

IFN Policy Paper nr 68, 2014

Slänger vi pengarna i sjön? Resurser, reformer och utbildningspolitiska slutsatser

Gabriel Heller Sahlgren

Slänger vi pengarna i sjön? Resurser, reformer och utbildningspolitiska slutsatser

Gabriel Heller Sahlgren¹

Rapporten är skriven på uppdrag av Timbro.

¹ Forskningschef vid Centre for Market Reform of Education i London och affilierad forskare vid Institutet för Näringslivsforskning (IFN).

Innehåll

1. Introduktion	1
2. Två teoretiska ramverk	3
a. Det akademiska ramverket	3
b. Det utbildningspolitiska ramverket.....	4
3. Effektstorlek	6
4. Vad säger forskningen?	7
a. Hanushek vs Krueger – ett avstamp i analysen	7
b. Ett meningslöst räknande	8
c. Metodologiskt starka studier	9
Fältexperiment.....	9
Studier som analyserar variationen i resurser vid ett tillfälle	14
Studier som analyserar reformer och förändringar över tid.....	30
d. Är resurser mer effektiva bland elever i lägre årskurser?	49
5. Diskussion och slutsatser	50
Referenser	52

1. Introduktion

Frågan om huruvida ökade resurser på egen hand leder till högre resultat i skolan är en av de mest omdebatterade inom både utbildningsforskningen och politiken. Den är också högaktuell i Sverige i en valrörelse där både alliansregeringen och oppositionen har lovat att skjuta till mer resurser, för att minska klasstorleken bland elever i lågstadiet. Detta är politikernas svar på de sjunkande resultaten i internationella undersökningar som Pisa och Timss. Det verkar alltså råda konsensus i den svenska politiken om att mer pengar till skolan i sig är ett effektivt sätt att höja resultaten.

Stämmer detta? Den här rapporten diskuterar den mest relevanta nationalekonomiska forskningen om effekterna av resurser inom skolan samt vilka utbildningspolitiska slutsatser som bör dras. På grund av metodologiska svårigheter att separera orsakssamband från korrelationer är det inte enkelt att på ett övertygande sätt analysera effekterna av resurser. Det är därför viktigt att studierna som analyseras använder bra metoder.

Den akademiska frågan om huruvida resurser spelar roll "allt annat lika" är intressant, men av relativt liten betydelse för utbildningspolitiken. Reformerna som ökar/minskar resurserna har nämligen oavsiktliga konsekvenser som förändrar andra variabler, vilket gör att "allt annat förändras". Därför är det också viktigt att fokusera på den mest policyrelevanta forskningen.

Forskningen som använder metoder som på bästa sätt isolerar effekten av resurser "allt annat lika" är blandad. Resurser leder ibland till bättre resultat, ibland inte. Vissa studier kommer fram till olika svar beroende på vilka elever och vilket land de analyserar och det finns inget riktigt mönster som kan förklara detta. Dessutom varierar effektstorleken inom och mellan studier – de flesta finner faktiskt inga eller små positiva effekter av resurser. Detta är viktigt att ta hänsyn till eftersom den huvudsakliga frågan nämligen inte bör vara om resurser leder till bättre resultat, utan snarare om de är kostnadseffektiva. Trots att metodologierna som används är vedertagna är det fortfarande viktigt att notera kvarstående problem som gör att de utbildningspolitiska implikationerna är osäkra, oavsett de oavsiktliga konsekvenserna som uppstår i och med reformer.

En ofta citerad svensk studie finner stöd för att mindre klasser på mellanstadiet under 1980-talet och början på 1990-talet genererade bättre resultat och påverkade elevernas långsiktiga framtidsutsikter positivt (Fredriksson, Öckert och Oosterbeek 2013). Vid närmare anblick visar det sig dock att dessa resultat är mer osäkra än vad som rapporterats. Den publicerade studien analyserar endast elever i rektorsområden med en skola, vilket minskar urvalet med 80 procent och gör att studiens resultat inte nödvändigtvis kan extrapoleras till alla elever. När författarna använder det fullständiga urvalet och/eller ändrar metodologin något förändras vissa resultat kraftigt och det finns även tecken på att föräldrar flyttar för att undvika större klasser, vilket riskerar överdriva effekterna (Fredriksson, Öckert och Oosterbeek 2012, 2014). Att basera nationell politik på studien är därmed riskabelt. Dessutom finner två andra studier – varav en använder en liknande metodologi som den förstnämnda – inget stöd för att små

klasser ökar de svenska resultaten i Timss, vilket naturligtvis är viktigt givet betydelsen av internationella prov i debatten (Altinok och Kingdon 2012; Woessmann 2005).

Men även om ovannämnda forskning entydigt skulle indikera att utökade resurser är ett kostnadseffektivt sätt att höja resultaten i skolan – vilket alltså inte är fallet – är de utbildningspolitiska implikationerna fortfarande långt ifrån säkra. Genomgående reformer har som sagt ofta oavsiktliga konsekvenser som kan leda till att effekterna av ökade resurser minskas kraftigt eller försvinner helt. Det märkligaste i den svenska debatten är just avsaknaden av en diskussion kring forskningens utbildningspolitiska relevans.

Och bilden av forskningen som analyserar reformer som förändrar resurstilldelningen är inte speciellt positiv. I detta fall är effekten också blandad, vilket i detta fall i ännu högre utsträckning betyder små eller obefintliga effekter i genomsnitt. Detsamma gäller studier som analyserar förändringar i resurser över tid, antingen direkt eller via reformer som indirekt påverkar resursfördelningen.

Dessa resultat nedslående för de som menar att lösningen är ofokuserade och genomgående reformer som utökar resurserna för alla elever. Även om vissa elever ibland kan tjäna på resursökningar betyder det inte att man bör ösa pengar över alla elever.

Slutsatsen är därför att både regeringens och oppositionens fokus på mindre klasser i skolväsendet för alla elever riskerar att bli en dyr flopp. Innan genomgående öknings i resurser genomförs måste incitamentsstrukturen inom skolan förbättras.

Givet forskningsläget, som långt ifrån stödjer idén att resursökningar är ett effektivt sätt att öka elevernas prestationer, vore det bättre att genomföra ett policyexperiment och testa effekten av ökade resurser innan man genomför dyra nationella reformer som oftast inte har gett noterbara effekter i andra länder. Genom att använda policyexperiment kan man också särskilja effekterna av ökade resurser bland olika elever och därmed avgöra för vilka som resurserna bör höjas (och för vilka de kanske bör minskas). På det sättet kan man sedan genomföra riktade reformer som kan bli mer framgångsrika.

Även om extra resurser ibland kan leda till vissa positiva effekter bland vissa elever är det huvudsakliga problemet inom det svenska skolväsendet dåliga incitament, inte bristen på resurser. För att utökade resurser ska leda till väsentligt mycket bättre resultat inom skolan i genomsnitt krävs alltså en bättre incitamentsstruktur. Om politikerna är intresserade av att förbättra skolan mer än marginellt är det därför viktigt att de först förändrar incitamentsstrukturen innan de börjar ösa mer pengar över skolan.

2. Två teoretiska ramverk

I debatten om resursernas påverkan på resultaten kan man urskilja två teoretiska ramverk, ett akademiskt och ett utbildningspolitiskt. Även om dessa naturligtvis överlappar används de här som idealtyper för att påvisa en snedvridning som just nu präglar den svenska debatten: man ignorerar oftast forskningens utbildningspolitiska relevans.

I den rent akademiska diskussionen är man framförallt intresserad av huruvida resurser har en kausal effekt på resultaten och i sådana fall hur stor denna är. Visserligen diskuteras ofta policyslutsatser, men frågeställningen har framförallt varit om resurser "allt annat lika" producerar högre resultat. Betoningen ligger alltså på forskningens "interna validitet", det vill säga om studien är tillräckligt tillförlitlig för att man ska kunna dra slutsatser om orsakssamband.

I den utbildningspolitiska debatten är man istället framförallt intresserad av frågan huruvida reformer som ökar/minskar resurserna kommer att leda till högre/lägre resultat. Här är frågan om forskningens "externa validitet" väldigt viktig, det vill säga om de orsakssamband som påvisas i en studie kan generaliseras till andra kontexter och elever. För att en studie ska vara utbildningspolitisk relevant krävs både intern och extern validitet.

a. Det akademiska ramverket

Anledning till varför högre resurser teoretiskt sett kan leda till högre kvalitet inom skolan "allt annat lika" är ganska självklar. Om vi antar att skolor strävar efter att maximera resultaten på ett effektivt sätt bör mer pengar leda till högre resultat. Argument som används i debatten är exempelvis att man med mer pengar kan höja lärarlönerna – som i ett längre perspektiv kan leda till mer högkvalitativa lärare – eller minska klasstorleken, vilket kan ge bättre studiemöjligheter och mer tid för individuell undervisning. Med färre elever i klassrummet kan lärare få mer arbetsro och tid att fokusera på enskilda elevers behov. Det finns naturligtvis en mängd andra teoretiska mekanismer varför resurser kan leda till bättre resultat.

Men det finns även skäl att ifrågasätta om mer resurser per automatik gör någon större skillnad. Det är nämligen långt ifrån säkert att skolor strävar efter att maximera resultaten på ett effektivt sätt. Antagandet gäller om skolor agerar som vinstmaximerande företag på vanliga marknader med konkurrens och fri prissättning. Men i de flesta skolsystem både finansierar och producerar offentliga myndigheter utbildning, vilket gör antagandet orimligt (Hanushek 2006).

Inom det svenska skolväsendet råder visserligen konkurrens, vilket bör ge incitament bland skolor att vara kostnadseffektiva, men detta beror naturligtvis på vilka incitament som råder på marknaden. Det finns en rad problem i det svenska utbildningssystemet som har gjort att incitamentsstrukturen är

undermålig (Heller Sahlgren 2013). Om skolor inte har incitament att prestera bättre kan istället högre resurser därför leda till mer resursslöseri.²

Den teoretiska poängen här är alltså inte att resurser inte spelar någon roll alls för resultaten, utan snarare att undermåliga incitament och andra faktorer kan minska eller uttradera de positiva effekterna (Hanushek 2003, 2008). Det står alltså inte klart huruvida ökade resurser inom skolan teoretiskt sett nödvändigtvis bör leda till speciellt mycket bättre resultat, om man inte samtidigt genomför andra förändringar som förbättrar incitamentsstrukturen.

b. Det utbildningspolitiska ramverket

Men även om högre resurser, allt annat lika, genererar högre resultat är detta inte speciellt relevant för den utbildningspolitiska debatten. Reformen som ökar resurserna för nämligen även med sig oavsiktliga konsekvenser som man måste ta hänsyn till. Ett exempel på detta är sannolikheten att den totala lärarkvaliteten minskar om man anställer fler lärare. Priset för mindre klasser eller högre lärartäthet kan då vara att fler elever får sämre lärare, vilket i sin tur har negativa effekter på elevernas resultat. Skolor kan också förändra andra rutiner när de får mer pengar, vilket också kan ha negativa effekter på resultaten.

Samtidigt kan naturligtvis skolor, föräldrar och elever också anpassa hur hårt de jobbar till förändrade resurser över tid. Ökade/minskade resurser kan göra att dessa aktörer minskar/ökar sin arbetsbörda eller på andra sätt kompenserar för förändringen som sker, vilket kan göra att effekten minskar eller försvinner (Houtenville och Smith Conway 2008; Das m.fl. 2013). Om aktörer inom skolan förändrar sin arbetsinsats för att ta hänsyn till resursförändringar kan det helt enkelt vara slöseri att öka resurserna eftersom resultaten ändå inte påverkas.³

Detta gäller självklart till viss mån även studier som på ett trovärdigt sätt analyserar effekterna av resurser vid en specifik tidpunkt, men det är ofta svårt för föräldrar att veta vilken resursfördelning som gäller när man använder den variationen. Även när föräldrar helt enkelt vet att deras barn får mer eller mindre resurser – vilket (troligtvis) är fallet i många klasstorleksstudier – är detta ofta oväntat, vilket gör att de inte hinner ändra sitt beteende.⁴ Detta gäller inte vid reformer, om alla är medvetna om förändringen som sker.

² Även fast forskningen visar på små positiva effekter av konkurrens i Sverige, medelstora effekter på Timss-resultaten, och inga effekter på kostnaderna (Böhlmark och Lindahl 2013) är det långt ifrån säkert att dagens konkurrens är tillräcklig för att säkerställa att högre resurser används på ett kostnadseffektivt sätt givet systemets utformning.

³ De totala välfärdseffekterna kan naturligtvis fortfarande vara positiva, vilket i sådana fall innebär att offentliga resurser till viss del tränger undan de resurser som aktörer inom skolan spenderar i form av deras arbetsbörda.

⁴ Detta är speciellt fallet i studier som använder regler för hur stora klasserna får vara för att analysera orsakssamband. Det är i många fall svårt för föräldrar att veta om tröskeln kommer att nås, möjligtvis undantaget i en skolmarknad som ger stora möjligheter till föräldrar att välja skola och till skolor att anta sina elever (se Urquiola och Verhoogen 2009).

Det är visserligen lättare för skolor, föräldrar och elever att kompensera för gradvisa förändringar i resurser än stora skillnader som uppstår i ett längre perspektiv (Gibbons och McNally 2013). Men effekten beror också på om aktörerna vet om att de får mer eller mindre resurser – är dessa införstådda i vilka förändringar som sker kan de lättare anpassa sitt beteende, vilket kan leda till att effekten minskar eller helt uttraderas (Das m.fl. 2013). Eftersom reformer tenderar att diskuteras öppet ger detta aktörerna stora möjligheter att anpassa sitt beteende, vilket inte nödvändigtvis är fallet när det gäller skillnader i resurser som beror på komplicerade beräkningar som föräldrarna inte förstår.

Detta är bara ett par exempel – som noteras i forskningen som diskuteras i den här rapporten finns det många andra. Det är väldigt svårt att veta vilka oavsiktliga konsekvenser som skapas av reformer som skjuter till mer resurser. Med andra ord kan man inte extrapolera resultat från välgjorda studier som gör allt för att isolera påverkan av resurser "allt annat lika", just eftersom reformer gör att "allt annat inte är lika". Det är därför väldigt viktigt att beslutsfattare är medvetna om vilka studier som särskiljer vad som är akademiskt relevant och vad som är utbildningspolitiskt relevant.

I slutändan spelar det dessutom ingen roll om det finns vissa positiva effekter av att öka resurserna – frågan är snarare om resultaten överväger kostnaderna. Om en reform kostar jättemycket pengar men endast ger små resultat bör den inte genomföras. Reformerna måste självklart vara kostnadseffektiva för att vara försvarbara.

Tyvärr har den svenska debatten ofta ignorerat forskningens utbildningspolitiska relevans. Socialdemokraterna (2013) släppte exempelvis förra året en rapport som ämnade analysera forskningen om klasstorlek och elevers prestationer. Den går igenom ett par rapporter från Skolverket och ett (selektivt) urval av studier. Slutsatsen är att små klasser är positiva för elevers resultat, framförallt för yngre elever, och att effekten är starkast för elever från lägre socioekonomisk bakgrund.

Men förutom att rapportens urval av studier är selektivt – vilket visas tydligt i forskningen som diskuteras i den här rapporten – finns det knappt någon analys alls av relevansen för partiets utbildningspolitiska slutsatser, speciellt inte när det handlar om studierna som finner positiva effekter. Detta är häpnadsväckande. Utbildningspolitiken kräver mer än en ytlig analys av forskningen om resurser. Den kräver en grundläggande analys av de utbildningspolitiska implikationerna – och detta kräver i sin tur att man separerar välgjorda och policyrelevanta studier från andra studier som är välgjorda men inte lika relevanta. Det kräver också ett fokus på effektstorleken och huruvida den eventuellt positiva effekten väger tyngre än kostnaderna.

Till sist är det värt att påpeka ett annat relevant problem som gäller metodologiskt starka studier: att de oftast fokuserar på en liten grupp elever bland vilka resursfördelningen kan antas vara slumpmässig. Om resurserna inte har liknande effekter för hela studiepopulationen kan man inte dra slutsatser gällande hur andra elevgrupper skulle svara på en liknande resursfördelning.

Bara för att det finns en positiv effekt av att gå från 30 elever till 15 elever i klassrummet behöver det exempelvis inte betyda att det också finns en positiv effekt av att gå från 25 till 15 elever i klassrummet. Detta är viktigt eftersom många studier analyserar förändringar i klasstorlek som har liten utbildningspolitisk relevans, just eftersom de analyserar effekterna av kraftiga minskningar (ofta halveringar) i klasstorleken, som är extremt dyra att genomföra i verkligheten. Forskare extrapolerar då istället ofta till vad effekten vore givet att effekten är linjär, vilket alltså inte alltid är självklart. Detta är viktigt att ta hänsyn till när det gäller de utbildningspolitiska implikationerna av studierna.

3. Effektstorlek

Resursbaserade reformer bör alltså inte endast bedömas efter huruvida de har några effekter alls, utan snarare efter hur starka dessa effekter är. Effektstorlek är väldigt viktigt i sammanhanget eftersom reformerna per definition kostar pengar. Påståendet att skolan är ovärderlig och måste få kosta vad den kostar är ett vanligt argument till försvar för att minsta lilla resultatförbättring är värd kostnaden, oavsett vad denna är. Men detta är ett dåligt argument. Detsamma skulle kunna sägas om andra skattefinansierade tjänster, exempelvis sjukvården, och pengar som spenderas på skolan skulle naturligtvis kunna göra större nytta i andra sektorer. I slutändan framstår detta troligtvis som ganska självklart för de flesta läsare. Säg att vi ökar resurserna med 100 procent, medan Pisaresultatet endast ökar med 5 poäng – att påstå att detta vore ett effektivt sätt att spendera skattemedel eftersom skolan "måste få kosta" framstår som okänsligt mot de sjuka patienter vi hade kunnat rädda för dessa pengar. Offentliga resurser växer inte på träd och politiker måste därför maximera skattebetalarnas avkastning för varje skattekrona. Effektstorleken är alltså viktig för att reformer ska kunna rättfärdigas.

För att avgöra effektstorleken och kunna jämföra den mellan olika studier rapporteras den ofta i form av hur många standardavvikelser (sd) bättre eller sämre elever presterar när de har tillgång till högre resurser. Det finns ingen vedertagen måttstock gällande vad som är en svag eller stark effekt, men inom utbildningsforskningen i USA definieras ofta effekter på $\pm 0,1$ sd som "små", $\pm 0,2$ - $0,3$ som "medelstora" och $\pm 0,5$ eller högre som "stora" (Greene, Peterson och Du 1997; Rainey 2011). Effektstorleken kan sedan sättas i förhållande till den internationella standardavvikelsen på individnivå i de internationella proven Pisa och Timss, som är 100 poäng. Exempelvis motsvarar en effekt på 0,05 sd alltså 5 poäng högre resultat i de viktigaste internationella proven.⁵ I den här rapporten noterar jag därför effektstorleken i form av hur många standardavvikelser högre resultat som produceras av en specifik procentuell ökning i resurserna eller ungefär 5 färre elever i klassrummet när det gäller studier som analyserar klasstorlek. Detta är nämligen den minskning som idag föreslås av regeringen och oppositionen för lågstadiet.

⁵ Som jag noterar i avsnitt 4.d lär sig elever dock mer under tidiga år, vilket gör att samma effektstorlek mätt i standardavvikelser ger starkare effekt i högre årskurser jämfört med lägre årskurser, om man fokuserar på hur mycket elever lär sig under ett givet år.

Det är viktigt att notera att effektstorleken bör rapporteras på elevnivå – variationen mellan skolor och skoldistrikt/kommuner är mycket lägre, vilket gör att effekten överdrivs om man beräknar den på aggregerad nivå. I de studier där författarna rapporterar effektstorleken baserad på högre nivå omvandlar jag den därför till elevnivå.⁶ Detta är viktigt eftersom andra litteraturanalyser tenderar att ignorera att vissa studier rapporterar effektstorleken på skol- eller distriktsnivå (se t.ex. Gibbons och McNally 2013). Det bör dock också betonas att ju högre nivå elevers resultat aggregeras på, desto högre är sannolikheten att resursernas effekter inflateras (Hanushek, Rivkin och Taylor 1996). Studier som analyserar data på elevnivå är därför att föredra och det är viktigt att notera detta när man tolkar forskningen som aggregerar data på högre nivåer.

4. Vad säger forskningen?

Vad säger då forskningen om effekterna av resurser? Nedan diskuteras existerande forskning med utgångspunkt från två omstridda forskningsöversikter. Studierna diskuteras både vad gäller de akademiska och utbildningspolitiska ramverken som noterades ovan, för att på bästa sätt kunna avgöra vad politiker bör lära av dem.⁷

a. Hanushek vs Krueger – ett avstamp i analysen

Frågan om mer pengar leder till högre presterande elever är omstridd bland forskare. Ett naturligt avstamp i den akademiska debatten är två ofta citerade artiklar av Hanushek (2003) och Krueger (2003), i vilka ekonomerna debatterar hur forskningsresultaten bör tolkas. Båda författarna fokuserar framförallt på studier från USA, där man över de senaste decennierna investerat kraftigt i högre lärartäthet och mindre klasser för att höja resultaten. Mellan 1960 och 2010 ökade den genomsnittliga lärartätheten i amerikanska offentliga skolor med hela 39 procent och klasstorleken i låg- och mellanstadiet minskade med ungefär 23 procent. Kostnaderna för dessa förändringar var naturligtvis väldigt höga (Chingos 2013).

Hanushek analyserar först trender över tid och finner att 17-åringars resultat i matematik och läsförståelse endast var marginellt högre 1999 jämfört med 1970, medan resultaten i naturvetenskap är mycket lägre.⁸ Självklart kan detta bero på att elever kom från andra förhållanden under den tidigare perioden, eller att

⁶ Liksom Dee (2012) dividerar jag effektstorleken på skolnivå med 3,2 för att konvertera den till elevnivå. När det gäller effektstorleken på skoldistriktsnivå, vilket en del amerikanska studier analyserar, rapporterar Goolsbee och Guryan (2006) att standardavvikelsen på skoldistriktsnivå är ungefär 15-20 procent mindre än den på skolnivå i Kalifornien. Jag dividerar därför effektstorleken på distriktsnivå med 3,84 för att få fram effektstorleken på individnivå.

⁷ Rapporten fokuserar på studier som analyserar industriländer – nyligen har dock forskare analyserat existerande studier från utvecklingsländer och konstaterat att resurser generellt inte ökar resultaten (Glewwe m.fl. 2013).

⁸ Resultaten har för övrigt inte förändrats i princip alls mellan 1999 och 2011 heller (<http://nationsreportcard.gov/>).

andra krav på skolor drev upp kostnader som inte kan förväntas producera högre resultat. Även om det är svårt att analysera detta visar Hanushek att ingen av dessa förklaringar är övertygande.

Det är dock svårt att veta om resultaten hade varit sämre utan resursökningarna, eftersom många andra förändringar i det amerikanska samhället skulle kunna motverka eventuellt positiva effekter. Även om författaren menar att resultaten ifrågasätter antagandet att mer resurser i form av ökad lärartäthet leder till högre resultat i genomsnitt, bör man därför vara skeptisk till denna slutsats.

Samtidigt sammanställer Hanushek även nationalekonomiska studier som har analyserat effekterna av ökade resurser i olika former. Han finner att endast 20 procent av studier som analyserar effekterna av lärartäthet/klasstorlek, lärarlöner och resurser per elev finner statistiskt signifikanta positiva effekter. Liknande resultat återfinns när författaren analyserar studier utanför USA.

