

HANS BONESRØNNING
Førsteamanuensis ved Institutt for samfunnsøkonomi, NTNU



Utforming av utdanningspolitikken – Hva kan økonomene bidra med?*

Tradisjonelt har økonomer ikke vært aktive premissleverandører i utdanningspolitikken. Dette er i ferd med å endres. Dagens styringsutfordringer, som i stor grad handler om hvordan øke humankapitaloppbyggingen uten ytterligere ekspansjon av utdanningssystemet, appellerer til økonomer. I denne artikkelen argumenteres for at styring via ressursinnsatsen bør forkastes til fordel for resultatstyring. God resultatstyring er vanskelig å etablere, og økonomene har viktige roller å spille i prosessen med å få velfungerende styringssystemer på plass.

1. INNLEDNING

Innflytelsesrike grupper utenfor utdanningssektoren uttrykker bekymring for utdanningssektorens funksjonsmåte. Dels dreier dette seg om at sektoren drar stadig mer ressurser, men kanskje først og fremst om at humankapitaloppbyggingen, målt ved elevresultater, er for dårlig. Dette har ført skole og utdanning helt til toppen av den politiske dagsorden. Spørsmålet er om det er mulig – ved endringer i styringssystemene – å få i pose og sekk, både lavere kostnader og bedre elevresultater.

Utfordringene i utdanningssystemet appellerer ikke bare til politikere, men også til økonomer. Utdanningsøkonomi

er således et raskt voksende forskningsfelt, og mange økonomer opptrer som rådgivere for utdanningspolitikere. Spesielt er dette tilfelle i USA. Europeiske økonomer kommer noe nølende etter. Økonomenes inntreden i skolesektoren er kontroversiell blant mange sektoraktører. Lærernes fagforeninger og mange pedagoger verner om det eksisterende systemet og uttrykker skepsis til økonominspirerte endringsforslag. Forsvarerne av status quo er mange og er meget godt organisert, og de har noen poenger som vi ikke skal avvise uten videre. Ett er påpekingen av at det ikke fins trylleformularer som leder til en bedre utdanningssektor. Et annet er at økonomenes alternativer til status quo er dårlig underbygd. Konfrontert

* Artikkelen er basert på et foredrag forberedt for Det nasjonale forskermøtet i samfunnsøkonomi i 2004. Takk til Bjarne Strøm og en anonym fagkonsulent for nyttige kommentarer til tidligere utkast.

med slike innvendinger vil det etter min mening være en riktig strategi for økonomer å insistere på en *informert diskusjon* om styring av utdanningssektoren. Det betyr først og fremst å insistere på at sektoren åpner seg og genererer data som kan brukes til å evaluere systemenes funksjonsmåte. Jeg skal imidlertid i dette innlegget la forholdet til pedagogprofesjonen, lærernes fagforeninger og andre økonomeksterne forsvarere av det etablerte system ligge, og i stedet rette blikket innover mot oss selv.

Blant økonomer synes det å være bred aspekt for at utdanningsøkonomi representerer en relevant utvidelse av faget, men det er få underavdelinger i faget som opplever kortere avstand mellom fag og politikk og som (følgelig?) kan vise til mer indre uenighet. Formålet med dette innlegget er derfor å diskutere om – og i tilfelle i hvilken grad – utdanningsøkonomer kan gi entydige faglige råd til politikerne om hvordan de kan oppnå bedre elevresultater, og bedre produktivitet, i skolesektoren.

Opplegget for diskusjonen er som følger. Først identifiseres sentrale elementer i eksisterende styringsopplegg som det er bred enighet om, og som nesten ingen ønsker å røre. Deretter identifiseres styring via ressursinnsatsen som det mest problematiske elementet i det etablerte systemet. Avslutningsvis diskuteres resultatstyring som alternativ til styring via ressursinnsatsen.

2. BRED ENIGHET OM NOEN GRUNNLEGGENDE SPØRSMÅL

Det er bred enighet blant økonomer om at grunnutdanningen bør være offentlig finansiert. Begrunnelsene for dette standpunktet varierer noe. Noen peker på positive eksternaliteter, for eksempel at utdanning fører til at individer viser større engasjement i samfunnet, eller at utdanning legger forholdene til rette for raskere økonomisk vekst gjennom innovasjon. I hvilken grad disse argumentene har noe for seg er vanskelig å si: Her er god empiri mangelvare. Det synes som flere legger vekt på at offentlig finansiering bør foretrekkes fordi det gir alle barn like muligheter til utdanning uavhengig av foreldrenes inntekt.

I den grad prinsippet om offentlig finansiert grunnutdanning utfordres, er det av personer som mener at foreldre vil vise mye større engasjement dersom de betaler

for barnas utdanning direkte, og at denne effekten er så sterk og viktig at skolepenger er å foretrekke.

Et annet spørsmål det er relativt stor enighet om er hva skolen bør drive med. De fleste økonomer synes å mene at basisferdigheter i lesing, skriving, og rekning er det viktigste. Historisk sett ligger det nok mye pragmatisme bak dette synspunktet. I sin tid krevde etableringen av en empirisk forskningstradisjon en venstresidevariabel; en målte da det som kunne måles, dvs. en gjennomførte tester for å få kartlagt elevenes kunnskapsnivåer. Den økonomiske intuisjonen tilsa at dette ikke var helt galt: Siden skolene allerede brukte mye tid på å måle læringsresultater i disse fagene, ble sannsynligvis mesteparten av undervisningsressursene også brukt her. Målene hadde altså stor indre validitet. Etter hvert har det kommet empiriske analyser som viser at slike ferdigheter er viktige – og stadig viktigere – elementer i humankapitalen. Målene har følgelig voksende ekstern validitet.

