

Lars-Erik Borge, SØF/NTNU og Marte Rønning, SSB

Bidrar økt ressursbruk i grunnskolen til bedre elevprestasjoner?

Innledning

Utdanningsforskere har i mange år vært av den oppfatning at økt ressursbruk i skolen, enten i form av redusert klassestørrelse eller økte utgifter per elev, har liten effekt på elevprestasjoner og læringsutbytte. Denne oppfatningen er hovedsakelig basert på Hanusheks (1986) innflytelsesrike gjennomgåelse av litteraturen på dette området. Hanusheks oppsummering er at det er store kvalitetsforskjeller mellom skoler, men at disse forskjellene ikke er et resultat av forskjeller i utgift per elev, klassestørrelse eller lett observerbare kjennetegn ved lærere (for eksempel utdanning) eller skoler (for eksempel organisering av undervisningen). Lærere og skoler er av vesentlig betydning, men det som betyr noe, er vanskelig observerbart og omtales som «teacher skills» og «school efficiency».

I løpet av 1990-tallet ble man i større grad klar over at det er betydelige metodiske utfordringer knyttet til å estimere den kausale effekten av økte ressurser på elevprestasjoner. Med (positiv) kausal effekt menes at økt ressursbruk forårsaker bedre prestasjoner. Mot slutten av 1990-tallet ble derfor Hanusheks forskning utfordret av studier som benytter såkalte sosiale eksperimenter, eller kvasieksemperimenter, for å identifisere den kausale effekten av redusert klassestørrelse på elevprestasjoner (Krueger 1999, Angrist og Lavy 1999).

Nyere studier innenfor utdanningsøkonomisk forskning benytter mer avanserte metoder i sitt forsøk på å avdekke den kausale effekten, men uten at dette har resultert i noen konsensus om hva som er effekten av økt ressursbruk på elevprestasjoner. De fleste av de nyere studiene rapporterer at elevprestasjonene øker dersom man reduserer klassestørrelsen (Angrist og Lavy 1999, Booser og Rouse 2001, Browning og Heinesen 2007, Krueger 1999, Krueger og Withmore 2001, Urquiola 2006), men størrelsen på effekten varierer betydelig. Andre studier (Hoxby 2000, Leuven, Oosterbeek og Rønning 2008) finner derimot ingen effekt av redusert klassestørrelse på elevprestasjoner. Mangelen på konsensus innebærer videre at de nyere studiene ikke gir klare implikasjoner for politikktutforming. Den betydelige variasjonen i de estimerte effektene betyr at det er svært usikkert hvorvidt redusert klassestørrelse er et særlig effektivt virkemiddel for å oppnå bedre elevprestasjoner. Dette gjelder selv når det utelukkende fokuseres på de arbeidene som viser at redusert klassestørrelse gir bedre elevprestasjoner.

Effekt av klassestørrelse på elevprestasjoner er neppe universell, men kan variere med sammensetningen av elevmassen, egenskaper ved lærere og skoler, organisering av undervisningen, nasjonale og kommunale styringssystem, og så videre. Wößmann og West (2006) har undersøkt om effekten av klassestørrelse varierer mellom

skolesystemer i elleve ulike land. De finner at effekten varierer betydelig, og at land der redusert klassestørrelse bidrar til bedre elevprestasjoner, er kjennetegnet ved at lærerne har relativt svak kompetanse. Forfatterne tolker dette som at kompetente lærere håndterer store og små klasser like bra, mens mindre kompetente lærere har problemer med å håndtere store klasser. I så fall kan det være mer effektivt å øke lærernes kompetanse framfor å redusere klassestørrelsen. Denne tolkningen av resultatene kan være kontroversiell, men det bør være nokså ukontroversielt at effekten av klassestørrelse kan variere avhengig av institusjonelle forhold. Én umiddelbar implikasjon av dette er at man skal være forsiktig med å basere politikktutforming på resultater fra andre land, i hvert fall hvis skolesystemene er svært forskjellige.

Vi har to formål med denne artikkelen. Det første formålet er å presentere og illustrere forskningsdesign og identifikasjonsstrategier som kan benyttes for å avdekke hvordan elevenes ferdigheter påvirkes av økt ressursbruk i skolen. Det andre formålet er å gjøre rede for nyere norske studier som benytter moderne metoder for å avdekke kausale effekter av økt ressursbruk. Artikkelen er organisert på følgende måte: Vi gjør først rede for de sentrale metodiske og statistiske utfordringene, og det begrunnes hvorfor enkle korrelasjons- og regresjonsanalyser ikke er tilstrekkelige. Deretter viser vi hvordan sosiale og naturlige eksperimenter kan benyttes til å identifisere kausale effekter på en mer troverdig måte. Avslutningsvis gir vi en oppsummering og tolkning av resultatene fra de norske studiene.

Leseren bør gjøres oppmerksom på at deler av artikkelen forutsetter noe kjennskap til statistiske metoder og derfor kan være noe vanskelig tilgjengelig. Dette gjelder delen om regresjonsanalyse med kontroll-

variabler og den mer tekniske omtalen av instrumentvariabelmetoden i boks 1. Det er fullt mulig å følge framstillingene uten å lese disse delene. Lesere som ønsker en dypere forståelse (utover det som gis i dette kapitlet), anbefales å se nærmere på originalbidragene.