Krueger (2003) håller dock inte med om slutsatsen att pengar inte spelar någon roll. Han menar först och främst att man bör ge lika stor vikt till studier snarare än enskilda estimat i studier. Detta förändrar dock huvudresultaten väldigt lite. Visserligen är genomsnittsstudien positiv i författarens metaanalys, men detta bygger på att han buntar samman studier som finner statistiskt signifikanta effekter med de som finner icke-statistiskt signifikanta effekter. Krueger menar även att man bör ta hänsyn till studiernas kvalitet, vilket han gör genom att väga dem efter tidskriftens ranking. Men när han använder denna teknik finner han fortfarande att endast 35 procent av studierna visar statistiskt signifikanta positiva effekter av resurser.⁹ Återigen kan han bara komma fram till andra slutsatser genom att bunta samman studier med statistiskt signifikanta resultat tillsammans med studier som finner icke-signifikanta resultat. Trots den retoriska poängen att kvalitet bör värderas högre förändrar Kruegers resultat därför inte slutsatsen nämnvärt.¹⁰

b. Ett meningslöst räknande

Problemet med både Kruegers och Hanusheks studier är att det är meningslöst att räkna antalet positiva eller negativa studier/estimat, oavsett om man väger högkvalitativa studier tyngre än studier med låg kvalitet. Studier med dåliga metoder bör nämligen inte räknas med överhuvudtaget. Det är ytterst svårt att på ett övertygande sätt analysera effekterna av högre resurser och därför bör inte studier som inte på ett någorlunda tillförlitligt sätt kan separera orsakssamband från korrelationer räknas med i analysen.¹¹

⁹ Naturligtvis är det också värt att påpeka att i både Hanushek och Krueger ignorerar effektstorleken i de studier som finner positiva effekter.

¹⁰ För andra problem med Kruegers (2003) analys, se Hanushek (2003).

¹¹ Jag skriver "någorlunda" eftersom det inte finns en godkänd standard för vad som är acceptabelt – det är upp till forskaren att utvärdera huruvida resultaten är tillförlitliga. I den här uppsatsen diskuterar jag alla studier jag har kunnat hitta, som med acceptabla metoder försöker komma åt orsakssambanden. När det finns problem med analysen diskuterar jag även detta och noterar huruvida slutsatserna är tillförlitliga.

Det finns flera skäl till varför omvända orsakssamband kan skapa problem. Ett exempel är att skolor med högpresterande elever har större resurser, inklusive mindre klasser, vilket kan skapa en illusion av att resurser spelar en stor roll, även fast detta i sådana fall kan förklaras av eleverna som går i dessa skolor. Små klasser tenderar att vara populära bland föräldrar i de flesta länder och mer motiverade föräldrar kan därför välja skolor (antingen direkt eller via bostaden) med små klasser.¹² En annan möjlighet är naturligtvis att lågpresterande/stökiga elever får mer resurser (och därför går i mindre klasser), vilket skulle kunna skapa en illusion av att mer resurser har en negativ påverkan på resultat. Mycket riktigt är resursfördelning kompensatorisk i Sverige, vilket indikerar att detta problem är relevant här (Fredriksson och Vlachos 2011). Det är därför viktigt att fokusera på studier som använder bra metoder för att kunna komma åt orsakssambanden.

c. Metodologiskt starka studier

Givet diskussionen ovan fokuserar den här rapporten på studier som analyserar effekten av resurser med metoder som använder bra metoder för att kunna separera orsakssamband från korrelationer. Men som beskrevs i det andra kapitlet är det inte tillräckligt för politiker att veta att studierna är välgjorda eftersom de även måste vara utbildningspolitiskt relevanta. Nedan diskuteras därför forskningen både vad gäller dess interna samt externa validitet, vilket är viktigt för att dra utbildningspolitiska slutsatser.

Fältexperiment

Inom nationalekonomi föredras randomiserade fältexperiment för att forskarna med säkerhet ska kunna säga något om orsakssamband mellan två variabler, i det här fallet resurser och elevers prestationer. Fältexperiment har också potential att både ge hög intern validitet samt möjligheter för forskare att på egen hand utforma premisserna för studien. Detta ökar i sin tur möjligheterna för forskarna att manipulera situationen på ett sätt som gör att den externa validiteten ökar.

Project STAR

Det är därför anmärkningsvärt att det inom forskningen som analyserar effekterna av resurser på skolresultaten endast finns ett enda storskaligt experiment i modern tid. Detta är "Project STAR", som genomfördes i Tennessee i mitten på 1980-talet för att analysera om klasstorlek påverkar elevers resultat i sista året på förskolan samt de tre första åren i lågstadiet. En hel del studier har på olika sätt analyserat detta experiment.

¹² Detta är faktiskt högst sannolikt när det gäller USA som är det enda land av 18 som studeras av West och Woessmann (2006), där resurstilldelningen är regressiv och sämre presterande elever tenderar att gå i skolor med större klasser.

STAR-experimentet analyserade dagisbarn som följdes över fyra år upp till årskurs tre. Vissa placerades i klasser med 13-17 elever, vissa i klasser med 22-25 elever och vissa i klasser med 22-25 elever och en extra lärarassistent. I genomsnitt hade de små klasserna alltså ungefär 8 eller 32-36 procent färre elever.

Generellt rapporterar forskare positiva effekter från STAR-experimentet. Krueger (1999) visar att effekterna av att gå i en mindre klass gör att eleverna presterar bättre totalt sett. Effektstorleken uppgår till ungefär 0,22 standardavvikelser (sd) i genomsnitt efter fyra år. Det finns ingen effekt av att ha en lärarassistent. Intressant nog är effekten av mindre klasser överlägset starkast (0,12 sd) första året som eleverna går i små klasser, med endast väldigt små växande effekter därefter (0,035 sd per år).¹³ Men inte ens effektstorleken på 0,12 sd under det första året är speciellt imponerande givet den stora kostnaden att minska klasserna med 8 elever. Nyligen har forskning också stärkt slutsatsen att det inte finns några dynamiska effekter av klasstorleken i STAR ytterligare – att gå i en liten klass efter årskurs 1 påverkar generellt inte resultaten positivt (Ding och Lehrer 2010). Detta är märkligt och indikerar att klasstorlekens påverkan inte kan generaliseras till elevernas studier generellt, vilket ofta görs i utbildningspolitiska sammanhang.

Som Hanushek (2003) menar indikerar detta också att de vanliga argumenten som används för att försvara klasstorleksminskningar – exempelvis mer individualiserad undervisning och bättre studiemöjligheter – troligtvis inte kan förklara de positiva effekterna. Eftersom elevernas första år i en liten klass också var första året i en ny skola är det mycket möjligt att effekten berodde på det senare istället för att de gick i mindre klasser.

Andra studier har därefter analyserat hur effekterna skiljer sig bland olika elever och i ett längre perspektiv. Dessa visar att effekten var mycket större bland afroamerikanska elever jämfört med vita elever och att fattiga elever tenderade att tjäna något mer än andra elever. Intressant nog tjänade högpresterande elever också mer på att gå i mindre klasser än lågpresterande elever. I årskurs 4-8, efter att experimentet avslutats, finns effekten kvar, men den minskar med 50-75 procent och är i årskurs 8 inte längre statistiskt signifikant (Jackson och Page 2013; Schanzenback 2006/2007). Redan dessa resultat indikerar alltså att effekterna av små klasser i STAR-experimentet framförallt gäller afroamerikanska elever och att den positiva effekten försvinner med tiden.

Dessutom står det klart att STAR-experimentet inte var speciellt väl genomfört, vilket gör att ovanstående resultat är skakiga (Hanushek 2003). Endast 48 procent av eleverna som påbörjade experimentet i dagis var kvar efter tredje klass, när experimentet avslutades. Dessutom står det inte klart om lärare av olika kvalitet tilldelades slumpmässigt till klasser med olika storlek.¹⁴ Om bättre lärare tilldelades till de små klasserna blandas effekten av lärarkvalitet ihop med

¹³ Se Chingos (2013).

¹⁴ De randomiserades efter observerbara egenskaper, såsom utbildning och erfarenhet, men sådana egenskaper har oftast ingen påverkan på lärarnas förmåga att höja elevers resultat (Hanushek och Rivkin 2012).

effekten av mindre klasser. Intressant nog fanns det också endast positiva effekter av små klasser när eleverna undervisades av erfarna lärare (Mueller 2013), vilket också indikerar att effekterna av små klasser mycket riktigt beror på lärarna. Till råga på allt bytte ungefär 10 procent av eleverna klass med olika storlekar under experimentets gång (Ding och Lehrer 2010). STAR-experimentet led alltså av problem som i sin tur kan påverka resultaten.

Intressant nog har nyare forskning, där problemet med avhopp från experimentet inte är lika relevant, funnit relativt små effekter av mindre klasser på långsiktiga utfall. Dynarski, Hyman och Schanzenbach (2011) finner att mindre klasser ökade sannolikheten att elever läser vidare på universitet med 2,7 procentenheter (7 procent). Dessa effekter är dock koncentrerade bland afroamerikanska och fattiga elever – det finns inga effekter bland vita elever och elever som inte är fattiga.¹⁵ Sannolikheten att elever tar en två-årig högskoleexamen ökar samtidigt med 0,7 procentenheter (26 procent), men endast bland elever i skolor med en hög andel fattiga. Det finns dock ingen generell statistiskt signifikant effekt på sannolikheten att eleverna tar en kandidatexamen eller högre examen.¹⁶

Samtidigt finner Chetty m.fl. (2011) att de mindre klasserna inte hade någon genomsnittlig effekt på sannolikheten att elever gick på universitet någon gång innan de var 27 år gamla och kvaliteten på universitetet de läste vidare på, mätt i hur mycket eleverna som går där tjänar när de är 28 år. Däremot ökade de sannolikheten att eleverna läste på universitet vid 20 års ålder med 1,8 procentenheter. Detta motsvarar en effektstorlek på 0,04 sd, en väldigt liten effekt som dessutom endast är svagt statistiskt signifikant.

Dessutom verkar effekten på sannolikheten att eleverna läste på universitet vid 20 års ålder vara helt koncentrerad till pojkar (2,6 procentenheter/0,06 sd) – flickor tjänar ingenting på mindre klasser när det gäller detta mått. Likaså är effekten koncentrerad till icke-fattiga elever (3,6 procentenheter/0,08 sd) och sannolikt även svarta elever, även om detta är mer osäkert.¹⁷ När det gäller sannolikheten att elever läste på universitet någon gång innan de var 27 år gamla är effekten positiv bland svarta elever (5,3 procentenheter/0,11 sd) och fattiga elever (3,9 procentenheter/0,08 sd). När det gäller universitetskvalitet finns det endast positiva effekter bland svarta elever och pojkar som indikerar att de gick på universitet där eleverna hade ungefär 250 dollar (0,06 sd) högre genomsnittliga löner vid 28 års ålder än motsvarande elever som gick i större

¹⁵ Det senare är viktigt för den utbildningspolitiska relevansen i Sverige – se avsnittet "Hela USA" i "Forskning som analyserar andra reformer", som diskuterar Jackson, Johnson och Persicos (2014) studie, för mer diskussion kring detta.

¹⁶ Författarna rapporterar inte standardavvikelsen och jag kan därför inte omvandla resultaten till detta mått.

¹⁷ Onlineappendixen visar att effekten bland svarta elever var 2,72 procentenheter men inte statistiskt signifikant (standardfel: 2,04) och bland vita 1,07 procentenheter men inte statistiskt signifikant (standardfel: 1,1).

klasser.¹⁸ Redan detta indikerar att de långsiktiga effekterna av mycket mindre klasser inte är speciellt starka och att de varierar beroende på elevgrupp.¹⁹

Men kanske viktigast av allt är att det inte finns någon påverkan på elevernas löner eller inkomster senare i livet i genomsnitt, när eleverna är 25-27 år gamla. Däremot finns det faktiskt en statistiskt signifikant *negativ* effekt av mindre klasser bland flickor som indikerar att dessa får 1 130 dollar per år (0,07 sd) lägre i lön när de är 25-27 år gamla på grund av att de gick i mindre klasser.²⁰ Enligt STAR-experimentet *förlorade* alltså flickor på att gå i mindre klasser i ett längre och ekonomiskt perspektiv. Ingen av de andra elevgrupperna påverkades när det gäller lönerna. Detta är ett resultat som tidigare har ignorerats, både bland akademiker och bland media, troligtvis för att författarna endast rapporterar det i ett onlineappendix utan att notera det i uppsatsen.

Problem med de långsiktiga studierna

De långsiktiga resultaten är dock långt ifrån säkra på grund av icke-slumpmässiga dödsfall bland eleverna. Muenning, Johnson och Wilde (2011) finner att sannolikheten att eleverna dog före 2008 bland elever som gick i mindre klasser var högre än bland eleverna som gick i större klasser. Sannolikheten för att eleverna dog innan 2008 ökade med hela 58 procent jämfört med elever som gick i större klasser, en effekt som var koncentrerad bland pojkar, icke-fattiga elever och vita/asiatiska elever. Detta är inte bara ett problem för att eleverna dog, vilket är ett viktigt utfall i sig, men också eftersom dödsfall har samma effekt som avhopp för experimentets validitet.

Intressant nog finner Chetty m.fl. ingen effekt på sannolikheten att eleverna dör med data från Social Security Administration (SSA). De har tillgång till 95 procent av alla elevers dödsintyg, medan Muenning, Johnson och Wilde har tillgång till 99 procent med data från National Death Index. Enligt amerikanska epidemiologer är National Death Index den mer tillförlitliga källan, även om det är svårt att veta helt säkert att detta driver skillnaderna i resultaten.²¹ En sak

¹⁸ Författarna rapporterar endast standardavvikelse för hela urvalet och jag har därför rapporterat effektstorleken beräknat efter denna.

¹⁹ Inte heller finns det genomsnittligt positiva effekter på sannolikheten att eleverna har gift sig, att de är husägare, eller att de bor i områden med en hög andel människor med universitetsutbildning. Författarna finner positiva effekter på ett specifikt index som inkluderar dessa variabler, men denna drivs helt av en annan variabel. Små klasser ökar nämligen sannolikheten att elever har ett pensionskonto med 0,06 sd. Författarna analyserar inte skillnader mellan olika grupper elever förutom på det totala indexet och dessa indikerar inga positiva effekter bland flickor eller fattiga elever. Dock verkar mindre klasser minska sannolikheten att bli gravid i tonåren bland flickor (se Schanzenbach 2006/2007).

²⁰ Detta kan inte förklaras av att flickor som gick i mindre klasser läste vidare på universitet i högre utsträckning, vilket skulle kunna göra att de har lägre löner vid 25-27 års ålder på grund av mindre arbetserfarenhet. Det finns nämligen ingen effekt av små klasser på sannolikheten att flickor läser vidare på universitet överhuvudtaget.

²¹ Se Fillenbaum, Burchet och Blazer (2009), Huntington m.fl. (2013) och Sesso, Paffenbarger och Lee (2000) för forskning och diskussion gällande de båda databaserna. Chetty m.fl. (2011) hade inte tillgång till Muenning, Johnson och Wildes (2011) publicerade studie, utan endast ett manuskript, vilket kanske förklarar varför de rapporterar deras data fel: Chetty m.fl. menar att Muenning, Johnson och Wilde endast finner 141 dödsfall bland de som deltog i STAR innan 2008,

som kan göra detta är att Chetty m.fl. analyserar dödsfall före 2010, medan Muenning, Johnson och Wilde analyserar dödsfall före 2008. Det senare är viktigare eftersom Chetty m.fl. analyserar data gällande de långsiktiga resultaten till och med 2007. Över den period som Chetty m.fl. analyserar långsiktiga resultat är det alltså möjligt att sannolikheten att elever i mindre klasser dog vara högre än sannolikheten att elever i större klasser dog. Detta är ett viktigt resultat i sig men indikerar också att vi bör vara försiktiga när det gäller tolkningen av de existerande långsiktiga resultaten från STAR-experimentet.²²

Med andra ord är de långsiktiga effekterna förvånansvärt svaga och blandade givet en så stor klasstorleksminskning. Dessutom är resultaten inte alls säkra eftersom det verkar som att elever från de små klasserna dog i högre utsträckning. Men även om vi ignorerar detta står det klart att STAR-experimentets externa validitet är väldigt begränsad. Experimentet genomfördes i 1980-talets Tennessee och fokuserade på skolor i innerstadsområden med utsatta elever.²³ Oavsett om experimentet funnit starka positiva effekter i ett långsiktigt perspektiv och om det hade genomförts perfekt hade den utbildningspolitiska relevansen för att genomföra resursökningar för andra slags elever varit låg.

Fältexperiment som genomfördes innan andra världskriget

Project STAR är det enda fältexperimentet i modern tid. Däremot genomfördes många experiment innan andra världskriget. Rockoff (2009) rapporterar resultaten från dessa och visar att de generellt *inte* finner några positiva effekter av klasstorlek, trots att de verkar ha genomförts på korrekt sätt. Det är naturligtvis svårt att dra alltför stora slutsatser av dessa experiment eftersom de genomfördes för så länge sedan, men författaren drar slutsatsen att STAR-experimentet inte är tillräckligt för att vi ska kunna veta huruvida små klasser faktiskt bidrar till bättre resultat.

Det bör även betonas att det finns generella problem med fältexperiment som kan minska deras utbildningspolitiska relevans ytterligare. Detta på grund av att man inte kan säkerställa att aktörer – exempelvis elever, lärare och föräldrar – ändrar sitt beteende på grund av resurstilldelningen. Detta gör att experiment

medan siffran i författarnas publicerade studie är 146. Chetty m.fl. finner 154 dödsfall innan 2008 med färre observationer. Givet skillnaderna mellan databaserna är det långt ifrån säkert att de extra 8 dödsfallen handlar om ökad kvalitet i matchningen mellan elev och döds-certifikat, som Chetty m.fl. menar. Dessutom visar andra rapporter att 750 amerikaner registreras som avlidna hos SSA varje månad trots att de lever (Fahrenthold 2013). I september 2006 tvingades exempelvis SSA att "återuppliva" 23 366 människor som felaktigt hade registrerats som döda mellan januari 2004 och september 2005 (Johnson och Amons 2008). Mellan maj 2007 och april 2010 rapporterades också 36 657 felaktiga dödsfall (Ellis 2011). Dock registrerar SSA dödsfall som sker utomlands, vilket National Death Index inte gör (Hill 2001). Totalt sett verkar det dock orimligt att lita fullt ut på resultaten som Chetty m.fl. presenterar.

²² Wilde m.fl. (2011) finner också andra långsiktiga resultat, men dessa är inte tillförlitliga eftersom de endast lyckas matcha 54 procent av eleverna som deltog i STAR.

²³ På grund av beslutsprocessen gällande vilka skolor som skulle delta i experimentet var det stark övervikt på skolor i utsatta områden (Hanushek 2003).

lider av potentiella "Hawthorneeffekter".²⁴ Dessa kan aldrig uteslutas – eftersom resurstilldelningen i sig kan orsaka förändringar i skolors, elevers, föräldrars och lärares beteende – men de kan helt klart förvärras av att aktörerna vet om att eleverna jämförs med en kontrollgrupp och därför vet vad som står på spel i experimentet. Man kan exempelvis tänka sig att både föräldrar och lärare, som tenderar att föredra små klasser, jobbar extra hårt för att visa att de har rätt – och att de på det sättet försöker påverka beslutsfattare att minska klasstorleken för alla grupper efter att experimentet är avslutat. Samtidigt kan effekterna naturligtvis gå åt andra hållet om aktörerna börjar jobba hårdare som svar på att de får mindre resurser.

Det står därför klart att vi inte kan dra alltför stora utbildningspolitiska slutsatser, speciellt inte i Sverige, från vare sig det enda stora fältexperimentet i modern tid eller de som genomfördes innan andra världskriget. Nedan diskuteras istället resultat från så kallade naturliga experiment, samt resultat från studier med andra kvasiexperimentella metoder, där forskare använder alternativa strategier för att på bästa sätt analysera slumpmässig variation i resurstilldelningen.

Studier som analyserar variationen i resurser vid ett tillfälle

En hel del studier har analyserat resurseffekterna vid en specifik tidpunkt med olika metoder för att utröna orsakssamband. Fördelen är här att resursskillnaderna ofta är väldigt stora, vilket ökar sannolikheten att man hittar positiva effekter. Men detta är snarare en nackdel från en utbildningspolitisk synvinkel. Som nämntes i avsnitt 2.b är så stora resursförändringar som analyseras ofta alldeles för dyra för att kunna genomföra i praktiken. Och eftersom effekterna mycket väl kan vara icke-linjära är det riskabelt att extrapolera resultaten till mindre resursökningar som inte är lika dyra.

Hursomhelst diskuteras studierna i detalj, med speciellt fokus på de svenska resultaten. Sammantaget är denna forskning blandad: en liten del av forskningen finner meningsfulla effekter av resurser, medan en hel del studier endast finner små effekter och andra inte någon effekt alls. Detta beror delvis på variationen i resurser och hur stor den är, vilket land och vilka elever som analyseras, samt vilken metod som används. Det finns också skillnader när det gäller metodologin som gör att vissa studier på bättre sätt kan analysera effekterna av resurser "allt annat lika".

Sverige

I en studie som har fått stort genomslag i den svenska debatten analyserar Fredrikson, Öckert och Oosterbeek (2013) elever som gick i mellanstadiet på

²⁴ Krueger (1999) försöker analysera detta i STAR-experimentet genom att estimerar effekten av mindre klasser i kontrollgruppen, men detta kan inte göras inom ramen för experimentet. Eftersom vi inte har någon aning varför klasserna är mindre i delar av kontrollgruppen kan vi inte heller dra orsakssamband från denna beräkning.

1980-talet och början på 1990-talet. De använder en regel som stipulerade att det som mest fick vara 30 elever i en klass för att få fram slumpmässig variation i klasstorleken. Denna regel stipulerade att högst 30 elever fick gå i samma klass – vilket exempelvis innebar att skolor med 31 elever i samma årskull delades upp i två klasser, en med 15 elever och en med 16. Med 61 elever blir det tre klasser och med 91 elever blir det fyra klasser etc. Författarna utnyttjar denna regel och jämför skolor som hade något färre elever än vad gränsen stipulerade med de som hade något fler elever, vilket därför ledde till mycket mindre klasser. På detta sätt kan de analysera den kausala effekten av stora förändringar i klasstorleken.

Författarna finner positiva effekter av små klasser på akademiska och icke-akademiska utfall i både ett kortsiktigt och ett långsiktigt perspektiv. Resultaten indikerar att en klasstorleksminskning med 5 elever under tre år i mellanstadiet ökar (1) elevernas kognitiva förmåga med 0,17 sd när de är 13 år, (2) elevers icke-kognitiva förmåga med 0,13 sd när de är 13 år, (3) resultaten när eleverna är 16 år med 0,12 sd, (4) antal års utbildning med 0,10 sd, (5) sannolikheten att ta en kandidatexamen ökar med 38 procentenheter, (6) elevers framtida inkomster med 5,85 procent och löner med 3,78 procent jämfört med medelvärdet. Dessutom (7) ökar manliga elevers icke-kognitiva förmåga i mönstringsprovet med 0,15 sd.

