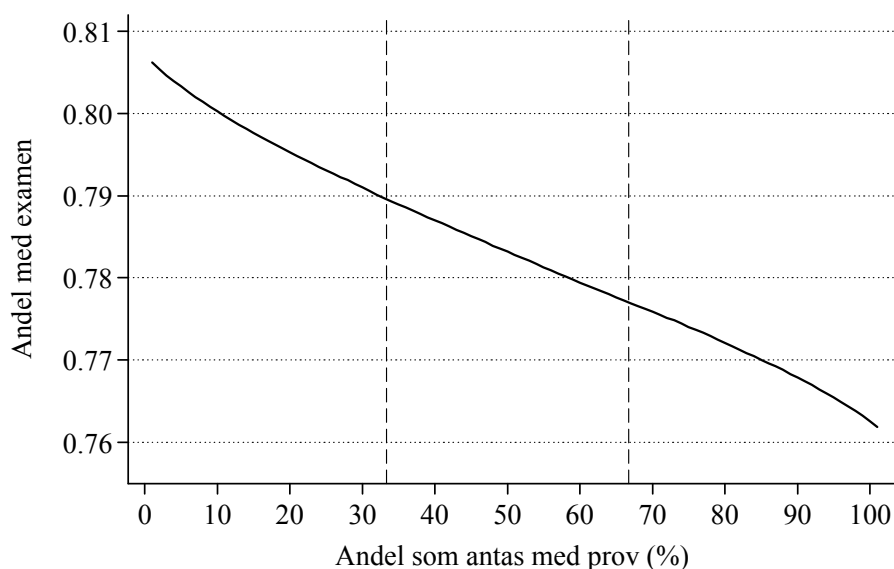
¹⁰⁵ Antagningsreglernas utformning kommer förstås att påverka hur många – och vilka – som väljer att skriva högskoleprovet. Genom att bara studera personer som skrivit provet bortser vi från detta problem.

¹⁰⁶ Vi saknar uppgift om sökande, vilket annars skulle ha varit en bättre studiepopulation.

¹⁰⁷ Eftersom många av de sökande med höga gymnasiebetyg också har bra provresultat, kommer gruppen av studenter som antas att påverkas av i vilken ordning de olika urvalsinstrumenten används. Vi redovisar därför genomsnittet för de studenter som skulle antas om betyg används först respektive om prov används först.

stor. I genomsnitt skulle en 10-procentig ökning av andelen som antas med högskoleprovet förväntas minska andelen som tar examen med ungefär 0,5 procent. Om man enbart antog sökande med hjälp av högskoleprovet skulle omkring 76 % ta examen inom 5 år.

Figur 10.6. Andel som förväntas ta examen om olika andel antas med prov



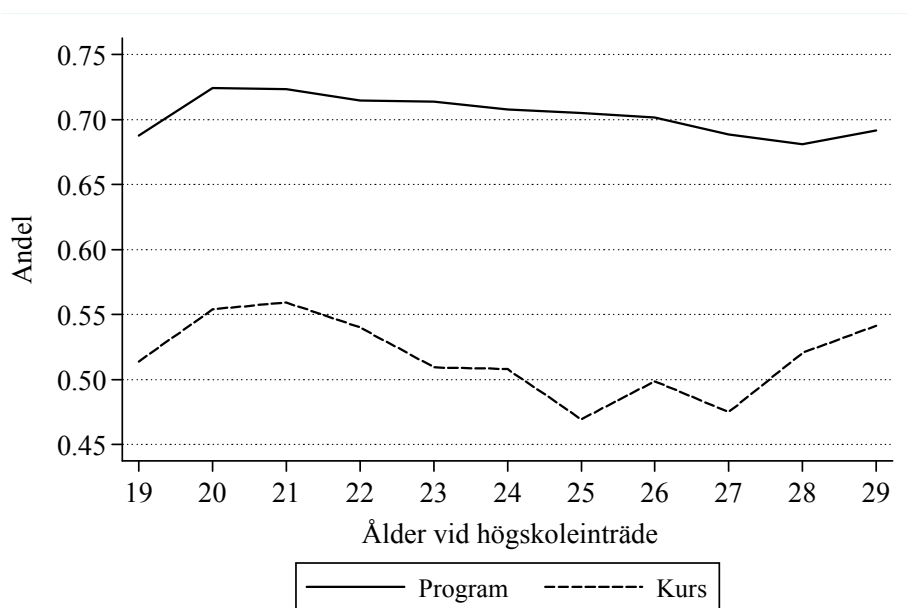
Not: Figuren visar skattningar av andel som uppnår minst kandidatexamen (examen eller poäng) efter fem år på högskolan, om olika andel av de sökande skulle antas via högskoleprovet istället för via betyg

Ett slopande av högskoleprovet som urvalsgrund skulle således kunna medföra en ökning av examenssannolikheten med som mest 5 procentenheter. I praktiken finns dock begränsningar för hur stor andel som får antas med betyg respektive prov. Enligt högskoleförordningen (1993:100, Kap 7, 13§) måste minst 1/3 men högst 2/3 av platserna tilldelas på basis av högskoleprovet. Dessa begränsningar illustreras av de streckade linjerna i Figur 10.6. Högskolorna har således ett manöverutrymme som motsvarar en förbättring av andelen som förväntas ta examen med omkring 1,5 procentenheter. Om högskolorna gavs möjlighet att helt slopa högskoleprovet som urvalsinstrument skulle examenssannolikheten kunna öka med ytterligare 1,5–2 procentenheter. Även om

detta inte är särskilt mycket, måste det finnas mycket goda skäl till varför man skulle vilja gå miste om denna effektivitetsvinst.

Utöver rena förmågeindikatorer har också arbetslivserfarenhet använts som urvalsinstrument till högskolan sedan högskolereformen 1977. Det var först i och med antagningen till högskolan hösten 2008 som möjligheten att meritera sig via arbetslivserfarenhet försvunnit. Det har dock funnits flera invändningar mot att låta arbetslivserfarenhet inverka på urvalet till högre studier. För det första är prognosvärdet begränsat. Det är mycket osäkert om arbetslivserfarenhet är en bra indikator på studieframgång (se t.ex. Henriksson och Wolming 1998). För det andra tenderar ett sådant system att öka medelåldern på studenterna. Därmed minskar kvarvarande tid på arbetsmarknaden liksom antalet år med förväntade inkomsttillskott. Troligen är också värdet på inkomstbortfallet under studietiden högre eftersom arbetslivserfarenhet antas generera högre produktivitet och lön. För det tredje kan ett system som belönar arbetslivserfarenhet få konsekvenser inte bara för dem som vid ansökningstillfället befinner sig på arbetsmarknaden, utan också för nyutexaminerade från gymnasieskolan. Det finns en risk att urvalsreglerna tvingar studenterna att arbeta några år för att kunna komma in på högskolan. En sådan fördröjningsmekanism är inte eftersträfvansvärd, eftersom produktionstillskottet blir lägre.

Figur 10.7. Ålder vid högskoleinträde och andel med examen



Not: Figuren visar andel som uppnår kandidatexamen inom fem år (examen eller poäng) för de som börjar på högskolan vid olika åldrar, givet kön, betyg, program/ämne, lärosäte och år.

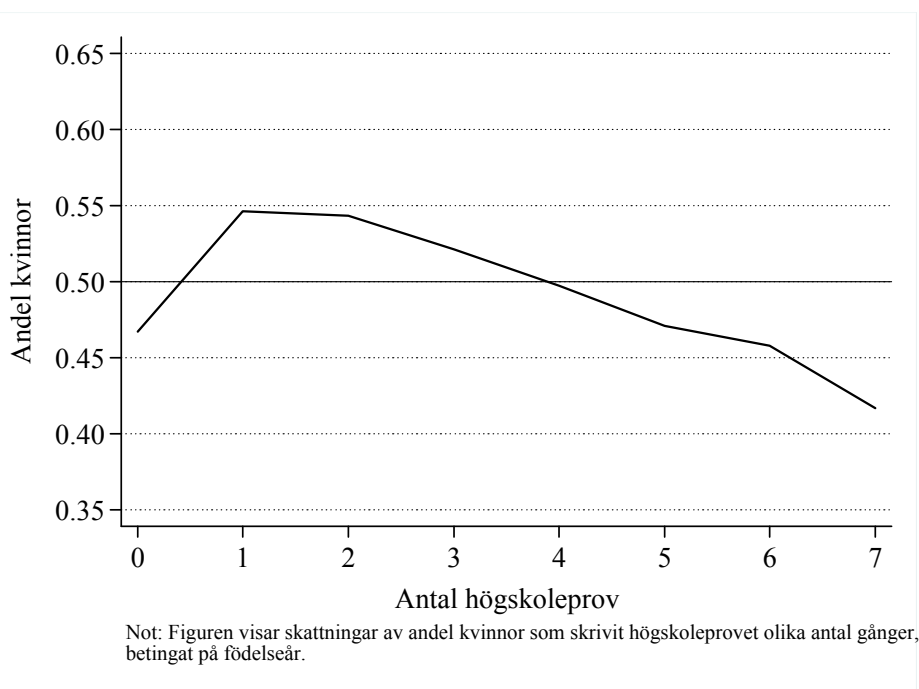
Figur 10.7 visar sambandet mellan ålder vid högskoleinträdet och andel som uppnått kandidatexamen inom 5 år. Analysen begränsas till personer som påbörjat högskolestudier under perioden 1993–2000 och tar hänsyn till skillnader i gymnasiebetyg, typ av högskolprogram/kurs, högskola och år. Analysen görs separat för personer som påbörjat sina högskolestudier på ett program respektive en kurs. Resultatet visar att det nästan inte finns något observerat samband mellan fördröjning av högskoleinträde och sannolikheten att slutföra studierna i tid. Det finns indikationer på att personer som börjat studera enstaka kurser vid en högre ålder har något lägre sannolikhet att uppnår en kandidatexamen inom fem år. Sammantaget antyder dock resultaten att längre arbetslivserfarenhet, vilken rimligtvis ökar med högre ålder, inte medför högre sannolikhet att avsluta högskolestudierna. Det omvända gäller dock inte heller, personer som väntar med högskoleinträdet tycks prestera i nivå med de studenter som påbörjar studierna direkt efter gymnasiet. Den resterande tiden på arbetsmarknaden kan dock vara kortare för äldre studenter, vilket får konsekvenser för utbildningens lönsamhet.

10.3.2 Påverkar urvalsinstrumenten vilka som går på högskolan?

Högskoleprovet introducerades som urvalsinstrument till högskolan i och med högskolereformen 1977. Ursprungligen användes det bara för personer som var behöriga till högskolan i kraft av ålder och arbetslivserfarenhet (de s.k. 25:4-orna), och som många gånger saknade fullständiga gymnasiebetyg. Från och med hösten 1991 är högskoleprovet öppet för alla studenter, och i dag antas ungefär 25 % av alla programnybörjare med hjälp av högskoleprovet. Ett av argumenten för att låta alla skriva högskoleprovet var att motverka den sociala selektionen (Wedman och Henrysson 1992). Även införandet av arbetslivserfarenhet som urvalsgrund hade liknande syften. Frågan är dock hur olika urvalsinstrument påverkat vilka grupper som antas till högskolan. Precis som i tidigare avsnitt fokuserar vi på övergångssannolikheterna med avseende på kön och socioekonomisk bakgrund. Först diskuteras högskoleprovet, därefter följer en granskning av arbetslivserfarenhet som urvalsinstrument.

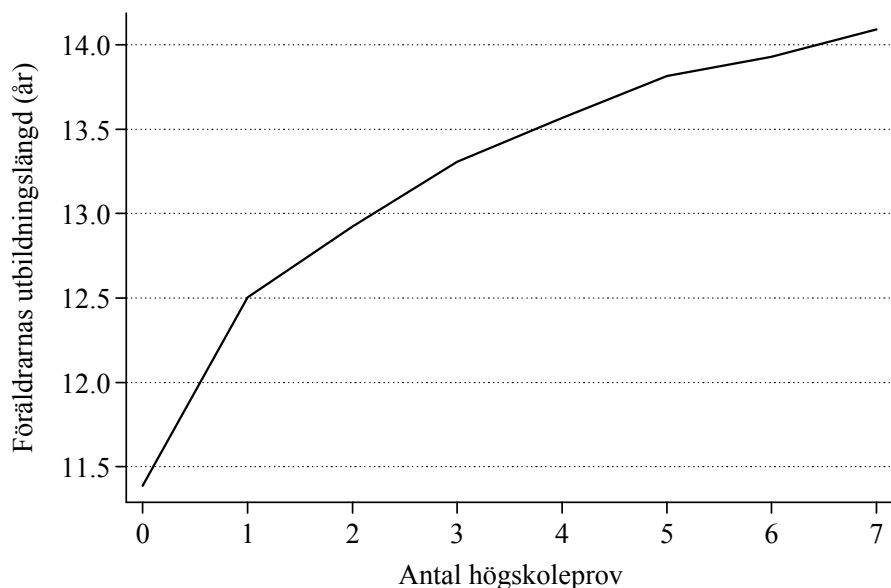
Vilka grupper som rekryteras till högskolan via högskoleprovet beror dels på vilka grupper som väljer att skriva högskoleprovet, dels på hur olika grupper presterar på provet. Vi studerar först vilka som skriver högskoleprovet, samt hur vanligt det är att olika grupper skriver om provet. Figur 10.8 visar andelen kvinnor som skrivit högskoleprovet olika många gånger under perioden 1985–2004; vi tar här bara hänsyn till skillnader i avslutningsår från gymnasiet. Kvinnor skriver högskoleprovet i något större utsträckning än män (omkring 47 % av kvinnorna har inte skrivit högskoleprovet alls). Män har dock högre sannolikhet att skriva många prov. Sammantaget finns inget tydligt könsmönster i provskrivning.

Figur 10.8. Andel kvinnor som skrivit olika antal högskoleprov



Skillnader i sannolikheten att skriva högskoleprovet för personer med olika familjebakgrund har studerats mer ingående. Gustafsson och Benjegård (1996) redovisade betydande skillnader i upprepat provtagande för personer med olika social bakgrund, där barn till tjänstemän hade mycket högre sannolikhet att skriva om högskoleprovet än barn till arbetare. Liknande resultat finns i Gustafsson m.fl. (2000). I Figur 10.9 studerar vi sambandet mellan antalet skrivna högskoleprov och föräldrarnas genomsnittliga utbildningslängd. Notera först att personer som skriver högskoleprovet minst en gång är en starkt selekterad grupp. I genomsnitt har föräldrar till dem som skriver högskoleprovet 1,6 års längre utbildning än föräldrar till dem som inte skriver provet (framgår inte av figuren). Föräldrarnas utbildningslängd stiger kontinuerligt med antalet skrivna prov. Skillnaden i föräldrarnas utbildningslängd mellan dem som inte skriver något prov alls och dem som skriver sju prov är mer än 2,5 år. Dessa resultat står sig även om man beaktar skillnader i gymnasiebetyg.

Figur 10.9. Antal skrivna högskoleprov och föräldrarnas utbildningslängd



Not: Figuren visar skattningar av föräldrarnas genomsnittliga utbildningslängd för personer som skrivit högskoleprovet olika antal gånger, betingat på kön och födelseår.

Konsekvenserna av den starka sociala selektionen till högskoleprovet förvärras ytterligare av att resultatet på högskoleprovet förbättras med antalet provtillfällen. Flera svenska studier har visat att ökningen i provresultat är störst den första gången provet skrivits om; några studier har funnit att resultaten ökar även vid fortsatt provtagande (Henriksson 1985, 1995; Henriksson och Henriksson 1991; Henriksson och Bränberg 1994; Gustafsson och Westerlund 1994; Gustafsson och Benjegård 1996; Cliffordson 2004). Dessa resultatförbättringar skulle dels kunna bero på att personerna lär sig skriva högskoleprovet (träningseffekt), del på att de faktiskt lär sig saker med tiden (ålderseffekt). Om resultatökningarna speglar verklig förbättring av de kunskaper och färdigheter högskoleprovet ska mäta – vilka i sin tur är tänkta att förutsäga studenternas studieframgång – är upprepad provskrivning inget större problem. I själva verket är det just sådana färdigheter som högskoleprovet ska fånga upp. Om skillnaderna i resultat mellan provtillfällen däremot helt eller delvis beror på att studenterna lär sig att skriva högskoleprovet, försämras högskoleprovets prediktionsvärdet samtidigt som skillnaderna mellan olika grupper tilltar.

Cliffordson (2004) skattade effekten av upprepat provtagande (träningseffekt) genom att jämföra personer som vid samma ålder skrivit olika antal prov. Hon beaktade också skillnader i bl.a. grundskole- och gymnasiebetyg. Hon fann att upprepat provtagande ökade resultatet på provet med i genomsnitt 0,25 *d*. Ålderseffekten skattades genom att relatera förändringen i provresultat – för de som skrivit två eller tre prov – till tidsintervallet mellan proven. Även i detta fall kontrollerade Cliffordson för skillnader i betyg. Hon fann då att ett års ålder förbättrade provresultaten med 0,13–0,19 *d*.