Korrelasjon versus kausalitet

Korrelasjon gir uttrykk for samvariasjonen mellom prestasjoner og ressursbruk. Vi kan tenke oss at vi har data for elevprestasjoner (for eksempel eksamenskarakterer) for elever i grunnskolen og tilhørende ressursbruk i skolen (for eksempel klassestørrelse). Dette er alt vi trenger for å beregne korrelasjonen. Leuven, Oosterbeek og Rønning (2008) studerte elevene som ble uteksaminert fra den norske grunnskolen i 2002 og 2003. En enkel sammenlikning viser at de som går i store klasser, gjennomgående oppnår bedre eksamenskarakterer enn de som går i små klasser. Det er altså en negativ korrelasjon mellom prestasjoner og ressursbruk (målt ved klassestørrelse).

Den negative korrelasjonen innebærer at elever eksponert for høy ressursbruk (små klasser) i gjennomsnitt presterer dårligere enn elever eksponert for lav ressursbruk (store klasser). Dette betyr ikke nødvendigvis at det er en negativ kausal effekt av ressursbruk – at de relativt svake prestasjonene blant elever i små klasser er forårsaket av at de er eksponert for høy ressursbruk. Den negative korrelasjonen kan like gjerne oppstå som følge av at høyt utdannede foreldre (i større grad enn lavere utdannende foreldre) gjennomgående bor i urbane områder der ressursinnsatsen per elev typisk er lavere enn i rurale strøk. Dersom foreldrenes utdanningsnivå er positivt korrelert med barnas prestasjoner (slik mange analyser indikerer), kan dette resultatet heller reflektere foreldrenes utdanningsnivå enn den ressursbruken elevene er eksponert for. Den negative korrelasjonen kan alternativt skyldes

kompenenserende ressursbruk. Med kompenenserende ressursbruk menes at skoleeiere (kommuner) legger til rette for små klasser i områder med antatt svake elever, eller at skoleledere (rektorer) plasserer svake elever i små klasser og sterke elever i store klasser.

Elevprestasjoner og ressursbruk er positivt korrelert dersom elever som er eksponert for høy ressursbruk, presterer bedre enn elever som er eksponert for lav ressursbruk. En positiv korrelasjon kan ikke nødvendigvis tolkes som en kausal effekt – at økt ressursbruk forårsaker bedre elevprestasjoner. Positiv korrelasjon kan like gjerne være et resultat av at foreldre som prioriterer barnas utdanning høyt, velger å bosette seg i områder med høy ressursbruk. Disse foreldrene vil trolig bidra til barnas læring på mange andre måter enn ved valg av bosted, og barna vil sannsynligvis gjøre det bra på skolen selv om de ikke er eksponert for høy ressursbruk.

Disse eksemplene illustrerer at ressursinnsatsen ikke er tilfeldig fordelt på tvers av kommuner, skoler og elever, men påvirket av ulike valg som blir tatt av ulike aktører (det kan være sentrale og lokale skolemyndigheter, skoleledere, lærere, foreldre og eleven selv). Og i mangel på informasjon om disse valgene, må vi være forsiktig med å gi resultatene en kausal tolkning (selv om vi observerer en positiv eller negativ korrelasjon). Fravær av korrelasjon gir heller ikke grunnlag for å utelukke en kausal effekt.

Videre er det verdt å merke seg at det er kausale effekter som er av interesse i forbindelse med politikktutforming. Anta at en beslutningstaker, det kan være kommune eller stat, vurderer å øke ressursbruken for å forbedre elevprestasjonene, og det kun foreligger informasjon om korrelasjoner. Hvis vurderingen av dette tiltaket baseres på en positiv korrelasjon, og denne skyldes at foreldre selekterer seg til ulike

skoler eller kompenenserende ressursbruk, risikerer man at ineffektiv politikk blir implementert. Man øker ressursbruken selv om det ikke har noen effekt på elevprestasjonene. Dersom det er fravær av korrelasjon, men en positiv kausal effekt, risikerer man å gjøre en annen type feil: ikke å gjennomføre et tiltak som faktisk virker. Man unnlater å øke ressursbruken selv om økt ressursbruk bidrar til bedre prestasjoner. For å avdekke kausale effekter må det benyttes mer avanserte statistiske metoder, og analysene må baseres på en klar forståelse av forhold som påvirker elevprestasjonene.

Regresjonsanalyse med kontrollvariabler

I analyser av effekten som ressursbruk har på elevprestasjoner, er det vanlig å ta utgangspunkt i en såkalt produktfunksjonstilnærming. Denne tilnærmingen innebærer at elevprestasjonene betraktes som et resultat av ressursinnsatsen eleven eksponeres for, andre kjennetegn ved lærerne og skolen, og egenskaper ved eleven. Begrepet produktfunksjonstilnærming er vel innarbeidet i litteraturen, men kan gi inntrykk av at utdanningsøkonomer betrakter skolen som en pølsefabrikk – at det er like enkelt å produsere gode elevprestasjoner som gode pølser. Dette er en feilaktig og nokså karikert beskrivelse av fagfeltet. Innenfor den utdanningsøkonomiske litteraturen legges det stor vekt på å poengtere forskjellene mellom utdanningsprosesser og de produktfunksjoner som studentene møter i innføringsbøker i økonomi (se for eksempel Hanushek 1986, s. 1142). Måling av produksjon (elevprestasjoner) og ressursinnsats (særlig lærerkvalitet) vies stor oppmerksomhet.