Det finns dock ingen effekt på sannolikheten att eleverna arbetar senare i livet, eller på kognitiv förmåga i mönstringsprovet bland män. Författarna finner heller inga skillnader beroende på elevens bakgrund förutom att sannolikheten att eleven arbetar senare i livet ökar bland elever från lägre socioekonomisk bakgrund (medan den faktiskt *minskar* bland de från medelhög socioekonomisk bakgrund), och att effekten när det gäller löner endast gäller elever från hög socioekonomisk bakgrund.

Det står klart att studien är väl genomförd. Men det finns även saker som gör att man kan ifrågasätta dess utbildningspolitiska relevans. Först och främst fokuserar författarna endast på elever i rektorsområden som har en skola. Detta öppnar för möjligheten att resultaten inte kan extrapoleras till andra elever, vilket mycket väl kan vara fallet som diskussionen nedan visar. För det andra förändras vissa av resultaten kraftigt på grund av små förändringar i metodologin. För det tredje finns det stöd för att föräldrar flyttar för att undvika större klasser, vilket riskerar överdriva effekterna.

Intressant nog skiljer sig resultaten som författarna presenterar i online-appendixen där det större urvalet används (elever som går i rektorsområden med endast en skola motsvarar endast 20 procent av urvalet). Där är effekten inte statistiskt signifikant i genomsnitt, förutom när det gäller akademiska resultat när eleverna är 13 år samt för deras löner senare i livet. Att de totala inkomsterna inte påverkas är speciellt noterbart, speciellt eftersom urvalet där löner observeras är 47 procent mindre. De presenterar dock inte hur resultaten på mönstringsprovet skiljer sig. När det fullständiga urvalet analyseras finns

alltså inga statistiskt signifikanta effekter på de flesta utfall, medan punktestimaten är liknande.²⁵

Dock presenterar författarna inte de heterogena effekter som de finner i rapporten som den publicerade studien baseras på (Fredriksson, Öckert och Oosterbeek 2012). Denna använder endast det fullständiga urvalet av elever och en något annorlunda metodologi.²⁶ Likt den publicerade studiens onlineappendix finner författarna i denna rapport (1) inga genomsnittliga effekter på antal års utbildning, sannolikheten att eleven tar en kandidatexamen, sannolikheten att eleven arbetar i vuxen ålder, samt fullständiga inkomster i vuxen ålder. De finner här heller (2) inga effekter i genomsnitt på mönstringsprovet, vare sig det gäller kognitiv förmåga eller icke-kognitiv förmåga.

Dessutom finner de (3) framförallt positiva effekter bland elever från familjer med högre inkomst, oavsett om de analyserar kortsiktiga utfall eller långsiktiga utfall. Bland elever från lägre socioekonomisk bakgrund finns *ingen* effekt förutom en svagt statistiskt signifikant effekt på kognitiv förmåga i mönstringsprovet. Effekten (4) varierar också mellan flickor och pojkar. Det finns inga effekter bland flickor på kognitiv och icke-kognitiv förmåga eller akademiska resultat när de är 13 och 16 år gamla. När det gäller dessa resultat är effekterna koncentrerade bland pojkar. Bland flickor finns det däremot en positiv effekt på antalet års utbildning, samt en svagt statistiskt signifikant effekt på sannolikheten att de tar en kandidatexamen samt på deras inkomster senare i livet. Det finns dock inga effekter på deras löner eller sannolikheten att de arbetar. Bland pojkar finns det istället endast en effekt på löner – det finns inga effekter på antalet års utbildning, sannolikheten att de tar en kandidatexamen, totala inkomster eller sannolikheten att de arbetar. Detta skiljer sig markant från den publicerade studien där få skillnader hittas.

Det står också klart att urvalet i denna uppsats är något större (max 30 818 elever) än det fullständiga urval som presenteras i online-appendixet till den publicerade studien (max 29 371 elever). Detta beror på att författarna har utelämnat elever i små skolor i den publicerade studien.²⁷ Detta och/eller den något annorlunda metodologin påverkar resultaten i olika riktning beroende på vilken variabel som analyseras, även när det gäller de genomsnittliga resultaten: (1) effekten på kognitiv förmåga bland 13-åringar är mindre än *hälften* av den i online-appendixet till den publicerade studien; (2) effekten på icke-kognitiv förmåga bland 13-åringar är mindre men är marginellt statistiskt signifikant,

²⁵ Författarna menar att detta beror på att metodologin är mer relevant i rektorsområden med en skola, vilket diskuteras mer nedan.

²⁶ Skillnaden är att författarna antar att klasstorleksregelns effekter varierar vid de olika trösklarna (30, 60, 90 elever etc.), I rapporten från 2012 antar de inte explicit att effekten varierar kring de olika trösklarna, men inkluderar istället fler polynomer för antalet elever som kontroll. Dessutom analyserar de effekten av att gå i en skola med antalet elever över/under en tröskel i studien från 2013, medan de analyserar effekten av klasstorleken som tröskeln förutspår (givet antalet elever) i rapporten från 2012.

²⁷ Detta bekräftades av Öckert i ett telefonsamtal 22 augusti 2014. Utelämningen berodde på metodologiska skäl och i onlineappendixet inkluderades eleverna från små skolor igen (där författarna använder samma metodologi som i rapporten från 2014, som diskuteras nedan).

vilket den inte är i online-appendixet till den publicerade studien; och (3) effekten på resultaten när eleverna är 16 år är statistiskt signifikant, vilket den inte är i online-appendixet. Poängen är att resultaten verkar fluktuera även beroende på ganska små skillnader i urvalets storlek och/eller en något annorlunda metodologi.

Metoden för hur författarna analyserar skillnaderna mellan olika elevgrupper skiljer sig också åt. I den publicerade studien använder författarna en så kallad interaktionsmodell där effekten av små klasser multipliceras med en indikator gällande elevers bakgrund/kön. Men i rapporten från 2012 använder de istället separata modeller för de olika elevgrupperna. Till skillnad från interaktionsmodellen tillåts här effekten av alla kontrollvariabler också att variera beroende på elevens bakgrund/kön.²⁸

Att skillnader i urvalet/metodologin spelar roll står klart i författarnas nya studie som fokuserar på vad som förklarar den positiva effekten av små klasser (Fredriksson, Öckert och Oosterbeek 2014). I denna rapport används endast urvalet från rektorsområden med en skola, men återigen inkluderas elever från små skolor liksom rapporten från 2012 och separata modeller för elever från hög och låg socioekonomisk bakgrund används. De använder också en något annorlunda metodologi jämfört med huvudmodellerna i tidigare studier.²⁹ Här finner författarna att små klasser endast är positiva för 13-åringars kognitiva förmåga bland elever från låg socioekonomisk bakgrund, vilket de alltså inte finner i den publicerade studien från 2013.

Kanske än mer häpnadsväckande är att detta är *totalt motsatt* vad de finner i rapporten från 2012, som också använder separata modeller för elever från låg och hög socioekonomisk bakgrund – men med en något annorlunda metodologi och det fullständiga urvalet. Som noterades ovan finner den uppsatsen att endast elever från *hög* socioekonomisk bakgrund tjänar på små klasser. Detta är bekymmersamt eftersom det indikerar att resultaten för olika elevgrupper varierar beroende på urvalet och/eller små skillnader i metodologin.³⁰

När vi synar resultaten i appendixet i rapporten från 2014 ser vi också att de genomsnittliga resultaten även skiljer sig jämfört med den publicerade studien. Här finns nämligen ingen effekt på antal års utbildning, sannolikheten att eleverna tar en kandidatexamen eller på icke-kognitiv förmåga vid 13 års ålder. Båda studierna fokuserar på det begränsade urvalet i rektorsområden med en

²⁸ Det bör dock betonas att definitionen av låg och hög socioekonomisk bakgrund skiljer sig mellan de rapporterna, jämfört med den publicerade studien från 2013.

²⁹ Skillnaden är att de här explicit antar att effekten av klasstorleksregeln varierar vid de olika trösklarna (30, 60, 90 elever etc.), vilket de inte gör i rapporten från 2012 (där de istället inkluderar fler polynomer för antalet elever som kontroll). Dessutom analyserar de effekten av att gå i en skola med antalet elever över/under en tröskel, medan de i rapporten från 2012 analyserar effekten av klasstorleken som tröskeln förutspår (givet antalet elever). Skillnaden med den publicerade studien från 2013 är endast att de använder separata indikatorer vid varje tröskel för att ta hänsyn till varierande effekter på ett något mer effektivt sätt.

³⁰ I telefonkorrespondens menade Öckert att det framförallt beror på skillnader i metodologi, eftersom deras strategi fokuserar på elever i rektorsområden med endast en skola oavsett vilket urval som används. Detta diskuteras mer nedan.

skola. Det verkar alltså som att författarnas något annorlunda metodologi och inkludering av små skolor i analysen har betydelse också för genomsnittresultaten.

Det är inte helt klart om det är skillnaderna i metodologi eller urval som driver skillnaderna i resultat mellan studierna. Författarna menar att de använder det mindre urvalet för att de vill undvika att strategin blir irrelevant i skoldistrikt med fler än en skola. De menar också att eftersom punktestimaten är lika i det större urvalet, men mindre precisa, är det sannolikt att skillnaderna beror på att metodologin inte fungerar lika bra i rektorsområden med fler än en skola. Metodologin förutsätter att klasstorleken påverkas när antalet elever i fjärde klass i rektorsområdet når taket som stipuleras av klasstorleksregeln. Men om det finns flera skolor i rektorsområdet är detta inte säkert – just eftersom eleverna är spridda på olika skolor – och då fungerar inte strategin. I telefonkorrespondens menade Öckert att metodologin i princip endast är relevant i rektorsområden med en skola och att studien inte skulle existera om de endast analyserade rektorsområden med fler än en skola. Men stämmer detta fokuserar metodologin *endast* på elever i rektorsområden med en skola – även när det fullständiga urvalet analyseras.³¹ Det är i sådana fall inte konstigt om punktestimaten inte förändras speciellt mycket, men det går då heller inte att avgöra om klasstorlekseffekterna omfattar elever i rektorsområden med fler än en skola.³²

Intressant nog är resultaten dock faktiskt oftast mer eller lika precisa i rapporten från 2012, med det fullständiga urvalet (inklusive de små skolorna), som i den publicerade studien och rapporten från 2014 – även om man endast jämför med det mindre urvalet. Detta verkar framförallt vara ett resultat av skillnaden i metodologi.³³ Men det är dock sannolikt att kombinationen urval och metodologi driver skillnaderna totalt sett.

Implikationerna av detta är följande: antingen (1) varierar resultaten beroende på huruvida det begränsade eller fullständiga urvalet används; (2) gäller

³¹ Detta är på grund av att metodologin endast kan estimerar effekten bland elever där klasstorleksregeln faktiskt gör att klasstorleken varierar beroende på antalet elever i rektorsområdet. I de fall den inte gör det plockas inte effekterna upp, oavsett vilket urval man analyserar.

³² I Tabell A.1 i den publicerade studien presenterar författarna deskriptiv statistik och finner inga större skillnader totalt sett mellan genomsnittet i det begränsade och totala urvalen. Först och främst överdrivs troligtvis avsaknaden av skillnader eftersom författarna jämför genomsnittet av variablerna i områden med en skola med genomsnittet i *alla* områden (snarare än genomsnittet i områden med fler än en skola). För det andra är föräldrarnas snittinkomst 4 procent lägre i det fullständiga urvalet och deras genomsnittliga antal års utbildning är också 0,2–0,3 år lägre även i denna jämförelse. För det tredje kan rektorsområden med ansvar över fler än en skola också skilja sig från de med ansvar över en skola på sätt som inte kan observeras, vilket kan påverka klasstorlekens effekter. Exempelvis kan ledarskapet vara annorlunda beroende på hur många skolor och elever området har ansvar över. Det är alltså långt ifrån säkert att resultaten kan extrapoleras till andra elever.

³³ I Tabell AV i den publicerade studiens onlineappendix redovisar författarna känslighetsanalyser med en liknande metodologi som i rapporten från 2012, men med det begränsade urvalet, och finner liknande precision i de flesta av (men inte alla) resultaten som de som redovisas i Tabell 3 och 4 i rapporten från 2012.

resultaten endast elever i rektorsområden med fler än en skola (oavsett vilket urval som används); (3) drivs skillnaderna av den något annorlunda metodologin och/eller (4) att författarna inkluderar/exkluderar små skolor. Poängen är dock densamma: det är oklart hur robusta resultaten är och svårt att basera politik på.

En sista poäng är att författarna i rapporten från 2014 finner att föräldrar, speciellt de från låg socioekonomisk bakgrund, tenderar att flytta ut ur rektorsområdet oftare om de går i större klasser – och att de elever som flyttar mycket riktigt går i mindre klasser som ett resultat av detta. Detta är ett problem eftersom det indikerar att mer motiverade föräldrar undviker större klasser genom att flytta från rektorsområdet, vilket i sin tur kan överdriva effekterna av mindre klasser.³⁴

Den publicerade versionen av författarnas studie – som under 2014 har fått stor spridning i media och politisk debatt – fokuserar alltså inte på klasstorleksminskningar bland den stora majoriteten elever som gick i mellanstadiet i rektorsområden med fler än en skola på 1980-talet och början på 1990-talet. Vissa av resultaten förändras kraftigt, speciellt gällande effekterna på olika elevgrupper, beroende på urval och metodologiska detaljer.³⁵ Dessutom finner de tecken på att elever i större utsträckning flyttar till andra rektorsområden när klasserna är större. Sammantaget står det klart att studien inte bör användas som huvudsakligt argument för de som vill minska klasstorleken, i varje fall över hela landet – osäkerheten i studien är helt enkelt alltför stor.³⁶

³⁴ Författarna menar att det är oklart om skolbyte är positivt netto, trots minskningen i klasstorlek, och refererar till forskning som finner negativa effekter. Men uppsatsen som refereras indikerar faktiskt att skolbyten i många fall inte alls har negativa effekter i ett kortsiktigt perspektiv och istället är positivt för resultaten åren efteråt (Hanushek, Kain och Rivkin 2004). Problemet är att denna uppsats metodologi troligtvis inte är tillräcklig för att komma åt orsakssambanden samt att den inte explicit separerar frivilliga från påtvingade skolbyten. En annan studie som på bättre sätt kommer åt orsakssambanden finner istället positiva effekter av frivilligt skolbyte och negativa effekter av påtvingade skolbyten (Schwartz och Stiefel 2012). Det är alltså sannolikt att frivilliga skolbyten, exempelvis på grund av att föräldrar vill undvika stora klasser, överdriver effekten av klasstorlek.

³⁵ Ett annat problem, som diskuteras senare i rapporten, är att annan forskning har funnit att ökade resurser och mindre klasser leder till högre betygsinflation och att lärare manipulerar resultaten i större utsträckning (Angrist, Battistin och Vuri 2014; De Witte, Seys och Solondz 2014). Detta påverkar troligtvis inte de långsiktiga resultaten i den svenska studien direkt, men däremot finns möjligheten att de påverkas indirekt. Den första kunskapskontrollen som analyseras är nämligen prov i slutet på mellanstadiet som administrerades av elevernas lärare (Svensson 1971). Detta öppnar för att resultaten kan bero på manipulering, av anledningar som diskuteras längre fram, vilket i sin tur kan ha kaskadeffekter som påverkar eleverna senare i livet (se t.ex. Murphy och Weinhardt 2013). Högre prestationer i lägre åldrar som beror på annat än högre kunskaper kan påverka resultaten och i slutändan även ekonomiska utfall positivt senare i livet också. Även om detta naturligtvis är spekulativt går det inte att utesluta att de positiva effekter av små klasser som faktiskt är någorlunda robusta inte delvis beror på dessa effekter.

³⁶ Det finns även andra problem med studien, exempelvis att effekten kan ha med annat än klasstorlek att göra. Mindre klasser kan exempelvis leda till mer åldersintegrerade klasser, eftersom skolor kan försöka spara resurser, vilket i sig kan ha en effekt på resultaten. Mycket riktigt finner författarna i deras nya rapport att mindre klasser ökar sannolikheten att eleverna går i åldersintegrerade klasser (Fredriksson, Öckert och Oosterbeek 2014). Skolor med mindre klasser kan även använda mer nivågruppering. Om elever i skolor med mindre klasser

Med en annorlunda ansats analyserar Lindahl (2005) 556 sjätteklassare i 16 skolor i Stockholm som skrev ett prov under vårterminen i femte klass samt i början av höstterminen och i slutet på vårterminen i sjätte klass. Genom att analysera skillnaden i förändring i provresultat under sommaren med förändringen i resultat under skolåret kan författaren kontrollera för hur mycket elever lär sig på egen hand, när klasstorleken inte bör ha någon effekt. Han finner positiva effekter av små klasser: 5 färre elever i en klass producerar 0,8-0,13 sd högre resultat.³⁷ Studien ignorerar dock att lärare som är olika bra inte slumpmässigt placeras i klasser med olika storlek. Dessutom är det naturligtvis svårt att dra alltför stora slutsatser av resultaten på grund av det lilla urvalet som är begränsat till Stockholm.

Det är därför intressant att notera två andra studier som analyserar effekten av klasstorlek i olika länder, inklusive Sverige, när det gäller resultaten i Timss 1995 och 2003. Woessmann (2005) analyserar resultaten i Timss 1995 bland 13- och 14-åringar i 15 europeiska länder och använder två olika strategier. För Sverige används samma strategi som Fredriksson, Öckert och Oosterbeek (2013) tillämpar. Författaren finner inga effekter på Timss-resultaten i matematik i Sverige. Effekten är nästan statistiskt signifikant i en modell, men även om vi räknar med den är effektstorleken väldigt liten: en halvering av klasstorleken skulle förbättra resultaten i Timss med 11 poäng (0,11 sd). Det är anmärkningsvärt att en liknande metodologi som används av Fredriksson, Öckert och Oosterbeek (2013) inte genererar positiva effekter i Timss 1995. En svaghet med studien är dock att Woessmann (2005) inte analyserar om föräldrar eller skolor manipulerar klasstorleksregeln, eller om skoldistriktet kompenserar svagare elever med lägre klasstorlek genom att förändra skolornas upptagningsområden, vilket kan påverka resultaten i olika riktningar.³⁸

I en färskare studie analyserar Altinok och Kingdon (2012) Timss 2003 med en annorlunda strategi som går ut på att jämföra elevers resultat vid samma tillfälle, men i olika ämnen där klasstorleken skiljer sig åt. De analyserar om skillnaden i elevers klasstorlek i matematik och naturvetenskap påverkar skillnaden i resultaten i dessa ämnen. Detta gör att de håller icke-observerbara skillnader mellan elever konstanta. Resultaten indikerar att effekterna av klasstorlek är minimala i Sverige. Detta gäller både när författarna antar att effekten är linjär och när de antar att effekten är icke-linjär.

nivågrupperades i högre utsträckning handlar författarnas resultat även om effekterna av nivågruppering, inte endast klasstorlek. Författarna finner att skolor med mindre klasser visserligen inte separerade elever efter bakgrund, men detta betyder inte att de inte nivågrupperade efter förmåga, vilket författarna inte kan analysera.

³⁷ Detta beror på om man räknar standardavvikelsen efter förändringen i provresultat eller efter resultaten i vårterminen i årskurs 6.

³⁸ Woessmann (2005) analyserar effekterna på skolnivå, medan Fredriksson, Öckert och Oosterbeek analyserar effekterna på rektorsområdesnivå. I fjärde klass har de senare författarna funnit tecken på att rektorsområdena förändrade upptagningsområdena för att gynna elever med lägre socioekonomisk bakgrund så att dessa fick lägre klasstorlek. Detta kan bidra till att minska effekten av små klasser. Det är dock oklart om detta även sker på högstadienivå och om detta inte kontrolleras med bakgrundsvariablerna som Woessmann (2005) inkluderar.

Det bör noteras att studien använder en annorlunda metod som kan kritiseras om det finns skäl att tro att hög- och lågpresterande elever sorteras in i klasser med olika storlekar i olika ämnen. Precis som med Lindahls (2005) studie är ett annat problem att lärare med olika kvalitet kan placeras i klasser med olika storlek beroende på ämne. Författarna tar hänsyn till dessa problem genom att endast analysera skolor som inte använder nivågruppering.³⁹ De kontrollerar även för en proxy gällande elevers ämnesspecifika förmåga, baserat på elevens självskattade inlärningsförmåga i ämnet. De håller även konstant lärarnas ålder, erfarenhet, utbildning och kön. Detta förändrar inte resultaten nämnvärt.

Den sista analysen är dock viktigast för att resultaten ska vara trovärdiga. Författarna använder där nämligen variationen i den faktiska klasstorleken som kan förklaras av skolans genomsnittliga klasstorlek i varje ämne (oavsett vilken årskurs de går i). De kan därmed kontrollera för sortering av elever till olika klasser inom olika skolor och att skolor kan ändra resurstilldelningen till olika årskurser. Effekten av mindre klasser är nu marginellt *negativ* i Sverige, men effektstorleken är i princip betydelselös. På grund av den sistnämnda analysen ökar förtroendet för att författarna analyserar orsakssamband.

Vad förklarar skillnaderna mellan forskningen om internationella prov och de andra studierna? En förklaring kan delvis vara att Altinok och Kingdons (2012) studie analyserar skillnader i klasstorlek generellt, medan Fredriksson, Öckert och Oosterbeek (2013) endast fokuserar på skillnader i klasstorlek som uppstår runt klasstorleksregeln, vilket inte nödvändigtvis är utbildningspolitiskt relevant. Woessmann (2005) finner dock inga effekter med en liknande metodologi i Timss, medan Lindahl (2005) inte utnyttjar klasstorleksregeln men ändå finner positiva effekter. En annan möjlig förklaring är naturligtvis att Timsstudierna analyserar klasstorlekseffekter med ett nationellt urval av 13- och 14-åringar på 1990- och 2000-talet, vilket skiljer sig kraftigt från Fredriksson, Öckert och Oosterbeeks (2013) fokus på mellanstadieelever i rektorsområden med en skola under 1980-talet och början på 1990-talet, samt Lindahls (2005) fokus på ett urval av Stockholms eleverna.

Vi vet i slutändan inte vad som gör att effekterna skiljer sig åt, men det är hursomhelst noterbart att forskningen inte finner positiva effekter av mindre klasser i Timss, vilket tillsammans med Pisa nu spelar en viktig roll i den svenska debatten.

Chile

Skillnaderna mellan det svenska skolsystemet på 1980-talet/början på 1990-talet och idag är värt att betona ytterligare för att förstå forskningens relevans. I och med skolpengssystemet och kommunaliseringen förändrades incitamenten bland aktörerna i skolan. Urquiola och Verhoogens (2009) studie är relevant i detta sammanhang eftersom de analyserar Chiles skolmarknad som till viss del liknar den svenska. Genom att använda en regel för hur stora klasserna får vara

³⁹ Intressant nog rapporterar 66 procent av skolorna att de använder nivågruppering matematik, medan siffran i naturvetenskap endast är 10 procent.

finner de först positiva effekter av små klasser i friskolor i grundskolan. Detta visar sig dock bero på att föräldrar från högre socioekonomisk bakgrund sorterar in sig i skolor med små klasser och när de kontrollerar för detta försvinner effekten helt.⁴⁰ Författarna gör framförallt en metodologisk poäng, men resultaten indikerar även att effekten av klasstorlek kan vara annorlunda när man tillåter skolval, vilket ger större möjligheter för föräldrar att reagera på förändringar i resurstilldelningen.