Hva skolen for øvrig driver med, eller bør drive med, synes det å være få klare oppfatninger om blant økonomer. Dette er viktige spørsmål, og følgelig et viktig forskningsfelt. Som antydnet ovenfor handler det om et vidt spekter av mulige sammenhenger, fra hvordan skolen bidrar til innovasjon i produksjonen til hvordan skolen bidrar til å redusere kriminalitet. En grundigere diskusjon faller utenfor rammen for dette innlegget.

De aller fleste land har utdanningssystemer som kjenne-tegnes ved offentlig finansiering og stor vekt på produksjon av basiskunnskaper og -ferdigheter. Utover dette har styringssystemene flere felles kjennetegn: I de fleste land er skolene eid av det offentlige, og styring skjer via ressursinnsatsen, dvs. ved at timetall i fagene, lærerinnsats per elev, krav til lærernes kvalifikasjoner, og i mange tilfeller også arbeidsmåter fastlegges sentralt/nasjonalt gjennom politiske vedtak. Profesjonene og skolebyråkratene er de sentrale aktørene, med lite rom for aktører på etterspørselssida. Det er økende tvil om dette er det beste systemet, og i mange land introduseres nå alternative styringsformer. Spesielt utfordres styring via ressursinnsatsen.

Den økonomisk-empiriske litteraturen som analyserer styring via ressursinnsatsen har levert viktige argumenter til denne debatten. Her gis en kort gjennomgang av hovedkonklusjonene.

3. HVA VET VI OM STYRING GJENNOM RESSURS-INNSATSEN?

Det hele startet med den såkalte Colemanrapporten fra 1966 (Coleman med flere, 1966). Sosiologen James Colemans oppdrag var å undersøke empirisk om skolen kunne brukes til å redusere raseforskjellene i USA ved å styrke utdannelsen for den svarte befolkningen.

Colemans konklusjon var deprimerende for politikere: Skolen kunne ikke påvirke utfallene. Colemans spissformulering er velkjent: «It's all in the family». Konklusjonen var kontroversiell i 60-årene, og er det fortsatt i dag – etter at resultatene er etterprøvd gang på gang i løpet av de snart 40 årene som har gått siden 1966. For en noe grundigere presentasjon av resultatene er det hensiktsmessig å ta utgangspunkt i den produksjonsmodellen som ligger til grunn for den empiriske forskningen. Modellen er – i alle fall tilsynelatende – svært enkel og består kun av den såkalte produktfunksjonen for utdanning:

$$O_{it} - O_{it*} = f(F_i^{(t-t*)}, P_i^{(t-t*)}, S_i^{(t-t*)}) + e_{it}$$

der O_{it} og O_{it*} er elev i's prestasjoner på tidspunkt t og t^* hhv., $F_i^{(t-t*)}$ er en vektor av familiekarakteristika som påvirker eleven i tidsrommet mellom t og t^* , $P_i^{(t-t*)}$ er karakteristika ved medelevene (peer group effekter), $S_i^{(t-t*)}$ er skoleinnsatsfaktorer, og e_{it} et restledd. Modellen er satt opp på value added form, som er den mest vanlige formen å bruke i dag (de tidligste studiene benyttet en nivåform).

Å benytte betegnelsen produktfunksjon for denne likningen er på mange måter misvisende. Likningen forstås nok best som en redusert form som følger fra en underliggende strukturmodell. Jeg kommer tilbake til strukturmodellen senere, men foreløpig er det produktfunksjonen som gjelder.

Colemanrapporten rapporterer signifikante koeffisienter for flere av elementene i F- og P-vektorene, men ingen signifikante effekter for elementene i S-vektoren, dvs. ingen signifikante effekter av klassestørrelse eller lærernes formelle kvalifikasjoner som utdanning og undervisningserfaring. Det er bakgrunn for de nevnte konklusjonene.

Det mest kontroversielle spørsmålet i produktfunksjonslitteraturen dreier seg om klassestørrelsen betydning for elevprestasjonene. Klassestørrelsen er den enkeltfaktor som i størst grad bestemmer enhetskostnadene. Samtidig

har klassestørrelsen lenge og av mange vært betraktet som kanskje det viktigste styringsinstrumentet med hensyn på kvalitet. Troen på klassestørrelsen som viktig determinant for undervisningskvalitet står sterkt blant politikere og folk flest. Blant økonomer er det en opphetet faglig strid om redusert klassestørrelse lar seg forsvare som virkemiddel i utdanningspolitikken.

Hvorfor har vi ikke etter snart 40 år med empiriske undersøkelser klart å avklare dette spørsmålet?

Etter min mening er den viktigste årsaken at det handler om små effekter, som på grunn av simultanitetsproblemer er godt skjult i data. Videre kan en mulig årsak være at klassestørrelsen betyr noe for noen elevgrupper, og under gitte betingelser (som vi ikke fullt ut kjenner), men at mindre klasser langt fra er en sikker vei til bedre elevprestasjoner.

Mange fremragende økonomer har i løpet av de siste fem til ti år vist stort engasjement i denne saken. Den som interesserer seg for klassestørrelsens betydning har derfor tilgang til mye god og innovativ empirisk forskning. Nærliggende eksempler er Angrist og Lavy (1999) og Hoxby (2000). Utgangspunktet for begge analyser er at ikke-eksperimentelle data for klassestørrelse og elevprestasjoner med stor sannsynlighet reflekterer toveis kausalitet enten fordi skolemotiverte foreldre oppsøker små klasser eller fordi skoleledere plasserer lavt presterende elever i små klasser. For å løse disse problemene utnytter Angrist og Lavy eksogen variasjon i klassestørrelse generert av interaksjonen mellom elevtilgang og en regel om maksimal klassestørrelse. Hoxby utnytter eksogen variasjon i klassestørrelse generert ved tilfeldig variasjon over tid i populasjonen av skolebarn. Selv om dette er empirisk forskning av høy kvalitet, har den foreløpig ikke brakt oss nærmere konsensus. Angrist og Lavy finner at elevprestasjonene er bedre i små klasser, for noen elevgrupper (fjerde- og femteklassinger), men ikke for alle (tredjeklassinger), Hoxby finner effekter av klassestørrelse «precisely estimated to zero».