Produktfunksjonstilnærmingen kan uttrykkes ved følgende regresjonsmodell:

$$Y_i = R_{s(i)}\alpha + X_i\beta + S_{s(i)}\varphi + \eta_i \quad (1)$$

Relasjon (1) uttrykker at skoleprestasjonene til elev i , Y_i , bestemmes av følgende forhold: ressursinnsatsen eleven er eksponert for, $R_{s(i)}$ ($s(i)$ identifiserer skolen til elev i); observerbare kjennetegn ved eleven, X_i (det kan være elevens kjønn, alder og innvandrerbakgrunn og foreldrenes inntekt, utdanning og sivile status); observerbare kjennetegn ved elevens skole, $S_{s(i)}$ (slik som skolestørrelse og lærernes utdanning, kjønn og erfaring); og et restledd η_i . I empiriske anvendelser er det vanlig å omtale X_i og $S_{s(i)}$ som kontrollvariabler. Parametere α representerer den kausale effekten av økt ressursbruk på elevens skoleprestasjoner (det er denne vi er interessert i), mens at β og φ er effektene av henholdsvis elev- og skolekjennetegn på prestasjonene.

Utstyrt med et innføringskurs i økonometrisk analyse, vil det være naturlig å estimere relasjon (1) ved bruk av vanlig minste kvadraters metode (OLS – Ordinary Least Squares). Og gitt at ingen viktige variabler er utelatt fra modellen, vil OLS gi et konsistent anslag på den kausale effekten av ressursinnsats (α).

Studien til Leuven, Oosterbeek og Rønning (2008) viser at inkludering av kontrollvariabler som fanger opp kjennetegn ved eleven og skolen, kan ha stor betydning for resultatene. Mens en enkel korrelasjonsanalyse viser en positiv sammenheng mellom klassestørrelse og elevprestasjoner, gir regresjonsanalysen motsatt resultat. Estimater fra regresjonsanalysen er imidlertid verken statistisk eller økonomisk utsagnskraftig, og kan derfor tolkes som at det ikke er noen systematisk sammenheng mellom klassestørrelse og elevprestasjoner etter at det er kontrollert for elev- og skolekjennetegn.

Det at estimatet til klassestørrelse endrer seg når det kontrolleres for observerbare kjennetegn ved eleven og skolen, tyder på at disse er korrelert med klassestørrelseva-

riabelen. Dette illustrerer viktigheten av å kontrollere for andre forhold (enn ressursbruk) som kan tenkes å variere systematisk med ressursbruken (jmfør diskusjonen i forrige avsnitt). En skal likevel være forsiktig med å gi OLS-estimatet på α en kausal tolkning. Det skyldes at vi ikke kan være hundre prosent sikker på at alle relevante variabler inngår i modellen. Selv om registerdata og annen statistikk gir grunnlag for å inkludere et bredt sett av forklaringsvariabler, er det i praksis umulig å fange opp alle relevante forhold som burde ha vært kontrollert for. For eksempel inneholder ikke administrative registre informasjon om *hvorfor* enkelte elever blir plassert i små klasser, mens andre blir plassert i store klasser. Andre eksempler på slike faktorer som ikke kan observeres på en enkel måte, og som derfor vanskelig lar seg inkludere som kontrollvariabler, er elevens motivasjon og adferd, foreldrenes oppmuntning og holdning til skolearbeidet, lærernes ferdigheter og egenskaper ved de ulike rektorene (listen kan lett gjøres lengre). I den grad slike ikke-observerbare forhold varierer systematisk med forklaringsvariablene ($R_{s(i)}$, X_i og $S_{s(i)}$ i relasjon (1)), vil ikke OLS-estimatet på α kun gi uttrykk for den kausale effekten, men også fange opp effekten(e) av utelatte variabler på elevprestasjoner. For eksempel, dersom en skole systematisk velger å plassere elever med "«dårlig oppførsel»" i små klasser, og «dårlig oppførsel» er negativt korrelert med elevprestasjoner, kan man risikere å undervurdere den kausale effekten av klassestørrelse på elevprestasjoner (gitt at man ikke har informasjon om dette).

Kort oppsummert, på grunn av at relevante forklaringsvariabler vanskelig lar seg måle (og inkludere som kontrollvariabler), vil bruk av OLS kunne gi skjeve estimat (avvik mellom estimert og kausal effekt). Dette er kilden til de metodiske utfordringer knyttet til å avdekke den kausale effekten av ressurser på elevprestasjoner ved hjelp av OLS. Fordi skjevhetene kan gå i ulik retning, er

det vanskelig å vite om vi overvurderer eller undervurderer den kausale effekten. Mer sofistikerte tilnærminger er nødvendige for å nærme oss den kausale effekten, og i det følgende gir vi oversikt over moderne metoder som benyttes i den utdanningsøkonomiske litteraturen.

Løsning I. Eksperimenter

Den ideelle evalueringssituasjonen er å dele den aktuelle elevpopulasjonen inn i to grupper ved hjelp av loddtrekning, og øke ressurstilgangen til kun den ene gruppen. På denne måten blir det rent tilfeldig hvilke elever som eksponeres for høy ressursbruk, og hvilke elever som eksponeres for lav ressursbruk. De to gruppene vil (forventningsmessig) ha de samme observerbare og uobserverbare kjennetegn. Og den kausale effekten av økt ressursbruk kan estimeres som differansen i elevprestasjoner (mellom de to gruppene) relativt til differansen i ressursinnsats.