Danmark

I Danmark analyserar Krassel och Heinesen (2014) ett frivilligt 10:e år i grundskolan, som utnyttjas av ungefär 50 procent av eleverna. Det extra året är framförallt till för elever som antingen är akademiskt lågpresterande eller som ännu inte har bestämt sig för om de vill läsa ett yrkesprogram eller akademiskt program på gymnasiet. Författarna använder en regel som stipulerar hur stora klasserna får vara, samtidigt som de kontrollerar för elevers resultat i årskurs 9 och skolspecifika effekter, och finner små positiva effekter av mindre klasser. Resultaten indikerar att 5 färre elever i en klass ökar genomsnittsresultaten i danska, engelska och matematik med 0,04 sd under det tionde året, en liten effekt för en resursökning i den storleksklassen. Författarna finner heller inga skillnader mellan olika elevgrupper. Med andra ord ger studien inget starkt stöd för att resurser ger noterbart högre resultat. Ett problem är dock att de använder klasstorleksregeln som appliceras i årskurs 1-9 – även om denna ändå verkar användas de facto i årskurs 10 – eftersom detta ökar sannolikheten att variationen i klasstorleken som analyseras inte skiljer slumpmässigt mellan skolor.

Andra studier har fokuserat på de långsiktiga effekterna av mindre klasser och högre lärartäthet i högstadiet. Browning och Heinesen (2007) analyserar elever i åttonde klass med en liknande metod som studien ovan och finner väldigt blandade resultat. Oftast finns faktiskt ingen statistiskt signifikant effekt av mindre klasser eller högre lärartäthet. Även om vi endast fokuserar på de modeller där författarna finner statistiskt signifikanta effekter är dessa inte heller speciellt stora: 5 färre elever i klassrummet ökar antalet års utbildning med ungefär 0,04 sd och sannolikheten att eleverna går ut gymnasiet med 0,05 sd. Samtidigt ökar 25 procent färre elever per lärartimme (beräknad på veckobasis) antalet års utbildning med 0,13 sd och sannolikheten att elever går ut gymnasiet med 0,14 sd. Här verkar alltså lärartätheten vara viktigare än klasstorleken, vilket inte stämmer överens med STAR-experimentets resultat. Effekterna är heller inte märkvärdiga givet hur stora resursökningar som skulle krävas för att genomföra dem. Och återigen är det viktigt att betona att effekterna försvinner i de flesta modellspecifikationerna.⁴¹ Det står också klart att effekterna skiljer sig beroende om författarna inkluderar bakgrundsvariabler, vilket inte borde vara fallet om maxstorleksregeln faktiskt skapar slumpmässig

⁴⁰ I andra kontexter kan självklart lägre presterande elever allokeras till mindre klasser, vilket gör att effekterna av små klasser kan underdrivas (Cohen Zada, Gradstein, Reuven 2011).

⁴¹ Dessa är faktiskt de vanligaste specifikationerna i studier som använder regler gällande klasstorleken för att analysera dess effekter.

variation i klasstorleken. Detta kan delvis bero på att föräldrar kan förutse klasstorleken, vilket underminerar författarnas metodologi.

En annan studie som bättre tar hänsyn till detta problem analyserar ett större urval av elever och finner andra resultat. Bingley, Jensen och Vibeke (2005) menar att klasstorleksregeln kan förutses av vissa föräldrar eftersom klasstorleken korrelerar över tid inom skolor. De fokuserar därför på syskon som går i samma skola men med klasser med olika storlek, vilket gör att de till stor del kan kontrollera för effekterna av skillnader i föräldrars förmåga att planera efter klasstorleksregeln. De finner att klasstorleken och lärartätheten i årskurs 8 har små positiva effekter. Resultaten indikerar att 5 färre elever i klassrummet ökar antalet års utbildning med 0,03 sd och 25 procent färre elever per lärarledda timme ökar antalet års utbildning med 0,05 sd. Författarna analyserar också om effekterna överväger kostnaderna och kommer fram till att kostnaderna alltid är högre, även under de mest gynnsamma förhållanden. Det bör dock betonas att denna kostnadseffektivitetsanalys är relativt enkel och resultaten beror naturligtvis på vilka (svårverifierbara) antaganden som görs.

England

I England analyserar Gibbons, McNally och Viarengo (2012) effekterna av totala resurser på 11-åringars resultat, framförallt bland skolor i fattiga stadsområden. Författarna utnyttjar att olika skoldistrikts sätt att allokerar pengar gör att väldigt lika skolorna som ligger nära varandra kan få olika resurser. Genom att analysera skolor som ligger nära gränsen för skoldistriktet kan de analysera orsakssambandet mellan högre resurser och resultat. De finner att en 20 procents ökning av resurserna leder till en ökning med 0,14 sd i matematik, 0,17 sd i naturvetenskap och 0,11 sd i engelska.

Dessa effekter återfinns endast bland skolor med en andel elever från låg socioekonomisk bakgrund som är högre än genomsnittet, bland skolor med en andel vita elever som är högre än genomsnittet och bland skolor där genomsnittsresultaten när eleverna är 7 år är lägre än genomsnittet. Däremot verkar alla elever som går i dessa skolor, oavsett bakgrund, tjäna på högre resurser, även om effekterna oftast är större bland fattiga elever och icke-vita elever. Med andra ord indikerar detta att endast skolor i de mest utsatta områdena tjänar på mer resurser.⁴² Det bör också betonas att urvalet är begränsat till elever i storstadsområden, som generellt kommer från lägre socioekonomisk bakgrund, så resultaten kan inte generaliseras till andra områden. Faktum är att 60 procent av urvalet är från London, jämfört med 14 procent i det nationella genomsnittet.

En poäng som är värd att betona är att det är väldigt svårt för föräldrar att veta att vissa skolor får mer pengar än andra på grund av den komplicerade resursfördelningsmekanismen – och ännu svårare att veta vilka implikationer

⁴² Intressant nog är det inte lärarresurser som förklarar dessa resultat – mer pengar leder faktiskt till att en *mindre* andel spenderas på lärare (lönerna bestäms dock centralt och påverkas inte). Istället verkar pengarna spenderas på annat, exempelvis IT-resurser och läromedel.

högre resurser har givet den kompensatoriska resursfördelningen – vilket författarna också noterar. Detta är en bra för en akademisk analys, eftersom det hjälper dem att analysera effekten av resurser "allt annat lika", men är inte lika relevant för reformer som explicit ökar resurserna för vissa skolor. Som betonas i kapitel 2 kan föräldrar reagera på resursökningar om de är medvetna om dem.

En annan färsk brittisk studie fokuserar på lärarlöner. Greaves och Sibietta (2014) utnyttjar att lärarlöner, som sätts på nationell basis i England, varierar något mellan London och resten av landet. De analyserar 11-åriga elever som går i skolan väldigt nära den geografiska gränsen för London och resten av England. Detta gör att dessa elever och skolorna de går i är väldigt lika varandra. De finner inga långsiktiga effekter av ungefär 5 procent högre lärarlöner på resultaten och författarna menar att det troligtvis är den stora variationen i resurser som förklarar Gibbons, McNally och Viarengos (2012) resultat. Detta exemplifierar att linjära antaganden om resursers påverkan på resultaten inte nödvändigtvis är försvarbara.⁴³

Frankrike

Den franska forskningen om resurser är blandad, men är också problematisk från metodologisk synvinkel. Bland elever i årskurs 3 finner Piketty (2004) positiva effekter när han analyserar en de facto klasstorleksregel. Resultaten indikerar att 5 färre elever i klassrummet i årskurs 2 leder till 0,16 sd högre resultat i matematik och 0,11 sd bättre resultat i franska. Effekterna är starkast bland elever från lägre socioekonomisk bakgrund samt lågpresterande elever. Dessa är starkare effekter än vad som återfinns i de flesta studier. Det bör här noteras att effekterna är likartade, om än något mindre, när författaren inte använder klasstorleksregeln och istället endast kontrollerar för bakgrund och tidigare provresultat. Detta tyder på att problemet med sortering, likt det som diskuteras ovan när det gäller Chile, kan skapa problem för författarens ansats.

Ett annat problem är att författaren inte kontrollerar för effekterna av hur stora skolorna är. Detta är viktigt eftersom bättre elever kan i större utsträckning söka sig till populära skolor, vilket i sin tur kan göra att de oftare behöver ha fler klasser. I sådana fall kan effekten bero på att eleverna i mindre klasser också går i bättre skolor, inte på att mindre klasser leder till bättre resultat.

Till sist är det problematiskt att författaren inte analyserar en faktisk regel för hur stora klasserna fick vara, som faktiskt inte finns i Frankrike, utan istället endast ser observerade förändringar i klasstorleken runt ett specifikt elevantal som ett tecken på att en inofficiell regel existerar. Detta är ett problem eftersom skolor i sådana fall kan välja om de vill ha fler klasser, vilket i sin tur gör att beslutet att ha små klasser inte alls är slumpmässigt. På grund av dessa problem är det svårt att dra alltför starka slutsatser från studien.

⁴³ Som föregående fotnot betonade är det dock värt att påpeka att lärarlöner inte förklarar Gibbons McNally och Viarengos (2012) resultat.

Samma sak gäller Piketty och Valdenaires (2006) forskning som använder samma metodologi bland lågstadie- och högstadieelever. Författarna finner att 5 färre elever skulle generera ett 0,1 sd högre meritvärde i högstadiet och även att mindre klasser har liknande effekter i lågstadiet som Piketty (2004) finner. Återigen är det dock svårt att tolka dessa resultat på grund av de problem som noterades ovan.

Att problemen kan påverka resultaten bekräftas av ytterligare en studie i vilken Bressoux, Kramarz och Prost (2009) fokuserar på klasstorlekseffekter bland tredjeklassare som har lärare med låg erfarenhet. Författarna använder en liknande metodologi som de andra studierna: med större elevantal bör skolor ha mindre klasser och författarna antar att denna variation i elevantal är slumpmässig. De finner först positiva effekter på ungefär 0,15 sd med 5 färre elever i klassrummet. Intressant nog försvinner effekten när de endast analyserar skolor med mellan 29 och 45 elever i årskurs tre, runt den gräns som ofta leder till mindre klasser. Detta är anmärkningsvärt eftersom det är just denna variation som kanske kan anses vara slumpmässig. Dock är det återigen viktigt att betona att det inte finns någon klasstorleksregel i Frankrike, vilket gör att resultaten inte är helt pålitliga alldeles oavsett dessa problem.⁴⁴

I den senaste studien antar Gary-Bobo och Mahjoub (2013) att elevantalet per årskull är slumpmässig – när de kontrollerar för elevantalet i hela skolan – och använder denna variabel tillsammans med den informella klasstorleksregeln för att få slumpmässig variation i klasstorleken. Resultaten indikerar att 5 färre elever i klassrummet minskar sannolikheten att eleverna får gå om årskurs 6 med 7 procentenheter och sannolikheten att eleverna får gå om årskurs 7 med 10 procentenheter.⁴⁵ Det finns ingen effekt alls i årskurs 8 eller årskurs 9. Författaren kan dock inte ta hänsyn till årliga förändringar i skolkvaliteten som i sin tur kan påverka antalet elever i olika årskurser och därför också klasstorleken. Studien lider även av problemet att det inte finns någon faktisk regel gällande klasstorleken över landet, vilket gör att resultaten återigen bör tolkas med försiktighet. Detta kan kanske förklara varför det inte finns några effekter alls i årskurs 6 när författarna endast använder den informella klasstorleksregeln för att få slumpmässig variation i klasstorleken, samtidigt som de kontrollerar för antalet elever i skolan.

Israel

Metoden att använda klasstorleksregler för att analysera effekterna av små klasser är idag vedertagen. Men pionjärerna i denna bemärkelse var Angrist och Lavy (1999). I israeliska offentliga skolor används en regel som stipulerar att det får vara som mest 40 elever i en klass. Detta skapar kraftig variation i klasstorleken för elever som går i skolor runt denna tröskel. Författarna finner

⁴⁴ Författarna använder även en annan metodologi där de kan analysera om det finns en relation mellan initiala provresultat och klasstorlek, vilket också Piketty gör. Men detta är inte ett slutgiltigt test eftersom avsaknaden av en relation kan bero på annat. Att effekten försvinner när de endast fokuserar på skolor med mellan 29 och 45 elever är ett tecken på detta.

⁴⁵ Med de data som anges kan jag inte beräkna effektstorleken i standardavvikelser.

en positiv effekt av klasstorleken bland elever i femte klass, men ingen robust effekt i fjärde klass och ingen effekt alls bland elever i tredje klass. I femte klass ger en klasstorleksminskning med 5 elever ungefär 0,11 sd högre resultat i läsförståelse och 0,07 sd i matematik. Men resultaten har alltså ingen påverkan i årskurs 3 eller 4. Detta är inte speciellt imponerande. Trots detta refereras studien som ett solklart exempel på hur bra metodologi visar att små klasser är bättre. Men detta stämmer alltså inte alls, utan studien finner istället blandade resultat.

Italien

Som fallet Chile indikerar kan små klasser mycket väl ha olika effekter beroende på institutionerna i skolsystemet. Detta stöds ytterligare av forskning från Italien. I en mycket intressant studie analyserar Angrist, Battistin och Vuri (2014) effekten av klasstorlek i italienska låg- och mellanstadieskolor. De använder samma metod som Fredriksson, Öckert och Oosterbeek (2013) och finner väldigt små effekter på provresultat: en klasstorleksminskning med 5 elever ökar resultaten med 0,03 sd i matematik och 0,02 sd i språk. Intressant nog skiljer sig resultaten över landet. I norra Italien är effekten 0,02 sd i matematik, medan det inte finns någon effekt alls i italienska. I de fattigare södra delarna är den istället 0,07 sd i matematik och 0,05 sd i italienska. Med unik data visar dock författarna att hela effekten av klasstorleken förklaras av manipulering av resultaten. När de tar hänsyn till detta finns inga effekter någonstans i landet. Små klasser verkar alltså öka sannolikheten att lärarna fuskar eller inte lägger ner tillräckligt mycket tid på rättningen och detta leder i sin tur till högre resultat. Men kunskaperna ökar alltså inte.⁴⁶

Resultaten kan visserligen vara begränsade till den italienska kontexten, men de indikerar att det kan vara svårt att tolka resultaten från andra länder, trots starka forskningsstrategier. Om mindre klasser leder till mer manipulering av resultaten är det inte säkert att effekterna som forskarna finner beror på att eleverna lära sig bättre, vilket är speciellt relevant för länder som Sverige där rättningen av elevers prov sker lokalt.

Japan

Metoden som introducerades av Angrist och Lavy (1999) har även applicerats i den japanska kontexten. Liksom i Israel får det vara max 40 elever i ett klassrum, vilket Akabayashi och Nakamura (2013) utnyttjar för att analysera den kausala effekten av klasstorlek på resultaten bland elever i årskurs 6 och 9 i Yokohama, Japans näst största stad. Författarna finner endast en svag positiv effekt av små klasser i japanska bland elever i årskurs 6, som indikerar att 5 färre elever i klassrummet ökar resultaten med ungefär 0,03 sd på elevnivå. Det finns dock

⁴⁶ Författarna menar att kostnaden att fuska ökar linjärt med klasstorleken. En annan förklaring till varför mindre klasser leder till mer manipulering av resultaten är att lärare i dessa klasser vill visa att de förbättrar sig med högre resurser. Som diskuteras i nästa avsnitt finns det stöd från Nederländerna att mer resurser ökar betygsinflationen där (De Witte, Geys och Solondz 2014).

ingen påverkan på matematikresultaten. Bland elever i årskurs 9 finns det heller ingen effekt oavsett ämne. Den genomsnittliga betydelsen av klasstorlek är alltså liten. Dessutom verkar det endast finnas en positiv effekt bland skolor som ligger i rikare områden. Där indikerar resultaten att 5 färre elever ökar resultaten i japanska och matematik med 0,05 sd. Detta är dock knappast heller imponerande.

Nederländerna

Samma metod har även använts i Nederländerna. Dobbelsteen, Levin och Oosterbeek (2002) analyserar elever i årskurs 4, 6 och 8 och finner inga effekter av mindre klasser. De menar att effekten av mindre klasser även gör att elevsammansättningen förändras i klassrummet, vilket kan påverka eleverna. De finner att antalet elever som har liknande IQ har en positiv effekt på resultaten och att homogeniteten när det gäller den genomsnittliga IQ-nivån ökar med större klasser.⁴⁷ Samtidigt gör detta endast att effekten av mindre klasser blir marginellt statistiskt signifikant i språk i sjätte klass, med en effektstorlek som indikerar att 5 färre elever i klassrummet ökar elevens resultat med 2,4 percentilpoäng.⁴⁸ Det finns dessutom fortfarande ingen påverkan i aritmetik och inte heller i något ämne i fjärde eller åttonde klass. Skillnaderna i klasstorlek är inte lika stora som i Sverige eller Israel, vilket kanske också kan förklara skillnaderna.

Norge

I Norge har man analyserat både effekterna av klasstorlek och antalet lärartimmar. När det gäller klasstorlek och resurser använder Leuven, Oosterbeek och Rønning (2008) samma metod som den som användes av Fredriksson, Öckert och Oosterbeek (2013) i Sverige. Även i Norge har man en regel som stipulerar att max 30 elever får undervisas i en klass, men istället för mellanstadiet analyserar författarna elever i högstadiet. De finner inga effekter alls på resultaten på ett externt prov som skrivs i nionde klass. De analyserar även om klasstorleken i låg- och mellanstadiet – där maxgränsen är 28 elever – har någon påverkan på elevernas resultat i högstadiet men finner inga skillnader. Med samma metodologi kontrollerar författarna även om antalet lärarledda timmar – vilket kommuner spenderar mest pengar på – påverkar resultaten, men finner ingenting som tyder på detta.

En annan studie av Iverson och Bonesrønning (2013) analyserar effekten av klasstorlek i lågstadiet med samma metod och finner väldigt små effekter: 5 färre elever i klassrummet leder till 0,02 sd högre resultat. Intressant nog finns inga effekter bland elever med högutbildade föräldrar. Däremot leder 5 färre elever i klassrummet till en ökning i resultaten med 0,02 sd bland elever med

⁴⁷ Notera att detta går emot det klassiska argumentet om kamrateffekter som flitigt används i den svenska debatten, vilket inte är konstigt eftersom dessa inte är robusta (Angrist 2013).

⁴⁸ Författarna presenterar ingen deskriptiv statistik så jag kan inte omvandla resultaten till standardavvikelser.

lågutbildade föräldrar och med 0,04 sd bland skilsmässobarn. Dessa resultat är inte heller robusta utan beror helt och hållet på förändringen i klasstorlek från 28 till 14-17 elever, där effekten växer till 0,05 sd respektive 0,11 sd för fem färre elever i klassrummet. En klasstorleksminskning från 28 till 19-21 elever har däremot ingen effekt alls. För att det överhuvudtaget ska finnas någon (liten) effekt alls krävs alltså halveringar av klasstorleken vilket är extremt dyrt.⁴⁹

En annan norsk studie analyserar resurser i sig snarare än klasstorlek eller lärarresurser. Hægeland, Raaum och Salvanes (2012) analyserar resultaten på högstadiet men använder en annorlunda metod: de jämför kommuner som har vattenfall med liknande kommuner som inte har det. De som har vattenfall får högre fastighetsskattsintäkter via vattenkraftverken – i Norge tar inte alla kommuner ut fastighetsskatt. Författarna visar att kommunerna med vattenfall också spenderar mer pengar på skolan och menar att denna variation kan användas för att mäta effekterna av resurser inom skolan. De finner positiva effekter på resultaten som indikerar att en ökning av resurserna med ungefär 20 procent skulle öka resultaten med 0,16 sd. De finner även att en ökning med antalet lärarledda timmar med 20 procent leder till en förbättring i resultaten med 0,46 sd.

Det finns dock problem med studien. Den viktigaste invändningen är att den inte kan utesluta att det finns systematiska, icke-observerbara skillnader mellan kommunerna med vattenfall och de som dessa jämförs med. Om detta stämmer och skillnaderna påverkar både resultat och resurser är resultaten inte tillförlitliga. Ett exempel är att endast ungefär 50 procent av kommuner tar ut fastighetsskatter, vilket är precis en sådan variabel som kan både påverka resursfördelningen och resultaten. Tidigare finner mycket riktigt Fiva och Rønning (2008) att fastighetsbeskattning i sig påverkar resultaten i Norge positivt på grund av incitamentseffekter även när man kontrollerar för resurstilldelningen. Med tanke på att Leuven, Oosterbeek och Rønning (2008) inte finner positiva effekter när de analyserar samma resursvariabel (antalet lärarledda timmar) men med en mer vedertagen metodologi är studien därför inte tillförlitlig.

USA

I USA har Dee och West (2011) jämfört åttondeklassares resultat i olika ämnen, där undervisningen sker med olika klasstorlekar, vilket håller konstant skillnader mellan elever som påverkar resultaten. Eftersom eleverna undervisas av samma lärare i de olika ämnena hålls även lärares kvalitet konstant. De finner inga effekter i genomsnitt av mindre klasser. Dock finner de små positiva effekter i skolor i stadsområden som indikerar att 5 färre elever i klassrummet ökar resultaten med 0,06 sd. Samtidigt finner de en positiv påverkan på icke-kognitiva färdigheter som relaterar till engagemang i skolan. Effektstorleken är ungefär

⁴⁹ Samtidigt är det intressant att Leuven, Oosterbeek och Rønning (2008) inte finner positiva effekter av små klasser i låg- och mellanstadiet på resultaten i högstadiet. Även om det finns positiva effekter av små klasser i lågstadiet verkar dessa alltså försvinna med tiden.

densamma som för provresultat. Med andra ord är effekterna av klasstorlek liten även i denna studie.

Flera länder

Som beskrevs i avsnittet om Sverige – där metodologin beskrivs mer utförligt – har både Woessmann (2005) och Altinok och Kingdon (2012) studerat Timss-resultat bland 13- och 14-åriga elever i 15 respektive 47 länder. Woessmann (2005) använder klasstorleksregler och finner endast statistiskt signifikanta negativa effekter i Island och Spanien. Eftersom resultaten är relativt precisa kan han utesluta stora effekter i alla länder. Med andra ord verkar klasstorlek inte vara speciellt viktigt i de länder som studeras.

Altinok och Kingdon (2012) fokuserar istället på hur skillnader i klasstorlek i matematik och naturvetenskap påverkar skillnaderna i resultat i dessa ämnen. På detta sätt kan de hålla skillnader mellan elever som påverkar resultaten i både matematik och naturvetenskap konstant. De finner negativa effekter i 14 av länderna, men dessa är oftast små. De största effekterna återfinns i Chile där resultaten indikerar att resultaten ökar med ungefär 0,09 sd med en klasstorleksminskning på 5 elever. Generellt är effekten starkare i utvecklingsländer. I vissa länder återfinns också en icke-linjär påverkan, där klasstorleken först är positiv och sedan negativ – eller först negativ och sedan avtagande. Brytpunkten varierar också kraftigt mellan länder. Återigen verkar alltså mindre klasser inte vara ett speciellt kostnadseffektivt sett att öka resultaten i internationella prov.

Sammanfattning

Sammanfattningsvis står det klart att forskningen som analyserar elever vid ett tillfälle är väldigt blandad och att effekterna beror på vilket land som analyseras samt vilken metodologi som används. Vissa studier finner inga effekter, vissa finner små effekter och ett fåtal finner större effekter. Det enda fältexperimentet som genomförts i modern tid finner vissa positiva effekter av resurser – som är koncentrerade till det första året eleverna går i en liten klass – men har hittills inte funnit positiva effekter på elevernas löner. Experimentet var dessutom dåligt genomfört och fokuserade på fattiga elever i Tennessee, vilket minskar dess utbildningspolitiska relevans ytterligare. De experiment som genomfördes innan andra världskriget visar generellt inte positiva effekter av mindre klasser, men det är naturligtvis också svårt att dra starka utbildningspolitiska slutsatser från dessa.