Enkelte hevder med styrke at det trengs eksperimentelle data for å avklare spørsmålet. Det fins egentlig bare ett slikt datasett, generert i det såkalte STAR-eksperimentet i Tennessee i perioden 1985- 89. I dette eksperimentet ble 6-åringer og lærere tilfeldig fordelt mellom små klasser

(13-17 elever) og store klasser (22-25 elever). Flere forskere, blant andre Krueger og Whitmore (2001), har påvist at elevene i de minste klassene lærte mer enn elevene i de største klassene – og at det var langsiktige effekter både i form av større sannsynlighet for å ta collegeforberedende tester og å prestere bedre på slike tester. Ser vi nærmere på funnene oppdager vi at hele den positive effekten som mindre klasser har på testresultater påløper det første året, at disse effektene er til stede i bare 40 av 79 skoler, og de langsiktige effektene drives av svarte elever. Ser vi nærmere på utformingen eksperimentet oppdager vi at det ikke er «rett frem» å gjennomføre eksperimenter av denne typen. Viktigst i så måte er at over 50 prosent av elevene som initialt deltok måtte erstattes underveis, og alle deltagere godt visste hva de deltok i.

Kombinasjonen av litt «rare» funn og svakheter i gjennomføringen gjør at det ikke er enighet om at dette ene eksperimentet gir avgjørende evidens for at redusert klassestørrelse er fornuftig politikk. Det er imidlertid bred enighet blant utdanningsøkonomer om at STAR er et eksempel til etterfølgelse. Vi trenger flere eksperimenter!

Et påtrengende spørsmål er *hvorfor* redusert klassestørrelse ikke systematisk gir bedre elevresultater. I en tynn teoretisk litteratur er Lazear (2001) det nyeste bidraget. Lazears utgangspunkt er at undervisning er et kollektivt gode, men at det oppstår en trengselseffekt når enkeltelever forstyrrer undervisningen (det kan være bråk eller spørsmål fra enkeltelever som klassen for øvrig kjenner svaret på). Lazear viser at trengsels effekter (bråk) gir lite variasjon i undervisningstid – og dermed i elevprestasjoner – i området rundt de mest vanlige klassestørrelsene. Lazear viser også at optimal klassestørrelse reduseres når sannsynligheten for at elevene vil bråke, øker. Altså at klassestørrelsen vil være endogen til elevprestasjonene. Lazears teori vil kunne være vanskelig å selge til skolens folk, fordi forståelsen i de kretser synes å være at undervisningen i stor grad foregår som individuell veiledning. Dvs. at undervisning i hovedsak er et privatgode. Aksepterer vi dette som en riktig beskrivelse av virkeligheten i norsk skole er det overraskende at analyser av ferske data fra grunnskolen gir svake og insignifikante effekter av klassestørrelse (Bonesrønning, 2003).

Jeg kommer litt tilbake til teorier om klassestørrelsen senere. Min lesing av den empiriske litteraturen er at Coleman ikke tok mye feil når det gjelder denne variabelen. Det er imidlertid i dag bred enighet om at Coleman bommet når det gjaldt lærernes betydning.

Modellen som gir grunnlag for å hevde at Coleman tok feil ser slik ut.

$$O_{it} - O_{it^*} = f(F_i^{(t-t^*)}, P_i^{(t-t^*)}) + \sum_j T_{ij} + e_{it}$$

der T_{ij} er en indikatorvariable som viser om eleven har lærer j i perioden mellom t og t^* .

Estimering av denne modellen (se for eksempel Rivkin med flere, 2001) identifiserer sterke og signifikante forskjeller i lærereffektivitet. (De som tenker seg om her vil oppdage at estimering av den ovenstående likningen ikke uten videre vil separere lærereffekter fra andre effekter som opptrer på klassenivå. Det betyr at empiriske analyser som ønsker å separere ut lærereffektene må ha tilgang til data med en del spesielle egenskaper.)

Status med hensyn til lærere er derfor i dag at lærerne betyr mye for elevens kunnskapsøking, men at de viktigste karakteristika ved effektive lærere er ukjente, først og fremst er det ingen sterk sammenheng mellom lærereffektivitet og lærernes utdanning og undervisningserfaring. Variasjonen i lærerkvalitet genererer produktivitetsforskjeller mellom skoler, men også betydelig variasjon i læringsutbytte mellom klasser innenfor den enkelte skole. Det ligger betydelige utfordringer for den empiriske forskningen i å utforme hypoteser om hva som skaper gode lærere, og teste disse ved hjelp av velegnede statistiske metoder. Viktigst i vår sammenheng er imidlertid at de refererte funnene gir grunnlag for den kanskje mest grunnleggende kritikken som kan rettes mot eksisterende styringsordninger: Sertifisering av lærere er i dag et nøkkelledd for å sikre kvaliteten i skoletilbudet, men denne ordningen forhindrer ikke betydelig variasjon i lærerkvalitet. Det etablerte lønnssystemet som bygger på logikken i eksisterende styringssystem, oppfordrer aspirerende lærere til å ta en lang utdanning, men oppfordrer ikke til innovativ lærerpraksis som sikter mot å forbedre undervisningen. Videre, den etablerte styring gjennom ressursinnsatsen fokuserer ikke elevenes læringsutbytte, og har ingen etablerte mekanismer som muliggjør læring fra

beste praksis. Annerledes uttrykt, måten vi styrer utdanningssektoren har minst to viktige svakheter. Lærernes insentiver til innovasjon er svake, og mulighetene for å lære av beste praksis er dårlige. (Jeg kommer senere tilbake til diskusjonen om lærernes insentiver *bør* være svake).

