

”Value added-indikatoren: Et nyttig verktøy i kvalitetsvurdering av skolen?”

Kortversjon av SSBs rapport 42/2011

Behov for value added-indikatorer på grunn av økt interesse for skolens resultat kvalitet

De fleste OECD-land har de siste årene lagt økende vekt på skolens resultat kvalitet. I Norge publiserte Utdanningsdirektoratet allerede i 2005 skolebidragsindikatorer på Skoleporten, og årets rapport, **”Value added-indikatoren: Et nyttig verktøy i kvalitetsvurdering av skolen?”**, bygger videre på blant annet analysene som ble gjort den gangen.

Value added-indikatorer er i prinsippet mer informative enn andre resultatmål som uttrykk for skolens bidrag til elevenes læring. Value added-indikatorer skiller seg fra andre indikatorer ved at de også benytter informasjon om elevenes resultater på et tidligere tidspunkt. De korrigerer dermed for viktige forskjeller mellom skoler med hensyn til elevsammensetning som ikke bør fanges opp av indikatorer for skolens bidrag. **Value added-indikatoren er mulig å tolke som skolens bidrag til endring i kunnskaper i tidsrommet mellom de to målepunktene.**

Det er ikke bare forskjeller i elevgruppene som gjør det vanskelig å sammenligne skoler ved å se på gjennomsnittskarakterer på skolenivå. Nedenfor følger en presentasjon av utfordringer og muligheter ved bruk av value added-indikatorer for kvalitetsvurdering av skoler. Presentasjonen bygger i sin helhet på SSBs rapport **”Value added-indikatoren: Et nyttig verktøy i kvalitetsvurdering av skolen?”**

Hva skaper resultatforskjeller mellom skoler?

Gjennomsnittresultatene på skolenivå kan forklares ut fra tre hovedfaktorer:

- **Skolens bidrag til læring**
Formålet med å beregne skolens resultater uavhengig av faktorer skolene ikke kan gjøre noe med, er å komme nærmere en kvantifisering av forskjeller mellom skolene i deres bidrag til elevenes læring. Value added-indikatorer sier imidlertid ingen ting om hvilke skolefaktorer, for eksempel forskjeller i ressursbruk, som betyr mest for resultatforskjellene.
- **Elevenes forutsetninger og bakgrunn**
Familiebakgrunn – slik den måles ved hjelp av tilgjengelig informasjon - kan forklare ca 30 prosent av variasjonen i karakterene mellom enkeltelever for grunnskoler resultatene i Norge. Studier viser også en klar sammenheng mellom enkeltelevers resultater på ulike trinn i utdanningen, men igjen er det ingen naturlov at gode eller dårlige karakterresultater ett år fører til tilsvarende resultater neste år, og det finnes mange eksampler på det motsatte.

Elever med relativt lik bakgrunn går ofte på samme skole. Det finnes flere årsaker til det, men konsekvensen er at mye av karakterforskjellene mellom skolene er påvirket av elevsammensetningen.

- **Tilfeldig variasjon**

Beregningene av skolens resultater baserer seg på skårene fra samtlige av skolens elever, og det kan derfor i første omgang virke merkelig å fokusere på tilfeldig variasjon og usikkerhet. Tanken er imidlertid at det interessante ikke primært er skolens resultat ett bestemt år, men mer hva som kjennetegner skolens resultater over tid. I et slikt langsiktig perspektiv er en viktig kilde til usikkerhet antallet elever ved skolen. Skoler med et lavt elevtall vil typisk få større variasjon i resultatene mellom enkeltår enn større skoler.

Tilsvarende kan prosessene som fører fram til standpunkt karakterer og eksamens karakterer føre til tilfeldig variasjon. Mens standpunkt karakterer er basert på mange vurderinger over tid, er eksamen mer avhengig av elevens dagsform. På den andre siden er det langt fra sikkert standpunkt karakterene settes ut fra samme skala i og mellom skolene. Bruk av opp til tre eksterne sensorer per besvarelse sikrer bruk av felles skala på en helt annen måte for eksamens karakterene.

Hvordan kan vi tallfeste skolens bidrag til elevenes læring?

Hovedutfordringene når vi skal tallfeste skolens bidrag til elevenes læring er følgelig å korrigere for elevsammensetningen på skolen samt ta høyde for tilfeldig variasjon. Det finnes to måter å beregne såkalte skolebidragsindikatorer, men i prinsippet er de to framgangsmåtene like. Forskjellen ligger i mengden elevkjennetegn som tas med i beregningene.