Givet problemen med fältexperimenten vore det naturligtvis bra om andra metodologiskt starka studier som jämför elever vid en tidpunkt entydigt visade hur resurser påverkar resultaten. Men så är det inte. Studier som använder precis samma metodologi har ofta olika effekter beroende på vilket land/system och vilka elevgrupper som analyseras, vilka årskurser eleverna går i, samt hur stora resursskillnaderna är. Ibland har resurser en större effekt, ibland en mindre effekt, och ibland ingen effekt alls. Det är värt att notera att resultaten

spretar och det finns lite som tyder på att vissa årskurser tjänar mer på resurser än andra generellt.

Det är också värt att betona forskningen från Italien som finner att hela effekten av klasstorlek kan förklaras av manipulation av resultaten, som ökar med mindre klasser. Detta är ytterst relevant i ett land som Sverige, där lärarna rättar sina egna elevers prov och sätter vilka betyg de vill. Även om resurser skulle producera mycket högre resultat är det långt ifrån självklart att detta också ökar kunskaperna i lika stor utsträckning.

Generellt verkar det dock som att studier som analyserar stora skillnader som uppstår på grund av svårmanipulerade regler och komplicerade formler för resursfördelningen *ibland* har större effekter på *vissa* elever, medan de som använder andra starka ansatser finner små eller inga effekter. Även om vi skulle bortse från att studierna inte kan svara på om reformer som ökar resurserna skulle leda till bättre resultat står det därför klart de utbildningspolitiska implikationerna av forskningen är långt ifrån klara.

Studier som analyserar reformer och förändringar över tid

För att kunna analysera huruvida reformer som ökar resurserna kommer att ha positiva effekter är det som sagt dock viktigt att även diskutera forskningen som fokuserar just på reformer som förändrar resurstilldelningen. En annan ansats som använts är att analysera resursförändringar över tid, antingen direkt eller via reformer som indirekt påverkar dem. Dessa studier är också blandade, men i detta fall betyder det en generellt obefintlig eller svag påverkan på resultaten i genomsnitt, vilket ytterligare indikerar att den utbildningspolitiska relevansen av forskningen i det förra avsnittet bör ifrågasättas ytterligare. Denna kan som sagt inte ta hänsyn till de oavsiktliga konsekvenser som kan uppstå i och med reformer.

Klasstorleksreformer

Låt oss börja med genomgående klasstorleksreformer. Dessa har endast analyserats i Kalifornien och Florida, som båda genomförde stora klasstorleksminskningar i hopp om att detta skulle leda till bättre resultat. Kalifornien genomförde 1996 en minskning av klasstorleken i lågstadiet från i genomsnitt 30 till 20 elever, vilket kostade 1 miljard dollar om året. Florida gick ännu längre och reglerade år 2002 klasstorleken för alla elever med ett författningstillägg till en kostnad av 22 miljarder dollar över nio år. Vad var resultaten?

Kalifornien

I Kalifornien finner Jepsen och Rivkin (2009) att en minskning i klasstorleken med tio elever ledde till 0,06-0,09 sd högre resultat i matematik och 0,04-0,06 sd högre resultat i läsförståelse. Detta gäller dock på skolnivå där variationen är

mindre. När vi omvandlar resultaten till elevnivå uppgår effekten endast till ungefär 0,02-0,03 sd i matematik och 0,01-0,02 sd i läsförståelse, en väldigt liten effekt för en så stor klasstorleksminskning. Om vi antar att effekten är linjär – vilket alltså inte alls är givet – skulle en klasstorleksminskning med 5 elever, som både den svenska regeringen och oppositionen föreslår, att leda till 0,005-0,015 sd bättre resultat.

Men dessa estimat håller lärarkvalitet konstant. Som beskrevs i det andra kapitlet har reformer ofta oavsiktliga konsekvenser som bör tas med i beräkningen. Reformen betydde också att man var tvungen att anställa ett stort antal nya lärare och författarna finner att en lärare med endast ett års erfarenhet reducerar resultaten lika mycket som klasstorleksminskningen ökade resultaten. För dessa elever fanns alltså ingen effekt alls av klasstorleksminskningen.

Detta tyder dock på att effekterna blir starkare över tid när lärarna får mer erfarenhet. Författarna simuleringar gällande detta indikerar att den långsiktiga påverkan var 0,06 sd i matematik och 0,03 sd i läsförståelse på elevnivå. Om jag återigen antar att effekten är linjär skulle en klasstorleksminskning med 5 elever alltså producera ungefär 0,02 sd högre resultat i ett längre perspektiv. Återigen är detta extremt liten effekt givet vad reformen kostade.⁵⁰

Det fanns dock ytterligare oavsiktliga konsekvenser som författarnas simuleringar inte tar hänsyn till. Sims (2008) finner först och främst att reformen skapade incitament att skapa klasser med elever från olika årskurser. Detta hade också en större negativ effekt på resultaten bland elever som placerades i de blandade klasserna än de positiva effekterna av klasstorleksminskningen.⁵¹ I en annan studie finner samma författare att kraven att minska klasstorleken i lågstadiet även ökade klasstorleken med två elever i fjärde och femte klass (Sims 2009).⁵²

Med andra ord kan reformen i Kalifornien inte beskrivas som annat än ett stort och dyrt fiasko. Vissa av de oavsiktliga konsekvenserna hade kunnat undvikas med bättre utformning av reformen, men dessa är naturligtvis svårt att undkomma helt hållet. I slutändan vet vi faktiskt inte heller varför effekterna uteblev. Resultaten är exempelvis extremt magra även när förändringar i lärarkåren är kontrollerade för. Det är alltså troligt att andra oavsiktliga konsekvenser motverkade den positiva effekten av mindre klasser.

⁵⁰ En annan studie som endast utvärderar effekten av klasstorleksminskningen i dagis på resultaten i andra klass finner också väldigt små effekter med en något annorlunda metod (Funkhouser 2009).

⁵¹ Det är dock värt att notera att den generella forskningen gällande klasser med elever från olika årskurser inte nödvändigtvis är dåligt i alla situationer (Leuven and Rønning 2011; Thomas 2012).

⁵² Författaren finner även en negativ effekt på resultaten i matematik, men inte i språk, av de större klasserna i femte klass. Resultaten indikerar att 5 färre elever leder till 0,11 sd högre matematikresultat. Det finns dock inga effekter alls i fjärde klass. Dessa resultat är dock beroende av att skolors beslut att öka klasstorleken inte påverkas systematiskt av andra variabler som påverkar resultaten, vilket inte är troligt.

Florida

Men kanske lärde sig Florida något av Kaliforniens misslyckande? Detta är mycket möjligt eftersom den genomfördes sju år efteråt (och faktiskt efter att forskningen funnit magra effekter i Kalifornien).⁵³ Till skillnad från i Kalifornien – där man genomförde klasstorleksminskningen i ett hugg – minskades klasstorleken i Florida gradvis från och med 2003. En annan skillnad i Florida var att alla årskurser omfattades av reformen, inte bara lågstadiet. Taket för klasstorleken var 18 elever i dagis till och med årskurs tre, 22 elever i årskurs 4-8 och 25 elever i årskurs 9-12. Reformen gjorde att den genomsnittliga klasstorleken minskades från 23 till 16 elever i dagis till och med tredje klass, från 24 till 19 elever i årskurs 4-8, samt från 24 till 22 elever i årskurs 9-12.

Två studier har utvärderat Floridas reform. Den första fokuserar på effekterna bland elever i tredje till åttonde klass. Chingos (2012) jämför elever som påverkades mer av reformen eftersom de gick i skolor där klasstorleken var större innan den genomfördes, med elever som påverkades mindre eftersom de gick i skolor som redan hade så små klasser som reformen krävde. Mer specifikt analyserar författaren hur trenden i provresultat före och efter reformen skiljer sig mellan dessa grupper, både på skoldistriktsnivå och på skolnivå. Han finner inga effekter alls på resultaten bland elever i årskurs 6-8 i analysen på skoldistriktsnivå tre år efter reformen – det finns faktiskt en negativ effekt av mindre klasser i läsförståelse i årskurs 7. I studien är skillnaden i klasstorlek dock endast 1,5 elever i genomsnitt, men upp till 2,7 i årskurs 7. Detta kan troligtvis delvis förklara skillnaderna med de små effekter som forskare fann i Kalifornien.

Författaren finner heller inga effekter alls i analysen på skolnivå sex år efter reformen, vare sig när han fokuserar på elever i årskurs 6-8 eller elever i årskurs 3-5. Det finns faktiskt visst stöd för negativa effekter i läsförståelse i årskurs 3-5. Här är skillnaden i klasstorlek 3,6 elever i årskurs 3-5 och 2,1 i årskurs 6-8, vilket återigen är mycket mindre än i Kalifornien.

Dessa är magra resultat, vilket skulle kunna bero på att lärarkvaliteten sjönk. Det finns inga effekter på lärarnas erfarenheter i genomsnitt, men lärarlönerna minskade med två procent i skolor som påverkades av reformen och ökade med två procent i skolor som inte påverkades av reformen, vilket möjligtvis skulle kunna förklara avsaknaden på effekter.

Men detta antagande får inget stöd i den andra studien, som fokuserar på effekterna av reformen på lärarnas effektivitet. Dieterle (2013) använder en något annorlunda metodologi och noterar först den totala effekten på matematikresultaten. Han jämför trender i resultat bland skoldistrikt som hade större genomsnittsklasser än vad reformen krävde året innan den genomfördes, med trender i resultat bland skoldistrikt som inte påverkades av reformen eftersom de redan hade tillräckligt små klasser. Samma analys genomförs på skolnivå. Förutom en enskild effekt i distriktsanalysen på 0,03 sd två år efter

⁵³ Jepsen och Rivkins studie kom först ut som rapport 2002.

reformen finns inga effekter alls upp till fem år efteråt. På skolnivå finns det istället visst stöd för negativa effekter.

Studiens huvudfokus är dock huruvida förändringar i lärarkvaliteten – här mätt som hur mycket lärarna bidrar till hur elevernas resultatförbättringar över tid – kan förklara de magra resultaten. Författaren finner att kvaliteten på lärarna sjönk väldigt lite och att detta inte kan förklara varför effekterna av reformen på resultaten inte är större. Faktum är att lärarkvaliteten sjönk mindre i skolor som hade större genomsnittsklasser än vad reformen krävde året innan den genomfördes, jämfört med skolor där genomsnittsklasserna var lägre. Det verkar med andra ord som att de minskade lärarlönerna som Chingos (2012) finner inte påverkade lärares effektivitet negativt.

Att en reform som kostade 22 miljarder dollar att genomföra inte gav några positiva effekter alls kan inte beskrivas som annat än en stor flopp. Trots lärdomar från Kalifornien och en utformning som till viss del skulle kunna minska de oavsiktliga konsekvenserna lyckades politiker i Florida alltså inte producera bättre resultat med hjälp av mindre klasser.

Forskning som analyserar andra reformer

Förutom klasstorleksreformerna finns det även forskning som analyserar reformer som ökade/sänkte resurserna på andra sätt än via klasstorleksminskningar. Vissa studier analyserar reformer som fokuserar på underpresterande elever/skolor medan andra analyserar reformer som ökade resurserna för alla skolor. Lyckades dessa höja resultaten på ett effektivt sätt? Det här avsnittet går igenom reformer i olika länder och finner att de flesta faktiskt inte verkar ha ökat resultaten nämnvärt.

USA

Historiskt sett har amerikanska offentliga skolor finansierats lokalt, och variationerna i resurser har varit stora även inom olika delstater. Sedan 1970-talet har flera delstater dock genomfört reformer för att öka resurserna i skoldistrikt som historiskt sett har spenderat lite pengar på skolan. Ett par uppsatser analyserar om dessa reformer ledde till högre resultat. Andra uppsatser analyserar om mer resurser från den federala regeringen leder till förbättrade resultat, medan ytterligare studier fokuserar på alternativa reformer.

Kalifornien

De flesta studier har fokuserat på effekterna av högre resurser totalt sett. I Kalifornien analyserar dock Cellini, Ferreira och Rothstein (2010) investeringar i skolbyggnader som skedde på grund av folkomröstningar i olika skoldistrikt. Detta ledde till ungefär 10 procent högre resurser totalt sett. De finner generellt ingen effekt på resultaten, förutom efter sex år då effekten är marginellt

statistiskt signifikant i matematik bland tredjeklassare. Därefter försvinner den. Totalt sett finns det alltså väldigt lite stöd för att dessa investeringar ledde till högre resultat.

Kansas

Forskare har även utvärderat en reform som ämnade minska skillnaderna i resurser mellan olika skoldistrikt i Kansas. Deke (2003) finner att en resursökning med 20 procent leder till att sannolikheten att elever läser vidare på icke-traditionell högre utbildning (till exempel två-åriga program) ökar med 5 procent på skoldistriktets nivå.⁵⁴ Däremot har resurserna ingen påverkan alls på sannolikheten att eleverna läser vidare på traditionella universitetsutbildningar eller på avhoppet i gymnasiet. Neymotin (2010) analyserar förutom avhopp även provresultat i fjärde, sjunde och tionde klass efter ytterligare tillägg till reformen som förändrade resursfördelningen och finner inga effekter. Det bör dock betonas att studierna inte helt och hållet kan kontrollera för andra förändringar som skedde under perioden och som också kan tänkas påverka resultaten. De bör därför tolkas med försiktighet.

Kentucky

Samma sak gäller en rapport om resursernas betydelse i Kentucky. Clark (2003) analyserar en liknande reform som genomfördes 1990 och finner inga effekter bland åttonde klassare eller på ACT-provet, vilket fungerar som ett antagningsprov till universitetet. Författaren medger själv dock att det är svårt att separera effekterna av ökade resurser från andra förändringar som reformen skapade.

Maryland

I Maryland genomfördes en reform som ökade resurserna till skoldistrikt med en högre andel elever från låg socioekonomisk bakgrund. Chung (2013) använder reformen för att analysera om den ökade resurserna och minskade avhoppet på gymnasiet. Trots att reformen ledde till högre resurser finns det inga tecken på att denna ökning ledde till färre avhopp ett par år efter reformen.

Massachusetts

En annan reform som ämnade minska skillnaderna i resurser mellan olika skoldistrikt genomfördes i Massachusetts 1993, vilket omfördelade resurserna mellan skoldistrikt beroende på deras tidigare resurser, elevsammansättning, huspriser och invånarnas inkomstnivåer. Guryan (2001) jämför distrikt som ligger på olika sidor, men väldigt nära, en specifik gräns i formeln som bestämde

⁵⁴ Författaren rapporterar inte standardavvikelsen, vilket gör att jag inte kan konvertera effektstorleken till individnivå.

resursfördelningen. Han finner blandade effekter. I årskurs 4 finner han en positiv påverkan. Om vi omvandlar effekterna till elevnivå indikerar resultaten att motsvarar en 20 procent ökning av resurserna leder till en förbättring med ungefär 0,10 sd i genomsnitt i matematik, läsförståelse, naturvetenskap och samhällskunskap. Denna effekt försvinner dock med en något annorlunda modellspecifikation, vilket inte är förtroendeingivande. I åttonde klass finns inga effekter alls. Studien ger i helhet knappast därför starkt stöd för att resurser spelar speciellt stor roll för resultaten.

Michigan

Även i Michigan genomfördes en reform under 1990-talet, vilket innebar att vissa skoldistrikt ökade sina skolresurser väsentligt mycket snabbare än andra. Papke (2008) finner positiva effekter på andelen fjärdeklassare som når godkänt i matematik, medan Roy (2011) även analyserar resultaten i läsförståelse och finner positiva effekter på andelen elever som når godkänt. Omvandlar man effektstorleken till elevnivå är den ungefär 0,05-0,10 sd i läsförståelse och 0,10-0,14 sd i matematik av en ökning med 20 procent i resurserna. Det fanns dock inga effekter på ACT-provet, vilket är ett slags högskoleprov. Eftersom alla elever inte skriver detta prov lider dock dessa resultat troligtvis av selektionsproblem. Chaudhary (2009) analyserar både genomsnittsresultat och andelen som når godkänt och finner positiva effekter på provresultat i fjärdeklass: en ökning av resurserna med 20 procent skulle öka de genomsnittliga resultaten med ungefär 0,05 sd på elevnivå och sannolikheten att eleven når godkänt med 0,09 sd. Författaren finner dock inte effekt i årskurs 7.

Problemet med dessa studier är dock att de inte tar tillräckligt hänsyn till andra eventuella förändringar som skedde under reformperioden och som kan vara systematiskt relaterade till att vissa skoldistrikt fick mer pengar. En nyare studie av Hyman (2013) visar att effekterna i årskurs 4 försvinner när han kontrollerar för sådana förändringar.⁵⁵ Författaren analyserar endast resultaten i matematik, men effekten där försvinner helt och hållet. Han finner heller inga effekter på eleverna i sjunde klass och endast en marginellt statistiskt signifikant effekt på resultaten i årskurs 11, som motsvarar en ökning i resultaten med 14 procent av 20 procent högre resurser. Samtidigt indikerar resultaten att elevernas prestationer i andra året på gymnasiet samt sannolikheten att de läser vidare efter gymnasiet ökar med 14 procent av 20 procent högre resurser.⁵⁶ Det finns dock inga effekter på sannolikheten att eleverna når en eftergymnasial examen totalt sett, även om det finns en påverkan bland högpresterande elever.

Totalt sett finns det därför inget robust stöd för att reformen i Michigan genererade mer än marginellt högre resultat. Man bör även vara försiktig med att dra tydliga slutsatser eftersom det står klart att högre resurser även verkar

⁵⁵ Det bör också betonas att Hymans (2013) studie analyserar data på elevnivå, till skillnad från de tidigare studierna som analyserar data på skoldistriktsnivå. Som beskrivs i kapitel 3 tenderar mer aggregerade analyser att inflatera resultaten.

⁵⁶ Författarna rapporterar inte standardavvikelse så jag kan inte omvandla effektstorleken.

påverka andra variabler i en riktning som indikerar att resultaten i studierna är inflaterade på grund av problematisk metodologi.⁵⁷

New York

I New York analyserar Lee och Polachek (2014) folkomröstningar i de olika skoldistrikten, som bestämmer hur mycket pengar skolorna får. Omröstningar som faller politikernas förslag tenderar att leda till mindre resurser jämfört med året innan. Författarna jämför skoldistrikt där andelen som röstar är runt 50 procent och finner inget stöd för att ökade resurser ökar resultaten. Detta gäller för alla årskurser från 3 till 8, samt för sista året på gymnasiet. Det finns dock visst stöd för att andelen gymnasieavhopp minskade något, men effekten är inte stor: en ökning av resurserna med 20 procent skulle minska avhoppet med ungefär 0,03 sd på elevnivå.⁵⁸ Dessutom är inte heller denna effekt helt robust utan försvinner när författaren använder en något annorlunda statistisk metod.

Ett par andra studier har utvärderat effekterna av den amerikanska federala regeringens "Title 1"-program. Programmet ger pengar till skoldistrikt med större andelar elever från låg socioekonomisk bakgrund. Dessa studier finner inga eller till och med negativa effekter av ökade resurser i New York samt ett anonymt skoldistrikt (Matsudaira, Hoesk och Walsh 2012; Van der Klaauw 2008; Weinstein m.fl. 2009). Samtidigt är den totala resursökningen liten (ofta under 5 procent) eftersom delstaterna agerar strategiskt och minskar sin resurstilldelning till skolor som får federalt stöd. Detta gör att studierna endast kan analysera effekterna av små resursökningar.⁵⁹

Wyoming

I Wyoming genomfördes också en annorlunda reform som ökade skolresurserna per elev med 16 procent från ett år till ett annat.⁶⁰ Naturgaspriserna hade gått upp rejält och delstaten valde att spendera 200 miljoner dollar extra på skolan under skolåret 2006-07. Skoldistriktet fick välja hur de ville spendera de extra

⁵⁷ Exempelvis minskar andelen fattiga och afroamerikanska elever i distrikt som fick mer pengar på grund av reformen. Detta indikerar att mer bemedlade föräldrar flyttade till distriktet som spenderade mer pengar, vilket naturligtvis kan ha en positiv effekt på resultaten i sig. Hymans (2013) metodologi är starkast – och är den enda som analyserar resultaten på individnivå – men det är långt ifrån säkert att den är tillräcklig.

⁵⁸ Dessa resultat är beräknade efter medelvärdet av förändringarna i resurser samt standardavvikelsen av förändringar i andelen avhopp mellan 2004 och 2009 eftersom författaren inte rapporterar deskriptiv statistik i nivåer.

⁵⁹ Det finns ytterligare en studie som analyserar "Title 1"-programmet i Alabama, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee och Virginia. Cascio, Gordon och Reber (2013) finner att ökningarna i resurser mellan 1965 och 1969 verkade minska avhoppet bland vita 18- och 19-åringar men att de inte hade någon effekt bland svarta. Resultaten indikerar att en ökning med 20 procent minskar avhoppet med 2,64 procentenheter bland vita utan att ha några effekter på svarta elever. På grund av datatillgänglighet erkänner dock författarna att deras resultat är något spekulativa.

⁶⁰ Detta var dubbelt så mycket som ökningen i Wyoming de sex föregående åren tillsammans och 10 procentenheter högre än vad som skedde i USA i snitt under detta år.

pengarna, men framförallt användes pengarna till att höja lärarlönerna samt till att öka andelen lärare med mastersutbildning. Thompson (2012) finner att denna ökning av resurserna i princip inte hade någon effekt alls på fjärdeklassares provresultat i matematik och en liten negativ effekt på resultaten i läsförståelse. I årskurs 8 finns det en väldigt liten positiv effekt i matematik men återigen en svagt negativ effekt i läsförståelse.⁶¹ Fem år efter en reform som ökade resurserna så kraftigt återfanns alltså inga generella positiva effekter.⁶²

Hela USA

Det finns även ett par studier som har analyserat skoldistrikt över hela USA.⁶³ Hoxby (2001) fokuserar på reformer som ämnade producera en mer jämlik resursfördelning mellan skoldistrikt i olika delstater under perioden 1970-1990. Hon analyserar ungefär 40 000 elever och finner att ökade resurser på grund av reformerna inte hade några effekter på sannolikheten att eleverna hoppar av gymnasiet. Dock finns det visst stöd för att resurser minskar avhoppet i skoldistrikt som utan reformerna hade spenderat väldigt lite på skolan. Men resultaten är alltså blandade.

En färsk studie av Jackson, Johnson och Persico (2014) fokuserar istället på de långsiktiga effekterna av reformerna. De analyserar ett mindre urval av elever över hela USA under perioden 1967-2010 och hur de påverkas av specifika domstolsbeslut som leder till förändringar i resursfördelning mellan skoldistrikt. De finner i genomsnitt positiva effekter av resurser på sannolikheten att elever går ut gymnasiet och antal år av utbildning. Resultaten indikerar att en ökning av resurserna med 20 procent under hela skolgången skulle öka sannolikheten att ta gymnasieexamen med 17 procent och antal års utbildning med 6 procent.⁶⁴

Samtidigt finns det inga genomsnittseffekter av resurser på löner, familjeinkomster eller sannolikheten att vara fattig senare i livet. Ökningarna av resurser som skedde på grund av domstolsbeslut, och som ledde till högre sannolikhet att gå ut gymnasiet, var därför inte kostnadseffektiva i genomsnitt.