4. FAMILIENS ROLLER I DET ETABLERTE STYRINGSSYSTEMET

I de snart 40 årene som har gått siden Colemanrapporten har en lang rekke produktfunksjonsstudier og andre typer studier av familiens rolle i barnas kunnskapsproduksjon bekreftet foreldres/foresattes viktige rolle som innsatsfaktorer i kunnskapsproduksjonen. Vi har også fått noen indikasjoner på at foreldrenes betydning varierer mellom land. For eksempel viser analyser av PISA-data at sosio-økonomisk bakgrunn har større innflytelse på leseferdighetene til 14-15 åringer i Norge enn i mange andre land (se for eksempel OECD/UNESCO-UIS 2003). En rimelig hypotese er at mellom-land-variasjon i betydningen av familiebakgrunn reflekterer ulikheter i samspillet mellom hjem og skole, men dette vet vi lite om. Vi vet lite om hvordan aktører i hjemmet responderer på endringer i skolen, som økt ressursinnsats eller omlegginger av undervisningsteknologien. Og vi vet lite om hvordan skolen responderer på endringer i hjemmet som for eksempel økt utdanningsnivå hos foreldre, økt kvinnelig yrkesdeltagelse eller økt skilsmissemfrekvens.

I tillegg til produsentrollen er foresattes rolle å etterspørre undervisningstjenester, som de kan gjøre gjennom voice- eller exit-mekanismer. Det foreligger noe evidens for at foresatte er villige til å betale mer for bedre elevprestasjoner, for eksempel i form av høyere boligpriser (se for eksempel Black, 1999), men at de også ønsker høyere skoleutgifter. Det er imidlertid vanskelig å utlede sikre konklusjoner om konsumentatferd fra empirisk forskning fordi konsumentenes alltid har mangelfull informasjon om skolekvalitet. I intervjuer sier foreldre at de vektlegger akademiske prestasjoner høyest ved valg av skole, og vi har noe indirekte evidens om at skolene responderer på økt press fra etterspørselssiden ved forbedringer i produktivitet, først og fremst gjennom bedre elevprestasjoner.

Det eksisterende styringssystemet gir lite plass til foreldrene. Profesjonene sitter i førersetet, og foreldrene inviteres ikke aktivt med i produksjonsprosessen. Voicemekani-

smene er dårlig utviklet, først og fremst gis svært lite informasjon som muliggjør meningsfull foreldremedvirkning. Det kan sikkert mange lesere med karrierer som klassekontakter, bekrefte. Exitmulighetene er begrenset ved den såkalte naboskoleregelen som sier at foreldre kan velge skole ved å velge bosted.

Foreldre er tiltenkt mer aktive roller i alternative styrings-systemer. Dette kommer jeg til om litt. Nå er tiden kommet til å si noe om den underliggende strukturmodellen for kunnskapsproduksjon.

5. KUNNSKAPSPRODUKSJON – EN STRUKTUR-MODELL

Produktfunksjonstilnærmingen behandler skolen som en svart boks – vi putter noe inn i den ene enden og noe annet kommer ut i den andre. Vår kunnskap om det som skjer inne i boksen er beskjedent. Et forsøk på å si noe om dette kan ta utgangspunkt i den strukturmodellen som ligger bak den reduserte formen. Denne modellen er ikke vel etablert, men litteraturen gir flere eksempler på anvendelser av fragmenter av denne modellen – først og fremst benyttet til å diskutere hvordan endringer i styringsopp- legget vil påvirke anstrengelsen til lærere og elever. Her vil jeg ta utgangspunkt i det jeg anser må være kjernen i en slik modell: Samspillet mellom elev, lærer og foreldre. (Correa og Gruver (1987) kommer nærmest denne modellen ved at de diskuterer samspillet mellom lærer og elev.) Det antas at kunnskapsmengden som genereres i dette samspillet er gitt ved den anstrengelse (effort) som alle parter velger å yte i kunnskapsproduksjon:

$$O_{it} - O_{it^*} = f(e_i^{(t-t^*)}, a_i^{T(t-t^*)}, a_i^{P(t-t^*)})$$

der $e_i^{(t-t^*)}$, $a_i^{T(t-t^*)}$ og $a_i^{P(t-t^*)}$ er den anstrengelse eleven, læreren og foreldrene henholdsvis yter mellom tidspunktene t og t^* . Anstrengelsen fastlegges gjennom at alle parter antas å maksimere en nyttefunksjon med elevens kunnskapsøkning og egen fritid som argumenter. Merk at elevenes kunnskapsøkning blir å betrakte som et kollektivt gode i denne formuleringen. Stor kunnskapsproduksjon skjer dersom alle parter samtidig velger å yte stor anstrengelse i kunnskapsproduksjon, men det er neppe tvil om at elevens anstrengelse er den viktigste innsatsfaktoren. En kan tenke seg at kunnskapsproduksjon foregår selv om $a^T = a^P = 0$, men neppe dersom $e = 0$.

Hva kan denne modellen potensielt brukes til? Mye, men la meg her begrense meg til noen få eksempler.