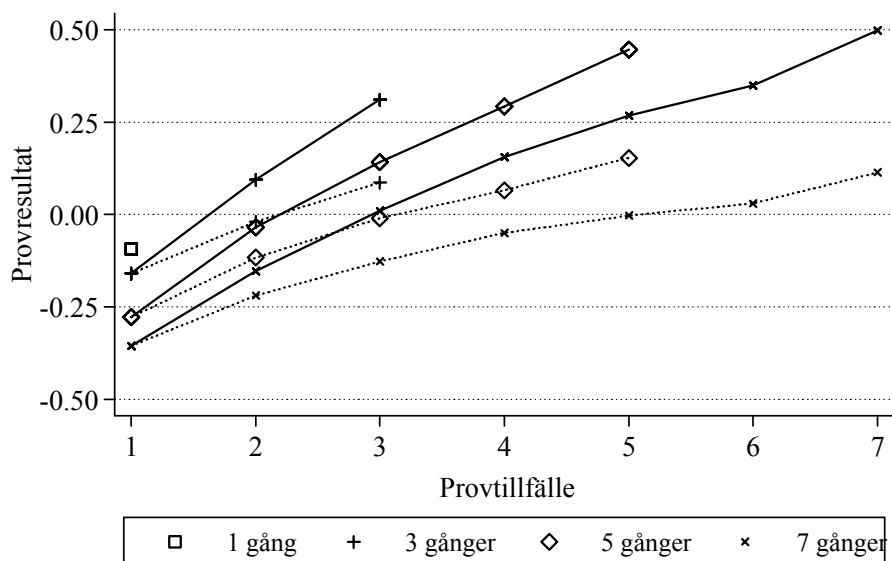
Vi har skattat effekten av upprepat provtagande för personer som skrivit högskoleprovet olika antal gånger under perioden 1980–2005. Analysen begränsas till prov skrivna före högskoleinträdet. Eftersom det inte är slumpmässigt hur många prov som personerna skriver, analyserar vi personer som skrivit samma antal prov gemensamt. De heldragna linjerna i Figur 10.10 visar standardiserat provresultat vid olika provtillfällen för personer som skrivit högskoleprovet en, tre, fem eller sju gånger. Notera först att ju fler gånger personer skriver provet, desto sämre var deras resultat det första provtillfället. Detta tyder på en negativ selektion av personer som skriver högskoleprovet flera gånger. Om man studerar förändringen i provresultat för samma grupp av individer, ser man en tydlig resultatförbättring på omkring 0,27 *d* mellan första och andra provtillfället. Även därefter är resultatökningen påtaglig, även om effekten tycks avta något med antalet provtillfällen. I genomsnitt ökade provresultaten mellan andra och tredje provtillfället med 0,20 *d*, och mellan tredje och fjärde tillfället med omkring 0,17 *d*.

För att försöka separera träningseffekten från ålderseffekten, har vi tagit hänsyn till de skrivandes ålder vid respektive tillfälle (ålderseffekt) samt deras genomsnittliga provresultat (selektionseffekt). Resultaten ges av de streckade linjerna i Figur 10.10. Det resulterar i betydligt svagare resultatutveckling med antalet skrivna prov. Ungefär hälften av resultatförbättringen mellan olika provtillfällen kan förklaras av selektions- och ålderseffekter. Det är dock notabelt att det finns betydande träningseffekter av upprepat provskrivning: skillnaden i resultat mellan första och andra provtillfället som beror på att individerna lär sig att skriva högskoleprovet uppmäts till ungefär 0,14 *d*. Träningseffekten för ytterligare prov är något lägre: 0,10 *d* från andra till tredje provtillfället, respektive 0,09 *d* från tredje till fjärde tillfället.¹⁰⁸ Eftersom barn till

¹⁰⁸ Våra skattningar av ålderseffekter på högskoleprovet tyder på att ett års ålder förbättrar resultaten på högskoleprovet med omkring 0,07 *d*.

högutbildade föräldrar i större utsträckning skriver om högskoleprovet tenderar dessa träningseffekter att förvärpa den sociala snedrekryteringen till högskolan.

Figur 10.10. Provresultat för de som skrivit högskoleprovet flera gånger



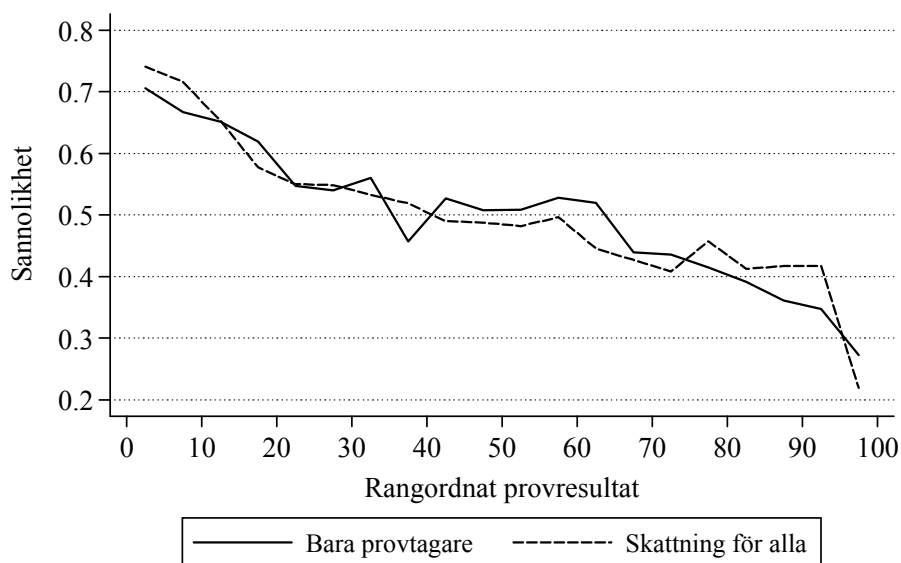
Not: Figuren visar skattningar av genomsnittligt standardiserat provresultat vid olika provtillfällen för de som skrivit högskoleprovet olika antal gånger, betingat på kön och provtidpunkt. Den prickiga linjen betingar också på personens genomsnittliga resultat och ålder vid provet.

Vi har nu studerat vilka personer som skriver högskoleprovet, och hur upprepat provdeltagande varierar med avseende på kön och familjebakgrund. För att kunna uttala oss om hur olika urvalsinstrument påverkar sammansättningen av dem som läser på högskolan, kommer vi nu att studera gruppskillnader i resultatet på högskoleprovet och skillnader mellan dem som är olika gamla (och har olika lång arbetslivserfarenhet) när de börjar på högskolan. För att undvika att fånga upp de effekter som upprepat provtagande har på resultaten, kommer analyserna av högskoleprovet att baseras på personernas första provtillfälle.

Högskoleprovet har ofta beskyllts för att gynna män, och män presterar onekligen betydligt bättre på högskoleprovet än vad kvinnor gör (se t.ex. Stage 1996). Vi har beräknat sannolikheten för kvinnor att uppnå olika högskoleprovsresultat med hjälp av UGU-data. Den heldragna linjen i Figur 10.11 visar

könsskillnaderna i provresultat för dem som skriver högskoleprovet. I genomsnitt befinner sig män 12 percentilranger högre upp i provfördelningen (framgår inte av figuren). Sannolikheten för en kvinna att återfinnas längst ned i provfördelningen är ungefär 70 %. Chansen för kvinnor att uppnå ett visst resultat sjunker sedan kontinuerligt ju längre upp i provfördelningen man kommer. Kvinnor har bara omkring 30 procents sannolikhet att uppnå de allra bästa provresultaten.

Figur 10.11. Sannolikhet för kvinnor att uppnå olika resultat på högskoleprovet



Not: Figuren visar sannolikheten att kvinnor uppnår det resultat på högskoleprovet som anges av x-axeln. Provresultaten har skattats separat för män och kvinnor utifrån förmåga i Åk 6, provresultat i Åk 6 och Åk 9, betyg i Åk 9, gymnasiebetyg och föräldrarnas utbildning.

De skillnader som finns mellan kvinnor och mäns resultat på högskoleprovet måste dock inte innebära att kvinnorna missgynnas av provets konstruktion. Eftersom det är frivilligt att skriva högskoleprovet är det möjligt att män som skriver provet är mer positivt selekterade än kvinnor som skriver provet. Reuterberg (1996) visade att män som skrev högskoleprovet har bättre observerade egenskaper (förmåga i åk 6 och högre betyg i åk 9) än dem som inte skrev provet. Samma sak gäller för kvinnor, men där var skillnaden inte lika stor.

Detta tyder på att urvalet av män som skriver provet är mer positivt än urvalet av kvinnor som skriver provet.

Skillnaden i selektion till högskoleprovet för män och kvinnor gör det svårt att uttala sig om könsskillnader i provresultat i hela befolkningen. Ett sätt att närma sig denna fråga är att skatta hur olika observerade egenskaper (betyg, provresultat, familjebakgrund, etc.) påverkar högskoleprovsresultatet för män respektive kvinnor. Om sambanden mellan de observerade egenskaperna och provresultaten är samma för provtagare som för dem som inte skriver provet, kan man tillskriva alla individer det provresultat som deras egenskaper förutsäger. Vi har skattat högskoleprovsresultatet för män och kvinnor separat med hjälp av den detaljerade information om tidigare skolprestationer som finns i UGU-materialet.¹⁰⁹ Den streckade linjen i Figur 10.11 visar att skillnaderna i provresultat mellan män och kvinnor inte kan förklaras av skillnader i urvalet till högskoleprovet med avseende på de variabler som vi beaktat. Kvinnor har fortfarande mer än 70 procents sannolikhet att nå de lägsta skattade provresultaten, medan de har mindre än 35 procents sannolikhet att nå de högsta skattade provresultaten. Denna analys tyder på att skillnaden i selektionen till högskoleprovet mellan män och kvinnor inte kan förklara de skillnader som finns i provresultat.¹¹⁰

På samma sätt som män tenderar att skriva bättre än kvinnor på högskoleprovet finns betydande skillnader efter familjebakgrund. Reuterberg (1996) samt Gustafsson och Benjegård (1996) har visat på stora skillnader i provresultat med avseende på föräldrarnas yrke och utbildning. Konsekvenserna av att använda högskoleprovet som urvalsgrund istället för betyg, beror dock på hur dessa skillnader förhåller sig till skillnaderna i gymnasiebetyg. Gustafsson och Westerlund (1994) fann att social bakgrund hade något mindre betydelse för högskoleprovsresultatet än för gymnasiebetygen.

Ett problem med att jämföra betydelsen av social bakgrund för olika urvalsinstrument, är att det är frivilligt att skriva högskoleprovet. Precis som i analysen av könsskillnader har vi därför valt att använda UGU-data för att skatta provresultat i hela befolkningen. Detta har dock en mindre betydelse för resultaten. Figur 10.12 visar föräldrarnas genomsnittliga utbildningslängd för personer med olika gymnasiebetyg respektive olika provresultat. Båda urvals-

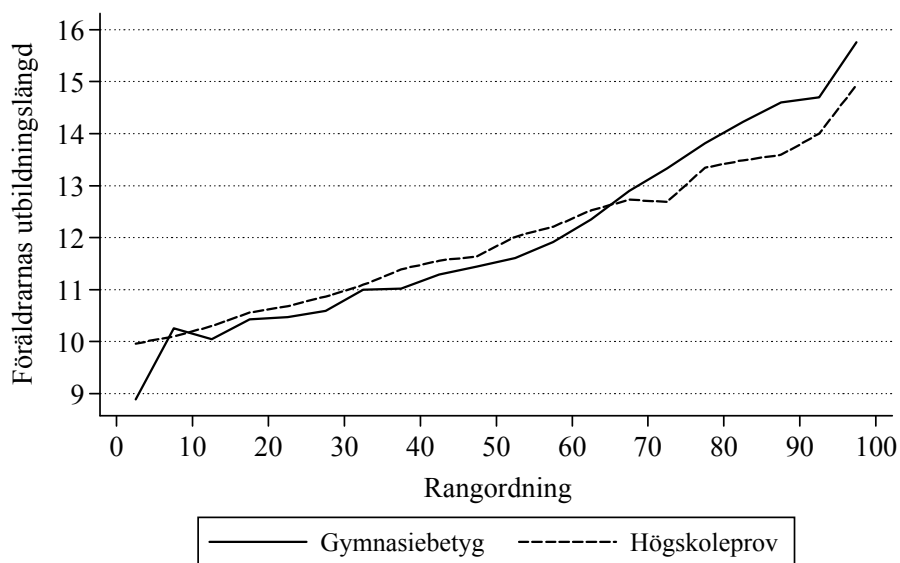
¹⁰⁹ Modellen förklarar ungefär 60 % av variationen i provresultat för både kvinnor och män.

¹¹⁰ Det absolut enklaste sättet att utreda frågan om gruppsskillnader i högskoleprovet vore att ge provet till ett slumpmässigt urval av personer (t.ex. nyutexaminerade från gymnasiet). Med tanke på den uppmärksamhet som skillnader mellan framförallt män och kvinnor har fått, och på att en sådan undersökning skulle vara enkel att genomföra, är det förvånansvärt att det inte har gjorts.

instrumenten är starkt relaterade till föräldrarnas utbildningslängd. Sambanden mellan familjebakgrund och de olika måtten på skolprestationer ser dessutom likartade ut. Uppväxtmiljö tycks således vara ungefär lika betydelsefullt för prestationerna på högskoleprovet som för betygen. Därför kan inte en övergång till högskoleprovet förväntas påverka den sociala snedrekryteringen till högskolan. Notera dock att denna slutsats bara avser de som skriver högskoleprovet en gång. Upprepat provtagande tenderar att öka den sociala selektionen.

Sammantaget tycks användandet av högskoleprovet som urvalsinstrument i stället för gymnasiebetyg minska den könsmässiga skillnaden i andelen som påbörjar högskolestudier. Det beror framförallt på att män gynnas av högskoleprovets konstruktion i förhållande till kvinnor. Högskoleprovet tenderar också att förvärpa den sociala snedrekryteringen till högskolan. Det beror huvudsakligen på att barn till högutbildade föräldrar har större sannolikhet att skriva högskoleprovet och att de skriver det många gånger; den sociala selektionen förvärras också av att man lär sig skriva bra på högskoleprovet.

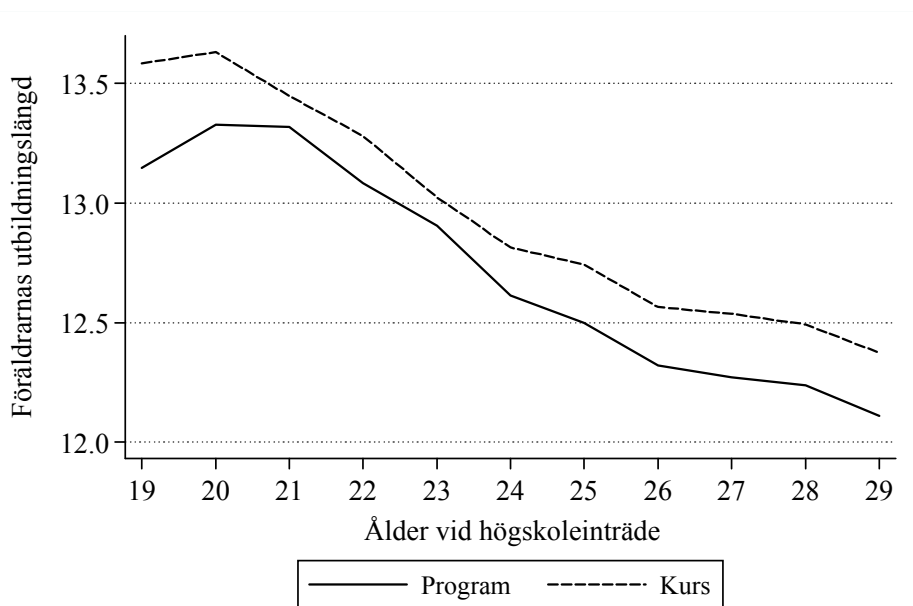
Figur 10.12. Familjebakgrund för personer med olika betyg respektive provresultat



Not: Figuren visar föräldrarnas genomsnittliga utbildningslängd vid olika gymnasiebetyg och högskoleprovresultat. Provresultaten och betygen har skattats utifrån förmåga i Åk 6, provresultat i Åk 6 och Åk 9 och betyg i Åk 9, både i nivå och interagerat med föräldrarnas utbildning.

Arbetslivserfarenhet har använts som urvalsgrund till högskolan fram till och med antagningen hösten 2008. Ett av argumenten för att använda arbetslivserfarenhet som urvalsinstrument har varit att minska den sociala snedrekryteringen till högskolan. Frågan är dock om personer med olika familjebakgrund väljer att påbörja sina studier vid olika ålder. Figur 10.13 visar föräldrarnas genomsnittliga utbildningslängd för personer som påbörjar sina studier mellan 19 och 29 års ålder. Vi har tagit hänsyn till skillnader i födelseår, gymnasiebetyg, lärosäte samt program/kurs. Den heldragna linjen visar sambandet mellan ålder och familjebakgrund för dem som påbörjar ett högskoleprogram, medan den streckade linjen visar dem som påbörjar en enstaka kurs på heltid. Ju äldre personerna är vid högskoleinträdet desto svagare är deras familjebakgrund. Användandet arbetslivserfarenhet (som rimligen stiger med tilltagande ålder) minskar således den sociala selektionen till högskolan. Det är därför möjligt att slopandet av arbetslivserfarenhet som urvalsgrund kan medföra ökade sociala skillnader på högskolan.