Men det visar sig att effekten varierar kraftigt mellan elever från fattiga familjer och de från icke-fattiga familjer. Faktum är att det inte finns några effekter *alls*

⁶¹ Författaren rapporterar inte standardavvikelsen för resultaten, men medelvärdet före reformen var 278,5-283 poäng (poängskalan är 0-500) i läsförståelse respektive matematik i åttonde klass och både den positiva och negativa effekten är ungefär en poäng. Resultaten indikerar alltså en extremt liten effekt.

⁶² Effekterna av höjda lärarlöner kan naturligtvis ta längre tid innan de visar sig, om mekanismen är att fler begåvade individer utbildar sig till lärare på grund av högre löner. En annan mekanism är dock att effektiva lärare söker sig till Wyoming istället för de andra delstaterna som används som kontrollgrupp, vilket i sig kan ha effekter på resultaten. Men detta verkar alltså inte ha skett.

⁶³ Tre andra studier analyserar reformer som förändrar resursfördelningen, men dessa rapporterar antingen inte effekten av resurser specifikt eller så fokuserar de på hur resurserna påverkade skillnader i resultat mellan elever från olika grupper istället för resultaten i sig (Card och Payne 2002; Downes 2002; Downes och Figlio 1998). Jag diskuterar därför inte dessa, men noterar att effekterna är blandade.

⁶⁴ Författarna rapporterar inte standardavvikelser och jag kan därför inte beräkna effektstorleken baserad på denna.

bland icke-fattiga familjer, vare sig när det gäller gymnasieexamen, antalet års utbildning eller ekonomiska utfall i genomsnitt. Det finns faktiskt tecken på att icke-fattiga elever som går igenom en längre tid i skolan med ökande resurser har *lägre* löner och familjeinkomster, som dock inte är statistiskt signifikant i genomsnitt. Eftersom den negativa effekten växer kraftigt över tid är det dock möjligt att den genomsnittliga effekten maskerar en statistiskt signifikant effekt på längre sikt. Bland icke-fattiga elever verkar ökade resurser som bäst vara meningslösa i ett långsiktigt perspektiv och de kan till och med vara negativa. Detta är ett remarkabelt resultat givet att författarna studerar effekterna av att öka resurserna under hela skolgången.

Däremot finns det en positiv påverkan på ekonomiska utfall bland fattiga elever. Resultaten indikerar att 20 procent mer resurser leder till en ökning av sannolikheten att fattiga elever tar gymnasieexamen med 29 procent, medan antalet års utbildning skulle öka med 7 procent. Samtidigt skulle lönerna öka med 24,6 procent, familjeinkomsterna med 52,2 procent och sannolikheten att falla i fattigdom minska med 19,7 procentenheter. Detta är betydelsefulla effekter. Som författarna noterar var resurserna under tidsperioden i genomsnitt 50 procent av vad de är idag, och på grund av avtagande effekter är det därför sannolikt så att det skulle krävas mycket högre resursökningar för att uppnå liknande effekter idag.

Men det är inte säkert varför löneeffekterna uppstår. Vi vet från nederländsk forskning att betygsinflationen kan öka med högre resurser bland fattiga elever (De Witte, Seys och Solondz 2014). Samtidigt indikerar italiensk forskning att manipulering av resultaten ökar med mindre klasser, vilket också är en slags betygsinflation (Angrist, Battistin och Vuri 2014). Eftersom gymnasieexamen i USA historiskt har berott på vilka betyg lärarna sätter, utan extern rättning, är det mycket möjligt att resursökningarna kan ha resulterat i betygsinflation – vilket i sin tur kan ha gett fördelar till de fattiga elever som fick mer resurser. De positiva ekonomiska effekterna som uppstod på grund av att eleverna i högre grad gick ut gymnasiet kan därför vara resultatet av en signaleffekt istället för en humankapitaleffekt. En signaleffekt uppstår exempelvis om arbetsgivare antar att elever som gått ut gymnasiet är mer högpresterande än de som inte gör det, och därför betalar högre löner till de förstnämnda, även om deras färdigheter inte skiljer sig. Detta kan vara fallet om betygsinflation förklarar resultaten. En humankapitaleffekt uppstår istället om eleverna som gått ut gymnasiet faktiskt är mer högpresterande än de som inte har det, vilket i sådana fall bör reflekteras i löneskillnader.

Att det inte finns några genomsnittliga effekter på löner eller inkomster stödjer tesen att resultaten kan förklaras av signaleffekter. Att det finns vissa tecken på negativa effekter på lönerna och inkomsterna bland icke-fattiga elever ger också ytterligare stöd för den.⁶⁵ Totalt sett verkar reformerna alltså ha gett upphov till ett nollsummespel när det gäller elevernas framtidsutsikter. Självklart kan

⁶⁵ Att det verkar finnas en viss (icke-robust) effekt av resurser på gymnasieavhoppen men inte på provresultaten i någon årskurs i New York stämmer också överens med detta (Lee och Polachek 2014).

signaleffekten vara viktig för eleverna som påverkas av den, men den innebär per definition också att andra elever förlorar på dem.

Men även om tesen inte stämmer är de utbildningspolitiska implikationerna för Sverige långt ifrån säkra. Andelen svenska barn i fattiga familjer är ungefär 10 procent enligt svenska mått mätt (SCB 2014).⁶⁶ Redan detta innebär att resursökningen inte skulle påverka speciellt många elever om vi extrapolerade effekterna. Men den svenska fattigdomsgränsen är högre i Sverige än i USA. I USA anses exempelvis ett barn med ett syskon och två föräldrar leva i fattigdom om familjen har en inkomst före skatt på eller under ungefär 12 000 kr per månad (IRP 2014). I Sverige är däremot gränsen för samma barn 19 800 kr i *disponibel inkomst* per månad (SCB 2014). Andelen fattiga i Sverige med det amerikanska måttet är alltså lägre än 10 procent. Det går därför inte att extrapolera effekterna på fattiga elever från USA till Sverige – det är sannolikt att andelen barn som skulle påverkas i svenska förhållanden skulle vara mycket mindre.

Men även om en extrapolering till den svenska situationen skulle vara försvarbar – givet att resultaten faktiskt förklaras av humankapitaleffekter vilket är alltså långt ifrån säkert – vore det ändå slöseri att öka resurserna för andra än den lilla andel av eleverna som är fattiga. På grund av att studien inte finner några långsiktiga effekter på ekonomiska utfall totalt sett står det också klart att de ökade resurser inte var kostnadseffektiva i genomsnitt. Istället bör man tolka resultaten som att resurser kanske kan spela roll för fattiga elever under vissa omständigheter, men att andra elever också verkar få betala för detta.

Belgien

Forskningen har även analyserat data från Belgien. År 2002 introducerades en reform i Flandern som ökade personalresurserna för skolor med elever från låg socioekonomisk bakgrund. Ooghe (2013) jämför skolor som precis ligger runt gränsen för att få mer resurser för att analysera den kausala effekten bland elever i lågstadiet. Författaren finner inga effekter av tre år med extra resurser i matematik, läsförståelse och stavning i genomsnitt när eleverna är 8 år. Inte heller finns robusta effekter när elever från olika bakgrund analyseras separat. Det är svårt att tyda hur stora resurserna faktiskt var, men författaren rapporterar att en skola som låg precis över gränsen för att få mer resurser skulle kunna minska klasstorleken med 4,5 elever i första klass. En resursökning motsvarande den klasstorleksminskning som både den svenska regeringen och oppositionen nu stödjer hade alltså inga robusta positiva effekter av reformen i Belgien.

England

I England har flera regeringar också försökt öka resultaten med extra resurser. En reform som specifikt fokuserade på mellanstadie-och högstadieskolor i

⁶⁶ Även om det är diskutabelt om fattigdom är ett absolut eller ett relativt begrepp är det relevanta begreppet här den absoluta fattigdomen, vilket är vad författarna analyserar.

utsatta områden var "Excellence in Cities". Denna introducerades i slutet på 1990-talet och ökade resurserna med 4,4 procent. Dessutom infördes "Pupil Learning Credits Scheme" i början på 2000-talet, en reform som gav ytterligare resurser till vissa skolor. Totalt sett ökade reformerna resurserna med ungefär 9 procent per år i genomsnitt. Machin, McNally och Meghir (2010) finner att reformerna hade en liten positiv påverkan på resultaten i matematik när eleverna var fjorton år, men inga effekter i engelska. Dessutom är effekterna i matematik små. Om vi antar att effekten av resurser är linjär innebär detta att 20 procent extra resurser skulle leda till en ökning med 0,04 sd, en väldigt liten effekt. Dessutom är den framförallt koncentrerad till högpresterande elever. Författarna genomför en enkel analys över reformens effektivitet, baserad på hur mycket eleverna kan komma att tjäna i ett längre perspektiv, och kommer fram till att kostnaderna var i princip lika höga som effekterna.

Frankrike

I Frankrike analyserar Bénabou, Kramarz och Prost (2009) en reform från 1982 som gav mer pengar till skolor i vissa utsatta områden. Pengarna användes framförallt till att ge bonusar till lärarna (oberoende av elevernas prestationer), samt i mindre utsträckning minska klasstorleken och öka antalet lärarledda timmar. Författarna jämför högstadiel elever som gick på dessa skolor 1980, innan reformen, med elever som gick på dessa skolor 1989. De finner inga effekter alls av de extra resurserna på sannolikheten att eleverna går ut skolan med fullständiga betyg eller att de klarar centralproven i slutat på gymnasiet.⁶⁷ Resurserna som tilldelades motsvarande i slutändan dock endast 5 procent extra – eftersom de spreds ut över många elever – och att effekten uteblev kan ha att göra med att resurserna helt enkelt inte var tillräckliga.

Just på grund av att resursökningen var liten avskaffades programmet och ersattes 2006 med en annan reform som ämnade koncentrera de högre resurserna till färre skolor. Detta betydde att dessa skolor fick i genomsnitt 10 procent högre resurser jämfört med genomsnittet i landet. 90 procent av de extra resurserna användes för att anställa fler lärare och lärarassistenter. Caille, Davezies och Garrouste (2013) och Befy och Davezies (2013) analyserar effekterna av den nya reformen och finner inga eller negativa effekter på elevers resultat. Det visar sig dock att reformen ledde till ökad skolsegregation via signaleffekter: elever från högre socioekonomisk bakgrund börjar undvika skolorna när de får högre resurser (Davezies och Garrouste 2014), en tydlig oavsiktlig konsekvens av reformen.

Även om det inte finns något stöd för att de ökade resurserna förbättrade resultaten är det viktigt att komma ihåg att sorteringseffekten gör det svårare att på ett säkert sätt utvärdera reformen. Dock är sorteringseffekten precis en oavsiktlig konsekvens som måste tas i beaktning av politiker som har

⁶⁷ Intressant nog ökade andelen unga lärare och andelen lärare utan lärarutbildning trots bonusarna kort efter att reformen genomfördes. Detta indikerar en potentiell oavsiktlig konsekvens av den fokuserade reformen: en stigmaeffekt som skrämde bort mer erfarna och utbildade lärare.

förhoppningar att reformer som ökar resurserna kommer att leda till bättre resultat.

Nederländerna

I en liknande studie analyserar De Witte, Geys och Solondz (2014) en nederländsk reform som genomfördes 2007 och som gav totalt 250 miljoner euro årligen över tio år till 40 fattiga områden med hög arbetslöshet. Hur mycket varje område fick berodde på olika faktorer. Även om pengarna fördelades mellan olika sociala program var förbättringar inom skolan det viktigaste målet med reformen. Författarna jämför trenderna i resultaten på gymnasiet före och efter reformen i dessa 40 områden med andra områden som inte fick extra pengar. De finner först inga effekter alls. Men när de kontrollerar för hur mycket pengar distriktet fick framträder en annan bild: resursernas påverkan på resultaten i centralproven är negativ upp till 1 250 euro per invånare, men ökar något sedan. Efter 2 000 euro extra per invånare blir marginaleffekt positiv. Men den är i verkligheten väldigt liten. Maxbidraget var 3 962 euro per invånare (ungefär 55 procent mer än den genomsnittliga elevkostnaden) vilket enligt resultaten endast ökade elevernas centralprovresultat med 5 procent jämfört med åren innan reformen genomfördes.⁶⁸ I genomsnittsdistriktet var effekten fortfarande negativ. Studien finner alltså inget stöd för att reformen ledde till bättre resultat.

Eftersom det inte finns några effekter alls av resurser på prov som rättas internt i skolorna menar författarna också att resurserna ökar betygsinflationen, eftersom de negativa effekterna på centralproven borde reflekteras i de internrättade proven. De menar att betygsinflationen uppstår på grund av att skolor har incitament att öka resultaten eller sänka nivån på proven för att visa att de förbättrar sig. Detta är naturligtvis viktigt i den svenska kontexten där elevers resultat bestäms av deras lärare. Att den har implikationer för hur många andra studier bör tolkas har redan diskuterats i samband med Jackson, Johnson och Persicos (2014) studie.

En annan studie bekräftar att effekterna av höjda resurser inte är positiva i Nederländerna. Leuven m.fl. (2007) analyserar två reformer som genomfördes år 2000. Den första gav extra pengar för personal till låg- och mellanstadieskolor med en hög andel fattiga elever. Den andra reformen gav extra pengar för datorer och mjukvara till låg- och mellanstadieskolor med en hög andel fattiga elever. Båda ökade resurserna med ungefär 10 procent. Författarna jämför skolor som låg precis över gränsen för att få subventionerna med de som låg precis under gränsen. De finner att reformerna inte hade positiva effekter på resultaten – faktum är att båda hade *negativa* effekter på vissa av utfallen medan andra inte är statistiskt signifikanta. De spekulerar att detta beror på att skolpengen i Nederländerna redan är viktad, och på grund av avtagande effekter är det mycket möjligt att skolor med en hög andel fattiga elever redan har nått brytpunkten för resursernas effekter.

⁶⁸ Författarna presenterar inte standardavvikelsen och jag kan därför inte beräkna effektstorleken med detta mått.

Ytterligare en studie fokuserar på en reform som gav högre resurser till elever med inlärningsvårigheter. de Haan (2012) analyserar effekten av att ge 56 procent mer resurser till gymnasieelever med inlärningsvårigheter än den genomsnittliga elevkostnaden. Skolorna får välja själva hur de ska spendera pengarna. Författarinnan finner positiva effekter på resultaten bland dessa elever. Om vi antar att effekten är linjär indikerar resultaten att 20 procent högre resurser genererar en effektstorlek på 0,06–0,08 sd i holländska samt 0,09–0,12 sd i matematik. Problemet med studien är dock att den använder en speciell metod för att komma åt orsakssambanden, som kräver specifika antaganden om hur eleverna som inte får extra resurser skulle ha presterat om de hade fått det samt hur inkomsten i grannskapet påverkar elevers resultat. Detta gör att studiens ansats inte är alltför tillförlitlig i utbildningspolitiska sammanhang.

Forskning som analyserar förändringar över tid

Förutom studier som analyserar reformer finns det även forskning som analyserar resursförändringar över tid, antingen direkt med alternativa metoder som kan komma åt orsakssambanden eller via reformer som fokuserar på något annat men som indirekt också skapar förändringar i resurstilldelningen. Under vissa omständigheter kan forskare använda dessa reformer som naturliga experiment. En fördel är att aktörer själva inte blir lika medvetna om resursförändringarna, vilket kan minimera att de agerar strategiskt på ett sätt som förstärker eller minskar resurseffekterna. Detta problem gör dock att det inte är helt självklart att effekter i dessa naturliga experiment kan extrapoleras till reformer som explicit ökar resurserna.

Sverige

I Sverige finns endast en studie som analyserar förändringar över tid. Fredriksson och Öckert (2008) fokuserar på kommunaliseringen, som ledde till resursförändringar som påverkade kommuner olika. Totalt sett minskade lärartätheten mellan 1991 och 2000 från 9,1 till 7,4 i grundskolan. Författarna beräknar att detta motsvarar en ökning i klasstorleken med fyra elever, från 21,8 elever per klass till 25,8 elever per klass. Neddragningarna skedde över hela landet, men mer i vissa kommuner än i andra.

Författarna finner att högre lärartäthet är positivt för resultaten, med liknande effekter bland olika elevgrupper, men effekten är liten. Resultaten indikerar att en ökning i lärartäthet med 10 procent leder till en ökning i meritvärdet med 0,02 sd. Den totala effekten av de minskade resurserna under 1990-talet är därför ungefär 0,03 sd totalt sett. Om vi istället beräknar resultaten efter en klasstorleksökning på 5 elever är totaleffekten 0,05 sd. Detta är också en liten påverkan givet att det skulle krävas en ökning av lärartätheten med ungefär 25 procent från dagens nivå.⁶⁹

⁶⁹ Detta baseras på antagandet att en ökning med lärartätheten med 2,13 procentenheter leder till fem färre elever i klasserna.

När författarna analyserar effekterna på resultaten i mönstringsprovet bland män är effekterna större och indikerar att resursförändringen på 1990-talet minskade resultaten med 0,06 sd. Detta betyder i sin tur att fem färre elever i klasserna skulle öka resultaten med 0,09 sd. Givet kostnaden är detta inte heller en speciellt stor effekt, speciellt eftersom det är resultatet av flera års minskande lärartäthet. Att effekterna fortfarande inte är speciellt starka totalt sett är därför uppseendeväckande.

Man bör också komma ihåg att studien endast analyserar effekterna av *minskade* resurser snarare än ökade resurser, vilket inte är samma sak eftersom elever och lärare kan reagera annorlunda på neddragningar jämfört med öknings. Det är därför svårt att säga om effekten hade varit likadan om de istället hade analyserat en reform som indirekt ökade lärartätheten mer i vissa kommuner jämfört med andra.

Dessutom står det klart att kommunaliseringen påverkade annat än resurserna och det är långt ifrån säkert att författarnas resultat endast fångar upp effekterna av resursförändringar. Många andra förändringar som kan påverka lärarkvaliteten olika i kommunerna kan också spela roll. Effekten av friskolereformen tas heller inte i beaktning. Samma sak gäller den ekonomiska krisen som elever under 90-talet levde igenom, som kan ha påverkat elever annorlunda i olika kommuner oavsett resurstilldelningen i skolan. Ett annat problem är återigen att resurserna kan öka betygsinflationen, vilket i sin tur också kan leda till kaskadeffekter och därmed öka resultaten senare i livet.⁷⁰ Sammantaget är det vanskligt att basera utbildningspolitiken på uppsatsen.

Danmark

I Danmark använder Heinesen (2010) en annorlunda metod för att analysera klasstorlekseffekter i årskurs 9. Författaren menar att föräldrar ofta väljer skola efter klasstorlek i grundämnen, men att de inte gör det i andra specifika ämnen, exempelvis B-språk. Han fokuserar därför på resultaten i franska och begränsar sig till skolor som har 20 eller färre elever som läser franska. Genom att analysera förändringar i klasstorleken över tid inom skolan använder författaren variation i klasstorleken som inte kan förutses när elever och föräldrar väljer skola. Detta eftersom denna drivs av slumpmässig variation i antalet elever som väljer franska istället för tyska. Han finner positiva effekter av små klasser som knappt påverkas alls av kontroller för elevernas bakgrund, elevantal i skolan eller elevers resultat i andra ämnen. Den positiva påverkan indikerar att 5 färre elever i klassrummet genererar ungefär 0,12 sd högre resultat i franska. Effekten är störst bland pojkar och lägre presterande elever.

Studien är väl genomförd, men kan fortfarande överdriva eller underdriva effekterna av klasstorlek.⁷¹ Även om det faktiskt är så att variationen i

⁷⁰ För exempel på kaskadeffekter, se Murphy och Weinhardt (2013).

⁷¹ Exempelvis kan elever byta språk tidigt i årskurs 7, vilket författaren inte observerar, och om detta påverkas av klasstorleken är det inte säkert att resultaten stämmer. Vissa elever som föredrar små klasser kan byta till tyska redan i årskurs 7, vilket kan göra att effekterna är

klasstorlek är slumpmässig är det oklart vilka utbildningspolitiska slutsatser som bör dras. Först och främst analyseras endast resultaten i franska och effekterna kan inte extrapoleras till andra ämnen. Därför analyseras ett litet urval av alla elever – endast 33 314 av 154 191 elever gick i en skola som erbjöd franska och av dessa läste endast 8 582 elever ämnet. Det är inte alls säkert att effekterna är likadana för den stora majoriteten elever som väljer ett annat språk. Författaren försökte analysera effekterna av klasstorlek i tyska men på grund av databegränsningar är resultaten väldigt känsliga beroende på modellspecifikation. Den utbildningspolitiska relevansen av studien är därför långt ifrån klar.

England

I England ökade resurserna med 40 procent mellan 2000 och 2007. Ett par studier har analyserat om dessa förändringar ledde till förbättringar i resultat. Holmlund, McNally och Viarengo (2010) analyserar dessa förändringar genom att kontrollera för en stor mängd variabler, inklusive skolspecifika effekter samt skolspecifika trender i resultaten. De finner att en ökning med resurserna med 1 000 pund (vilket motsvarar 33 procent) leder till positiva effekter i engelska, matematik och naturvetenskap på 0,03-0,05 sd bland mellanstadieelever. Detta är en väldigt liten effekt givet kostnaderna.

Med i princip samma metod analyserar Nicoletti och Rabe (2012) effekterna av resurser i högstadiet och finner att en resursökning på 1 000 pund (vilket motsvarar 25 procent) resulterar i en effekt på 0,02-0,04 sd bland elever i sista året på högstadiet. När de håller konstant familjeeffekter – vilket i praktiken innebär att de analyserar syskon – minskar dock effekterna till 0,02 sd för alla ämnen. Återigen är effekterna alltså väldigt små givet kostnaderna. I en nyare uppsats finner författarna dessutom att effekterna ökar linjärt efter hur högpresterande eleverna var i mellanstadiet. En ökning med resurserna med 25 procent ökar resultaten bland de högst presterande eleverna med 0,09 sd, medan effekten bland de lägst presterande eleverna är 0,03 sd. Genomsnittseffekten är 0,05 sd (Nicoletti och Rabe 2013). Detta är fortfarande en liten total effekt.

En annan intressant studie från England studerar effekterna av IT-investeringar i skolan. År 2001 förändrades formeln för hur dessa resurser fördelades mellan skoldistrikt, vilket skapade vinnare och förlorare. Machin, McNally och Silva (2007) analyserar effekterna av denna reform, medan de håller andra resursförändringar konstant, och finner positiva effekter på mellanstadieelevers resultat i engelska, och till viss del i naturvetenskap, men inte i matematik. Författarna analyserar andelen elever som når godkänt i skoldistriktet och

överdrivna om eleverna som byter presterar sämre. Intressant nog är klasstorleken i tyska väldigt svagt korrelerad med klasstorleken i franska, vilket gör att det är möjligt att strategiskt välja språk efter klasstorlek. Men eftersom författaren finner att kontroller för elevernas bakgrund eller prestationer i andra ämnen inte har någon större påverkan på effekten av klasstorlek, trots att de har enskilt starka effekter på resultaten i sig, verkar detta inte vara speciellt viktigt. Ett annat problem är dock att lärare i skolor med färre elever som lär sig franska kan ha sökt sig dit medvetet, vilket kan korrelera med deras kvalitet.

finner att 100 procent högre IT-resurser leder till en ökning i andelen elever som når godkänt med 4 procent efter tre år i engelska och naturvetenskap, jämfört med året innan reformen.⁷² Effekten är alltså blandad, men heller inte speciellt stor.