Det mest kjente (og kanskje eneste «large scale») sosiale eksperimentet som har blitt utført, er «The Tennessee Student-Teacher Achievement Ratio Experiment», kjent som «Project STAR». I dette eksperimentet som startet opp skoleåret 1985/86 i Tennessee, USA, ble barneskoleelever og deres lærere i løpet av de fire første årene i grunnskolen tilfeldig plassert i klasser av ulik størrelse: en lærer fordelt på 22-25 elever; en lærer og en assistent fordelt på 17-24 elever; en lærer fordelt på 13-16 elever. Ved å analysere dataene fra «Project STAR» finner Krueger (1999) at elever i mindre klasser scorer høyere på standardiserte tester enn elever i store klasser. Prestasjonene øker med 4 prosentpoeng det første året eleven er i en liten klasse og med 1 prosentpoeng for hvert påfølgende år. I en oppfølgingsanalyse finner Krueger og Whitmore (2001) at redusert klassestørrelse i grunnskolen kan ha langvarige effekter. Sannsynligheten for å ta en

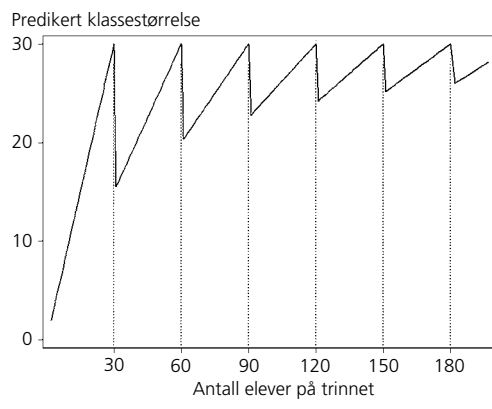
«college-entrance exam» er høyere for elever som har gått i små klasser, enn for elever som har gått i store klasser.

Sosiale eksperimenter i stor skala er lite utbredt, og forskere har derfor jaktet på såkalte naturlige eksperimenter (eller kvasieksperimenter). Naturlige eksperimenter er et resultat av at ulike institusjonelle forhold bidrar til tilfeldig variasjon i den sentrale forklaringsvariabelen (i dette tilfellet ressursinnsatsen) som kan benyttes til å identifisere den kausale effekten. I norske anvendelser har klassedelingsregelen og kommunale kraftinntekter blitt benyttet som naturlige eksperimenter for å identifisere den kausale effekten skoleressurser har på elevprestasjoner.

Klassedelingsregelen

Klassedelingsregelen (som i Norge ble opphevet i 2003) er et eksempel på et institusjonelt forhold som kan generere tilfeldig variasjon i ressursbruken i grunnskolen. Ifølge denne regelen kunne ingen klasser i norske ungdomsskoler bestå av flere enn 30 elever. Regelen skaper dermed en diskontinuitet mellom antall elever og klassestørrelse. Figur 1 illustrerer sammenhengen mellom klassestørrelse

Figur 1. Klassedelingsregelen i norske ungdomsskoler



og antall elever (per trinn) som følger av klassesdelingsregelen. Den predikerte klassestørrelsen er lik antall elever så lenge antall elever er mindre eller lik 30. Hvis antall elever øker fra 30 til 31, deles elevene automatisk inn i to klasser. Den ene klassen vil ha 15 elever og den andre 16, slik at gjennomsnittlig klassestørrelse blir 15,5. Så lenge antall elever er mindre eller lik 60 (og større enn 30), er det to klasser på trinnet, og gjennomsnittlig klassestørrelse øker med økende antall elever. Men i det elevtallet øker fra 60 til 61, øker også antall klasser fra to til tre (20, 20 og 21 elever per klasse). Samme resonnement gjelder når antall elever går fra 90 til 91, fra 120 til 121, fra 150 til 151, og så videre. Ettersom disse «knekkpunktene» er generert av en sentralisert regel og ikke av valg tatt av foreldre og andre skoleledere, kan man utnytte disse til å avdekke den kausale effekten av redusert klassestørrelse på elevprestasjoner.

Skolene som befinner seg i et lite intervall rundt de ulike knekkpunktene, har omtrent lik skolestørrelse. Derimot avhenger klassestørrelsen av om skolen befinner seg like over eller like under et av knekkpunktene. Videre er det rimelig å tenke seg at det er nokså tilfeldig om en elev befinner seg i en skole som ligger like over eller like under et av disse knekkpunktene. Man kan derfor argumentere for at de elevene som befinner seg i et lite intervall rundt disse knekkpunktene, er like på de fleste måter – bortsett fra at noen går i små klasser, mens andre går i store klasser. En annen måte å si det på er at utfallet til elevene i de store klassene gjenspeiler det utfallet som elevene i de små klassene ville ha prestert dersom de hadde gått i de store klassene. Effekten av redusert klassestørrelse på elevprestasjoner kan estimeres som differansen i elevprestasjonene til de elevene som befinner seg like under og like over disse knekkpunktene.

Mange land har en slik regel for klassesdeling, og de første til å utnytte den var Angrist og Lavy (1999) for israelske skoler. Etter det har den blitt brukt av en rekke andre forskere: Urquiola (2006) for Bolivia; Piketty (2004) og Gary-Bobo og Mahjoub (2006) for Frankrike; Browning og Heinesen (2007) og Bingley mfl. (2005) for Danmark; Wößmann (2005) for Danmark, Frankrike, Tyskland, Hellas, Island, Irland, Norge, Spania, Sverige og Sveits; Bonesrønning (2003), Leuven, Oosterbeek og Rønning (2008) og Bonesrønning og Iversen (2009) for Norge. I det følgende gir vi en nærmere beskrivelse av de norske studiene.