Figur 10.13. Ålder vid högskoleinträde och föräldrarnas utbildningslängd



Not: Figuren visar föräldrarnas utbildningslängd för dem som börjar på högskolan vid olika ålder, givet kön, betyg, program, ämne, lärosäte och födelseår.

10.3.3 Finns en koppling mellan urvalsinstrument och avkastning?

Vilken koppling finns mellan urvalsinstrumenten och avkastning på högskolestudier? På en väl fungerande utbildningsmarknad – utan externa effekter eller kreditrestriktioner – kommer de personer med högst förväntad avkastning på högskoleutbildning att vidareutbilda sig. Samma krav skulle man vilja ställa på urvalsinstrumenten, det vill säga att de som antas till högskolan också är de med högst avkastning. Ett mer modest krav är att avkastningen på utbildning åtminstone inte ska falla med högre betyg, bättre högskoleprovsresultat, etc. Vi ska i det här avsnittet sammanfatta de fåtal studier som direkt eller indirekt behandlar denna frågeställning. Vi presenterar också egna skattningar av avkastningen på utbildning för personer med olika betyg, högskoleprovsresultat respektive ålder vid högskoleinträdet.

Studier av skillnader i avkastningen på utbildning för personer med olika egenskaper (såsom gymnasiebetyg) möter i princip samma metodologiska utmaningar som studier av den genomsnittliga avkastningen på utbildning (se avsnitt 5.1). Det är emellertid möjligt att mäta *skillnaden* i avkastning mellan exempelvis personer med olika bra gymnasiebetyg, under förutsättning att skillnaderna i andra egenskaper (som forskaren inte kan observera) för personer med olika lång utbildning är desamma för dem som har låga respektive höga betyg. Med andra ord, så länge den uppmätta avkastningen på utbildning är lika missvisande för dem som har låga betyg som för de som har höga betyg, kan man uttala sig om skillnaden i avkastning mellan grupperna. Om, å andra sidan, personer med låga betyg som väljer att vidareutbilda sig exempelvis är mer motiverade än personer med höga betyg som läser vidare, kommer den uppmätta skillnaden i avkastning inte att motsvara den verkliga skillnaden.¹¹¹

En annan komplikation med att mäta hur avkastningen på högskoleutbildning varierar med olika urvalsinstrument, är att antagningen sker bland dem som sökt en specifik studieinriktning vid en viss högskola. Eftersom människors preferenser för olika utbildningar och studieorter kan ha en direkt inverkan på avkastningen på utbildning, kan jämförelser av avkastningen på utbildning för personer som sökt olika utbildningar vara missvisande för betydelsen av olika urvalsinstrument. Ett angränsande problem är att urvalet till olika program och högskolor sker med hjälp av de urvalsinstrument som ska utvär-

¹¹¹ Till skillnad från studier som försöker mäta den kausala effekten av utbildning på inkomsten, är frågeställningen nu istället vilka studenter som har mest att vinna på att antas till högskolan. Därmed är den totala skillnaden i avkastningen på utbildning mellan exempelvis de som har låga respektive höga gymnasiebetyg av intresse, oavsett om personer med olika bra betyg också skiljer sig andra avseenden (familjebakgrund, kön, motivation, etc.).

deras. Om de populäraste inriktningarna och studieorterna också ger högst avkastning – pga. hög efterfrågan på arbetsmarknaden, bättre utbildningskvalitet eller bättre studiekamrater – kan det leda till en överskattning av skillnaden i avkastning för personer med exempelvis olika höga betyg. För att beakta båda dessa problem skulle man helst vilja studera personer som sökt samma utbildning vid samma lärosäte.

Slutligen kan det uppstå problem att skatta skillnader i avkastning för olika urvalsinstrument, om avkastningen på utbildning skiljer sig åt mellan olika nivåer. Orsaken till det är att urvalsindikatorerna påverkar vilka som personer som kan avancera i utbildningssystemet. Därmed kommer fördelningen av utbildningsnivåer att vara olika bland dem som har olika värden på urvalsinstrumenten. Personer med exempelvis låga gymnasiebetyg kommer att vara överrepresenterade på lägre utbildningsnivåer, medan personer med höga betyg kommer att vara överrepresenterade på högre utbildningsnivåer. Om inte hänsyn tas till dessa skillnader finns risk att skillnaden i avkastning för personer med olika värden på urvalsinstrumenten fångar upp skillnader i avkastning på olika utbildningsnivåer.

Det finns ett fåtal svenska studier som undersöker förekomsten av skillnader i avkastning på utbildning för personer som har olika betyg, testresultat eller ålder vid högskoleinträdet. Kjellström (1999) använde UGU-data för personer födda 1948 och 1953. Han visade att avkastningen på utbildning i början på 1990-talet inte ökade med uppmätt förmåga i åk 6; i många fall var avkastningen lägre för dem med högre förmåga än för dem med lägre förmåga. Analysen avser dock alla utbildningsnivåer, och är inte begränsad högskoleutbildning. Dessutom analyseras ett tidigt förmågetest som inte används vid antagningen till högskolan. Det är därför svårt att överföra detta resultat till skillnader i avkastning på högskoleutbildning för olika urvalsinstrument.

Nordin (2008) studerade skillnader i avkastning på utbildning för personer med olika resultat på det militära inskrivningsprovet. Studien begränsas till män födda 1965–71 med en årsinkomst över 80 000 kronor år 2001. Han analyserade hur avkastningen på utbildning varierar med uppmätt kognitiv förmåga, och fann att avkastningen ökade med stigande förmåga. Resultaten påverkas inte mycket när kontroller för familjebakgrund lades till modellen. Nordin beaktade också skillnader i studieinriktning (10 kategorier) för att se om resultaten berodde på att personer med olika förmåga läste olika inriktningar. Eftersom han samtidigt kontrollerade för om personerna avslutat gymnasiet samt hur många poäng de tagit på högskolan, är resultaten dock mycket svårtolkade.

Vi ser flera problem med Nordins studie. För det första använder han ett mått på individernas kognitiva förmåga som uppmätts vid 18–19 års ålder. Eftersom utbildningsval påverkas av kognitiva förmåga, samtidigt som utbildning förbättrar förmågorna (se avsnitt 5.2), finns en växelverkan mellan båda dessa variabler. Nordin uppmärksammade detta problem, och hävdade att det medför en underskattning av avkastningen på utbildning för personer längst ner i fördelningen. Det skulle därmed innebära att den stigande avkastningen på utbildning delvis är en artefakt, vilket han dock menar att olika känslighetsanalyser dock inte ger något stöd för. Vi tror dock att det är mycket svårt att på förhand uttala sig om vilka effekter detta problem ger, och att några enkla tester av resultatens trovärdighet inte finns.¹¹² För det andra beaktar han bara mycket grova skillnader i utbildningens kvalitet (10 inriktningskategorier). Eftersom personer med högre kognitiv förmåga kan tänkas söka sig till utbildningar och studieorter med högre kvalitet, och ha högre sannolikhet att bli antagna där, är det möjligt att interaktionseffekterna delvis fångar skillnader i utbildningskvalitet. Nordins resultat måste därmed tolkas med stor försiktighet.

Holmlund m.fl. (2008) studerade arbetsmarknadseffekten av studieuppehåll före högskolestudierna. Först beskrev de hur förekomsten av uppskjutna högskolestudier förändrats över tiden. De visade att andelen som gått direkt till högskolan efter gymnasiet har ökat kraftigt under början av 1990-talet, för att sedan avta något. När de analyserade vad som påverkar sannolikheten (och längden på) studieuppehållen, fann de att män, personer med låga betyg, de med svagare familjebakgrund och de som gått kortare gymnasieprogram hade större sannolikheten att komma in på högskolan senare. Det är därmed viktigt att beakta dessa skillnader för att kunna uttala sig om hur uppskjutna studier påverkar avkastningen på utbildning.

Analysen av arbetsmarknadseffekterna av uppskjutna högskolestudier baseras på personer födda 1966–72, som tagit ut examen från högskolan, och som började sina högskolestudier med olika långt uppehåll. De beaktade också skillnader i bland annat familjebakgrund och gymnasiebetyg, samt detaljerad information om studiernas inriktning. I genomsnitt ledde ett års uppskjutna högskolestudier till 2,9 (2,1) procents lägre inkomster vid 30 (35) års ålder. Med hjälp av skattningar för åldrarna 20–40 år, beräknade de sedan effekterna av uppskjutna högskolestudier över hela livscykeln. De fann då att ett års se-

¹¹² Ett alternativt tillvägagångssätt vore att beakta skillnader i utbildningslängd vid testtillfället, och bara studera skillnader i avkastning på utbildning som införskaffats efter mönstringen (huvudsakligen högskoleutbildning).

nareläggning av högskolestudierna medförde en sänkning av livsinkomsterna motsvarande 20 % av de genomsnittliga inkomsterna vid 40 års ålder. Att skjuta upp studierna är således förknippat med stora kostnader, både för individen och för samhället.

Vi har inte hittat några svenska studier som analyserat skillnader i avkastning på högskoleutbildning för personer med olika gymnasiebetyg eller högskoleprovsresultat. Därför presenterar vi i Tabell 10.2 egna skattningar av avkastningen på utbildning för olika urvalsinstrument. Analysen begränsas till personer med slutbetyg från gymnasiet, som påbörjat högskolestudier och med en årsinkomst över 135 000 kronor 2007. Analyserna av betyg och ålder vid högskoleinträdet baseras på personer födda 1966–70 (som är 37–41 år gamla år 2007), medan analysen av högskoleprovet baseras på personer födda 1973–77 (30–34 år) som skrivit högskoleprovet. Vi skattar dels den genomsnittliga avkastningen på högskolan, dels hur avkastningen på högskoleutbildning varierar med olika urvalsinstrument.

Tabell 10.2. Inkomstsamband för högskoleutbildning och olika urvalsinstrument

	(1)	(2)	(3)	(4)	(5)	(6)
Utbildningslängd	0,025*** (0,001)	0,018*** (0,001)	0,025*** (0,001)	0,020*** (0,001)	0,019*** (0,001)	0,018*** (0,001)
Gymnasiebetyg	0,101*** (0,001)	0,033*** (0,002)	. .	0,055*** (0,002)	. .	0,026*** (0,002)
Högskoleprov	. .	. .	0,004*** (0,001)	-0,007*** (0,001)	. .	. .
Ålder vid högskolestart	. .	. .	. .	. .	-0,016*** (0,001)	-0,013*** (0,001)
Utbildningslängd×Gymnasiebetyg	0,008*** (0,001)	0,001** (0,001)	. .	0,003*** (0,001)	. .	-0,000 (0,001)
Utbildningslängd×Högskoleprov	. .	. .	0,001* (0,001)	-0,001 (0,001)	. .	. .
Utbildningslängd×Ålder vid Högskolestart	. .	. .	. .	. .	-0,001*** (0,000)	-0,001*** (0,000)
Kontroll för inriktning×lärosäte		X	X	X	X	X
Antal personer	88 901	88 901	113 271	113 271	88 901	88 901

Noter: Tabellen visar regressionskattningar av sambandet mellan olika egenskaper och logaritmerad inkomst 2006. Alla analyser baseras på personer med slutbetyg från gymnasiet, som påbörjat högskolestudier och med en årsinkomst över 135 000 kronor 2006. Analyserna i kolumn (1) – (2) samt (5) – (6) begränsas till personer födda 1966-70. Analyserna i kolumnerna (3) – (4) begränsas till personer födda 1973-77 och som skrivit högskoleprovet. Gymnasiebetygen och högskoleprovsresultaten har standardiserats. Interaktionsvariablerna har beräknats genom att först dra bort medelvärdet från varje variabel, vilket innebär att skattningarna för variablerna var för sig avser genomsnittssambandet. Alla modeller kontrollerar för utbildningsnivå, kön och födelseår. Skattingen av utbildningslängd avser genomsnittseffekten av ett års utbildning, och kommer från en specifikation där utbildningslängd uttrycks i antal år.

För att undvika att skillnader i avkastning för personer med t.ex. olika höga betyg ska fånga skillnader i preferenser eller skillnader i avkastning mellan program och lärosäten, begränsas analysen till personer som påbörjat sina högskolestudier. Därmed är det möjligt att jämföra personer som påbörjat samma program/ämne vid samma högskola vid samma tillfälle. De skillnader i utbildningslängd som vi studerar har uppkommit på grund av att personer läst olika länge på högskolan. Vidare tar vi hänsyn till skillnader i avkastning på olika utbildningsnivåer, genom att tillåta att varje utbildningsnivå har en egen inkomsteffekt. Slutligen tar vi hänsyn till eventuella mätfel i utbildningsvariabeln, genom att utnyttja sambandet mellan utbildningsnivå och erhållna högskolepoäng. De skattningar vi presenterar avser den genomsnittliga avkastningen på högskoleutbildningen samt hur den varierar för personer som har olika betyg, högskoleprovsresultat samt ålder vid högskoleinträdet.

Kolumn (1) i Tabell 10.2 visar resultaten om man inte tar hänsyn till vilket program eller vilket lärosäte som studierna påbörjas. Avkastningen på högskoleutbildning uppmäts då till ungefär 2,5 %, vilket är nära de skattningar som Öckert (2010) presenterade (se kapitel 5). Vad som är av större intresse här är dock hur olika urvalsinstrument värderas på arbetsmarknaden. Första kolumnen visar att en standardavvikelse högre gymnasiebetyg är förknippad med 10 % högre inkomster. Interaktionen mellan gymnasiebetyg och utbildningslängd är liten men positiv: en standardavvikelse högre betyg medför en ökning av avkastningen på högskoleutbildning med 0,8 %. Detta kan tyda på att en fördelning av högskoleplatser på basis av gymnasiebetyg också medför att de som har mest att vinna på att studera också ges möjlighet att läsa på högskolan. Resultaten är också i linje med vad Nordin (2008) fann, även om avkastningen inte ökar särdeles mycket med gymnasiebetygen.

Kolumn (2) i Tabell 10.2 tar hänsyn till skillnader i påbörjat program/ämne, högskola samt tidpunkt. Det får två konsekvenser för skattningarna: sambandet mellan betyg och inkomster minskar kraftigt och högre betyg är bara svagt förknippat med högre avkastning på högskoleutbildning. Resultaten i kolumn (1) kan således till stor del förklaras av att personer med höga betyg läser utbildningar med hög avkastning, snarare än att avkastningen skulle öka betydligt med gymnasiebetyg.

Kolumn (3) i Tabell 10.2 studerar sambandet mellan högskoleprovet och inkomster. Precis som högskoleprovet har ett svagt samband med studieframgång, tycks det heller inte var förknippat med högre inkomster. Avkastningen på utbildning tycks dock öka svagt med högskoleprovsresultatet; en standard-

avvikelse bättre provresultat medför 0,1 % högre avkastning. I kolumn (4) tas också hänsyn till skillnader i gymnasiebetyg vilket dels gör att provresultatet är förknippat med lägre inkomster, dels att avkastningen på utbildning inte längre tycks öka med provresultat. Inte heller vad gäller högskoleprovet tycks platserna på högskolan ges till de med högst avkastning.

Kolumn (5) och (6) i Tabell 10.2 visar hur ålder vid högskolestudiernas början påverkar inkomsterna i allmänhet och avkastningen i synnerhet. Precis som Holmlund m.fl. (2008) fann, är en senare studiestart förknippad med betydligt lägre inkomster: att skjuta upp studierna ytterligare ett år medför ett inkomstbortfall på omkring 1,7 %. Detta beror antagligen på att personerna har hunnit spendera kortare tid på arbetsmarknaden efter studierna, samtidigt som arbetslivserfarenhet före högskolan inte värderas lika högt. Estimaten i Tabell 10.2 visar också att avkastningen på utbildning avtar något med tilltagande ålder vid högskoleinträdet, men effekten är mycket liten. Kolumn (6) beaktar även skillnader i gymnasiebetyg. Det minskar inkomstbortfallet av ett senarelagt högskoleinträde lite grand, antagligen för att personer som kommer in på högskolan senare i allmänhet har lägre gymnasiebetyg. Resultatet att avkastningen på högskoleutbildning avtar med ålder vid högskoleinträdet kvarstår dock.

Sammantaget tycks inte urvalsinstrumenten till högskolan välja ut de personer som har mest att vinna på att läsa vidare. Med undantag för de antagningsregler som ger företräde till äldre personer samt till dem med längre arbetslivserfarenhet, gäller dock inte det omvända heller; avkastningen minskar inte med urvalsinstrumenten. Eftersom avkastningen inte tycks skilja sig med avseende på urvalsgrund, borde urvalsinstrumentens lämplighet snarare baseras på deras prognosförmåga samt vilka grupper som väljs ut till högskolan.

Vi har i det här avsnittet visat att gymnasiebetyg är en bra indikator på studieframgång, även om denna urvalsgrund tenderar att befästa den könsmässiga och sociala selektion till högskolan. Vi har också visat att högskoleprovet har flera oönskade egenskaper; det förutsäger inte studieframgång utöver vad gymnasiebetygen gör och de förstärker social selektion till högskolan framförallt genom möjligheterna att skriva provet flera gånger. Dessutom tycks arbetslivserfarenhet eller ålder inte utgöra en god urvalsgrund till högskolan, dels eftersom det inte förutsäger studieframgång, dels eftersom personer som skjuter upp högskolestudierna får ett betydande inkomstbortfall.