- For det første til å generere hypoteser om egenskaper hos gode lærere: Modellen fremhever lærernes evne til å påvirke elevens innsats som en potensiell suksessfaktor: Kanskje er det slik at gode lærere er kjennetegnet ved at de organiserer undervisningen og etablerer belønningssystemer (for eksempel ved bruk av karakterskalaen) slik at de tvinger eller oppmuntrer stor innsats fra elevene? Eller at faglige dyktige lærere ikke er blinde for at elevene kan respondere på god undervisning eller stor læreranstrengelse med å redusere sin egen innsats?
- For det andre til å belyse klassestørrelsens (manglende) betydning. Vi må da først foreta en utvidelse av antall elever i modellen fra en til mange. Gitt at undervisning er et privatgode (noe den ikke alltid er) vil endringer i klassestørrelsen i første omgang kunne påvirke lærerinnsats per elev. En reduksjon i klassestørrelsen fra 30 til 20 elever vil gi en økning i lærerinnsats per elev med noe mindre enn 5 minutter per dag. Det kan diskuteres om dette er mye eller lite, men uansett må lærerinnsatsen ha en høy marginal avkastning for at denne økningen skal slå mye i elevens kunnskapsnivå. I tillegg kommer altså substitusjonsproblemer: Det er ikke usannsynlig at eleven responderer på økt innsats fra lærerens side med å redusere egen anstrengelse. I så fall er det svært usikkert om redusert klassestørrelse bidrar til økt kunnskapsproduksjon.
- For det tredje kan modellen brukes til å belyse samspillet mellom hjem og skole: Foreldre kan velge å respondere på økt innsats fra skolens side med å redusere egeninnsatsen. Eller foreldre kan jakte på skoler med høy kvalitet fordi de ønsker å komplementere egen stor innsats. Listen over mulige interaksjoner kan enkelt utvides.

De fleste teorimodeller som diskuterer alternative styringssystemer har til felles at de fokuserer en av aktørenes (primært lærernes eller elevenes) anstrengelse. Problemstillingen er da hvordan ulike institusjonelle ordninger (som for eksempel bruk av standarder, bruk av adgangsbegrensning til høyere utdanning, eller bruk av eiendomsskatt til å finansiere skolen) påvirker denne aktørens anstrengelse. Det antas implisitt at de andre aktørenes

anstrengelse er konstant. Styrken til disse modellene er dermed at de eksplisitt fokuserer de viktigste innsatsfaktorene, aktørenes anstrengelse, i kunnskapsproduksjon. Svakheten kan være at de er for partielle.

6. RESULTATSTYRING

Produktfunksjonsstudiene viser at styring via ressursinnsatsen har klare svakheter, i den forstand at økt ressurstilførsel i form av flere og bedre utdannede lærere ikke systematisk gir bedre resultater, og videre at det er stor variasjon i oppnådde resultater for den samme ressursinnsatsen. Spørsmålet er om det mulig å etablere andre styringssystemer som sikrer bedre og mer jevn kvalitet i tilbudet. Det nærliggende alternativet er resultatstyring. Resultatstyring vil enkelt si at inntektene/belønningen til aktørene i skolen følger fra oppnådde resultater. Resultatstyring gir dermed sterkere incentiver – som kan lede til mer ønsket atferd, men også mer uønsket atferd. Utfordringen er derfor å utforme systemene slik at produktiviteten forbedres.

Dette er ingen enkel oppgave, og spesielt ikke når kunnskapsgrunnlaget om resultatstyring i utdanningssektoren er spinkelt. Hva bør vi som økonomer gjøre i en slik situasjon? Her diskuteres flere muligheter.

I hvilken grad kan vi gi anbefalinger basert på økonomisk teori?

Resultatstyring krever i første omgang at de ønskede resultatene identifiseres og kvantifiseres. Vi oppdager da raskt at skolen ikke er som i den enkle anstrengelsesmodellen og karakterisert ved at den produserer ett gode, men er karakterisert ved mange målsettinger, hvor mye ikke er målbart. Det vil ikke være mulig å etablere resultatindikatorer som fanger opp alle skolens målsettinger. Økonomisk teori gir klare prediksjoner om at aktørene vil omallokere ressursene til det som måles.

Aktørene vil i tillegg kunne velge ulike typer dysfunksjonell atferd, dvs. atferd som leder til bedre skårer på indikatorene uten at reelle forbedringer har funnet sted. Dette fanges heller ikke opp av den enkle modellen presentert foran. Modellutvidelsen kan bestå i at vi lar aktørene velge mellom to typer anstrengelse; funksjonell og dysfunksjonell, der sistnevnte er kjennetegnet ved at den ikke gir samfunnsøkonomiske gevinster. Begge typer

anstrengelse gir bedre skårer på indikatorene, men den dysfunksjonelle kan ha større marginal avkastning for lærerne.

Teorier av disse typene, som har vært mye vektlagt i faglitteraturen, kan leses på minst to måter. Som en advarsel mot å ta i bruk resultatstyring, eller som en påminning om at det må lages sofistikerte informasjonssystemer som tar utfordringene med mange målsettinger og dysfunksjonell atferd på alvor.

Hva slags insentivsystemer har en å velge mellom? En relativt mild variant er å offentliggjøre skoleresultater, dvs. etablere et offentlig informasjonssystem, og la det bli med det. Som økonomer tror vi at et slikt opplegg vil kunne disiplinere skolens folk direkte gjennom det potensielle ubehaget ved offentliggjøring, og indirekte ved at foreldre i større grad tar bostedsbeslutninger basert på bedre informasjon om skolekvalitet.