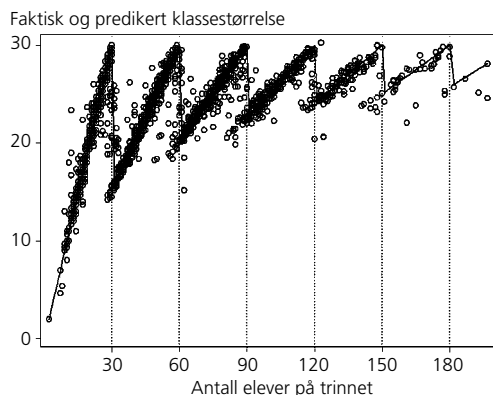
Leuven, Oosterbeek og Rønning (2008) analyserer effekten av redusert klassestørrelse på elevprestasjoner i 10. klasse ved å utnytte regelen for maksimal klassestørrelse. Som et mål på elevprestasjoner benyttes eksamensresultatene i grunnskolen, og som et mål på ressursinnsats benyttes gjennomsnittlig klassestørrelse (som er beregnet ved å dele antall elever per trinn på antall klasser per trinn). Ettersom denne regelen sluttet å eksistere fra og med skoleåret 2003/04, består analysene av avgangselevene i grunnskolen i skoleårene våren 2002 og våren 2003 (eksamenskarakterer på individnivå er ikke tilgjengelig for tidligere kull). Selv om klassesdelingsregelen ikke lenger praktiseres i Norge, har den fortsatt relevans. Elevene er fortsatt organisert i grupper, og gruppestørrelse er en sentral del av organiseringen. Selv om det ikke lenger er en klar regel for fastsettelse av gruppestørrelsen, vil analyser basert på den tidligere klassesdelingsregelen gi viktig informasjon om effektene av endret gruppestørrelse. Det er imidlertid en ulempe at det ikke er mulig å benytte denne identifikasjonsstrategien på nyere data.

En viktig forutsetning for å kunne utnytte klassesdelingsregelen er at den faktisk blir

brukt av skolene. Av figur 2 ser vi at ser vi at skolenes tilpasning er i tråd med denne forutsetningen. Faktisk klassestørrelse samsvarer i stor grad med predikert klassestørrelse. Ved å utnytte variasjonen i klassestørrelse i et lite intervall (± 5) rundt knekkpunktene finner Leuven, Oosterbeek og Rønning (2008) at det ikke er noen forskjell i eksamensresultater mellom skoler over og under knekkpunktene. Tolkningen av dette funnet er at redusert klassestørrelse ikke bidrar til bedre elevprestasjoner (målt ved eksamenskarakterer).

En innvending mot å benytte klassedelingsregelen er at mange små skoler utelates fra analysen. Dette kan være særlig problematisk i et land som Norge, der bosettingsmønsteret er veldig spredt-bygd, og der mange skoler bare har en klasse per trinn. Leuven, Oosterbeek og Rønning (2008) tar hensyn til dette ved å utføre separate analyser for elever som befinner seg i skoler der det predikerte antall klasser er lik én. Dette er typisk små skoler som befinner seg i rurale områder, og det er rimelig å anta at foreldre som bor i disse skolekretsene, velger å la barna gå i disse skolene på grunn mangel på andre

Figur 2. Forholdet mellom faktisk og predikert klassestørrelse i norske ungdomsskoler (skoleårene 2001/02 og 2002/03)



Kilde: Leuven, Oosterbeek og Rønning (2008). Data hentet fra GSI.

alternativer. Videre kan ikke rektorer og andre skoleledere bevisst plassere elever i store eller små klasser (ettersom det kun er én klasse). Man kan derfor argumentere for at disse skolene er mindre utsatt for uobserverbar sortering av elever, og at det derfor er mindre grunn til å tro at estimatene er skjeve. Heller ikke i disse små skolene finner Leuven, Oosterbeek og Rønning (2008) at redusert klassestørrelse bidrar til bedre elevprestasjoner. Urquiola (2006) var den første som utnyttet denne estimeringsstrategien.

Kraftinntekter

Hægeland, Raaum og Salvanes (2008) tar utgangspunkt i at mange norske kommuner mottar betydelige kraftinntekter i form av naturressursskatt (på kraftproduksjon), eiendomsskatt (fra kraftforetak) og konsesjonskraftinntekter. Bakgrunnen for de betydelige kraftinntektene er et politisk ønske om at deler av gevinsten ved kraftutbygginger skal tilfalle utbyggingskommunen. Videre er det slik at eiendomsskatt og konsesjonskraftinntekter ikke omfattes av inntektsutjevningen i inntektssystemet, det vil si at kommunens rammeoverføringer fra staten ikke reduseres som følge av høy eiendomsskatt og høye konsesjonskraftinntekter. Dette har resultert i at mange kraftkommuner har et svært høyt inntektsnivå sammenliknet med andre kommuner. Det høye inntektsnivået gir seg utslag i høy ressursbruk og gode tjenester på de fleste områder, også i grunnskolen.

Ettersom kraftinntektene i hovedsak er knyttet til naturgitte forutsetninger (nærhet til elver og fossefall) og utvikling av teknologi for om lag hundre år siden, kan man argumentere for at det er relativt tilfeldig hvilke kommuner som er kraftkommuner. Med andre ord – disse naturgitte forholdene skaper et naturlig eksperiment som kan utnyttes til å studere effekten av ressursinnsats i skolen på elevprestasjoner når den

sentrale antakelsen er at den høye ressursinnsatsen ikke fanger opp andre forhold som samtidig påvirker elevprestasjonene.