10.4 Avkastning på högskolestudier vid olika lärosäten

Det finns stora löneskillnader mellan personer som läst olika utbildningsinriktningar vid olika lärosäten. Frågan är vad dessa löneskillnader beror på. En möjlig förklaring är att personer som läst olika utbildningar har förvärvat olika mycket produktiva kunskaper som i sin tur värdesätts på arbetsmarknaden. Löneskillnaderna skulle därmed ha orsakats av utbildningarnas produktiva innehåll. En annan – inte helt uteslutande – förklaring, är att personer som läst olika utbildningar redan från början hade olika produktiva egenskaper. Löneskillnaderna skulle därmed, åtminstone delvis, ha orsakats av ursprungliga skillnader mellan individer som gått olika utbildningar. Den empiriska utmaningen är att fastställa hur stor del av löneskillnaderna som har orsakats av skillnader i utbildningsinriktning och lärosäte, och hur stor del som beror på andra skillnader mellan personer som genomgått olika utbildningar.

Tre olika metoder används för att mäta kausala effekter av utbildningskvalitet på olika arbetsmarknadsutfall (jämför kapitel 5): (1) studier som beaktar observerade egenskaper, (2) studier som jämför syskon och tvillingar samt (3) studier som utnyttjar institutionella begränsningar. Gemensamt för alla tre ansatser är att de försöker isolera och mäta effekterna av skillnader i val av lärosäte som kan antas vara mer eller mindre slumpmässiga.

Måtten på utbildningskvalitet varierar i litteraturen. En del studier jämför studenter som läst vid olika lärosäten. Därmed fångas alla skillnader mellan olika utbildningsinstitutioner. Andra studier använder olika mått på studentkvalitet (betyg, högskoleprovsresultat, sökande per plats, etc.) och utbildningskvalitet (lärartäthet, andel disputerade, vetenskapliga publiceringar, undervisningstid, samverkan med näringsliv, etc.). Den internationella litteraturen om utbildningskvalitet studerar vanligen olika kvalitetsindikatorer medan de flesta svenska studierna jämför studenter som tagit examen från olika lärosäten.

10.4.1 Studier som beaktar observerade egenskaper

Urvalet till olika lärosäten sker dels genom individers val, dels genom antagningssystemet. Som diskuteras i kapitel 5 är det svårt att veta vilka faktorer som påverkar människors val av lärosäte, samtidigt som flera av de faktorer som vi tror är betydelsefulla kan vara svåra att få information om eller att mäta. Vad gäller urvalet av sökande till högskolan har vi betydligt bättre kunskap om

vilka faktorer som är viktiga; genomsnittsbetyg från gymnasiet och resultat på högskoleprovet.¹¹³ Dessa faktorer är också enkla att mäta och beakta.

Det finns flera internationella studier som försöker mäta den kausala löneeffekten av att gå på olika lärosäten genom att beakta skillnader i familjebakgrund och förmåga (se exempelvis Datcher-Loury och Garman 1995; Monks 2000; Dale och Krueger 2003; Black och Smith 2004; Black m.fl. 2005; Black och Smith 2006; Reynolds 2007). Även en del svenska studier finns, såsom Gustafsson (1996), Gartell och Regnér (2002, 2005) samt Eliasson (2006). Ett genomgående resultat från de svenska studierna är att löneskillnaderna mellan lärosäten minskar när fler faktorer som påverkar utbildningsvalen beaktas.

Eliasson (2006) finner, till skillnad från tidigare svenska studier, väldigt små löneskillnader mellan olika lärosäten när han beaktar skillnader i genomsnittsbetyg från gymnasiet samt andra faktorer. Ett problem med hans studie är dock att uppföljningsperioden är relativt kort, och att löneskillnader mellan lärosäten mycket väl kan uppkomma senare i arbetsmarknadskarriären.

Vi har valt att göra egna skattningar av sysselsättnings- och inkomstpremierna för studier vid olika lärosäten. Analysen baseras på personer födda 1965–73 som påbörjade sina högskolestudier för 26 års ålder, och avgränsas till inriktningar som finns representerade på alla lärosäten. Högskolestuderande jämförs med personer som inte läst på högskolan alls. Vi skattar sambandet mellan genomsnittligt antal studieår på det lärosäte där studierna påbörjades och sannolikheten att vara sysselsatt år 2006 samt log-inkomster för de som är sysselsatta. Till skillnad från de flesta tidigare studier tar vi således hänsyn till att studielängden kan variera systematiskt mellan olika lärosäten. Alla modeller kontrollerar för föräldrarnas utbildningslängd, gymnasiebetyg, samt fixa effekter för interaktionen mellan kön, födelseår, födelselän och gymnasielinje.

Den första kolumnen i Tabell 10.3 visar skillnader i sysselsättning för studenter som läst vid olika lärosäten. Ett års högskolestudier är förknippat med 1,8 procentenheter högre sannolikhet att arbeta i genomsnitt. Sambandet varierar i styrka mellan lärosäten; sambandet är starkast (ca. 2,5 procentenheter per år) för studenter vid universiteten i Karlstad, Växjö och Örebro samt de regionala högskolorna; sambandet än något svagare (1,3–1,7 procentenheter) för studerande vid Uppsala, Lund, Göteborg och Luleå. I de flesta fall är skillnaderna mellan dessa två grupper av lärosäten statistiskt säkerställda.

¹¹³ I Sverige har vi länge haft ett enhetligt antagningssystem som baseras på ett fåtal urvalsindikatorer. Situationen är långt mer komplex i många andra länder, där högskolorna själva sätter upp (hemliga) antagningskriterier (t.ex. i USA).

Tabell 10.3. Inkomstsamband för högskoleutbildning och olika urvalsinstrument

Lärosäte (Andel studenter, %)	Sysselsättning	Inkomster
Uppsala universitet (6,2)	0,013*** (0,002)	0,010*** (0,002)
Lunds universitet (10,8)	0,017*** (0,002)	0,011*** (0,002)
Göteborgs universitet (14,2)	0,016*** (0,002)	0,004*** (0,002)
Stockholms universitet (15,4)	0,021*** (0,002)	0,019*** (0,002)
Umeå universitet (4,9)	0,018*** (0,003)	0,009*** (0,003)
Linköpings universitet (9,3)	0,020*** (0,002)	0,015*** (0,002)
Luleå tekniska universitet (5,8)	0,014*** (0,002)	0,015*** (0,002)
Karlstads universitet (4,4)	0,025*** (0,003)	0,010*** (0,003)
Växjö universitet (4,7)	0,024*** (0,003)	0,012*** (0,003)
Örebro universitet (2,8)	0,024*** (0,004)	0,016*** (0,005)
Regionala högskolor (21,5)	0,026*** (0,002)	0,013*** (0,002)
Antal personer	147 256	124 491

Noter: Tabellen visar regressionsskattningar av sambandet mellan antal år vid olika lärosäten och sysselsättning samt logaritmerad inkomst år 2006. Analyserna baseras på personer födda 1965-73 med slutbetyg från studieförberedande gymnasielinjer. Högskolestuderande som påbörjat studierna före 26 års ålder jämförs med personer som inte läst på högskolan alls. Analysen begränsas till utbildningsinriktningar som finns representerade på alla högskolor. Personer med en årsinkomst över 135 000 kronor betraktas som sysselsatta, och inkomstanalyserna begränsas till sysselsatta. Alla modeller kontrollerar för föräldrarnas utbildningslängd, gymnasiebetyg, samt fixa effekter för interaktionen mellan kön, födelseår, födelseän och gymnasielinje. Effekten av gymnasiebetyg tillåts variera med gymnasielinje. Regionala högskolor utgörs av högskolorna i Borås, Dalarna, Gävle, Halmstad, Jönköping, Kalmar, Kristianstad, Mälardalen, Skövde, Sundsvall/Härnösand, Trollhättan/Uddevalla och Östersund. Stockholms universitet omfattar också KTH, KI och Lärarhögskolan, medan Göteborgs universitet också omfattar Chalmers tekniska högskola.

Den sista kolumnen i Tabell 10.3 presenterar skillnader i avkastning mellan olika lärosäten för dem som arbetar. Precis som Eliasson (2006) fann så tyder vår analys på att inkomstskillnaderna mellan olika lärosäten är relativt små. Högst avkastning (1,5–1,9 %) har studenter vid Universiteten i Stockholm, Linköping, Luleå och Örebro, medan avkastningen är något lägre (0,4–1,0 %) för studenter vid universiteten i Uppsala, Göteborg, Umeå och Karlstad. I en del, men inte alla, fall är skillnaden i avkastningen mellan olika lärosäten statistiskt säkerställd.

Ett allmänt problem med vår analys, liksom med tidigare studier som beaktar observerade skillnader, är att det uppmätta sambandet mellan lärosäte och inkomst bara kan ges en kausal tolkning under förutsättning att *alla* faktorer som påverkar både val av lärosäte och inkomst har beaktats. Vanligtvis har man bara tillgång till grundläggande uppgifter om personers uppväxtmiljö och tidigare studieprestationer. Rimligtvis kan även andra faktorer vara betydelsefulla för utbildningsbesluten.

10.4.2 Studier som jämför syskon (tvillingar)

För att beakta ytterligare faktorer som kan påverka utbildningsvalen mäter ett fåtal studier avkastningen på att läsa vid olika högskolor för syskon (t.ex. Lindahl och Regnér 2005). Syskon delar vanligen uppväxtmiljö och i genomsnitt halva genuppsättningen. Andra studier försöker mäta kausala effekter av utbildning genom att studera tvillingar, och då helst enäggstvillingar eftersom de är genetiskt identiska (se exempelvis Behrman m.fl. 1996). Man kan därmed konstanthålla alla ärftliga egenskaper som påverkar både utbildning och lön.

Lindahl och Regnér (2005) studerade inkomstskillnader för syskon födda 1951–64 med examen från olika högskolor. När de jämförde personer som studerat vid olika högskolor och tog hänsyn till skillnader i studiernas nivå och inriktning fann de att de som tagit examen från något av de äldre universiteten (Uppsala, Lund, Stockholm, Göteborg eller Umeå) hade omkring 7 % högre inkomster än de som studerat vid en annan högskola. Om de istället studerade skillnader i inkomster mellan *syskon inom samma familj* som studerat vid olika lärosäten, sjönk effekten av att läsa vid ett universitet till omkring 4 %. Detta tyder på att en stor del av inkomstskillnaderna mellan personer som läst vid olika lärosäten kan förklaras av skillnader i arv och uppväxtmiljö.

Frågan är om syskon- och tvillingstudierna mäter kausala effekter av utbildningskvalitet. Huvudinvändningen mot studier som beaktar icke-observerade skillnader är densamma som för studier som beaktar observerade skill-

nader, nämligen obeaktade skillnader mellan personer som gör olika utbildningsval. Ett flertal forskare har kritiserat syskon- och tvillingstudierna, för att många viktiga faktorer som påverkar utbildningsval och lön inte har beaktats (se t.ex. Neumark 1999; Bound och Solon 1999). Tvillingstudier tampas också med en svårlöst paradox. Å ena sidan är fördelen med att studera tvillingar att de är *lika*, å andra sidan kan vi bara mäta effekten av att läsa vid olika högskolor för tvillingar som väljer *olika*. För att avkastningen på utbildningskvalitet i tvillingstudier ska kunna tolkas kausalt, måste man anta att alla skillnader i val av lärosäte inom ett tvillingpar är slumpmässiga. Frågan är förstås varför lika tvillingar väljer olika, och om det finns obeaktade skillnader i produktiva egenskaper som ger upphov till skillnaderna i utbildningsval.

10.4.3 Studier som utnyttjar begränsningar i utbildningsvalen

Den tredje typen av studier väljer ett helt annat angreppssätt. Istället för att försöka mäta och beakta observerade och icke-observerade skillnader mellan personer som läst vid olika högskolor, försöker de utnyttja situationer då utbildningsval (av olika anledningar) är begränsade. Det kan exempelvis handla om skillnader i kostnaden att läsa vid olika högskolor, såsom avståndet till högskolor av olika kvalitet eller skillnader i studieavgifter mellan högskolor (Card 1995; Brewer m.fl. 1999). Om dessa begränsningar enbart påverkar sannolikheten att läsa vid lärosäten av olika kvalitet, men *inte* har någon direkt inverkan på personernas inkomster, kan man relatera löneskillnaderna för personer med och utan begränsningar till motsvarande skillnader i sannolikhet att läsa vid högskolor av olika kvalitet.

Öckert (2010) är den enda svenska studie som utnyttjar institutionella begränsningar i utbildningsvalen för att mäta effekter av utbildningskvalitet. Han använde uppgifter om sökande till högskolan 1982, och jämför inkomsterna för sökande som precis klarat respektive missat antagningsgränsen till högskolor och linjer av hög kvalitet till motsvarande inkomstskillnad för sökande till högskolor med lägre kvalitet. Han finner dock inga effekter av utbildningskvalitet på inkomster.

Sammantaget tycks avkastningen på högskoleutbildning inte skilja sig särdeles mycket mellan olika högskolor. Om man jämför studenter som tagit examen vid olika lärosäten finns betydande inkomstskillnader. Dessa minskar betydligt om hänsyn tas till skillnader i gymnasiebetyg eller om man utnyttjar institutionella begränsningar i utbildningsvalen. Eftersom studierna fångar upp

såväl utbildningskvalitet som studentkvalitet, tyder de också på att kamrat-effekterna på högskolan är förhållandevis små.

En allmän brist med tidigare studier är att de ofta jämför arbetsmarknadsutfall för personer som tagit examen från olika lärosäten (undantaget är Öckert 2010). Detta kan säkert motiveras i länder där andelen som fullföljer sina studier är hög. I Sverige, däremot, är andelen som tar ut examen relativt låg; ungefär 70 % av dem som påbörjar ett utbildningsprogram har tagit ut en examen 11 år senare. Motsvarande siffra för fristående kurser är 15 % (SCB 2005). Både ur individens och ur samhällets synvinkel är det viktigt att studera dem som inte fullföljer. Dessutom kan utbildningens kvalitet i sig påverka hur stor andel som tar ut sin examen. Studenter vid lärosäten med hög kvalitet kanske har lägre sannolikhet att avbryta sina studier av studierelaterade orsaker, samtidigt har de också större chans att få ett arbete utan att de har hunnit ta ut en examen. Andelen som blir klara är således också en typ av kvalitetsmått, även om nettoeffekten är oklar. En studie av grundutbildningens kvalitet bör därför följa personer som påbörjar sina studier vid inom samma eller jämförbara utbildningsprogram med vid olika lärosäten.

10.5 Avkastning på högskolestudier med olika inriktning

Ett relativt utforskat område – både i Sverige och internationellt – är hur avkastningen på utbildning varierar mellan olika utbildningsinriktningar. Det finns en litteratur som behandlar i vilken utsträckning studenter vid olika inriktningar får arbete inom utbildningens ”målyrke” (se avsnitt 10.1.3 om överutbildning). När det gäller rena avkastningsstudier är dock kunskapen betydligt mer begränsad. I detta avsnitt presenterar vi egna skattningar av sambanden mellan utbildningsinriktning, sysselsättning och inkomster.

Analysen av arbetsmarknadssituationen för studenter som läst olika utbildningsinriktningar baseras på personer födda 1965–73 som påbörjade sina högskolestudier före 26 års ålder. Bara studenter som påbörjat sina studier vid någon av universiteten eller vid de regionala högskolorna ingår i analysen (jämför avsnitt 10.4.1). Högskolestuderande jämförs med personer som inte läst på högskolan alls. Vi skattar sambandet mellan genomsnittligt antal studieår på den inriktning där studierna påbörjades och sannolikheten att vara sysselsatt år 2006 samt log-inkomster för de som är sysselsatta. Alla modeller kontrollerar för föräldrarnas utbildningslängd, gymnasiebetyg, samt fixa effekter för interaktionen mellan kön, födelseår, födelselän och gymnasielinje.

Tabell 10.4. Sysselsättning och inkomster för högskolestudier med olika inriktning

Lärosäte (Andel studenter, %)	Sysselsättning	Inkomster
Förskollärare/fritidspedagog (7,2)	0,021*** (0,002)	-0,049*** (0,001)
Grundskollärare (8,5)	0,033*** (0,001)	-0,038*** (0,002)
Ämneslärare (6,3)	0,025*** (0,001)	-0,029*** (0,001)
Läkare/Tandläkare (3,0)	0,028*** (0,002)	0,060*** (0,003)
Civilingenjör (43,1)	0,013*** (0,001)	0,034*** (0,001)
Biolog/Kemist/Fysiker (2,1)	0,003 (0,003)	-0,006* (0,003)
Systemutvecklare/Programmerare (3,0)	0,020*** (0,002)	0,048*** (0,002)
Jurist (4,9)	0,019*** (0,002)	0,050*** (0,002)
Psykolog/Socionom (2,1)	0,028*** (0,002)	-0,021*** (0,002)
Samhällsvetare/Kulturvetare (2,9)	0,011*** (0,002)	0,001 (0,003)
Civilekonom (17,0)	0,012*** (0,001)	0,056*** (0,001)
Antal personer	156 932	132 832

Noter: Tabellen visar regressionsskattningar av sambandet mellan antal år vid olika högskoleinriktningar och sysselsättning samt logaritmerad inkomst år 2006. Analyserna baseras på personer födda 1965-73 med slutbetyg från studieförberedande gymnasielinjer. Högskolestuderande som påbörjat studierna före 26 års ålder jämförs med personer som inte läst på högskolan alls. Analysen begränsas till de lärosäten som återfinns i Tabell 10.3. Personer med en årsinkomst över 135 000 kronor betraktas som sysselsatta, och inkomstanalyserna begränsas till sysselsatta. Alla modeller kontrollerar för föräldrarnas utbildningslängd, gymnasiebetyg, samt fixa effekter för interaktionen mellan kön, födelseår, födelselän och gymnasielinje. Effekten av gymnasiebetyg tillåts variera med gymnasielinje.