Forsterkning av insentivene utover dette kan i hovedsak gjøres på to måter. En kan legge et belønningssystem på toppen av informasjonssystemet, dvs. å la skolens inntekter følge av oppnådde resultater. Eller en kan åpne for fritt skolevalg, hvor første steg ville være å oppheve nabo-skoleregelen. Det er gode grunner til å tro at et belønningssystem ville kunne gi mer av alt; både bedre elevresultater og mer dysfunksjonell atferd. Fritt skolevalg vil kunne gi mer disiplinering av aktørene i skolen, kanskje kombinert med mindre dysfunksjonell atferd enn hva et belønningssystem tenderer til å skape fordi foreldre – i større grad enn ved et belønningssystem – sannsynligvis vil basere seg på en helhetsvurdering av skolens produksjon. På den andre siden vil imidlertid fritt skolevalg kunne gi mer segregering. Graden og karakteren av segregering vil være avhengig av hvordan systemet utformes: Hvem får velge? Skal alle ha like store vouchere? Osv..

Disse eksemplene er nok for å illustrere et hovedpoeng: Teorien gir ikke klare anbefalinger om bruk av resultatstyring eller hvordan insentivsystemer bør utformes. Teorien er likevel verdifull, først og fremst fordi den rekker langt med hensyn til å identifisere hvilke gevinster og problemer som kan oppstå ved ulike typer reformer.

Teori lyser opp noen felter, mens andre blir liggende i skyggen. Etter min oppfatning har utdanningsøkonomer

så langt ikke i tilstrekkelig grad tatt innover seg hva det betyr at kunnskapsproduksjon er karakterisert ved at kundene er innsatsfaktorer i produksjonen. Kundene kan være elever eller foreldre. Den enkle interaksjonsmodellen mellom lærer og elev som inneholder dette elementet, identifiserer potensielt viktige substitusjonsproblemer. Noen eksempler er allerede nevnt, som at elever kan respondere på god undervisning med å redusere egeninnsatsen, eller at foreldre kan tenkes å jakte på gode undervisningstilbud fordi de søker etter muligheter til å redusere egen innsats i kunnskapsproduksjonen. Overser vi slike effekter kan vi lett komme til å avskrive styringsordninger som egentlig har stort potensial, eller anbefale ordninger som ikke har nevneverdig potensial. Det er trivielt at manglende teori kan bidra til at vi gir dårlige råd til politikere.

Fins empiri som vi kan bruke til å informere politikere?

I USA har reformer med resultatstyring (accountability) stått øverst på dagsorden i de fleste delstater en ti års tid nå, og er i senere tid påskyndet av President Bushs føderale reform «No child left behind» fra 2001. Utformingen av reformene varierer mye mellom statene; men de inneholder alle ett eller flere av elementene nevnt ovenfor.

Overalt er omfattende testing av elevene i sentrale skolefag kjernen i systemet. Noen steder har testresultatene konsekvenser for elevene, for eksempel ved at det stilles krav om en viss minimumsprestasjon for opprykk til neste klassetrinn. Mest vanlig er at testresultatene benyttes som grunnlag for å beregne skolekvalitet og/eller identifisere lærernes bidrag til skolekvalitet.

Empiriske analyser tyder så langt på at accountability virker slik vi tror det vil virke: Systemet har bidratt til noe bedre resultater (se Hanushek og Raymond, 2003), men det har også gitt oss «the test score inflation», dvs. raske forbedringer i testresultater uten at dette reflekterer en tilsvarende forbedring i elevenes kunnskapsnivå. Framgangen tilskrives i stedet mer eller mindre dysfunksjonell atferd i skolene, som at det drives intense testforberedelser, at lærerne hjelper elevene med å besvare oppgavene, at de lavest presterende elevene holdes borte fra testene osv..

Det er strid om hvilke implikasjoner disse resultatene bør få. Noen ser dette som eksempler som underbygger

påstanden om at skolen ikke bør eksponeres for sterke insentiver, andre tolker dette som barnesykdommer som kan fjernes ved bedre tester, bedre rutiner for prøvetaking, bedre mål på skolens bidrag til elevprestasjonene osv.. Etter min mening er det åpenbart mye å hente på bedre utforming av testene. Så langt har disse i for stor grad vært konsentrert på «lave kunnskapsnivåer», som har gjort det opportunt for lærere å drille enkle ferdigheter, kanskje på bekostning av dypere forståelse. I tillegg har kontrollapparatet knyttet til testavvikling vært svakt. En har ikke tatt ordentlig innover seg hva det betyr å drive med såkalt «high stake testing» – testing der mye står på spill for lærerne. Det er fortsatt et åpent empirisk spørsmål om en makter å redusere gevinstene ved dysfunksjonell atferd tilstrekkelig til at reelle produktivtetsgevinster kan oppnås, dvs. slik at lærerne velger den tunge veien som er å forbedre undervisningen. Endelig må vi innse at suksessindikatorer som er i bruk ofte ikke er gode nok, dvs. indikatorene fanger ikke i tilstrekkelig grad opp skolens bidrag til elevenes læringsutbytte. Mange steder brukes ikke de best tilgjengelige indikatorene, samtidig har de best tilgjengelige indikatorene et forbedringspotensial. Dvs. her har både politikere og forskere en jobb å gjøre.

Hva er erfaringene fra systemer som har forsøkt å utnytte konkurransen mellom skoler til å høste produktivtetsgevinster, slik som vouchersystemer og systemer med charter schools?

Vouchererfaringene kommer hovedsakelig fra Milwaukee og Chile. I Milwaukee ble fra 1990 vouchere gitt til lavinntektsfamilier. Hoxby (2003) dokumenterer at de skolene som har flest elever som benytter vouchere for å komme seg vekk, opplever betydelig produktivtetsvekst – primært gjennom forbedring i elevresultatene. Milwaukee-forsøkene illustrerer også et annet poeng: At segregeringen avhenger av hvordan voucherene utformes. I Milwaukee har vouchere gitt mindre inntektssegregering primært fordi lavinntektsfamilier (som er de eneste som får velge) har valgt seg til skoler hvor foreldre har høyere gjennomsnittsinntekt.