De fleste kommuner med betydelige kraftinntekter per innbygger har relativt få innbyggere og et spredt bosettingsmønster. Hægeland, Raaum og Salvanes (2008) etablerer derfor en kontrollgruppe for hver av de 126 kommunene som er klassifisert som kraftkommuner. Kontrollgruppen konstrueres som de 5 kommunene (blant de som ikke er klassifisert som kraftkommuner) hvor beregnet utgiftsbehov per elev (basert på kostnadsnøkkelen i inntektssystemet) er mest likt den aktuelle kraftkommunen.

I en første tilnærming sammenliknes kraftkommunene med sine kontrollgrupper. Elevprestasjoner måles ved eksamenskarakterer (korrigert for elevkarakteristika) for elever som gikk ut av grunnskolen våren 2003, og ressursbruken ved utgifter per elev og lærertimer per elev. En enkel sammenlikning av kraftkommuner og kontrollgrupper viser at kraftkommunene har signifikant bedre elevprestasjoner. Det er først og fremst kommunene med særlig høye kraftinntekter som bidrar til dette resultatet. En økning i ressursinnsatsen per elev med 10 prosent anslås å øke gjennomsnittskarakteren med 0,044 (karakterene er målt på en skala fra 1 til 6). Effekten kan alternativt illustreres som at 1 av 25 elever forbedrer sin prestasjon med én karakterenhet.

Løsning II. Instrumentvariabelmetoden

De analysene av naturlige eksperimenter som er omtalt foran, gjør ikke bruk av avanserte statistiske metoder. De er snarere basert på at elevene deles i to grupper, en som eksponeres for høy ressursbruk, og en annen som eksponeres for lav ressursbruk, og hvor det er rimelig å anta at fordelin-

gen av elever mellom de to gruppene er tilfeldig. De statistiske testene baseres på sammenlikning av gjennomsnittsverdier for de to gruppene. Analyseopplegget er med andre ord det samme som i kontrollerte eller sosiale eksperimenter, jamfør omtalen av «Project STAR». I forbindelse med naturlige eksperimenter vil et slikt analyseopplegg ofte ha den ulempen at analysen må baseres på relativt få observasjoner. I tilfellet med klassesdelingsregelen var analysen utelukkende basert på skoler i et lite intervall rundt knekkpunktene. Og selv om analysen basert på kraftinntekter omfattet et relativt stort antall kommuner (126 kraftkommuner med tilhørende kontrollkommuner), omfatter den en beskjeden andel av elevene siden små kommuner er kraftig overrepresentert. I studier med naturlige eksperimenter er det derfor også vanlig å benytte mer avanserte statistiske metoder.

Det vanligste er å estimere relasjon (1) ved bruk av den såkalte instrumentvariabelmetoden, og hvor det naturlige eksperimentet står sentralt i valg av instrument. Instrumentvariabelmetoden er en statistisk metode, som på en mest mulig effektiv måte utnytter de deler av variasjonen i data som kan antas å være tilfeldig. Enkelt forklart går instrumentvariabelmetoden ut på at man «byter ut» forklaringsvariabelen av interesse (i dette tilfellet ressursvariabelen) med et instrument, eller en «proxy» (stedfortredende) variabel. Dette instrumentet, eller proxyvariabelen, bør være høyt korrelert med ressursinnsatsen og ikke ha noen direkte påvirkning på elevprestasjonene (kun en indirekte påvirkning gjennom ressursinnsatsen). Det vises til boks 1 for nærmere beskrivelse av metoden. Norske studier som benytter metoden, er: Leuven, Oosterbeek og Rønning (2008); Hægeland, Raaum og Salvanes (2008); Bonesrønning (2003); Bonesrønning og Iversen (2008, 2009).

I det følgende gir vi en omtale av valg av instrument og resultater fra disse studiene.

Leuven, Oosterbeek og Rønning (2008) benytter også instrumentvariabelmetoden i sin studie, hvor instrumentet er basert på klassesdelingsregelen. Den predikerte klassestørrelsen, avledet fra klassesdelingsregelen og antall elever på trinnet, benyttes da som instrument for den faktiske klassestørrelsen. Data for samtlige skoler kan da utnyttes i estimeringen av relasjon (1), ikke bare skolene rundt knekkpunktene. Resultatene påvirkes lite av at utvalgsstørrelsen utvides på denne måten. Konklusjonen er fortsatt at redusert klassestørrelse ikke bidrar til bedre elevprestasjoner.

Bonesrønning (2003) er en tidlig studie som utnyttet klassesdelingsregelen til å estimere effekten av ressursbruk på prestasjoner. Analysen omfatter utvalgte skoler i fem fylker der elevene ble testet som en del av prosjektet. Fordi studien er basert på et begrenset antall skoler, vil det være relativt få skoler rundt knekkpunktene. Instrumentvariabelmetoden er derfor eneste farbare vei her. Bonesrønning finner at effekten av redusert klassestørrelse er relativt liten, men at den er statistisk utsagnskraftig for gutter.