Tabell 10.4 visar sambandet mellan utbildningsinriktning och studenternas arbetsmarknadssituation 15–20 år efter högskoleinträdet. Den första kolumnen beskriver skillnaderna i sysselsättning mellan personer som läst olika länge på olika inriktningar. Allra starkast är sysselsättningssambanden (2,5–3,3 procentenheter per utbildningsår) för grundskollärare, ämneslärare, läkare/tandläkare och psykologer/socionomer. Betydligt svagare samband mellan utbildningslängd och sysselsättning (0,3–1,3 procentenheter) finns för utbildningar för biologer/kemister/fysiker, civilingenjörer, samhällsvetare/kulturveta-re och civilekonomer.

Den andra kolumnen av Tabell 10.4 visar sambanden mellan studier med olika inriktning och inkomster för de som arbetar. Även här finns betydande skillnader. Högst avkastning på högskoleutbildning har läkare/tandläkare på 6,0 %. Därefter följer civilekonomer med en inkomstpremie på 5,6 % per studieår. Även jurister och systemutvecklare/programmerare har hög avkastning på sina studier. I botten ligger lärare med stora negativa inkomsteffekter för varje ytterligare studieår. Även psykologer/socionomer, biologer/kemister/fysiker och samhällsvetare/kulturveta-re har avkastningar som är negativa eller nära noll.

10.6 Sammanfattning

Vi har i det här kapitlet diskuterat några av de utbildningspolitiska avvägningarna vad gäller högre utbildning. I avsnitt 10.1 gick vi igenom litteraturen om överutbildning. Vi pekade på bristen på tydlig teoretisk förankring och svårigheterna med empiriska studier av överutbildning. Eftersom avkastningen på utbildning i Sverige har ökat sedan början av 1980-talet, tycks inte överutbildning vara ett ökande problem (vilket ibland hävdas).

Avsnitt 10.2 behandlade de faktorer som bestämmer studenternas efterfrågan på studieplatser. Trots omfattande ansträngningar tycks den sociala selektionen till högskolan inte ha förändrats nämnvärt under de senaste 20 åren. Däremot har andelen kvinnor ökat kraftigt, och idag är 60 % av alla studenter kvinnor. Tidigare forskning och egna analyser tyder på att den sociala selektionen till högskolan går att spåra ända tillbaks till de tidiga skolåren, och att skillnaderna i skolprestationer mellan dem som så småningom kommer att söka respektive inte söka till högskolan förstärks under hela grundskolan. Val av gymnasieprogram och gymnasiebetyg förklarar bara en liten del av den sociala selektionen till högskolan, vilket tyder på att insatser måste sättas in tidigt för att nya grupper ska kunna lockas till högskolan. Två tredjedelar av könsskill-

naden i sannolikheten att läsa på högskolan beror på skillnader i betygssättningen i åk 9 för elever med samma resultat på de nationella proven. För kvinnor och män med samma grundskolebetyg finns ingen skillnad i sannolikheten att läsa på högskolan.

Vi har gått igenom betydelsen av regionala högskolor och studiemedelsystemet för att bryta den sociala snedrekryteringen till högskolan. Både evidensen om studiemedelsystemets samt regionaliseringens rekryteringseffekter är tämligen svag: Ingen av de studier som baserats på svenska data har använt en trovärdig strategi för att mäta orsakssamband. Eftersom studiemedelsystemet och regionaliseringen av högskolan ofta framhålls som de kanske viktigaste åtgärderna för att motverka den sociala selektionen till högskolan i Sverige, är det angeläget med mer forskning på detta område.

Vi har ingående diskuterat och analyserat konsekvenserna av antagningssystemet till högskolan. Urvalet till högre studier sker huvudsakligen på basis av olika förmågeindikatorer. Vår genomgång visar att gymnasiebetygen kan prognostisera studenternas studieprestationer väl. En standardavvikelse högre gymnasiebetyg är förknippad med 14 procentenheters högre sannolikhet att ta en kandidatexamen efter 5 år på högskolan. Däremot är prognosvärdet för resultatet på högskoleprovet för personer med samma gymnasiebetyg obefintligt. Högskoleprovets dåliga prognosförmågan är synnerligen problematisk då 25 % av alla studenter antas via högskoleprovet. Våra beräkningar visar att andelen som tar examen skulle kunna öka med 1,5–2 procentenheter om högskoleprovet helt slopades som urvalsinstrument. Inte heller arbetslivserfarenhet (ålder) kan förutsäga studieframgång.

Valet av urvalsinstrument påverkar också sammansättningen av studenterna på högskolan. Antagning på basis av gymnasiebetyg eller högskoleprovsresultat medför att personer från studieovana hem kommer att vara underrepresenterade på högskolan, eftersom föräldrabakgrund har stor betydelse för barnens studieresultat. Högskoleprovet förvärrar den sociala selektionen till högskolan ytterligare. Det beror dels på att det är möjligt att skriva högskoleprovet flera gånger – en möjlighet som barn till högutbildade föräldrar utnyttjar i högre grad – och dels på att det sker betydande förbättringar i resultaten för varje nytt försök. Däremot tycks antagning som baseras på erfarenhet (ålder) att i större utsträckning attrahera barn till lågutbildade föräldrar. Att använda gymnasiebetygen som urvalsgrund medför också att andelen kvinnor blir högre, eftersom de i allmänhet presterar bättre i skolan än män. Högskoleprovet tycks å andra

sidan gynna män, varför användandet av detta prov vid antagningen kan tänkas minska kvinnornas dominans på högskolan.

Urvalskriterierna till högskolan betyder inte så mycket för avkastningen på högskoleutbildning. Personer med höga betyg eller goda provresultat tjänar inte så mycket mer på att utbilda sig än personer med sämre betyg eller resultat. Det stora undantaget gäller personer som antagits på basis av arbetslivserfarenhet (ålder). Äldre studenter har betydligt lägre avkastning på sina högskolestudier. Det beror antagligen dels på att deras uteblivna inkomster under studietiden är högre, dels på att värdet på arbetslivserfarenhet före högskolestudierna inte värderas så högt. Ett försenat högskoleinträdet med ett år är förknippat med 1,3 % lägre inkomster.

Vi har vidare diskuterat och analyserat skillnader i utbildningskvalitet mellan olika lärosäten. Även om lärosätet har viss betydelse för skillnader i sysselsättning och löner, så är dessa skillnader relativt små. Dessa analyser tyder således inte på betydande kvalitetsskillnader mellan svenska högskolor. Det tycks däremot finnas stora skillnader i avkastning mellan olika utbildningsinriktningar. En del utbildningar är förknippade med relativt höga avkastningar, medan andra inriktningar snarast har negativ avkastning. Skillnader i avkastning mellan olika utbildningsinriktningar är emellertid ett mycket utforskat område, varför fler studier är angelägna.

Referenser

- Behrman, J., M. Rosenzweig och P. Taubman (1996), College choice and wages: Estimates using data on female twins, *Review of Economics and Statistics*, 78, 672–685.
- Black, D. och J. Smith (2004), How robust is the evidence on the effects of college quality? Evidence from matching, *Journal of Econometrics*, 121, 99–124.
- Black, D. och J. Smith (2006), Estimating the returns to college quality with multiple proxies for quality, *Journal of Labor Economics*, 24, 701–728.
- Black, D., K. Daniel och J. Smith (2005), College quality and wages in the United States, *German Economic Review*, 6, 415–443.
- Boalt, G. (1947), *Skolutbildning och skolresultat för barn ur olika samhällsgrupper i Stockholm*. Stockholm: P.A. Norstedt & Söner.
- Bornmalm-Jardelöw, G. (1988), *Högre utbildning och arbetsmarknad*. Doktorsavhandling, Pedagogiska institutionen, Göteborgs universitet.
- Bound, J. och G. Solon (1999), Double trouble: on the value of twins-based estimation of the return to schooling, *Economics of Education Review*, 18, 169–182.
- Brewer, D., E. Eide och R.G. Ehrenberg (1999), Does it pay to attend an elite private college? Cross-cohort evidence on the effects of college type on earnings, *Journal of Human Resources*, 34, 104–123.
- Carneiro, P. och J. Heckman (2003), Human Capital Policy, ingår i J. Heckman och A. Krueger (red.), *Inequality in America: What Role for Human Capital Policies*. Cambridge MA: MIT Press.
- Cliffordson, C. (2008), Differential prediction of study success across academic programs in the Swedish context: The validity of grades and tests as selection instruments for higher education, *Educational Assessment*, 13, 56–75.
- Centrala studiestödsnämnden (2006), Uppföljning av 2001 års studiestödsreform, rapport.
- Centrala studiestödsnämnden (2008), CSN:s enkätundersökning av de studerandes ekonomiska och sociala situation, rapport.

- Cliffordson, C. (2004), Effects of Practice and Intellectual Growth on Performance on the Swedish Scholastic Aptitude Test (SweSAT), *European Journal of Psychological Assessment*, 20, 192–204.
- Dale, S. och A. Krueger (2002), Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables. *Quarterly Journal of Economics*, 117, 1491–1527.
- Datcher-Loury, L. och Garman, D. (1995), College selectivity and earnings. *Journal of Labor Economics*, 13, 289–309.
- Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp och G. Wagner (2006), Individual risk attitudes: New evidence from a large, representative, experimentally-validated survey, IZA DP 1730.
- Dryler, H. (1994), Etablering av nya högskolor - ett medel för minskad snedrekrytering, i R. Erikson och J. Jonson (red.), *Skola och sortering. Studier av snedrekrytering och utbildningens konsekvenser*. Stockholm: Carlssons förlag.
- Dynarski, S. (2000), Hope for whom? Financial aid for the middle class and its impact on college attendance, *National Tax Journal*, 53, 629–661.
- Dynarski, S. (2003), Does aid matter? Measuring the effect of student aid on college attendance and completion, *American Economic Review*, 93, 279–288.
- Eliasson, K. (2006), *College choice and earnings among university graduates in Sweden*, Avhandling, Umeå Economic Studies no. 693.
- Erikson, R. och J. Jonsson (1993), Ursprung och utbildning. Social snedrekrytering till högre utbildning. (SOU 1993:85). Stockholm: Fritzes.
- Erikson, R. och J. Jonsson (2002), Varför består den social snedrekryteringen?, *Pedagogisk forskning i Sverige*, 7, 210–217.
- Erikson, R., O. Nordström Skans, A. Sjögren, O. Åslund (2006), *Fritt inträde? Ungdomars och invandrades väg till det första arbetet*, Valfärdsrådets rapport 2006. Stockholm: SNS Förlag.
- Fredriksson, P. (1997), Economic incentives and the demand for higher education, *Scandinavian Journal of Economics*, 99, 129–142.

- Fägerlind, I. (1975). *Formal education och adult earnings: A longitudinal study on the economic benefits of education*. Stockholm: Almqvist & Wiksell.
- Gartell, M. och H. Regnér (2005). Sambandet mellan val av högskola och inkomster efter examen för kvinnor och män, IFAU Rapport 2005:12.
- Gartell, M. och H. Regnér (2002), Arbetsmarknaden för högskoleutbildade. Inkomstutveckling och geografisk rörlighet under 1990-talet, SACO, Stockholm.
- Groot W. och H. Maassen van den Brink (2000), Overeducation in the labour market: a meta-analysis, *Economics of Education Review*, 19, 149–158.
- Gustafsson J-E. och M. Benjegård (1996), Högskoleprovet, social bakgrund och upprepade provtagning, ingår i ” Ingår i Högskoleprovet - Genom elva forskares ögon, Högskoleverkets rapportserie 1996:22 R.
- Gustafsson, J-E., och A. Westerlund (1994). Socialgruppskillnader i prestationer på Högskoleprovet. I R. Erikson och J.O. Jonsson (red.) *Sorteringen i skolan*. Stockholm: Carlssons.
- Gustafsson, J-E. och S-E. Reuterberg (2000), Metodproblem vid studier av Högskoleprovets prognosförmåga – och deras lösning! *Pedagogisk Forskning i Sverige*, 5, 273–283.
- Gustafsson, J-E., A. Andersson och M. Hansen (2000), Prestationer och prestationsskillnader i 1990-talets skola, Institutionen för pedagogik och didaktik, Göteborgs universitet. PM 2000-02-13.
- Gustafsson, L. (1996), Vilken högskola är bäst? En empirisk analys av de svenska ekonomutbildningarna, Statistiska centralbyrån, Bakgrundsfakta till arbetsmarknads- och utbildningsstatistiken 1996:1.
- Hartog, J. (2000), Over-education and earnings: Where are we, where should we go?, *Economics of Education Review*, 19, 131–147.
- Henriksson, W. (1985). Effekter av övning och instruktion på testprestation, i W. Henriksson, S. Henrysson, C. Stage och I. Wedman (red.), *Prov för urval till högskolan*. SOU 1985: 57.
- Henriksson, W. och S. Henriksson (1991). Effekter av upprepade provtagning. PM 40. Pedagogiska institutionen, Umeå Universitet.

- Henriksson, W. (1995). Repeated test taking and the SweSAT. PM 13. Pedagogiska institutionen, Umeå Universitet.
- Henriksson, W och K. Bränberg (1994), The effects of practice on the Swedish Scholastic Aptitude Test (SweSAT). *Scandinavian Journal of Educational Research*, 38, 129–148.
- Henriksson, W. och S. Wolming (1998), Academic performance in four study programmes: a comparison of students admitted on the basis of GPA and SweSAT scores, with and without credits for work experience, *Scandinavian Journal of Educational Research*, 42, 135–150.
- Holm, E. och N. Häggström (1972), Regional rekrytering till universitet och högskola under 1960-talet, i SOU 1972:23.
- Holzer, S. (2009), *University Choice, Equality, and Academic Performance*, Doktorsavhandling, Växjö universitet.
- Holmlund B, Q. Liu och O. Nordström Skans (2008), Mind the gap? Estimating the effects of postponing higher education, *Oxford Economic Papers*, 60, 683–710.
- Holmlund, H., M. Lindahl och E. Plug (2008), The causal effect of parent's schooling on children's schooling: A comparison of estimation methods, IZA DP 3630.
- Häkkinen, I. och R. Uusitalo (2003), The effect of a student aid reform on graduation: A duration analysis, Working Paper Series 2003:8, Department of Economics, Uppsala University.
- Högskoleverket (1997), Grundläggande högskoleutbildning: Politik och planering eller den osynliga handen i full verksamhet?, Högskoleverkets skriftserie 1997:2S.
- Högskoleverket (2001), Högskoleprovets prognosvärde. Samband mellan provresultat och framgång första studieåret vid civilingenjörs-, jurist- och grundskolläraryrkesutbildningarna, Rapport 2001:19.
- Johansson, M. och K. Katz (2007), Wage differences between women and men in Sweden - the impact of skill mismatch, IFAU Working Paper 2007:13.
- Kane, T. (1995), Rising public college tuition and college entry: How well do public subsidies promote access to college? NBER WP 5164.

- Kjellström, C. och H. Regnér (1999a), The Effects of Geographical Distance on the Decision to Enroll in University Education, *Scandinavian Journal of Educational Research*, 43, 335–348.
- Kjellström, C. och H. Regnér, (1999b), Does Distance to University Affect Enrollment Decisions? Evidence from Data on Three Cohorts of Sweden, ingår i C. Kjellström, *Essays on Investment in Human Capital*, Swedish Institute for Social Research – Dissertation Series No. 36.
- Korpi, T. och M. Tåhlin (2009), Educational mismatch, wages, and wage growth: Overeducation in Sweden, 1974–2000, *Labour Economics*, 16, 183–193.
- le Grand, C., R. Szulkin och M. Tåhlin (2001), Lönestrukturens förändring i Sverige, i J. Fritzell, M. Gähler och O. Lundberg (red.) *Välfärd och arbete i arbetslöshetens årtionde*, SOU 2001:53. Stockholm: Fritzes.
- le Grand, C., R. Szulkin och M. Tåhlin (2004), Överutbildning eller kompetensbrist? Matchning på den svenska arbetsmarknaden 1974–2000, ingår i M. Bygren, M. Gähler och M. Nermo (red.) *Familj och arbete. Vardagsliv i förändring*. Stockholm: SNS Förlag.
- Lindahl, L. och H. Regnér (2005), College choice and subsequent earnings. Results using Swedish sibling data. *Scandinavian Journal of Economics*, 107, 437–457.
- McPherson, M. och M. Shapiro (1991), Does student aid affect college enrolment? New evidence on a persistent controversy, *American Economic Review*, 81, 309–318.
- Monks, J. (2000), The returns to individual and college characteristics. Evidence from the National Longitudinal Survey of Youth. *Economics of Education Review*, 19, 279–289.
- Neumark, D., (1999), Biases in twin estimates of the return to schooling. *Economics of Education Review*, 18, 143–148.
- Nielsen, H.S., T. Sørensen och C. Taber (2008), Estimating the effect of student aid on college enrollment: Evidence from a government grant policy reform, NBER WP 14535.