Chile har hatt et vouchersystem helt siden 1981. Ordningen går ut på at myndighetene tilbyr vouchere til alle elever som ønsker å begynne i private skoler. Resultatet er ikke uventet at andel elever i private skoler

har økt, fra 20 prosent i 1981 til 40 prosent i 1988. Hsieh og Urquiola (2003) viser at det er middelklassen som har forlatt offentlige skoler til fordel for de private, og at det ikke kan spores forbedringer i elevprestasjonene i Chile i 20-årsperioden. Det eksisterer to syn på denne analysen. Det ene er at analysen ikke er metodisk holdbar: Hsieh og Urquiola mangler data om førsituasjonen, og fritt skolevalg er gjennomført i fullskala fra dag en. Under slike forhold er det vanskelig å etablere troverdige evalueringer. Det andre synet er at analysen gir nyttig kunnskap om hva som kan gå galt med fritt skolevalg: Chilenske foreldre har gått etter elevsammensetningen ved skolene, og skolene har konkurrert om de attraktive elevene, og ikke vært opptatt av å øke sin produktivitet. Viktig i denne forbindelse er at de private skolene er gitt muligheter til å drive «cream skimming» gjennom inntakspolitikken. En mulig lærdom fra Chile kan derfor være at ikke nok er gjort for å hindre skolene i å konkurrere på den enkle måten som er å heve nivået på inntatte elever. En annen lærdom kan være at ikke nok er gjort for å sikre at foreldre gjør informerte valg gjennom offentlig publisering av såkalte «value added» indikatorer.

Charter schools (tilsvarer omtrent friskoler hos oss) representerer en annen måte å etablere mer konkurranse i utdanningsmarkedet. Charter schools har raskt fått betydelig omfang i flere amerikanske stater, og friskoler har hatt rask vekst i Sverige. Evalueringene av dette tiltaket består som oftest i å studere om etablering av charter schools/friskoler har positive produktivitetseffekter på nærliggende offentlige skoler. Hoxby (2003) viser at offentlige skoler responderer på konkurranse fra charter schools med å forbedre elevprestasjonene – men først når konkurransen har nådd et visst nivå (6 prosent eller mer av elevene går til charter schools). Ahlin (2003) viser at svenske friskoler kan ha liknende effekter på offentlige skoler i Sverige.

Hvordan skaffe mer empiri?

All empiri om alternative styringssystemer som er referert ovenfor kommer fra politisk initierte reformer. Strategien med å basere seg på evalueringer av gjennomførte reformer har åpenbare svakheter. En er at reformene ikke nødvendigvis er relevante for å løse problemene (vi har latt andre definere problemene og foreslå løsninger), en annen at de typisk omfatter mange typer tiltak og gjennomføres i full skala.

Det fins en annen tilnærming: Småskalaeksperimenter som utformes av forskere. Det mest kjente er som nevnt ovenfor STAR i Tennessee, der elever og lærere i de første årene i grunnskolen ble allokert tilfeldig til klasser av ulik størrelse. Selv om det er betydelig strid om evalueringene, er det bred enighet om at STAR er et eksempel til etterfølgelse. Victor Lavy og Josh Angrist (se for eksempel Lavy (2002) eller Angrist og Lavy (2002)) har initiert en rekke småskalaforsøk med ulike skolepolitiske virkemidler i Israelsk skole. Forskerne har hånd om utformingen: Ett virkemiddel prøves ut om gangen i liten skala. I prinsippet fordeles deltagende aktørene tilfeldig mellom en behandlingsgruppe og en kontrollgruppe. Både tradisjonelle virkemidler, som klassestørrelse og betydningen av lærernes utdanning (in-service training), og nye virkemidler som resultatavhengig avlønning av lærere og belønning til elever som oppnår gode resultater, er prøvd ut. Resultatene er nesten uten unntak positive, dvs. de finner at det meste virker: Bedre prestasjoner følger ved å gjøre klassene mindre, å gi lærerne mer opplæring på jobben, å utsette lærerne for resultatavhengig avlønning, og ved å gi bonus til lavt presterende elever som oppnår vitnemål. Disse forsøkene representerer eksempler på hvor nær vi kan komme eksperimenter utenfor laboratoriene. Den viktigste bekymringen er kanskje at Hawthorneeffekter kan gjøre seg gjeldende. Hawthorneeffekter oppstår når du gjennom å vie individer større oppmerksomhet oppnår at de arbeider mer/bedre. I vår sammenheng betyr det at korrelasjonene mellom elevprestasjoner og politikkvariablene avdekt i eksperimenter kan være spuriøse.

7. LÆRERREKRUTTERING

Argumentene for resultatstyring er først og fremst at slike systemer kan bidra til bedre ressursutnyttelse for gitt beholdning av lærere. Mange vil imidlertid hevde at det virkelig store potensialet for produktivitetsfremgang ligger i å rekruttere og beholde gode lærere. Det store gapet mellom de beste og dårligste lærerne i dagens system bidrar til å sannsynliggjøre at det ligger et stort potensial her.

De mest nærliggende grep er å satse på bedre lærerutdanning og høyere lærerlønn. Flere forhold tilsier imidlertid at slike virkemidler vil kunne gi bare små positive effekter. Produktfunksjonsstudiene har vist at lærerkvali-

teten ikke øker med lengden på utdanningen. En mulig årsak kan være at lærerutdanningen er for dårlig. Det er nærliggende å tro at denne utdanningen lider under systemsvakhetene i skolen for øvrig: Det fins ingen systemer som sikrer at en akkumulerer erfaring om hva som virker og ikke virker i undervisning, og dermed famler lærerutdanningen... Dersom dette er riktig betyr det også at lærerutdanningen ikke vil bli mye bedre før skolen har gjennomført nødvendige reformer. Dvs. de viktige reformene må først komme nedstrøms. Hva med høyere lønn? Høyere lønn vil kunne resultere i flere, men ikke nødvendigvis bedre, søkere til læreryrket. Høyere lønn vil i tillegg bidra til å redusere utskiftningen i læreryrket. Blant annet kan de dårligste lærerne velge å stå lenger i yrket enn de ellers ville gjort.