Att spendera mer pengar på teknologi är dock något annat än att öka resurserna generellt eftersom det även innebär en specifik förändring i hur lärare agerar med eleverna. Författarna finner att effekten framförallt beror på att de mest effektiva skoldistrikten fick mer pengar på grund av reformen och att dessa var i en stark position att kunna använda dem på ett effektivt sätt. Det verkar också som att det inte är tillgången till datorer som driver resultaten, utan snarare investeringar i uppdateringar och lärares kunskaper om hur tekniken kan användas.⁷³

Finland

Hur förändringar i resurstilldelningen påverkar resultaten har även analyserats i Finland. Häkkinen, Kirjavainen och Uusitalo (2003) använder en liknande ansats som Fredriksson och Öckert (2008) men analyserar effekten av minskade resurser på gymnasiet under 90-talets ekonomiska nedgång. De finner inga effekter alls. Skillnaderna kan delvis förklaras av att författarna analyserar förändringar från år till år, vilket skapar mindre variation i resurser jämfört med ansatsen i Sverige, som analyserar resultaten vid två tidpunkter med 9 års mellanrum. Dessutom analyserar författarna resursförändringarnas effekter på ett externt rättat prov i slutet på gymnasiet, medan Fredriksson och Öckert (2008) fokuserar på resursförändringar i grundskolans internt rättade prov och på mönstringsprovet bland män. Samtidigt återstår samma problem som noterades i den svenska studien: den ekonomiska nedgången och andra förändringar kan i sig ha påverkat elever i kommuner med olika resursförändringar olika.

Israel

I Israel analyserar Lavy (2012) en reform från 2004 som ändrade resurstilldelning från klassnivå till elevnivå. Resurstilldelningen viktades så att elever från lägre socioekonomisk bakgrund fick mer pengar. Detta gjorde att vissa skolor fick högre resurser efter reformen, medan andra fick lägre. Författaren finner att en ökning/minskning av undervisningsresurserna med 20

⁷² Eftersom författarna inte presenterar relevant deskriptiv statistik för att kunna beräkna effektstorleken har jag här använt nationell data. Andelen godkänt i engelska och naturvetenskap var 75 respektive 85 procent år 2000:

<http://webarchive.nationalarchives.gov.uk/20130401151655/http://www.education.gov.uk/researchandstatistics/statistics/allstatistics/a00196893/percentage-of-pupils-achieving-level-4-or-above-in>.

⁷³ Just på grund av att forskningen både analyserar resursökningar i sig och att resultaten inte beror på ökad tillgång till datorer eller internet gör att den är värd att notera i den här rapporten, till skillnad från forskningen om IT inom skolan som är blandad (se t.ex. Angrist och Lavy 2002; Carrillo, Onofa och Ponce 2011; Goolsbee och Guryan 2006).

procent ökar/minskar femteklassares genomsnittresultat i matematik, naturvetenskap och engelska med 0,06 sd på elevnivå. Det finns väldigt lite som tyder på att effekterna spillar över från ett ämne till ett annat. Han visar även resultat som indikerar att detta är de långsiktiga effekterna av ökade undervisningsresurser. Detta är återigen en liten påverkan givet kostnaderna. Studien är den enda som har data på hur resurser på klassrumsnivå påverkar resultaten, vilket gör att författaren kan analysera resursernas effekter på den nivå där den bör spela störst roll. Att studien inte hittar mer än marginella effekter är därför synnerligen intressant.

Norge

Den studie som analyserar klasstorleksregler i Norge – och som diskuterades i avsnittet om studier som analyserar variation i resurser vid ett tillfälle – använder även en av Hoxbys (2000) metoder, som fokuserar på slumpmässiga förändringar i befolkningsmängden som i sin tur påverkar klasstorleken. Att det finns fler elever på grund av att födelsetalen var ovanligt höga under ett år påverkar i sin tur antalet klasser som behövs, vilket påverkar klasstorleken. Eftersom befolkningsförändringarna är slumpmässiga har de ingenting med andra faktorer som påverkar klasstorleken att göra, vilket i sin tur möjliggör en analys av orsakssamband. Men Leuven, Oosterbeek och Rønning (2008) finner ingen effekt av klasstorleken på resultaten med den metoden heller. Författarnas slutsats att det inte finns några effekter alls av klasstorlek i Norge backas alltså upp av dessa helt skilda forskningsstrategier.⁷⁴

USA

Ovanstående metod som har använts för att analysera klasstorlekseffekter i Norge introducerades som sagt av Hoxby (2000), som analyserar effekten av klasstorlek i Connecticut med två olika strategier. Den första är densamma som den som har använts i Norge: genom att använda slumpmässiga förändringar i befolkningsmängden, som i sin tur påverkar klasstorleken, kan Hoxby (2000) analysera den kausala effekten av klasstorlek. Den andra strategin använder alla förändringar i elevantal över tid, vilket i sin tur triggar klasstorleksregler likt de som diskuterades i studier som jämför elever vid ett specifikt tillfälle. Skillnaden är alltså att Hoxby (2000) endast fokuserar på elever som går i samma skola vid olika tidpunkter och därför går i klasser med olika storlek. Detta gäller båda strategierna. Författarinnan finner inga effekter alls av mindre klasser bland elever i fjärde och sjätte klass, oavsett vilken metod hon använder. Resultaten varierar inte beroende på elevernas bakgrund. Hon visar också att effekten av att använda klasstorleksregeln på det sätt som ibland används i studier som diskuterades i avsnittet om studier som jämför elever vid ett tillfälle inte ger tillförlitliga resultat. Författarinnans resultat talar därför starkt emot idén att mindre klasser producerar bättre resultat.

⁷⁴ Detsamma gäller för övrigt i princip alla andra metoder som använts för att analysera klasstorlekseffekter, inklusive kombinationen av befolkningsförändringar över tid med regeln om hur stora klasserna får vara.

Nyligen har också Cho, Glewwe och Whitley (2012) använt Hoxbys (2000) första metod för att analysera effekterna av klasstorlek i Minnesota bland tredje och femte klassare. De finner väldigt små positiva effekter av mindre klasser: 5 färre elever i klassrummet skulle öka resultaten med 0,02-0,03 sd. Detta är väldigt magra effekter och indikerar att klasstorleksminskningar inte genererar mer än marginellt högre resultat.

En annan välgjord amerikansk studie har utvärderat effekterna av klasstorlek i Texas. Rivkin, Hanushek och Kain (2005) analyserar elever i årskurs 4-7 och fokuserar på skillnader i klasstorleken som elever upplever över tid, samtidigt som de kontrollerar för skolspecifika effekter varje år. De menar att skillnaderna i klasstorlek därför uppstår slumpmässigt på grund av att vissa elever byter skola – en effekt som också hålls konstant – samt förändringar i skoldistriktens regler gällande klasstorleken. De finner att klasstorleken påverkar elevers prestationer: 5 färre elever i klassrummet ökar matematikresultaten med 0,06 sd i årskurs 4, med 0,04 sd i årskurs 5 och med 0,02 sd i årskurs 6. Det finns dock inga effekter i årskurs 7. I läsförståelse är effekten 0,05 sd i årskurs 4 och 0,02 sd i årskurs 5, medan det inte finns några effekter i årskurs 6-7. Författarna finner ingenting som tyder på att effekterna skiljer sig mellan elever från olika bakgrund. Även om författarna finner positiva effekter av klasstorleken är de alltså inte speciellt starka.

Flera länder

Ett par studier har också analyserat elever i flera olika länder. Woessmann och West (2006) analyserar Timss 1995-data från 11 länder (vilket inte inkluderar Sverige) och jämför elever inom samma skola men i olika årskurser (7 och 8), samtidigt som de håller effekten av årskursen konstant. De menar att skillnaderna i klasstorlek mellan dessa årskurser kan anses vara slumpmässiga och presenterar olika test för att säkerställa att så är fallet. De finner en effekt på 0,14 sd i Frankrike och 0,13 sd i Island när det gäller matematik av 5 färre elever i klassrummet. I Grekland är effekten 0,12 sd i naturvetenskap, medan den i Spanien är 0,03 sd.⁷⁵ I de andra länderna finns det ingen effekt alls. De analyserar dock inte icke-linjära effekter.

Woessmann (2005) använder samma metodologi som Woessmann och West (2006), men inkluderar även skolor där det inte finns uppgifter om klasstorlek. Han använder istället beräkningar för att få fram klasstorlek för dessa skolor, samt kontrollerar för detta så att det inte driver resultaten. Han finner då endast en negativ effekt i Island i matematik. Även om resultaten varierar något är budskapet detsamma: effekten av mindre klasser är oftast marginell.

Författarnas metodologi är inte perfekt eftersom skolor kan minska klasserna i en årskurs om den är lägre presterande än den andra. Likaså kan lärare av olika kvalitet sorteras in i klassrum med elever som presterar på olika nivåer. Men

⁷⁵ Dessa baseras på den internationella standardavvikelsen som är 100 poäng. Om man beräknar effektstorleken efter den nationella standardavvikelsen blir effektstorleken något högre.

författarna presenterar resultat som indikerar att detta inte driver resultaten.⁷⁶ Man bör dock fortfarande vara något försiktig med tolkningen eftersom detta inte helt kan uteslutas. En mer trolig förklaring är dock att skillnaderna i klasstorleken är mindre än de radikala minskningar som observeras i studier som analyserar regler för hur stora klasserna får vara.

Till sist har även Denny och Oppedisano (2013) analyserat effekten av klasstorleken i USA och Storbritannien när det gäller resultaten i Pisa. De använder en liknande metod som Woessmann och West (2006) och finner negativa effekter av små klasser i Storbritannien men inga effekter i USA. De kan dock inte ta hänsyn till att skillnader i hur många som går på skolan över tid kan påverkas av resultaten och de genomför inte de kontroller som Woessmann och West (2006) genomför. Man bör därför tolka resultaten med försiktighet.

Sammanfattning

Det står alltså klart att även studierna som analyserar reformer och förändringar i resurser över tid är blandade. Majoriteten av studier finner små eller inga effekter alls av resurser, oavsett vilken slags metodologi som används och vilken slags resurs som analyseras.

Först och främst står det klart att de reformer som genomgående minskat klasstorleken i Kalifornien och Florida inte kan beskrivas som annat än stora floppar. Kostnaden var väldigt hög men resultaten var antingen obefintliga eller väldigt små. Detta är högst relevant för den svenska kontexten, där både regeringen och oppositionen nu har föreslagit att klasstorleken ska minska med 5 elever. Resultaten i dessa studier bör därför ses som en skarp varning till svenska politiker att tänka en gång till innan de genomför den reform som föreslagits.

Att resultaten inte påverkas nämnvärt av reformer som ökar resurser mer generellt står också klart. Flertalet reformer i USA verkar inte ha varit speciellt framgångsrika. Den mest positiva studien finner faktiskt inga genomsnittliga effekter på elevernas löner och inkomster senare i livet, trots att de fick väsentligt högre resurser under större delen av sin skolgång. Att det endast finns positiva effekter bland fattiga elever – både när det gäller utbildnings- och ekonomiska utfall – gör även att de utbildningspolitiska slutsatserna är oklara. Eftersom det inte finns genomsnittseffekterna på löner och inkomster, och eftersom det finns indikationer att icke-fattiga elever faktiskt fick lägre löner och inkomster, kan resultaten drivas av signaleffekter som uppstår på grund av betygsinflation. Som bäst indikerar studien att mer pengar till fattiga elever kan förbättra deras resultat, men att resurser för den stora majoriteten elever inte har någon effekt. Som värst indikerar den att majoriteten av elever får betala för att de fattiga får högre löner på grund av betygsinflation. Samtidigt finner andra studier i andra länder inte generellt att elever från låg socioekonomisk bakgrund

⁷⁶ De finner exempelvis inget stöd för att det finns skillnader mellan elever i olika årskurser eller att effekten av klasstorlek i skolor som ökar antalet elever från en årskurs till en annan skiljer sig från effekten i andra skolor.

tjänar på mer resurser, vilket gör det ännu svårare att dra starka utbildningspolitiska slutsatser från forskningen.

Att forskare har funnit att betygsinflationen ökar med högre resurser i Nederländerna är också viktigt i den svenska kontexten. Liksom den italienska studien om klasstorlekens effekter indikerar detta att högre resultat på grund av resurser inte alltid beror på högre kunskaper. I ett land som Sverige där resultaten bestäms internt är detta viktigt att komma ihåg.

När det gäller studier som analyserar resursförändringar över tid finns det heller inget konsekvent stöd för att mer pengar leder till nämnvärt bättre resultat. Det är också värt att betona att kvaliteten på studierna inte verkar vara avgörande för denna slutsats. Om något finner faktiskt studierna med starkast metodologi – de som använder Hoxbys metod för att analysera klasstorlekseffekter – ännu svagare effekter än metodologiskt svagare studier som kan plocka upp annat än resursernas påverkan på resultaten.

Det finns därför inget starkt stöd för att förändringar i resursfördelningen över tid påverkar resultaten mer än marginellt. Det finns vissa skillnader beroende på vilka elever och vilket land som analyseras, men generellt verkar inte resurser ha en konsekvent positiv och nämnvärd effekt som kan rättfärdiga de stora kostnaderna.

d. Är resurser mer effektiva bland elever i lägre årskurser?

Det verkar alltså inte som att genomgående reformer som ökar resurserna för alla elever är bra utbildningspolitik. Men är effekten större bland yngre elever? Det är numera populärt att hävda att resurser är mer effektiva i lägre årskurser. Detta är exempelvis tesen i Socialdemokraternas (2013) rapport om klasstorlekens effekter på resultaten. Idén bygger på James Heckmans teoretiska modeller som stipulerar att investeringar i utbildning när barnen är yngre ger högre avkastning än investeringar när barnen är äldre. Men som Gibbons och McNally (2013) visar är det empiriska forskningsstödet för denna tes inte speciellt starkt: "Problemet är att den empiriska forskningen gällande effekterna av resurser inte helt klart visar att det är lättare att åtgärda de små skillnaderna som existerar tidigt i livet än det är att åtgärda skillnader senare i livet" (Gibbons och McNally 2013, s. 25). Forskning som använder liknande metoder i samma kontext finner ofta liknande effektstorlekar bland elever i lägre och högre årskurser. Författarna menar dessutom att en mer noggrann läsning av Heckmans teoretiska modell gällande investeringar under livets gång indikerar att en mer balanserad approach är att föredra – effekterna av investeringar i unga år försvinner under barnets senare utveckling om man inte ökar resurserna där också.

Argumentet att tidiga interventioner är effektiva har även en empirisk basis. Men denna bygger framförallt på småskaliga experiment/program med fokus på utsatta elever i USA under mitten på 1900-talet. Det är därför svårt att dra slutsatsen från dessa att storskaliga interventioner skulle vara effektiva

(Schaeffer 2009). De senaste resultaten i den enda randomiserade utvärderingen som hittills har gjorts från "Head Start"-programmet i USA – det största amerikanska programmet för barn i förskoleåldern från låg socioekonomisk bakgrund och deras föräldrar (som kostar ungefär 47 000 kr per barn) – visar att de positiva effekterna försvann när barnen började skolan och alltså inte längre var med i programmet. Efter första klass presterade barnen som genomgick programmet lika bra som de som inte deltog. Efter tredje klass var resultaten också likadana. Detsamma gäller generellt andra utfall än akademiska resultat (Puma m.fl. 2012). Att det enda storskaliga fältexperiment som faktiskt genomförts för att testa Heckmans idéer inte haft några tydliga effekter efter att programmet avslutats är bekymmersamt för tesen att vi bör investera mer i yngre elever jämfört med äldre.⁷⁷

En liknande bild framträder i forskningen som har diskuterats i den här rapporten. Resultaten varierar väldigt mycket och det finns väldigt lite som tyder på att de yngsta eleverna tjänar mer på resurser. Detta är intressant eftersom eleverna lär sig mycket mer under ett år i lägre årskurser än vad de gör i högre årskurser, vilket indikerar att resurser faktiskt har starkare marginaleffekter på humankapitalackumuleringen i högre årskurser. Exempelvis har amerikanska forskare beräknat att elever i snitt ökar sina resultat med 0,97 sd i första klass, en effekt som sjunker under tiden i skolan. I årskurs 8 ökar resultat endast med 0,24 sd och i sista året på gymnasiet med endast 0,06 sd (Hill m.fl. 2007). Detta betyder att samma effektstorlek i yngre årskurser betyder mycket mindre i absoluta termer av hur mycket eleverna lär sig under året.

Idén att investeringar i lågstadiet skulle vara mer effektiva än investeringar i senare årskurser, som exempelvis Socialdemokraterna (2013) och många andra i den svenska debatten hävdar, verkar alltså vila på ganska skakig empirisk grund.

5. Diskussion och slutsatser

Sammantaget står det alltså klart att forskningen om relationen mellan resurser och resultat i skolan är långt ifrån så positiv som ofta hävdas i debatten. Först och främst är studierna som analyserar effekten av resurser vid ett tillfälle blandad. Resultaten beror på vilken variation i resurser och vilka länder som analyseras. Som beskrevs i avsnitt 4.c är den mest citerade svenska studien mer osäker än vad som hittills framkommit i debatten. Viktigt är också att två andra studier – varav en använder samma metodologi som den förstnämnda – inte finner positiva effekter av mindre klasser i Timss.

⁷⁷ Andra studier som använder kvasiexperimentella metoder finner positiva effekter i både ett kortsiktigt och långsiktigt perspektiv (se t.ex. Deming 2009; Ludwig och Miller 2007). Problemet är att det inte går att utesluta att resultaten beror på metodologiska brister, och det finns dessutom en rad andra problem som gör det svårt att extrapolera resultaten till andra sammanhang (se Besharov m.fl. 2011). Det är därför svårt att påstå att vi bör lita mer på dessa studier än på det enda storskaliga fältexperiment som genomförts på området. (Det är exempelvis noterbart att Deming [2009] finner positiva effekter på provresultat i lågstadiet, vilket alltså inte återfinns i den randomiserade studien.)

För det andra finner den mest utbildningspolitiskt relevanta forskningen som analyserar reformer och andra förändringar i resurserna över tid mindre och ännu mer blandade effekter. Det finns lite som tyder på att universella resursökningar är kostnadseffektiva i genomsnitt. Detta ger en ännu dystrare bild över politikernas möjligheter att öka resultaten enbart genom att slänga mer pengar på skolan.

All slags forskning ger alltså blandade resultat. Det står dock klart att *fler* studier som analyserar variation i resursfördelningen vid ett tillfälle finner meningsfulla effekter, jämfört med studier som analyserar reformer och förändringar i resurserna över tid. Detta är bekymmersamt eftersom huvudfrågan inte är om resurser påverkar elevers resultat överhuvudtaget, utan om reformer som ökar/minskar resurserna genererar högre resultat på ett kostnadseffektivt sätt. Reformerna påverkar många andra variabler som i sin tur påverkar effekterna av resurser markant.

Det finns även andra förklaringar till varför resurseffekterna skiljer sig åt. Gibbons och McNally (2013) menar att skillnaderna tyder på att små förändringar över tid inte har samma effekt eftersom aktörer inom skolan kan förändra sin egen input som svar på högre resurser. Men som denna rapport har visat finns det också reformer som förändrat resursfördelningen kraftigt och plötsligt men som ändå endast haft små effekter totalt sett efter flera år.

En annan förklaring skulle kunna vara att studierna som analyserar reformer och resursökningar över tid inte använder lika bra metoder. Det finns inte mycket som tyder på det. Visserligen varierar kvaliteten i metodologin bland studierna. Men även metodologiskt starka studier har funnit liknande resultat. Dessutom har jag visat att studier som analyserar reformer och som använder väldigt lika metoder som den mindre utbildningspolitiskt relevanta forskningen generellt finner svaga effekter av ökade resurser.

Jag menar att allt detta istället tyder på att en stor del av forskningen är ganska irrelevant för utbildningspolitiken. I sin strävan att analysera "allt annat lika" för att utröna om resurser spelar roll har många forskare ignorerat de oavsiktliga konsekvenser som uppstår vid reformer och förändringar i resurserna över tid. Det är därför av yttersta vikt att politiker fokuserar på den mest utbildningspolitiskt relevanta forskningen istället för att dra alla studier över en kam.

Ta klasstorleksreformerna i Kalifornien och Florida som exempel. Det är anmärkningsvärt att dessa reformer genomfördes med entusiasm efter studierna om Project STAR – det enda storskaliga fältexperimentet i modern tid – och att dessa floppade rejält. På ett liknande sätt används nu studien av Fredriksson, Öckert och Oosterbeek (2013) av svenska politiker när de försöker bjuda över varandra i sina respektive förslag om mindre klasser. Att politiker nu förlitar sig på en studie där resultaten varierar så pass mycket beroende på metodologiska detaljer och/eller urval – och som dessutom finner tecken på att effekten kan vara överdriven på grund av att motiverade föräldrar flyttar närmare skolor med mindre klasser – är märkligt från en evidensbaserad utbildningspolitisk

synvinkel, speciellt givet de blandade resultaten i resten av forskningen. Indikationerna på att en del av de effekter som funnits i annan forskning drivs av betygsinflation utgör ytterligare en varningsflagg.

Poängen är inte att resurser inte spelar någon roll alls, utan att ökade resurser inte kan spela en *viktig* roll i genomsnitt inom ramen för skolans incitamentsstruktur som den ser ut idag. Detta betyder dock inte att resurser aldrig kan vara viktiga även inom denna incitamentsstruktur. Riktade reformer till specifika elever kan under vissa omständigheter kanske öka resultaten, beroende på hur resursfördelningen ser ut för tillfället. Som Silva (2009, s. 8) argumenterar "verkar det rimligt att dra slutsatsen att 'pengar' *generellt* inte spelar roll, men detta är inte samma sak som att pengar *aldrig* spelar roll eller att pengar inte *kan* spela roll". Problemet är som Hanushek (2006) menar att det är väldigt svårt att veta när och under vilka omständigheter resurser har positiva effekter, vilket också ges stöd av forskningen som diskuterats här.

Eftersom litteraturen är så pass blandad som den är bör vi alltså inte anta att genomgående resursökningar i Sverige kommer att ha några större effekter. De är definitivt inte lösningen på skolans problem. Innan vi kan förvänta oss att ökade resurser bör leda till en mycket mer högkvalitativ skola måste först incitamentsstrukturen förändras. En skiss på hur incitamenten på skolmarknaden skulle kunna förbättras generellt diskuterar jag i en annan rapport (Heller Sahlgren 2013). Som beskrivits i den här studien indikerar exempelvis forskning från Italien och Nederländerna att mindre klasser och högre resurser generellt kan öka betygsinflationen. Att öka resurserna i dagens svenska kontext där lärare kan sätta betyg hur de vill riskerar alltså att spå på betygsinflationen ytterligare istället för att höja kunskaperna. Detta är bara ett exempel – innan vi har åtgärdat de stora strukturella problem som finns i dagens svenska skolsystem bör vi helt enkelt inte utöka resurserna på generell basis.

Ett bättre alternativ är att genomföra ett större policyexperiment för att utröna under vilka omständigheter och för vilka elever resurser kan vara ett bra sätt att öka resultaten i dagens svenska kontext. Då kan vi få reda på om vi bör ge mer pengar för vissa elever och om vi bör ge mindre till andra. Detta är ett mycket vettigare alternativ än att ösa mer pengar på alla elever – risken är nämligen stor att vi helt enkelt slänger pengarna i sjön.