- Nordin, M. (2008), Ability and rates of return to schooling – making use of the Swedish enlistment battery test, *Journal of Population Economics*, 21, 703–717.
- Nordin, M., I. Persson och D-O. Rooth (2008), Education-occupation mismatch: Is there an income penalty?, IZA DP 3806.
- Reuterberg, S-E. (1996), Den sociala bakgrundens betydelse för prestationer på Högskoleprovet. Ingår i Högskoleprovet – Genom elva forskares ögon, Högskoleverkets rapportserie 1996:22 R.
- Reuterberg, S.-E., and Svensson, A. (1992), *Social bakgrund : studiestöd och övergång till högre studier : delbetänkande Utredningen om den sociala snedrekryteringen till högre studier*. Statens offentliga utredningar, 1992:122. Stockholm: Allmänna förlaget.
- Reynolds, C.L. (2007), Where to attend? Estimating the effects of beginning at a two-year college, manuskript, University of Michigan.
- SCB (2005), Universitet och högskolor. Genomströmning och resultat i högskolans grundutbildning t.o.m. 2003/04, Statistiska meddelanden, UF 20 SM 0502.
- SCB (2009), Behöriga förstahandssökande och antagna till program efter inriktning höstterminerna 1998–2008, nedladdad 2009-08-24 (<http://www.scb.se/statistik/UF/UF0206/2008A02/Tabell%204.xls>).
- SOU 1948:42, *Betänkande och förslag angående studentens sociala stödåtgärder*, Ecklesiastikdepartementet.
- SOU 1973:2, *Högskolan. Betänkande av 1968 års utbildningsutredning*, Utbildningsdepartementet.
- Stage, C. (1996), Kvinnor och män. Ingår i Högskoleprovet – Genom elva forskares ögon, Högskoleverkets rapportserie 1996:22 R.
- Svensson, A., J-E. Gustafsson och S-E. Reuterberg (2001), Högskoleprovets prognosvärde. Samband mellan provresultat och framgång första studieåret vid civilingenjör-, jurist och grundskollärarytbildningarna, Högskoleverkets rapportserie 2001:19R.
- Turner, S. (1998), The vision and reality of Pell grants: Unforeseen consequences for students and institutions. I L. Gladieux (red.) *Memory*,

Reason, Imagination: A Quarter Century of Pell Grants. Washington, D.C.: College Board.

- van der Klaauw, W. (2002), Estimating the effect of financial aid offers on college enrolment: A regression-discontinuity approach, *International Economic Review*, 43, 1249–1287.
- Wedman, I. och S. Henrysson (1992), Överväganden om könsmässig och social snedrekrytering vid Högskoleprovets tillkomst och utformning. I *Betyg och högskoleprov för män och kvinnor*. (UHÄ-rapport 1992:03). Stockholm:
- Willingham, W., C. Lewis, R. Morgan, och L. Ramist (1990), Predicting college grades. An analysis of institutional trends over two decades. Princeton, N.J.: Educational Testing Service.
- Åberg, R. (2003), Unemployment persistency, over-education and the employment chances of the less educated, *European Sociological Review*, 19, 199–216.
- Öckert, B (2010), What's the value of an acceptance letter? Using admissions data to estimate the return to college, *Economics of Education Review*, 29, 504-516 .
- Öckert, B. och H. Regnér (2000), Högre utbildning i Sverige – En problemorienterad diskussion om utbildningssatsningar, Institutet för social forskning.

11 Avslutning

Detta kapitel börjar med att sammanfatta lärdomarna från forskningen. Sedan diskuterar vi fyra större frågor med utgångspunkt från denna sammanfattning: Är avkastningen på utbildning för låg i Sverige? Varför har svenska grundskoleelevers resultat försämrats sedan mitten av 1990-talet? Vilka kriterier bör ett välfungerande betygssystem tillfredsställa? Är skolan en kommunal eller central angelägenhet? Till sist kommer vi att diskutera förutsättningarna för forskning och utvärdering.

11.1 Sammanfattning av forskningsgenomgången

Som en utgångspunkt för denna forskningsgenomgång skissade vi en modell som i vårt tycke sammanfattar de utbildningspolitiska avvägningarna på ett intressant sätt. Individer har olika förmågor men också olika familjebakgrund. Utan utbildningspolitiska interventioner kommer barn med samma förmåga att få olika utbildning: barn med bättre ekonomiska förutsättningar får en högre utbildning än de med sämre förutsättningar. En uppgift för utbildningspolitiken är att motverka detta mönster. Det är dock viktigt att komma ihåg att det varken är realistiskt eller önskvärt att barn ska ha samma utbildning oavsett familjebakgrund. Föräldrar med olika utbildningsbakgrund har olika förutsättningar att hjälpa barnen med skolarbetet. Och det krävs interventioner som medborgarna knappast är beredda att acceptera för att utjämna sådana skillnader. Dessutom samvarierar barnens förmåga med familjebakgrund på ett sätt som gör att det inte är effektivt att sträva efter att eliminera alla skillnader. Vi kan alltså förvänta oss att familjebakgrunden har mycket stor betydelse för barns utbildningsutfall. Den relevanta frågan är hur mycket politiken ska och kan utjämna skillnaderna i utbildningsresultat för barn med olika familjebakgrund.

Modellen har implikationen att tidiga insatser, särskilt riktade mot barn med sämre ekonomiska förutsättningar, är det bästa sättet att stimulera både effektivitet och jämlikhet. Detta innebär dock inte att avkastningen på vuxenutbildning alltid är låg. Även i den stiliserade tankeram som modellen utgör finns ett argument för s.k. andrachansutbildning.

Med denna modell som bakgrund har vi gått igenom olika politikåtgärder från förskola till högre utbildning. Förskolan byggdes ut mycket kraftigt under framförallt 1970-talet. Denna utbyggnad hade sannolikt inte utbildningspolitiska motiv utan drevs framförallt av en ökad efterfrågan på barnomsorg i samband med kvinnors inträde på arbetsmarknaden. Kopplingen till sysselsättning

innebär att barn i förskolan tenderat att ha fördelaktigare familjeegenskaper än barn som inte gått på förskolan.

Från mitten av 1990-talet sker dock en viss inriktningsförändring. Ett första steg i denna process är att även barn till arbetslösa får möjlighet att gå i förskola. Ett ytterligare steg tas genom att en allmän förskola införs, till en början från 4 års ålder, men numera gäller allmän förskola från 3 års ålder. Förskolan har alltså över tiden kommit att nå en bredare grupp. Denna utveckling motverkas dock av introduktionen av vårdnadsbidraget som skapar incitament för föräldrar med sämre arbetsmarknadsutsikter att vara hemma med sina barn.

Förskolan skulle alltså kunna vara ett medel för att utjämna förutsättningsarna för barn med olika bakgrund. Finns det då någon evidens som tyder på att så är fallet? Tyvärr finns mycket få studier som trovärdigt skattat effekten av förskola, både generellt och för utsatta grupper, med hjälp av svenska data. Det finns dock resultat från våra grannländer som kan vara överförbara på svenska förhållanden. En färsk studie från Norge tyder på utbildningsnivån ökar som ett resultat av att man gått på förskolan och att effekterna är särskilt starka för barn som har lågutbildade mödrar (se Havnes och Mogstad 2009).

Effekter av förskolan framstår som ett mycket angeläget område för ytterligare forskning. Eftersom nästan alla barn idag har någon erfarenhet av förskola skulle det vara av stort intresse att studera betydelsen av verksamhetens innehåll: Vilken typ av förskolepedagogik främjar barnen på lång sikt? Det finns idag inga studier som kan sprida ljus över denna fråga.

En förklaring till bristen på svenska studier är att det saknats centralt insamlad information om vilka som deltar i förskoleverksamheten. Forskarna har fått förlita sig till specialundersökningar som Levnadsnivåundersökningen (LNU) och data från projektet Utvärdering genom uppföljning (UGU). Dessa datamängder innehåller dock endast rudimentär information om förskolegång vilket begränsat möjligheterna att skatta orsakssamband på ett trovärdigt sätt.

Det finns något flera svenska studier av betydelsen av antalet lärare och lärarnas egenskaper för barns skolutfall. De bästa studierna på området tyder på att antalet lärare per elev (eller mindre klasser) har en positiv effekt på elevprestationer. Det övergripande resultatmönstret tyder på att effekten är särskilt stor för yngre elever och för elever med sämre förutsättningar. Eftersom skillnaderna i ekonomiska förutsättningar är mindre i Sverige än i många andra industrialiserade länder kan den genomsnittliga effekten av att öka antalet lärare per elev vara något mindre i Sverige än i t.ex. USA.

De stora frågorna när det gäller satsningar på att öka lärartätheten är om intäkterna överstiger kostnaderna och om det finns tillräckligt med lärarkompetens för att öka lärartätheten med bibehållen undervisningskvalitet.

Det är inte bara lärartätheten som är viktig. Det har också stor betydelse vilken lärare eleven har: ungefär 10 % av variationen i resultat mellan elever har att göra med vilken lärare de undervisas av. Dessa lärareffekter har inte så mycket att göra med sådana egenskaper som vi regelmässigt observerar i data (t.ex. utbildningsnivå). Däremot verkar lärarnas uppmätta ämnesdidaktiska kunskaper vara av relativt stor betydelse. Även om detta kan tyckas vara en självklarhet är resultatet potentiellt betydelsefullt. En viktig fråga i detta sammanhang är i vilken utsträckning som lärarutbildningen ger dessa ämnesdidaktiska kunskaper. I denna fråga har vi dock mycket lite kunskap.

En betydande del av variationen i lärareffektivitet har sannolikt att göra med icke-observerade individegenskaper. Detta illustreras av en färsk studie av Rockoff m.fl. (2009) på amerikanska data. De undersöker om faktorer som kan observeras vid rekryteringstillfället förklarar lärareffektivitet. Författarna skiljer på faktorer som: (1) regelmässigt används vid rekryteringen (t.ex. behörighet och utbildningsnivå); och (2) inte används vid rekryteringen men skulle kunna mätas (t.ex. högskoleprovsresultat, uppmätt förmåga att lära ut ett ämne och andra personliga egenskaper). Det är bara den andra typen av faktorer som har en statistiskt säkerställd effekt på elevresultat. Och även om effekten finns där, så är den ganska liten.

Givet att det finns en betydande variation i lärareffektivitet som är svår att förklara med de egenskaper som normalt observeras, blir det viktigt att attrahera rätt individer till läraryrket och se till att de bra lärarna stannar kvar i yrket. Den stora policyfrågan är hur man ska åstadkomma detta. En del av svaret ligger säkert i lärarlönerna, men eftersom lönebildning är decentraliserad har inte beslutsfattare på central nivå någon möjlighet att direkt påverka dessa. Läraryrkets grad av attraktivitet påverkas också av andra faktorer än lönen, som arbetsförhållanden, fortbildningsmöjligheter och tillgängliga karriärgångar.

Länder med en sammanhållen grundskola har typiskt en lägre variation i elevprestationer. I länder där man övergått från ett differentierat system till ett sammanhållet system (t.ex. i samband med grundskolereformerna i Sverige, Finland och Norge) har sambandet mellan utfall i föräldra- och barngenerationen försvagats.

Den svenska grundskolan har blivit mindre sammanhållen i samband med att skolval har blivit en allt viktigare företeelse. Föräldrar kan välja mellan alla

typer av skolor, både kommunala och fristående skolor. Den konkurrens som kommer från fristående skolor verkar ha förbättrat genomsnittliga prestationer. De färskaste studierna på området tyder dock på att effekten är förhållandevis liten. Och det finns också viss evidens som tyder på att kostnaderna ökat. Skolval tycks även ha bidragit till att öka segregationen mellan t.ex. invandrare och inhemskt födda.

Ett decentraliserat system, med långtgående valmöjligheter, förutsätter systematisk uppföljning och tillsyn. En sådan verksamhet kan bidra med information som gör att barn och föräldrar kan fatta rationella val. En systematisk uppföljnings- och tillsynsverksamhet kan också sprida information om de goda exemplen och förhindra de dåliga exemplen. I detta avseende har utvecklingen sedan början av 1990-talet lämnat en del övrigt att önska. Tillsynen minskade snarast efter decentraliseringen. Så småningom påbörjades en inspektionsverksamhet som innebar att verksamheten vid en enskild skola granskades vart sjätte år i genomsnitt. Inspektionsverksamheten i Sverige har inte blivit utvärderad. Evidensen från England och Holland tyder på ömsom positiva och ömsom negativa effekter. Även om man ska ta dessa resultat på allvar så underskattar de värdet av en inspektionsverksamhet. Det finns också en preventiv effekt av inspektioner: risken för granskning kan påverka beteendet även om skolornas verksamhet inte granskas.

Vår uppfattning är att en systematisk tillsynsverksamhet borde baseras på regelbundna tester av elevers kunskaper (t.ex. tester i tredje, femte, och nionde klass som nyligen införts). Standardiserade tester är det effektivaste sättet att samla in enhetlig information från alla skolor (förutsatt att testerna mäter kunskaper på ett relevant sätt). Inspektionsverksamheten kan sedan fokusera på varför resultaten framstår som särskilt goda eller dåliga i enskilda skolor.

Även om det finns goda argument för en sammanhållen grundskola, är det rimligt att elever gör inriktningsval efter grundskolan. De negativa konsekvenserna av denna ”specialisering” mildras avsevärt om det finns en väl utbyggd vuxenutbildning såsom i Sverige. Vår läsning av evidensen på området är att de positiva aspekterna i form av att tillhandahålla en väg in på arbetsmarknaden via en (välfungerade) yrkesutbildning överväger de negativa konsekvenserna av en för tidig specialisering. Detta gäller särskilt för grupper med sämre studieförutsättningar.

Examensfrekvensen från gymnasiet minskade avsevärt i samband med att programgymnasiet och meritvärdessystemet introducerades under 1990-talet. Vi har redovisat en utvärdering av försöksverksamheten som föregick program-

gymnasiet och också analyserat de första åren med meritvärdessystemet. Även om de nya systemen kan innehålla en del barnsjukdomar, tyder resultaten på att en del av nedgången i andelen med slutbetyg ("examensfrekvensen") har att göra med det tredje året på yrkesutbildningen; denna reform har alltså minskat examensfrekvensen och då särskilt för grupper med sämre studieförutsättningar. Det tredje året på yrkesutbildningen har inte heller ökat sannolikheten att gå vidare till högre studier; detta är anmärkningsvärt då det var ett uttalat syfte med reformen.

Merparten av nedgången i examensfrekvensen har dock att göra med övergången från det relativa betygssystemet till meritvärdessystemet. Detta följer direkt av att meritvärdessystemet faktiskt innebär att betyget icke-godkänd delas ut.

En konsekvens av meritvärdessystemet är att bristen på studieframgång synliggörs på individnivå. En intressant fråga är hur individerna drabbas av detta. Vi har redovisat analyser som tyder på att inkomsteffekterna av meritvärdessystemet är negativa för elever som har påbörjat en yrkesutbildning. Vi tolkar detta resultat som att den minskade sannolikheten att avgå med slutbetyg i meritvärdessystemet gör att individer också bestraffas med lägre inkomster. Effekterna är återigen särskilt stora för elever med sämre studieförutsättningar.

Vi har inte studerat konsekvenserna av meritvärdessystemet på grundskolenivå. Likväl har vi svårt att se att det finns några avgörande skäl som talar mot att de kvalitativa resultaten på gymnasienivå gäller även på grundskolenivå. Som en följd av att elever slutar grundskolan utan ett godkänt slutbetyg har man inrättat det individuella programmet. Vi har inte sett några regelrätta utvärderingar av det individuella programmet, men faktum är att sannolikheten att ta sig vidare från det individuella programmet till ett annat program är låg. Det är sannolikt att samma typ av mekanismer verkar på grundskolenivå som på gymnasienivå. Elever som inte får ett godkänt slutbetyg på grundskolenivå riskerar att straffas ut, vilket sannolikt har negativa arbetsmarknadseffekter på längre sikt.

Vuxenutbildning kan ha hög avkastning för individer som har felinvesterat i yngre åldrar. Evidensen på detta område är inte så stark som på en del andra områden. Vår läsning av litteraturen är ändå att vuxenutbildning för individer med svagare förutsättningar har högre avkastning än breda satsningar. Det verkar även som att avkastningen på yrkesinriktade kurser är högre än avkastningen på teoretiska studier.

I likhet med förskolan har det skett en massiv utbyggnad av högskolan under de senaste 50 åren. Utbyggnaden av högskolan har huvudsakligen skett i regioner där det tidigare inte har funnits ett universitet. Regionaliseringen har dock sannolikt haft en ringa effekt på den sociala snedrekryteringen. Det finns mycket som tyder på att skillnaderna mellan barn med olika utbildningsbakgrund uppstår i förhållandevis tidig ålder. Om den sociala snedrekryteringen uppfattas som ett problem bör man sträva efter att rikta insatser mot barn till lågutbildade föräldrar redan i tidig ålder. Sådana åtgärder är sannolikt mer effektiva än att göra förändringar i t.ex. studiemedelssystemet.

En betydande andel av eleverna skriver högskoleprovet. Våra analyser visar att gymnasiebetyget har betydligt större förmåga att prognostisera andelen elever som tar en högskoleexamen än högskoleprovet. En del av förklaringen till detta är att högskoleprovet bara avspeglar sådana kunskaper och färdigheter som är möjliga att fånga med ett prov som genomförs under ett fåtal timmar, medan gymnasiebetyget utgörs av mätningar av ett brett spektrum av kunskaper och färdigheter under tre år. Sett ur detta perspektiv är det inte förvånande att betyget är ett betydligt precisare instrument för att förutsäga studieframgång än högskoleprovet.