Klassesdelingsregelen utnyttes også av Bonesrønning og Iversen (2009) i en analyse av nasjonale prøver på 4. trinn i skoleåret 2003/04. De finner at mindre klasser bidrar til bedre elevprestasjoner, men den kvantitative effekten er beskjeden. En reduksjon i klassestørrelsen på 8 elever forbedrer testresultatene med 3 prosent av et standardavvik. Videre finner de at effekten av klassestørrelse varierer mellom ulike grupper studenter. Mindre klasser bidrar til bedre prestasjoner for elever med foreldrene som kun har grunnskoleutdanning, men også for disse er den kvantitative effekten beskjeden. For de øvrige elevene er ikke effekten av redusert klassestørrelse statistisk signifikant.

Hægeland, Raaum og Salvanes (2008) benytter også instrumentvariabelteknikker i sin studie. Det er da kraftinntektene som benyttes som instrument for ressursbruken. Relasjoner tilsvarende (1) estimeres da med data for kraftkommunene og kontrollkommunene, samt med data for alle kommuner. Instrumentvariabelmetoden gir også som resultat at økt ressursbruk har en signifikant positiv effekt på elevprestasjonene. Den kvantitative effekten er sterkere enn ved sammenlikning av gjennomsnitt mellom kraftkommuner og kontrollkommuner. En økning i antall

Boks 1. Instrumentvariabelmetoden

Utelatte variabler fanges opp av restleddet. I den grad disse utelatte variablene er korrelert med de observerbare forklaringsvariablene i modellen, vil restleddet være korrelert med forklaringsvariablene. Dette er problematisk dersom man anvender OLS på relasjon (1). Instrumentvariabelmetoden adresserer nettopp dette problemet. Et gyldig instrument (Z_i) er en variabel som er korrelert med ressursbruken (R_{sit}), men som er ukorrelert med restleddet (η_i). Et gyldig instrument har med andre ord ingen direkte effekt på elevprestasjonene. Det påvirker kun prestasjonene indirekte gjennom ressursbruken. Det er enklest å forklare instrumentvariabelmetoden som tostegsprosedyre. I første steg foretas en regresjon med ressursbruk som avhengig variabel og instrumentet og elev- og skolekjennetegn som forklaringsvariabler. I andre steg estimeres relasjon (1) etter at den faktiske ressursbruken er erstattet med predikert ressursinnsats fra regresjonen i første steg. OLS benyttes som estimeringsmetode i begge steg. Denne tostegsprosedyren gir et konsistent anslag på den kausale effekten (α) dersom det valgte instrumentet tilfredsstillt kravene for å være et gyldig instrument. En mer utdypende beskrivelse av instrumentvariabelmetoden finnes hos Cameron og Trivedi (2005, s. 96-98).

lærertimer per elev med 10 prosent anslås å øke gjennomsnittskarakteren med opp mot 0,25. Det kan alternativt illustreres ved at 1 av 4 elever vil forbedre sin prestasjon med én karakterenhet. Dette må sies å være en svært sterk effekt.

Bonesrønning og Iversen (2008) studerer effekten av lærertetthet via resultatene på nasjonale prøver på 5. trinn i 2007. Som instrument for lærertetthet benytter de minstetimetallet i skolen. Ettersom beregningen av minstetimetallet er gitt av en formel basert på skolens elevtall og antall klasser beregnet etter den gamle klasseudelingsregelen, kan man argumentere for at det ikke fanger opp uobserverbare egenskaper ved relevante aktører som kan forårsake skjevheter i estimatene. Analyse-
ne viser at økt lærertetthet per elev har en positiv og signifikant effekt på prestasjonene, og det er spesielt jenter samt elever med lavt utdannede foreldre som tjener på at lærertettheten øker.

Avsluttende merknader

Kan økt ressursbruk i skolen bidra til økt læringsutbytte og bedre elevprestasjoner? Dette er et av de mest sentrale spørsmål i den utdanningspolitiske debatten, og også innenfor den utdanningsøkonomiske litteraturen. For å besvare dette spørsmålet er det nødvendig å avdekke den kausale effekten av ressurser på prestasjoner. Dette er ingen triviell oppgave. Enkle korrelasjons- og regresjonsanalyser er ikke tilstrekkelige, og et av formålene med denne artikkelen har vært å presentere statistiske metoder og identifikasjonsstrategier som i dag representerer forskningsfronten med hensyn til å avdekke kausale effekter av ressursbruk. Det sentrale forskningsdesign er utnyttelse av sosiale eksperimenter, mens instrumentvariabelmetoden er et viktig statistisk verktøy.

Det foreligger flere norske studier som benytter de mest moderne metodene for å estimere effekten av ressurser på elevprestasjoner, og vi har presentert resultatene fra disse. Som i den internasjonale litteraturen er ikke resultatene entydige. De varierer fra ingen effekt (Leuven, Oosterbeek og Rønning 2008), via små effekter for noen elevgrupper (Bonesrønning 2003, Bonesrønning og Iversen 2008, 2009), til svært sterke effekter (Hægeland, Raaum og Salvanes 2008). Så selv om metodene er blitt mer sofistikerte, er det fortsatt ikke etablert noen klar konsensus med hensyn til hvilken betydning økt ressursbruk har for elevprestasjoner og læringsutbytte. De sprikende resultatene har trolig sammenheng med at det benyttes ulike identifikasjonsstrategier og ulike mål på ressursbruk. Hægeland, Raaum og Salvanes (2008) argumenterer for eksempel for at identifikasjon basert på klasseudelingsregelen ikke tar hensyn til kompensierende ressursbruk rundt knekkpunktene. Klasser like under knekkpunktene kan bli tilført tilleggsressurser i form av ekstra lærere og/eller assistenter, og det har da mindre betydning om klassen ligger på den ene eller andre siden av knekkpunktene. Dette er absolutt en relevant innvending, men det er, spesielt for de laveste knekkpunktene, lite trolig at tilførselen av ekstra ressurser er så stor at klasser like under knekkpunktene oppnår samme ressursbruk per elev som dem like over. Leuven, Oosterbeek og Rønning (2008) prøver å imøtegå denne kritikken ved å benytte lærertimer per elev som mål på ressursbruk i stedet for klassestørrelse, men uten at dette påvirker resultatene.