De viktigste tiltakene vil sannsynligvis være å gjøre undervisningsjobben mer attraktiv for høyt produktiv arbeidskraft, samtidig som skoleledere får sterkere insentiver til å ansette de beste i køen av søkere. Annerledes uttrykt, begrunnelsene for å etablere alternative styringssystemer med vekt på insentiver og økt informasjonstilgang kan være å lage en skole som er attraktiv for potensielt dyktige lærere mer enn å øke anstrengelsen til de lærere som allerede er i systemet, og videre, at slike systemer forsterker skoleledelsens insentiver til å rekruttere de beste blant søkerne.

Dersom gevinstene primært kommer gjennom rekruttering av andre typer lærere betyr det at ting vil ta tid. I mellomtiden vil vi kunne oppleve en lang periode med betydelig strid om små kortsiktige effekter, og at de som forsvaret en omlegging av styringssystemene vil få en hard kamp med forsvarene av status quo.

8. AVSLUTNING

Mye tyder på at det etablerte styringssystemet i utdanningssektoren har utspilt sin rolle. Systemet er i økende grad uegnet fordi utfordringene i økende grad handler om å forbedre elevprestasjonene og produktiviteten. Alternativet er resultatstyring i en eller annen form, men det er ikke enkelt å utforme et resultatstyringssystem som fungerer etter hensikten. Sannsynligvis er det slik at djevelen sitter i detaljene, som betyr at økonomisk teori rekker bare et stykke på vei. Veien fram mot et mer velfungerende system må være preget av prøving og feiling. Dette er nettopp logikken bak resultatstyring:

Gjennom å etablere systemer som gir aktørene frihet til å eksperimentere og insentiver til å gjennomføre eksperimenter, i tillegg til å sikre gode evalueringer, vil vi kunne oppnå en skole som i større grad enn tidligere akkumulerer kunnskap om suksess og tiltrekker seg høyt produktiv arbeidskraft.

Økonomene har viktige roller å spille i denne prosessen. Økonomer har velegnet begrepsapparat for å tenke rundt styring (stat/market), ressursallokering og ikke minst, insentiver. Økonomer har de i tillegg de kvantitative metodene, og har lang trening fra andre områder i å analysere data av den type som skolesystemet nå etter hvert vil generere. Dette betyr at vi beveger oss inn på det som inntil nå i stor grad har vært fagpedagogenes enemerker. Fagpedagogene har lang erfaring i å analysere prosessene i skolen og klasserommene, og ofte med kvalitative, mer enn kvantitative, metoder. Det skulle tilsi at det foreligger en komplementaritet her, slik at begge parter har noe å tjene på samarbeid.

REFERANSER:

Ahlin, Å. (2003): «Does school competition matter? Effects of a large-scale school choice reform on student performance.» Working Paper 2003:3, Department of Economics, Uppsala

Angrist, J.D. og V. Lavy (1999): «Using Maimonides' rule to estimate the effect of class size on scholastic performance.» *Quarterly Journal of Economics*, 114(2), 533-575

Angrist, J.D. og V. Lavy (2002): «The effects of high school matriculation awards: Evidence from randomized trials.» NBER Working Paper no. 9389

Black, S.E. (1999): «Do better schools matter? Parental valuation of elementary education.» *Quarterly Journal of Economics*, 114(2), 577-599

Bonesrønning, H. (2003): «Class size effects on student achievement in Norway: Patterns and explanations.» *Southern Economic Journal*, 69(4), 952-965

Coleman, J.S., E.Q. Campbell, C.J. Hobson, J. McPartland, A.M. Mood, F.D. Weinfeld og R.L. York (1966): «Equality of Educational Opportunity.» U.S. Government Printing Office, Washington D.C.

Correa, H. og G.W. Gruver (1987): «Teacher-student interaction: a game theoretic extension of the economic theory of education.» *Mathematical Social Sciences*, 13, 19-47

Hanushek, E.A. og M.E. Raymond (2003): «Lessons about the design of state accountability systems.» I P.E. Peterson og M.E. West (forf.): «No child left behind? The politics and practice of accountability.» Washington DC: Brookings. 126-151

Hoxby, C.M. (2000): «The effects of class size on student achievement: new evidence from population variation.» *Quarterly Journal of Economics*, 115(3), 1239-1285

Hoxby, C.M. (2003): «School choice and school productivity (or could school choice be a tide that lifts all boats?)» I C.M. Hoxby (forf.): «The Economics of School Choice.» University of Chicago Press, Chicago.

Hsieh, C.T. og M. Urquiola (2003): «When schools compete, how do they compete? An assessment of Chile's nationwide school voucher program.» NBER Working paper no 10008

Krueger, A.B. og D.M. Whitmore (2001): «The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from Project STAR.» *Economic Journal*, 111, 1-28

Lazear, E. (2001): «Educational production.» *Quarterly Journal of Economics*, 116(3), 777-803

Lavy, V. (2002): «Evaluating the effect of teachers' performance incentives on pupils' achievements.» *Journal of Political Economy*. 1286-1317

Rivkin, S.G., E.A. Hanushek og J.F. Kain (2001): «Teachers, schools and academic achievement.» NBER Working Paper no 6691

OECD/UNESCO-uis (2003): «Literacy skills for the World of Tomorrow. Further results from PISA 2000.»