Svenska elever börjar läsa på högskolan förhållandevis sent. Man skulle kanske kunna hävda att detta faktum bidrar till att studenterna har gjort mer informerade utbildningsval än om de börjar studera på högskolan vid en yngre ålder. Det finns dock lite som tyder på att så skulle vara fallet: Inkomstförlusterna som är förenade med att fördröja starten på högskoleutbildningen är betydande. Ur både individens och samhällets synpunkt finns därför argument för att ta bort de restriktioner och de incitament som gör att individer senare lägger sina högskolestudier.

Regionaliseringen av högskolan föranleder också frågan om utbildningskvaliteten varierar över lärosäten. De studier som finns på området, och våra egna analyser, tyder på att lärosäteseffekterna är små. Skillnaderna i genomsnittsinkomster som finns mellan studenter som tar examen från olika högskolor har alltså mest att göra med att dessa studenter var olika redan i utgångsläget.

Vi har också redovisat skattningar av avkastningen på olika inriktningsval. Det finns en betydligt större variation mellan inriktningar än mellan lärosäten. Det är dock inte givet att denna variation har en kasual tolkning. Det är därför angeläget med ytterligare forskning på detta område.

En viktig fråga när det gäller högskoleutbildningen är om den är överdimensionerad. Det finns en stor litteratur som handlar om konsekvenserna av att hamna i yrken vars utbildningskrav inte motsvarar individens utbildningsnivå. Det finns en del problem med denna litteratur. Det är ändå rimligt att tro att överutbildning är ett problem för de individer som är överkvalificerade. När det gäller den större frågan om överutbildning är ett problem för samhällsekonomin är det värt att notera att samtidigt som förekomsten av överutbildning (såsom det mätts i litteraturen) ökat så har avkastningen på högskoleutbildning stigit. Detta talar för att överutbildning inte är ett problem för den genomsnittlige individen.

11.2 Policydiskussion

11.2.1 Är avkastningen på utbildning för låg i Sverige?

Internationella jämförelser av utbildningens lönepremie (se kapitel 4) tyder på att den är förhållandevis låg i Sverige. Våra analyser i kapitel 5 tyder på att ungefär hälften av denna lönepremie är en kausal effekt av ytterligare ett utbildningsår. Kan man utifrån dessa observationer hävda att avkastningen på utbildning är för låg i Sverige? Detta argument har inte varit ovanligt i den svenska debatten och diskussionen har då gällt avkastningen på högskoleutbildning (se t.ex. Storesletten och Zilibotti 2000; Söderström m.fl. 1992).

Innan vi går in på själva sakfrågan är det värt att göra två preciseringar som följer av att det är typiskt är bruttolöner som har studerats i de internationella jämförelserna. För det första behöver en låg (brutto)löneavkastning inte betyda att svenska skolor är dåliga på att förmedla kunskaper. De lönebildningsinstitutioner som vi har i Sverige kan göra att produktivitetsförbättringar inte får lika stort genomslag i ökade löner som i t.ex. de anglosaxiska länderna (se kapitel 5).

För det andra är inte enbart avkastningen såsom den mäts av bruttolöner relevant för individer som står i färd med att fatta ett utbildningsbeslut. Ur individernas synvinkel är det utbildningens effekter på inkomsten efter skatt och transfereringar som är primärt relevant. Dessutom torde individerna också ta hänsyn till utbildningens effekter på mer svårsmätbara aspekter, som intressantare arbetsuppgifter och bättre arbetsmiljö; i en färsk studie argumenterar Oreopoulos och Salvanes (2009) för att dessa ”icke-monetära” aspekter har åtminstone lika stora betydelse som de rent monetära aspekterna på utbildningsavkastningen.

Vad händer då med den privatekonomiska avkastningen på högskoleutbildning om man beaktar skatter och transfereringar (men bortser från aspekter som är svåra att mäta)? Edin och Holmlund (1995) har visat att om man även beaktar studiemedel så stiger avkastningen på högskoleutbildning i Sverige markant. De redovisar internräntekalkyler med och utan hänsyn till skatter och studiemedel. När de tar hänsyn till skatter och studiemedel stiger internräntan med 50 procent (från 7 till 11 %) för år 1991.

Ett generöst studiemedelssystem påverkar sannolikt lönerelationerna på arbetsmarknaden. Forskningsevidensen tyder på att löneskillnader efter utbildning till stor del bestäms av utbudet och efterfrågan på kvalifikationer – särskilt utbudet är relevant. Vi har noterat att även lönebildningsinstitutioner har betydelse, men även dessa institutioner påverkas av förhållanden på marknaden (se Edin och Topel 1997). Ett viktigt skäl till att lönepremien är låg i Sverige kan således vara att subventionsgraden är relativt hög i Sverige. Om vi tar högskoleutbildningen som exempel så gör frånvaron av avgifter och relativt generösa studiemedel att det relativa utbudet av högskoleutbildad arbetskraft är större än det annars skulle vara.¹¹⁴ Att avkastningen på utbildning är låg i Sverige har alltså delvis att göra med att subventionsgraden är hög.

Är avkastningen på utbildning för låg i Sverige? Eftersom löneskillnader på arbetsmarknaden till stor del styrs av efterfrågan och utbud, som i sin tur påverkas av utbildningspolitiken, är detta en oprecist formulerad fråga. En bättre fråga är: Subventionerar vi utbildning för mycket i Sverige? Det är mycket svårt att ta ställning till denna fråga. Man kan notera att höga subventioner kan utgöra ett problem. Studiemedlen ger t.ex. incitament att studera på högskolenivå men de ger inte med nödvändighet incitament till att avsluta utbildningen och omsätta de förvärvade kunskaperna på ett produktivt sätt.

11.2.2 Varför har svenska grundskoleelevers resultat försämrats?

Internationella undersökningar tyder på att elever i svenska skolor presterade bra i början och i mitten av 1990-talet. Detta gäller i alla ämnen, även i matematik där prestationerna traditionellt sett legat under det internationella genomsnittet eller varit genomsnittliga.

De undersökningar som genomförts efter 1995 tyder på avsevärda nedgångar. Svenska grundskoleelever är idag genomsnittliga i matematik och naturvetenskap. De ligger dock signifikant över snittet i läsförståelse, även om

¹¹⁴ Detta argument förutsätter att en ökad efterfrågan fått ett genomslag i ett ökat utbud, trots de antagningsrestriktioner som existerat inom den högre utbildningen.

man kan observera en nedgång även här. Resultatnedgången syns även på gymnasienivå. Trendmätningarna i TIMSS tyder på en markant nedgång i matematik och naturvetenskap – i storleksordningen 0,5–0,8 standardavvikelser – mellan 1995 och 2008 (se Skolverket 2009b).

Vid sidan av den genomsnittliga nedgången visar våra analyser av betygsutvecklingen att likvärdigheten i skolresultaten försämrats sedan 1998, med ökande skillnader mellan elever, skolor och kommuner, och ökande skillnader mellan elever med olika utbildningsbakgrund och kön.

Vad finns det för rimliga förklaringar till de negativa förändringarna i svenska grundskoleelevers prestationer sedan 1995? Ett sätt att ta sig an denna fråga är att utgå från de skillnader i resultatutveckling som finns mellan grundskoleelever i Sverige, Finland och Norge. Medan utvecklingen i Sverige varit negativ sedan 1995 uppvisar Finland en kontinuerlig positiv resultatutveckling från slutet av 1980-talet fram till de senast genomförda mätningarna. Norge deltog endast i ett fåtal av de tidiga internationella jämförelserna, men har deltagit i praktiskt taget samtliga undersökningar sedan 1995. Resultatmönstret i dessa sammanfaller helt med det svenska mönstret, med kraftiga försämringar i matematik och naturvetenskap. I det internationella perspektivet är dessa tre nordiska länder unika, med en tydlig trendmässig försämring för Norge och Sverige, och en tydlig trendmässig resultatförbättring för Finland. Hur kan då dessa mönster förklaras?

En hypotes som ofta framförs för att förklara de bättre resultaten i Finland är att andelen elever med utländsk bakgrund där är lägre än i Sverige. Som vi visat i kapitel 2 finns dock en stor resultatfördel för de finska eleverna i PISA 2006 även då vi tar hänsyn till migrationsbakgrund; resultatskillnaden är således inte en direkt effekt av att andelen elever med utländsk bakgrund är större i Sverige än i Finland. Den lägre andelen elever med utländsk bakgrund i Finland kan inte heller förklara den trendmässiga resultatförbättringen för de finska eleverna.

En andra tänkbar förklaring är att resurserna dragits ner i svensk skola sedan början av 1990-talet. Att resursutvecklingen i genomsnitt skulle vara en viktig förklaring motsägs dock av det skedde kraftiga resursneddragningar också i Finland i spåren av den ekonomiska krisen under 1990-talet.

En tredje möjlig förklaring har också att göra med resurser, men här ligger fokus på förändringen i fördelningen av resurser (Skolverket, 2009c). Även om vi konstaterat att resursutvecklingen i genomsnitt antagligen inte är en viktig förklaring, så kan inriktningen på resursinsatserna ha förändrats så att den

missgynnat elever med sämre studieförutsättningar. Skolverket (2009a) redovisade nyligen resultaten från en studie av hur resurser fördelas till elever med olika bakgrund. De konstaterar att i ett flertal av de undersökta kommunerna så fördelas resurser med ett givet belopp per elev. Detta kan tyda på att det kompensatoriska inslaget i resurstilldelningen minskat i samband med kommunaliseringen av skolan. Det fanns ett starkt kompensatoriskt inslag i det tidigare centralstyrda systemet, som alltså inte är lika tydligt i dagens decentraliserade system. Denna hypotes kan förmodligen delvis förklara den ökande variationen i elevresultaten och den försämrade likvärdigheten, men vår forskningsöversikt ger inte någon grund för att tro att resursers resultatpåverkan är tillräckligt stor för att den skall kunna förklara den genomsnittliga försämringen i Sverige.

En fjärde möjlig förklaring är den ökade segregeringen och differentieringen av eleverna i grundskolan (Skolverket, 2009c). Det framstår som rimligt att dessa faktorer delvis kan förklara den ökade variationen mellan skolor. Vår översikt ger också ett visst stöd för att alltför tidig differentiering leder till försämrade genomsnittresultat. De observerade försämringarna i Sverige är dock så stora att den ökade segregeringen och differentieringen endast kan förklara en mindre del av dessa.

En femte möjlig förklaring har att göra med skillnader i tillgången på lärarkompetens i Sverige och Finland. I anslutning till att grundskolan infördes i Finland vid början av 1970-talet byggdes också en mycket ambitiös 5-årig lärarutbildning på licentiatnivå upp, vilken även omfattade lärare för de lägre stadierna. Denna lärarutbildning har också förmått attrahera studerande med de högsta betygen från gymnasiet, och särskilda urvalsprov har utvecklats. Andelen lärare i den finska lärarkåren med denna lärarutbildning har sedan 1980-talet ökat trendmässigt. Den successiva förbättringen i de finska elevernas resultat samvarierar med denna ökade andel, och även om detta är en svag grund för en kausal slutsats är det svårt att hitta någon annan förklaring till de successivt förbättrade finska resultaten. Även i Sverige har lärarutbildningen förändrats, men mer i syfte att utveckla generell och flexibel lärarkompetens, än att utveckla specialistkompetens på forskningsgrund. Sedan början av 1990-talet har även andelen lärare utan pedagogisk högskoleutbildning ökat i Sverige. I princip skulle dessa förändringar delvis kunna förklara de svenska resultatförsämringarna men vi tror inte att detta kan vara en väsentlig förklaring till nedgången i resultat i Sverige. Grunden för denna slutsats är att förändrad lärarutbildning och försämrad rekrytering ännu endast i begränsad utsträckning haft genomslag i den totala lärarpopulationen. Det faktum att några motsva-

rande förändringar i lärarutbildning och andel behöriga lärare inte ägt rum i Norge, trots att mönstret av resultatförsämringar där är detsamma som för Sverige, talar också mot denna förklaring.

En sjätte möjlig förklaring är förändringar av arbetsformerna i skolan (Skolverket, 2009c). Bland annat som en följd av skrivningar i Lpo 94 har individualiseringen av skolarbetet ökat, och ett ökat ansvar för uppläggning och genomförande av skolarbetet har lagts på eleverna. Lärarens roll har också förändrats från direkt undervisning i helklass till handledning av elever som bedriver eget arbete. Intressant nog influerades Norge vid 1990-talets mitt av de svenska läroplansförändringarna och även där har en likartad förändring av arbetsformerna som i Sverige i riktning mot en högre grad av individualisering och en minskning av lärarledd undervisning ägt rum. Det faktum att resultatförsämringen i Norge sammanfaller i tid och omfattning med den svenska ger ytterligare stöd för denna förklaring.

Vilket forskningstöd finns för att nedgången i den lärarledda undervisningen är en väsentlig del i resultatförsämringen i svenska skolor? På ett allmänt plan finns förstås en stor mängd forskning som visar att lärarens undervisning är av stor betydelse för elevernas resultat (Hattie 2009), och när omfattningen av den lärarledda undervisningen minskar är det rimligt att förvänta sig att resultaten försämrar. Man kan också hänvisa till forskningen om resursers betydelse för studieresultat; en minskning av den lärarledda undervisningen till förmån för eget arbete är således liktydigt med en minskning av både lärartätheten och lärarkompetensen.

En färsk studie av Schwerdt och Wuppermann (2009) redovisar resultat som ligger i linje med detta resonemang. Studien tyder på att om andelen av undervisningstiden som läggs på undervisning i helklass ökar med 20 procentenheter (vilket ungefär motsvarar en standardavvikelse) så förbättras elevresultaten med 2 % av en standardavvikelse.¹¹⁵

En utvärdering från England pekar också i samma riktning. Machin och McNally (2008) undersökte effekterna av att förändra undervisningsättet i läs- och skrivförståelse – den s.k. literacy hour. Introduktionen av literacy hour innebar en övergång till ett mer systematiskt sätt att undervisa som utgick från den nationella läroplanen. Innan literacy hour undervisades elever sällan i

¹¹⁵ Studien är genomförd på den amerikanska delen av TIMSS 2003. För att hantera de problem som orsakas av att elever sorteras på lärare med olika undervisningsstil använder sig författarna av elevfixa effekter. I en strikt mening är det dock möjligt att andelen tid som läggs på undervisning i helklass avspeglar andra icke-observerade läraregenskaper. Lärare med en god förmåga att fånga elevernas intresse kanske t.ex. använder helklassundervisning i större utsträckning.

helklass. Efter literacy hour skedde detta i betydligt större utsträckning, undervisningen hade tydligare struktur, och variationen över skolor i sättet att undervisa läsförståelse minskade per konstruktion. Notera också att det som förändrades var sättet att undervisa, snarare än mängden undervisning i engelska. Machin och McNally (2008) visar att genomsnittseffekterna av literacy hour var positiva, de uppgick till ungefär 0,1 standardavvikelse. Läs- och skrivförståelsen ökade särskilt mycket bland lågpresterande elever; det finns inga effekter bland högpresterande elever.¹¹⁶

Resultaten i Machin och McNally (2008) innebär alltså att en övergång till ett mindre strukturerat arbetssätt skulle slå särskilt hårt mot elever med sämre studieförutsättningar. Det är rimligt att förvänta dessa typer av fördelnings-effekter, eftersom barn till högutbildade föräldrar har bättre möjligheter att få hjälp hemifrån (se också Clark 2009). De förändrade arbetsformerna kan därför förklara både den genomsnittliga resultatförsämringen och den minskade likvärdigheten.

Sammantaget bedömer vi att de ändrade arbetsformerna troligen är den viktigaste förklaringen till den genomsnittliga resultatförsämringen i Sverige. Vissa av de övriga förklaringar som vi diskuterat kan också ha haft betydelse, som försämringar i lärarkompetensen och det minskade inslaget av kompensatorisk resurstilldelning. Även vad gäller den minskade likvärdigheten är vår bedömning att de ändrade arbetsformerna har bidragit, men här är även det minskade inslaget av kompensatorisk resursallokering, och den ökade segregationen och differentieringen viktiga förklaringar.

11.2.3 Hur ska betygssystemet utformas?

Vi har presenterat en del fakta som tyder på att utformningen av dagens betygssystem är synnerligen problematisk. Ett syfte med meritvärdessystemet är att mäta kunskapsutvecklingen över tid. Meritvärdet har ökat kontinuerligt sedan det introducerades. Detta har skett samtidigt som svenska elevers position i de internationella mätningarna försämrats sedan 1995. Mätningar av svenska elevers spatiala och induktiva förmåga tyder inte heller på några förbättringar. I själva verket har den induktiva förmågan varit ungefär konstant mellan 1995 och 2005, medan den spatiala förmågan försämrats; se kapitel 2. Dessa observationer antyder en betydande betygsinflation i meritvärdessystemet, vilket också påvisats av Cliffordson (2004) samt Wikström och Wikström (2003).

¹¹⁶ Preliminära skattningar av effekterna av ett motsvarande initiativ på matematikområdet – ”the numeracy hour” – visar på mycket likartade effekter; se Machin och McNally (2009).