Referenser

- Akabayashi, Hideo and Ryosuke Nakamura. 2013. "Can Small Class Policy Close the Gap? An Empirical Analysis of Class Size Effects in Japan." *Japanese Economic Review* doi: 10.1111/jere.12017.
- Altinok, Nadir and Geeta Kingdon. 2012. "New Evidence on Class Size Effects: A Pupil Fixed Effects Approach." *Oxford Bulletin of Economics and Statistics* 74(2):203-234.
- Angrist, Joshua. 2013. "The Perils of Peer Effects." NBER Working Paper No. 19774, National Bureau of Economic Research, Cambridge, MA.

- Angrist, Joshua D., Erich Battistin, and Daniela Vuri. 2014. "In a Small Moment: Class Size and Moral Hazard in the Mezzogiorno." CEP Labour Market Workshop Paper, London School of Economics, London.
- Angrist, Joshua D. and Victor Lavy. 1999. "Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement." *Quarterly Journal of Economics* 114(2):533-575.
- Angrist, Joshua and Victor Lavy. 2002. "New Evidence on Classroom Computers and Pupil Learning." *Economic Journal* 112(482):735-765.
- Böhlmark, Anders and Lindahl Mikael. 2013. "Independent Schools and Long-Run Educational Outcomes Evidence from Sweden's Large Scale Voucher Reform." Working Paper 6/2013, Swedish Institute for Social Research (SOFI), Stockholm University.
- Bénabou, Roland, Francis Kramarz, and Corinne Prost. 2009. "The French zones d'éducation prioritaire: Much ado about nothing?" *Economics of Education Review* 28:345-356.
- Besharov, Douglas J., Peter Germanis, Caeli A. Higney, and Douglas M. Call. 2011. "Ludwig/Miller Head Start Evaluation." Report, School of Public Policy, University of Maryland, College Park, MD.
- Bingley, Paul, Vibeke M. Jensen, and Ian Walker. 2005. "The Effects of School Class Size on Length of Post-Compulsory Education: Some Cost-Benefit Analysis." Discussion Paper No. 1605, Institute for the Study of Labor (IZA), Bonn.
- Browning, Martin and Eskil Heinesen. 2007. "Class Size, Teacher Hours and Educational Attainment." *Scandinavian Journal of Economics* 109(2):415-438.
- Caille, Jean-Paul, Laurent Davezies, and Manon Garrouste. 2013. "Evaluation of the ambition success networks A regression discontinuity analysis." Conference Paper, Annual Meeting of the French Economic Association, Aix-en-Provence.
- Card, David and Abigail A. Payne. 2002. "School finance reform, the distribution of school spending, and the distribution of student test scores." *Journal of Public Economics* 83:49-82.
- Carrillo, Paul, Mercedes Onofa, and Juan Ponca. 2011. "Information Technology and Student Achievement: Evidence from a Randomized Experiment in Ecuador." IDB Working Paper Series IDB-WP-223, Inter-American Development Bank, Washington, DC.
- Cascio, Elizabeth U., Nora Gordon, and Sarah Reber. 2013. "Local Responses to Federal Grants: Evidence from the Introduction of Title I in the South." *American Economic Journal: Economic Policy* 5(3):126-159.
- Cellinie, Stephanie R., Fernando Ferreira, and Jesse Rothstein. 2010. "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design." *Quarterly Journal of Economics* 125(1):215-261.
- Chaudhury, Latika. 2009. "Education inputs, student performance and school finance reform in Michigan." *Economics of Education Review* 28:90-98.
- Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane W. Schanzenbach, and Danny Yagan. 2011. "How Does Your Kindergarten Classroom Effect Your Earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126:1593-1660.

- Chingos, Matthew M. 2013. "Class Size and Student Outcomes: Research and Policy Implications." *Journal of Policy Analysis and Management* 32(2):411-438.
- Cho, Hyaunkuk, Paul Glewwe, and Melissa Whitler. 2012. "Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools." *Economics of Education Review* 31:77-95.
- Chung, Il H. 2013. "Education Finance Reform, Education Spending, and Student Performance: Evidence From Maryland's Bridge to Excellence in Public Schools Act." *Education and Urban Society* DOI: 10.1177/0013124513498413.
- Clark, Melissa. 2003. "Education Reform, Redistribution, and Student Achievement: Evidence From the Kentucky Education Reform Act." Working Paper, Mathematica Policy Research.
- Cohen Zada, Danny, Mark Gradstein, and Ehud Reuven. 2009. "Class Size and the Regression Discontinuity Design: The Case of Public Schools." IZA Discussion Paper No. 4679, Institute for the Study of Labor (IZA), Bonn.
- Das, Jishnu, Stefan Dercon, James Habyariamana, Pramila Krishnan, Karthik Muralidharan, and Venkatesh Sundararaman. 2013. "School Inputs, Household Substitution, and Test Scores." *American Economic Journal: Applied Economics* 5(2):29-57.
- Davezies, Laurent and Manon Garrouste. 2014. "More harm than good ? Sorting effects in a compensatory education program." Unpublished Manuscript, Centre de Recherche en Économie et Statistique.
- de Haan, Monique. 2012. "The effect of additional funds for low-ability pupils: A nonparametric bounds analysis." CESifo Working Paper No. 3993.
- De Witte, Kristof, Benny Geys, and Catharina Solondz. 2014. "Public expenditures, educational outcomes and grade inflation: Theory and evidence from a policy intervention in the Netherlands." *Economics of Education Review* 40:152-166.
- Dee, Thomas. 2012. "School Turnarounds: Evidence from the 2009 Stimulus." NBER Working Paper No. 17990, National Bureau of Economic Research, Cambridge, MA.
- Dee, Thomas S. and Martin R. West. 2011. "The Non-Cognitive Returns to Class Size." *Educational Evaluation and Policy Analysis* 33(1):23-46.
- Deke, John. 2003. "A study of the impact of public school spending on postsecondary educational attainment using statewide school district refinancing in Kansas." *Economics of Education Review* 22:275-284.
- Deming, David J. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start." *American Economic Journal: Applied Economics* 1(3):111-134.
- Denny, Kevin and Veruska Oppedisano. 2013. "The surprising effect of larger class sizes: Evidence using two identification strategies." *Labour Economics* 23:57-65.
- Dieterle, Steven. 2013. "Class-size Reduction Policies and the Quality of Entering Teachers." Discussion Paper No. 2013-14, Scottish Institute for Research in Economics, Glasgow.

- Ding, Weili and Steven F. Lehrer. 2010. "Estimating Treatment Effects from Contaminated Multiperiod Education Experiments: The Dynamic Impacts of Class Size Reductions." *Review of Economics and Statistics* 92(1):31-42.
- Dobbelsteen, Simone, Jesse Levin, and Hessel Oosterbeek. 2002. "The causal effect of class size on scholastic achievement: distinguishing the pure class size effect from the effect of changes in class composition." *Oxford Bulletin of Economics and Statistics* 64(17):17-38.
- Downes, Thomas. 2002. "School Finance Reform and School Quality: Lessons from Vermont." Working Paper, Department of Economics, Tufts University, Medford, MA.
- Downes, Thomas A. and David N. Figlio. 1998. "School Finance Reforms, Tax Limits, and Student Performance: Do Reforms Level-Up or Dumb Down?." Discussion Paper 98-05, Department of Economics, Tufts University, Medford, MA.
- Ellis, Blake. 2011. "Social Security Wrongly Declares 14,000 People Dead Each Year." *CNN Money*, August 22.
- Fahrenthold, David A. 2013. "Agencies Can't Always Tell Who's Dead and Who's not, So Benefit Checks Keep Coming." *Washington Post*, November 3. http://www.washingtonpost.com/politics/agencies-cant-always-tell-whos-dead-and-whos-not-so-benefit-checks-keep-coming/2013/11/03/5e0b89f6-40be-11e3-a751-f032898f2dbc_story.html.
- Fillenbaum, Gerda G., Bruce M. Burchett, and Dan G. Blazer. 2009. "Identifying a National Death Index Match." *American Journal of Epidemiology* 170(4):515-518.
- Fiva, Jon H. and Marte Rønning. 2008. "The incentive effects of property taxation: Evidence from Norwegian school districts." *Regional Science and Urban Economics* 38(1):49-62.
- Fredriksson, Peter and Björn Öckert. 2008. "Resources and Student Achievement—Evidence from a Swedish Policy Reform." *Scandinavian Journal of Economics* 110(2):277-296.
- Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2011. "Long-Term Effects of Class Size." Discussion Paper No. 5879, Institute for the Study of Labor, Bonn.
- Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2012. "Long-Term Effects of Class Size." Working Paper 2012:5, Institute for Evaluation of Labour Market and Evaluation Policy, Stockholm.
- Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2013. "Long-Term Effects of Class Size." *Quarterly Journal of Economics* 128(1):249-285.
- Fredriksson, Peter, Björn Öckert, and Hessel Oosterbeek. 2014. "Inside the Black Box of Class Size: Mechanisms, Behavioral Responses, and Social Background." Discussion Paper No. 8019, Institute for the Study of Labor, Bonn.
- Fredriksson, Peter and Jonas Vlachos. 2011. "Reformer och resultat: Kommer regeringens utbildningsreformer att ha någon betydelse?" Rapport nr 2011/3, Finanspolitiska rådet, Stockholm.
- Funkhouser, Edward. 2009. "The effect of kindergarten classroom size reduction on second grade student achievement: Evidence from California." *Economics of Education Review* 28(3):403-414.

- Gary-Bobo, Robert J. and Mohamed-Badrane Mahjoub. 2013. "Estimation of Class-Size Effects, Using 'Maimonides' Rule' and Other Instruments: The Case of French Junior High Schools." Unpublished Manuscript, Paris School of Economics, Paris.
- Gibbons, Stephen and Sandra McNally. 2013. "The Effects of Resources Across School Phases: A Summary of Recent Evidence." CEP Discussion Paper No 1226, Centre for Economic Performance, London School of Economics, London.
- Gibbons, Stephen, Sandra McNally, and Martina Viarengo. 2012. "Does Additional Spending Help Urban Schools? An Evaluation Using Boundary Discontinuities." Working Paper, London School of Economics, London.
- Glewwe, Paul W., Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina 2013. "School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010." Pp. 13-64 in *Education Policy in Developing Countries*, edited by Paul Glewwe.
- Goolsbee, Austan and Jonathan Guryan. 2006. "The Impact of Internet Subsidies in Public Schools." *Review of Economics and Statistics* 88(2):336-347.
- Goolsbee, Austan and Jonathan Guryan. 2006. "The Impact of Internet Subsidies in Public Schools." *Review of Economics and Statistics* 88(2):336-347.
- Greaves, Ellen and Luke Sibieta. 2014. "Estimating the effect of teacher pay on pupil attainment using boundary discontinuities." IFS Working Paper W14/03, Institute for Fiscal Studies, London.
- Greene, Jay P., Paul E. Peterson, and Jiangtao Du. 1997. "The Effectiveness of School Choice: The Milwaukee Experiment." Working Paper, Program on Education Policy and Governance, Harvard University, Cambridge, MA.
- Guryan, Jonathan. 2001. "Does Money Matter? Regression-Discontinuity Estimates from Education Finance Reform in Massachusetts." NBER Working Paper No. 8269, National Bureau of Economic Research, Cambridge, MA.
- Hægeland, Torbjørn, Oddbjørn Raaum, and Kjell G. Salvanes. 2012. "Pennies from heaven? Using exogenous tax variation to identify effects of school resources on pupil achievement." *Economics of Education Review* 31:601-614.
- Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113:F64-F98.
- Hanushek, Eric A. 2006. "School Resources." Pp. 866-906 in *Handbook of the Economics of Education, Volume 2*. Elsevier.
- Hanushek, Eric A. 2008. "Incentives for Efficiency and Equity in the School System." *Perspektiven der Wirtschaftspolitik* 9:5-27.
- Hanushek, Eric A., John F. Kain, and Steven G. Rivkin. 2004. "Disruption versus Tiebout Improvement: The Costs and Benefits of Switching Schools." *Journal of Public Economics* 88:1721-1746.
- Hanushek, Eric A. and Steven G. Rivkin. 2012. "The Distribution of Teacher Quality and Implications for Policy." *Annual Review of Economics* 4:131-157.
- Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and the Estimated Effects of School Resources." *Review of Economics and Statistics* 78(4):611-627.
- Heinesen, Eskil. 2010. "Estimating class-size effects using within-school variation in subject-specific classes." *Economic Journal* 120:737-760.

- Heller Sahlgren, Gabriel. 2013. "Att skapa en fungerande skolmarknad." Rapport, Timbro, Stockholm.
- Heller Sahlgren, Gabriel. 2014. "Slänger vi pengarna i sjön? Resurser, reformer och utbildningspolitiska slutsatser." IFN Policy Paper, Institutet för Näringslivsforskning (IFN), Stockholm.
- Häkkinen, Iida, Tanja Kirjavainen, and Roope Uusitalo. 2003. "School resources and student achievement revisited: new evidence from panel data." *Economics of Education Review* 22:329-335.
- Hill, Mark E. 2001. "Re: 'Comparison of National Death Index and World Wide Web Death Searches.'" *American Journal of Epidemiology* 153(7):719.
- Hill, Carolyn J., Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey. 2007. "Empirical Benchmarks for Interpreting Effect Sizes in Research." MDRC Working Papers on Research Methodology, Manpower Demonstration Research Corporation, New York.
- Holmlund, Helena, Sandra McNally, and Martina Viarengo. 2010. "Does money matter for schools?" *Economics of Education Review* 20(6):1154-1164.
- Houtenville, Andrew J. and Karen Smith Conway. 2008. "Parental Effort, School Resources, and Student Achievement." *Journal of Human Resources* 43(2):437-453.
- Hoxby, Caroline M. 2000. "The Effects of Class Size on Student Achievement: New Evidence from Population Variation." *Quarterly Journal of Economics* 115(4):1239-1285.
- Hoxby, Caroline M. 2001. "All School Finance Equalizations Are Not Created Equal." *Quarterly Journal of Economics* 116(4):1189-1231.
- Huntington, Justin T., Mathew Butterfield, James Fisher, Daniel Torrent, and Mark Bloomston. 2013. "The Social Security Death Index (SSDI) Most Accurately Reflects True Survival for Older Oncology Patients." *American Journal of Cancer Research* 3(5):518-522.
- Hyman, Joshua. 2013. "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment." Job Market Paper, University of Michigan.
- IRP. 2014. "How is poverty measured in the United States?" Institute for Research on Poverty, Madison, WI. <http://www.irp.wisc.edu/faqs/faq2.htm>.
- Iversen, Jon M. V. and Hans Bonesrønning. 2013. "Disadvantaged students in the early grades: will smaller classes help them?." *Education Economics* 21(4):305-324.
- Jackson, C K., Rucker Johnson, and Claudia Persico. 2014. "The Effect of School Finance Reforms on the Distribution of Spending, Academic Achievement, and Adult Outcomes." NBER Working Paper No. 20118, National Bureau of Economic Research, Cambridge, MA.
- Jackson, Erika and Marianne E. Page. 2013. "Estimating the Distributional Effects of Education Reforms: A Look at Project STAR." *Economics of Education Review* 32:92-103.
- Jepsen, Christopher and Steven Rivkin. 2009. "Class Reduction and Student Achievement: The Potential Tradeoff between Teacher Quality and Class Size." *Journal of Human Resources* 44(1):223-250.

- Johnson, Alex and Nancy Amons. 2008. "'Resurrected,' but Still Wallowing in Red Tape." *NBC News*, februari 29.
http://www.nbcnews.com/id/23378093/ns/us_news-life/t/resurrected-still-wallowing-red-tape/#.U-vfh6jra7A.
- Krassel, Karl F. and Esil Heinsen. 2014. "Class-size effects in secondary school." *Education Economics* <http://dx.doi.org/10.1080/09645292.2014.902428>.
- Krueger, Alan B. 1999. "Experimental Estimates of Education Production Functions." *Quarterly Journal of Economics* 114(2):497-532.
- Krueger, Alan. 2003. "Economic considerations and class size." *Economic Journal* 113:F34-F63.
- Lavy, Victor. 2012. "Expanding School Resources and Increasing Time on Task: Effects of a Policy Experiment in Israel on Student Academic Achievement and Behavior." NBER Working Paper No. 18369, National Bureau of Economic Research, Cambridge, MA.
- Lee, Kyong-Gon and Solomon W. Polachek. 2014. "Do School Budgets Matter? The Effect of Budget Referenda on Student Performance." Discussion Paper No. 8056, Institute for the Study of Labor, Bonn.
- Leuven, Edwin, Mikael Lindahl, Hessel Oosterbeek, and Dinand Webbink. 2007. "The Effect of Extra Funding for Disadvantaged Pupils on Achievement." *Review of Economics and Statistics* 89(4):721-736.
- Leuven, Edwin, Hessel Oosterbeek, and Marte Rønning. 2008. "Quasi-experimental Estimates of the Effect of Class Size on Achievement in Norway." *Scandinavian Journal of Economics* 110(4):663-693.
- Leuven, Edwin and Marte Rønning. 2011. "Classroom Grade Composition and Pupil Achievement." IZA Discussion Paper No. 5922, Institute for the Study of Labor (IZA), Bonn.
- Lindahl, Mikael. 2005. "Home versus School Learning: A New Approach to Estimating the Effect of Class Size on Achievement." *Scandinavian Journal of Economics* 107(2):375-394.
- Ludwig, Jens and Douglas L. Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design." *Quarterly Journal of Economics* 122(1):159-208.
- Machin, Stephen, Sandra McNally, and Costas Meghir. 2010. "Resources and Standards in Urban Schools." *Journal of Human Capital* 4(4):365-393.
- Machin, Stephen, Sandra McNally, and Olmo Silva. 2007. "New Technology in Schools: Is There a Payoff?" *Economic Journal* 117:1145-1167.
- Matsudaira, Jordan D., Adrienne Hosek, and Elias Walsh. 2012. "An integrated assessment of the effects of Title 1 on school behavior, resources, and student achievement." *Economics of Education Review* 31:1-14.
- Mueller, Steffen. 2013. "Teacher experience and the class size effect — Experimental evidence." *Journal of Public Economics* 98:44-52.
- Murphy, Richard and Felix Weinhardt. 2013. "The Importance of Rank Position." CEP Discussion Paper No. 1241, London School of Economics.
- Neymotin, Florence. 2010. "The Relationship between School Funding and Student Achievement in Kansas Public Schools." *Journal of Education Finance* 36(1):88-108.

- Nicoletti, Cheti and Birgitta Rabe. 2012. "The effect of school resources on test scores in England." ISER Working Paper Series, No. 2012-13, Institute for Social and Economic Research, University of Essex.
- Nicoletti, Cheti and Birgitta Rabe. 2013. "School inputs and skills: complementarity and self-productivity." WP No. 2013-28, Institute for Social and Economic Research, University of Essex.
- Ooghe, Erwin. 2013. "Extra resources for disadvantaged pupils: who gains and what works?" Unpublished Manuscript, Department of Economics, KULeuven.
- Piketty, Thomas. 2004. "L'impact de la taille des classes et de la ségrégation sociale sur la réussite scolaire dans les écoles françaises : une estimation à partir du panel primaire 1997." Unpublished Manuscript, Paris.
- Piketty, Thomas and Mathieu Valdenaire. 2006. "L'impact de la taille des classes sur la réussite scolaire dans les écoles, collèges et lycées français: Estimations à partir du panel primaire 1997 et du panel secondaire 1995." Les Dossiers n173 , Ministère de l'Éducation Nationale, Paris.
- Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. "Third Grade Follow-up to the Head Start Impact Study - Final Report." OPRE Report 2012-45, Administration for Children and Families, U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation, Washington, DC.
- Rainey, Lydia. 2011. "Making Sense of Charter School Studies – A Reporter's Guide." Research Brief, National Charter School Research Project, Center on Reinventing Public Education, Washington, DC.
- Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, Schools, and Academic Achievement." *Econometrica* 73(2):417-458.
- Rockoff, Jonah. 2009. "Field Experiments in Class Size from the Early Twentieth Century." *Journal of Economic Perspectives* 23(4):211-230.
- Roy, Joydeep. 2011. "Impact of School Finance Reform on Resource Equalization and Academic Performance: Evidence from Michigan." *Education Finance and Policy* 6(2):127-167.
- SCB. 2014. "Ökade skillnader i barnfamiljers inkomststandard." Artikel. http://www.scb.se/sv_/Hitta-statistik/Artiklar/Okade-skillnader-i-barnfamiljers-inkomststandard/.
- Schaeffer, Adam B. 2009. "The Poverty of Preschool Promises: Saving Children and Money with the Early Education Tax Credit." Policy Analysis No. 641, Cato Institute, Washington, DC.
- Schanzenback, Diane W. 2006/2007. "What Have Researchers Learned from Project STAR?" *Brookings Papers on Education Policy* 205-28.
- Schwartz, Amy E. and Leanna Stiefel. 2012. "Moving Matters: The Causal Effect of School Mobility on Student Performance." Conference paper, Prepared for Annual Conference of the Association for Education Finance and Policy, Boston, MA.
- Sesso, Howard D., Ralph S. Paffenbarger, and I-Min Lee. 2001. "Comparison of National Death Index and World Wide Web Death Searches." *American Journal of Epidemiology* 152:107-111.

- Silva, Olmo. 2009. "Some Remarks on the Effectiveness of Primary Education Interventions." IZA Policy Paper No. 5, Institute for the Study of Labor (IZA), Bonn.
- Sims, David. 2008. "A strategic response to class size reduction: combination classes and student achievement in California." *Journal of Policy Analysis and Management* 27(3):457-478.
- Sims, David. 2009. "Crowding Peter to educate Paul: Lessons from a class size reduction externality." *Economics of Education Review* 28(4):465-473.
- Socialdemokraterna. 2013. "Mindre klasser ger högre resultat i skolan." Rapport, Stockholm.
- Svensson, Allan. 1971. "Relative Achievement: School performance in relation to intelligence, sex and home environment." PhD Dissertation, University of Gothenburg.
- Thomas, Jaime L. 2012. "Combination classes and educational achievement." *Economics of Education Review* 31(6):1058-1066.
- Thompson, Shane. 2012. "Funding Shocks to Public Schools and Student Outcomes." Unpublished Manuscript, University of Arizona.
- Urquiola, Miguel and Eric Verhoogen. 2009. "Class-Size Caps, Sorting, and the Regression-Discontinuity Design." *American Economic Review* 99(1):179-215.
- Van der Klaauw, Wilbert. 2008. "Breaking the link between poverty and low student achievement: An evaluation of Title I." *Journal of Econometrics* 142(2):731-756.
- Weinstein, Meryle G., Leanna Stiefel, Amy E. Schwartz, and Luis Chalico. 2009. "Does Title I Increase Spending and Improve Performance? Evidence from New York City." Working Paper, NYU Steinhardt, New York.
- West, Martin R. and Ludger Woessmann. 2006. "Which school systems sort weaker students into smaller classes? International evidence." *European Journal of Political Economy* 22:944-968.
- Wilde, Elizabeth T., Jeremy Finn, Gretchen Johnson, and Peter Muennig. 2011. "The Effect of Class Size in Grades K-3 on Adult Earnings, Employment, and Disability Status: Evidence from a Multi-center Randomized Controlled Trial." *Journal of Health Care for the Poor and Underserved* 22:1424-1435.
- Woessmann, Ludger. 2005. "Educational Production in Europe." *Economic Policy* 20(43):445-504.
- Woessmann, Ludger and Martin West. 2006. "Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS." *European Economic Review* 50:695-736.