Det er svært vanskelig å peke ut en studie som «bedre og mer riktig» enn de andre. Vi er imidlertid av den oppfatning at det må komme flere studier som dokumenterer betydelige positive effekter, før det er fornuftig å basere politikktutforming på at økt ressursbruk per elev vil gi vesentlig

forbedring av elevprestasjonene. Vår tolkning av de foreliggende studier er at økt ressursbruk neppe er et effektivt virkemiddel for å forbedre elevenes læringsutbytte. En viss forbedring av prestasjonene kan ikke utelukkes, men de potensielle gevinstene er trolig større ved å utnytte eksisterende ressurser bedre.

Vår vurdering av de foreliggende studier er altså at effekten av ressurser på elevprestasjoner i beste fall er liten. Det er imidlertid ikke opplagt hvordan dette skal tolkes. Én mulig tolkning er at elever, lærere og foreldre kompenserer stor klassestørrelse ved å yte en større innsats. En annen mulig forklaring er at lærerne ikke utnytter fordelen med mindre klasser til å bruke mer tid på den enkelte elev, men velger å undervise på mer eller mindre samme måte uavhengig av klassestørrelsen. Disse to tolkningene har svært ulike politikimplikasjoner. Fordi det er enklere å påvirke undervisningsmetoder enn elev- og foreldreinnsats, vil det også være enklere å oppnå effekt av redusert klasse- og gruppestørrelse dersom den andre tolkningen er riktig.

Dette leder oss over på et siste poeng: Effekten av økt ressursbruk er neppe universell, men vil blant annet avhenge av undervisningsorganisering og ledelse ved den enkelte skole, samt av rammebetingelser fastsatt av kommuner og nasjonale myndigheter. Det er et viktig tema for framtidig forskning å undersøke om og hvordan effekten av økt ressursbruk varierer med slike forhold.

Referanser

- Angrist, J.D., og Lavy, V. (1999): Using Maimonides' rule to estimate the effect of class size on scholastic achievement, *Quarterly Journal of Economic*, 114, 533-575
- Bingley, P., Jensen, V.M., og Walker, I. (2005): The effects of class size on educational attainment: Danish quasi-experimental evidence and evidence that controls for family, school and neighbourhood effects, Working Paper.
- Bonesrønning, H. (2003): Class size effects on student achievement in Norway: Patterns and explanations, *Southern Economic Journal* 69, 952-965.
- Bonesrønning, H., og Iversen, J.M.V. (2008): Suksessfaktorer i grunnskolen: Analyse av nasjonale prøver 2007, Rapport 05/08, Senter for økonomisk forskning, NTNU.
- Bonesrønning, H., og Iversen, J.M.V. (2009): Do students from disadvantaged backgrounds perform better in small classes? Mimeo, Institutt for samfunnsøkonomi, NTNU.
- Boozer, M., og Rouse, C. (2001): Intra-school variation in class size: Patterns and implications, *Journal of Urban Economics* 50(1), 163-189.
- Browning, M., og Heinesen, E. (2007): Class size, teacher hours and educational attainment, *Scandinavian Journal of Economics* 109(2), 415-438.
- Cameron, A.C., og Trivedi, P.K. (2005): *Microeconometrics: Methods and applications*, Cambridge University Press, New York.

- Gary-Bobo, R.J., og Mahjoub, M.B. (2006): Estimation of class-size effects using 'maimonides' rule': The case of french junior high schools, CEPR Discussion Paper No. 5754.
- Hanushek, E.A. (1986): The economics of schooling: Production and efficiency in public schools, *Journal of Economic Literature* 24, 1141–1177.
- Hoxby, C.M. (2000): The effects of class size on student achievement: New evidence from population variation, *Quarterly Journal of Economics* 115(4), 1239–1285.
- Hægeland, T., Raaum, O., og Salvanes, K.G. (2008): Pennies from heaven: Using exogenous tax variation to identify effects of school resources on pupil achievement, *IZA Discussion Paper* no. 3561, Bonn.
- Krueger, A.B. (1999): Experimental estimates of education production functions, *Quarterly Journal of Economics* 114(2), 497–532.
- Krueger, A.B. og Whitmore, D.M. (2001): The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star, *Economic Journal* 111, 1–28.
- Leuven, E., Oosterbeek, H., og Rønning, M. (2008): Quasi-experimental Estimates of the Effect of Class size on Achievement in Norway, *Scandinavian Journal of Economics* 110(4), 663-693.
- Piketty, T. (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, Working Paper.
- Urquiola, M. (2006): Identifying class size effects in developing countries: Evidence from rural Bolivia, *Review of Economics and Statistics* 88(1), 171–176.
- Wößmann, L. (2005): Educational Production in Europe, *Economic Policy*, 20(43):445–504.
- Wößmann, L., og West, M. (2006): Class-size effects in school systems around the world: Evidence from between-grade variation in times, *European Economic Review*, 5(3): 695-736.