I kapitel 2 visade vi även att betygsspridningen mellan skolor ökade dramatiskt i ämnen utan centralt bestämda betygskriterier efter att meritvärdessystemet infördes. Jämförbarheten över skolor har alltså påtagligt försämrats.

Till sist visade vi i kapitel 9 att nedgången i andelen elever med slutbetyg från gymnasiet i allt väsentligt har att göra med införandet av meritvärdessystemet. Än mer intressant är att nedgången i examensfrekvensen också har konsekvenser för individernas framtida arbetsmarknadsutfall, som påverkats till det sämre.

Vilka kriterier ska då ett rimligt utformat betygssystem tillfredställa? Denna grundläggande fråga analyserades redan i en statlig utredning från 1942; se SOU 1942:11. Utredningen, som särskilt fokuserade på folkskolebetygens funktion vid urval till realskolan, kom att förespråka ett relativt betygssystem där standardprov används som underlag för jämförelser mellan olika klassers prestationsnivåer. Riksdagen fattade beslut om ett sådant relativt betygssystem år 1949, och det infördes under de följande åren. I detta system användes den sjugradiga bokstavsskalan, vilken ersattes med den femgradiga sifferskalan vid grundskolans införande år 1962.

Vår uppfattning är att ett betygssystems grundläggande funktion är att rangordna elevers studieprestationer som gör att de är jämförbara oavsett i vilken skola eller vilket år man uppnått sitt betyg. Det är uppenbart att det nuvarande betygssystemet inte fyller denna funktion. En jämförbar rangordning av elever kräver att betygen (på klass- eller skolnivå) är förankrade i nationella provresultat (på klass- eller skolnivå).

Många av de problem som finns med dagens betygssystem härstammar från att ett instrument har tre olika funktioner. För det första används betygen för att utvärdera skolan på systemnivå. För det andra finns ambitionen att mäta kunskapsutvecklingen över tid. För det tredje är det meningen att betygen ska rangordna elever så att deras studieprestationer är jämförbara vid varje tidpunkt.

I själva verket behövs ett separat instrument för varje enskild funktion. Det finns välutvecklade metoder som gör att man kan mäta kunskapsutvecklingen över tid. Testerna som möjliggör dessa mätningar kan ges till ett förhållandevis litet urval av elever. Mätningarna av resultattrenderna i t.ex. TIMSS är exempel på dessa metoder. En huvudidé i dessa är att varje elev endast genomför en mindre del av en större mängd uppgifter, för att det skall vara möjligt att erhålla en adekvat representation av olika kunskaps- och färdighetsdomäner. För denna typ av mätningar finns en välutvecklad statistisk och psykometrisk

metodologi, med vars hjälp det är möjligt att erhålla goda populationsskattningar av resultatnivåer och förändringar i dessa.

För att elevers betyg ska vara jämförbara behöver de förankras i nationella provresultat. Detta ställer naturligtvis krav på utformningen av de nationella proven så att de testar relevanta aspekter. Det ställer också krav på att resultaten på de nationella proven inte inflateras, vilket kan motverkas med hjälp av objektiva bedömningsbara uppgifter, och att lärare på olika skolor byter och rättar varandras prov.

Utvärderingen av skolorna och skolsystemet kan i princip ske med hjälp av de nationella proven. Ett problem med dessa är dock att de innehåller ett begränsat antal uppgifter, och att det inte är möjligt att fånga alla väsentliga utfall med hjälp av de nationella proven. Detta gör att det kan krävas ett separat instrument för att bestämma utfall på systemnivå.

En återgång till tidigare ordning (ett förankrat relativt betygssystem) är antagligen inte realistiskt. Däremot är det klart att betygssystemet behöver förankras för att betygen ska uppfylla den grundläggande funktionen att rangordna elever för vidare antagning, t.ex. till högskoleutbildning. Ett välförankrat betygssystem har betydligt bättre egenskaper än andra kriterier för antagning till högre utbildning. I kapitel 10 visade vi att betygen har avsevärt bättre prognosförmåga än högskoleprovet. Betygen är resultatet av upprepade och flerdimensionella mätningar av elevers kunskaper, något som knappast kan åstadkommas med ett antagningsprov.

11.2.4 Är skolan en central eller kommunal angelägenhet?

Svenska elevers studieprestationer försämrades i samband med att ett stort antal systemgenomgripande reformer gjordes i svensk skola. En komponent av dessa systemreformer var att ansvaret för skolan decentraliserades till kommunerna.

Att dessa systemreformer orsakat den genomsnittliga resultatförsämringen torde vara omöjligt att leda i bevis. Däremot finns det en del som tyder på att systemreformerna orsakat en större resultatspridning mellan skolor. Vi har refererat till undersökningar som tyder på att fritt skolval leder till en ökad sortering. Undersökningen av Skolverket (2009a) antyder också att resursfördelningen på skolnivå idag har mindre kompensatoriska inslag vilket också kan förklara en ökad spridning i studieresultat.

Är en ökad spridning i resurser på skolnivå ett tillräckligt argument för att förorda centralisering? Egentligen inte eftersom man i sådana rekommendationer måste väga effektivitet mot jämlikhet. Däremot finns andra principiella

övervägningar som man bör beakta i detta sammanhang. Ett sådant övervägande är om utbildningspolitiken ska påverkas av kortsiktiga variationer i t.ex. budgetutrymme.

Det kommunala balanskravet innebär att kommunerna måste budgetera så att intäkterna överstiger kostnaderna. Detta innebär att kommunerna måste ta hänsyn till kortsiktiga svängningar i både det ekonomiska läget och storleken på barnkullarna när de bestämmer skolbudgeten.

Kortsiktiga överväganden bör inte påverka utbildningsresurser. Utbildningsbeslut är investeringsbeslut, dvs. de är långsiktiga och bör inte bero på cykliska fluktuationer i intäkter och utgifter. Detta resonemang leder till slutsatsen att det inte är lämpligt att utbildningsanordnaren lyder under ett strikt balanskrav.

Men ovanstående resonemang behöver naturligtvis inte betyda att ansvaret för skolan ska återföras till central nivå. Det finns andra tänkbara lösningar. Ett alternativ är att omformulera balanskravet så att kommunen i genomsnitt under en viss tidshorisont är i balans. Ett annat alternativ är att konstruera ett statsbidragssystem som kompenserar kommunerna för kortsiktiga svängningar i den ekonomiska utvecklingen.

11.3 Bättre förutsättningar för forskning och utvärdering

Utbildningspolitiken under 1990-talet kännetecknades av stora reformer: 1991 kommunaliserades skolan (till en början fanns ett sektorsbidrag till skolan kvar) samtidigt som den nya gymnasieskolan med treåriga yrkesutbildningar infördes; 1992 infördes fritt skolval och kommunerna ålades att finansiera fristående skolor; 1993 bakades sektorsbidraget till skolan in i det generella statsbidraget; 1996 infördes friare lönesättning för lärare; 1997 och 1998 reformerades betygssystemet. Det är påfallande hur lite evidens som verkar ha legat till grund för dessa systemgenomgripande reformer. En del av dessa reformer var kanske välmotiverade – andra kanske inte var det. På grund av brister i insamlingen av data och sättet på vilket reformerna genomfördes är det mycket svårt att i efterhand utvärdera dessa reformer.

När den statliga styrningen av skolan minskade så ökade inte insatserna för att följa upp och utvärdera. Detta är anmärkningsvärt då ett decentraliserat system ställer större krav på kontinuerlig uppföljning och utvärdering för att fungera väl. Standardiserade tester är ett oundgängligt hjälpmedel i en sådan uppföljning eftersom sådana ger information som är jämförbar över skolor.

En absolut nödvändig förutsättning för att resultaten av utbildningspolitiken ska kunna utvärderas är att data samlas in. I vissa avseenden är förutsättningarna goda för att utvärdera svensk utbildningspolitik. Till exempel finns det (genom SCB:s försorg) tillgång till hela befolkningens utbildningsnivå och utbildningsinriktning genom det s.k. Utbildningsregistret. Men i andra avseenden har förutsättningarna inte varit lika goda. Detta gäller särskilt information om utbildningsutfall innan 16 års ålder. Tills helt nyligen har det inte funnits centralt insamlad information om vilka grundskolor eleverna gått i förrän i årskurs 9, och det har funnits begränsat med information om deras studieresultat före 16 års ålder.

Dock har situationen förbättrats helt nyligen genom att beslut tagits om att skapa ett register med information om vilka skolor eleverna går i under hela grundskolan. Däremot sker ingen central insamling av vilka som går i förskolan. En annan förbättring är att elevers kunskaper testas vid en tidigare ålder. Tyvärr samlas inte den fullständiga informationen om elevernas resultat in centralt, utan bara information om eleverna klarat målen eller inte.

Tillgång till data är en nödvändig – dock inte tillräcklig – förutsättning för att skolan ska kunna utvärderas. För att kunna utvärdera politikförändringar krävs någon form av variation som gör att forskare kan göra trovärdiga jämförelser. Jämförelser av studieresultat över tid är mycket sällan trovärdiga, eftersom så mycket annat kan förändras över tiden som inte alls har att göra med förändringar i politiken. Därför bör politikförändringar genomföras med en tanke på att politiken ska kunna utvärderas.

Hur kan man se till att politiken blir utvärderingsbar? En grundregel är att se till att det vid varje tidpunkt finns grupper som dels fått ta del av en ny åtgärd medan andra inte har det. Till exempel kan större reformer introduceras stegvis (t.ex. i delar av landet). Det finns flera argument till att en sådan strategi är lämplig. Ett argument är att de genomförande myndigheterna behöver tid för att anpassa systemen till den nya politiken. Det är lättare att trimma organisationen i delar av landet än om politiken genomförs samtidigt i hela landet. Förutsättningarna för att införa den nya politiken på ett ordnat sätt förbättras således om reformerna implementeras stegvis i delar av landet. Ett annat argument för att införa reformer stegvis är att det allt som oftast inte är givet på förhand att en förändring är till det bättre. Med en stegvis implementering (såsom vid enhets-skolereformen) skapas förutsättningar för utvärdering.

Olika begränsningar i platsfördelningen till olika åtgärder kan också användas för att mäta effekterna av den förda politiken. Det kan exempelvis handla

om olika behörighetskrav, åldersbegränsningar, krav på tidigare studieprestationer eller minsta kötid. Tanken är då att kunna jämföra personer som befinner sig precis ovanför respektive nedanför dessa begränsningar. De är väldigt lika – exempelvis kan de vara nästan lika gamla, ha liknande tidigare studieresultat eller nästan lika lång kötid – men de får olika möjligheter att delta i åtgärden.

Mindre förändringar bör göras på försöksbasis. Ett exempel på en mindre förändring där man kunde ha väntat på utvärderingsresultat är införandet av förskoleklasser. Det finns också goda argument för att testa t.ex. pedagogiska metoder genom rena experiment. Det är mycket sällan som man a priori har en stark uppfattning om att en pedagogisk metod är bättre än en annan. Olika metoder kan testas i mindre skala under kontrollerade former.

Det är sällan som vi med säkerhet kan säga att politikförändringar otvetydigt är till det bättre. Därför bör man som regel säkerställa att reformer kan följas upp och utvärderas. Detta argument torde vara särskilt starkt på utbildningsområdet. Ett genomgående tema i denna rapport har varit att utbildningspolitiska insatser har långsiktiga konsekvenser. En felaktigt utformad utbildningspolitik riskerar följaktligen att få negativa effekter som kan vara mycket svåra att korrigera med andra insatser.

Referenser

- Clark, R.E. (2009), How much and what type of guidance is optimal for learning from instruction. I S. Tobias och T. M. Duffy (Red.). *Constructivist instruction. Success or failure?* New York: Routledge.
- Cliffordson, C. (2004) Betygsinflationen i de målrelaterade gymnasiebetygen, *Pedagogisk forskning*, 9, 1–14.
- Edin, P-A. och B. Holmlund (1995), The Swedish wage structure: The rise and fall of solidarity wage policy? I R. Freeman och L. Katz, *Differences and Changes in Wage Structures*. Chicago: University of Chicago Press.
- Edin, P-A. och R Topel (1997), Wage Policy och Restructuring: The Swedish Labor Market since 1960. I R. Freeman, R. Topel och B. Swedenborg (red.), *The Welfare State in Transition*. Chicago: University of Chicago Press.
- Hattie, J.A. (2009). *Visible learning. A synthesis of over 800 meta-analyses relating to achievement*. New York: Routledge.
- Havnes, T. och M. Mogstad (2009), No child left behind: Universal child care and children's long-run outcomes, Discussion Papers No. 582, Statistics Norway, Research Department.
- Machin, S. och S. McNally (2008), The literacy hour, *Journal of Public Economics*, 92, 1441–1462.
- Machin, S. och S. McNally (2009), The three Rs: What scope is there for literacy and numeracy policies to raise pupil achievement? Manuskript presenterat vid konferensen: Beyond the resource constraint: Alternative ways to improve schooling, Stockholm, maj 2009.
- Oreopoulos, P. och K. Salvanes (2009), How large are returns to schooling? Hint: Money isn't everything, NBER WP No. 15339.
- Rockoff, J., B. Jacob, T. Kane och D. Staiger (2009), Can you recognize an effective teacher when you recruit one? Under publicering i *Education Finance and Policy*.
- Schwerdt, G. och A. Wuppermann (2009), Is traditional teaching really all that bad? A within-student between subject approach, CESifo Working Paper no. 2634.

- Skolverket (2009a) Resursfördelning utifrån förutsättningar och behov? Rapport 330, Fritzes, Stockholm.
- Skolverket (2009b) TIMSS Advanced 2008, Rapport 336, Fritzes, Stockholm.
- Skolverket (2009c) Vad påverkar resultaten i svensk grundskola? Kunskapsöversikt om betydelsen av olika faktorer. Sammanfattande analys.
- Storesletten, K. och F. Zilibotti (2000), Education, educational policy and growth, *Swedish Economic Policy Review*, 7, 39–70.
- SOU 1942:11, *Betygsättningen i folkskolan*. Betänkande av 1939 års betygssakkunniga.
- Söderström, H.T. (red.) (1992), *Tillväxt utan gränser*, Konjunkturrådets rapport 1992. Stockholm: SNS Förlag.
- Wikström, C. och M. Wikström (2003), Grade inflation and school competition: An empirical analysis based on the Swedish upper secondary schools, manuskript, Institutionen för beteendevetenskapliga mätningar, Umeå universitet.

IFAU:s publikationsserier – senast utgivna

Rapporter

- 2010:1** Hägglund Pathric "Rehabiliteringskedjans effekter på sjukskrivningstiderna"
- 2010:2** Liljeberg Linus och Martin Lundin "Jobbnätet ger jobb: effekter av intensifierade arbetsförmedlingsinsatser för att bryta långtidsarbetslöshet"
- 2010:3** Martinson Sara "Vad var det som gick snett? En analys av lärlingsplatser för ungdomar"
- 2010:4** Nordström Skans Oskar och Olof Åslund "Etnisk segregation i storstäderna – bostadsområden, arbetsplatser, skolor och familjebildning 1985–2006"
- 2010:5** Johansson Elly-Ann "Effekten av delad föräldraledighet på kvinnors löner"
- 2010:6** Vikman Ulrika "Hur påverkar tillgång till barnomsorg arbetslösa föräldrars sannolikhet att få arbete?"
- 2010:7** Persson Anna och Ulrika Vikman "In- och utträdeseffekter av aktiveringskrav på socialbidragstagare"
- 2010:8** Sjögren Anna "Betygsatta barn – spelar det någon roll i längden?"
- 2010:9** Lagerström Jonas "Påverkas sjukfrånvaron av ekonomiska drivkrafter och arbetsmiljö?"
- 2010:10** Kennerberg Louise och Olof Åslund "Sfi och arbetsmarknaden"
- 2010:11** Engström Per, Hans Goine, Per Johansson, Edward Palmer och Pernilla Tollin "Underlättar tidiga insatser i sjukskrivningsprocessen återgången i arbete?"
- 2010:12** Hensvik Lena "Leder skolkonkurrens till högre lärlöner? – En studie av den svenska friskolereformen"
- 2010:13** Björklund Anders, Peter Fredriksson, Jan-Eric Gustafsson och Björn Öckert "Den svenska utbildningspolitikens arbetsmarknadseffekter: vad säger forskningen?"

Working papers

- 2010:1** Ferracci Marc, Grégory Jolivet och Gerard J. van den Berg "Treatment evaluation in the case of interactions within markets"
- 2010:2** de Luna Xavier, Anders Stenberg och Olle Westerlund "Can adult education delay retirement from the labour market?"
- 2010:3** Olsson Martin and Peter Skogman Thoursie "Insured by the partner?"

- 2010:4** Johansson Elly-Ann “The effect of own and spousal parental leave on earnings”
- 2010:5** Vikman Ulrika “Does providing childcare to unemployed affect unemployment duration?”
- 2010:6** Persson Anna och Ulrika Vikman “Dynamic effects of mandatory activation of welfare participants”
- 2010:7** Sjögren Anna “Graded children – evidence of longrun consequences of school grades from a nationwide reform”
- 2010:8** Hensvik Lena “Competition, wages and teacher sorting: four lessons learned from a voucher reform”

Dissertation series

- 2010:1** Johansson Elly-Ann “Essays on schooling, gender, and parental leave”
- 2010:2** Hall Caroline “Empirical essays on education and social insurance policies”