Skolebidragsindikatoren som ble publisert av direktoratet i 2005 var en såkalt tverrsnittindikator. Den kan tolkes som det gjennomsnittresultatet vi ville forventet at skolen skulle hatt, dersom alle elevene hadde en gjennomsnittlig familiebakgrunn. Her kontrollerer vi *indirekte* for elevenes forutsetninger ved å gi hver elev et kunnskapsnivå lik gjennomsnittet av det elevene med samme sosioøkonomisk bakgrunn har.

Årets rapport beregner value added-indikatorer. Denne formen for skolebidragsindikator kontrollerer for elevenes forutsetninger *direkte* ved å ta med elevenes tidligere resultatoppgjør. Value added-indikatoren kan derfor tolkes som skolens bidrag til elevenes læring i løpet av tiden som ligger mellom to måletidspunkt, relativt til andre skoler. Indikatorene sier med andre ord ikke noe om det absolutte bidraget fra skolene til endringene i elevenes kunnskap, men gir et bilde av forskjellene mellom skoler.

Systematisk gjennomføring av nasjonale prøver gir mulighet for å beregne value added-indikator på mellom- og ungdomstrinnet

Etter innføring av nasjonale prøver på 5. og 8. trinn høsten 2007 er det fra og med høsten 2010 mulig å beregne endringer i elevresultater fra to tidspunkt på mellom- og ungdomstrinnet. Det har tidligere bare vært mulig å beregne endringer over tid fra 10. trinn til henholdsvis Vg1, Vg2 og Vg3.

Kan valg av modell føre til ulike resultater for skolenes value added?

SSB har gjennomført analyser for å undersøke hvilken betydning valg av beregningsmodell har for resultatene / value added-indikatoren. Målet var å fange opp så mye som mulig av resultatvariasjonen utenfor skolens kontroll. I tillegg var det et poeng å få en så enkel value added-modell som mulig.

SSB undersøkte tre ulike forhold ved modellene:

- Hvordan kan vi best kontrollere for elevenes tidligere resultater?
- Hvor stor betydning har det å kontrollere for elevenes familiebakgrunn?
- Hvor stor betydning har forutsetningen om skoleeffekter (faste eller tilfeldige effekter) for resultatene?

En viktig implisitt antagelse i alle modellene er at skolene ikke selv strategisk påvirker hvem som svarer og ikke svarer på prøvene. Denne antagelse er spesielt viktig for value added-indikatorer basert på elevenes resultater på de nasjonale prøvene. Hvis enkelte skoler systematisk utelukker de svakeste elevene, vil grunnlaget for indikatorberegningene ikke være oppfylt.

Value added-indikator for mellomtrinnet: nasjonale prøver (NP) fra 5. trinn og 8. trinn

Analysen viser at ved å kontrollere for kjønn samt alle resultatene fra NP 5. trinn hver for seg, får vi en enkel og robust value added-indikator for mellomtrinnet. Analysen viser også at resultatet for value added-indikatoren i ubetydelig grad blir påvirket av ytterligere kontroll for elevenes familiebakgrunn. Det er med andre ord ikke behov for ytterligere kontroll for elevenes familiebakgrunn, noe som forenkler utregningen av denne typen indikator. Sist men ikke minst viste analysen at indikatoren heller ikke ble nevneverdig påvirket av måten skoleeffekten ble spesifisert på.

Value added-indikator for ungdomsskolen: nasjonale prøver (NP) fra 8. trinn og eksamen og standpunkt på 10 trinn

Den største forskjellen i analysene mellom mellomtrinnet og ungdomstrinnet er at det på ungdomstrinnet ikke er ett men to mulige resultatmål; eksamenskarakterer og standpunktkarakterer.

Resultatene av analysen er i stor grad tilsvarende som for mellomtrinnet. Analysen viser at elevenes tidligere resultater forklarer mye av variasjonen i individuelle resultater, og det har mindre betydning hvordan disse resultatene inkluderes. Det har videre liten betydning hvorvidt familiebakgrunnsvariabler inkluderes i value added-indikatoren eller ikke.

Valg av resultatmål har imidlertid betydning for indikatoren når vi kontrollerer for familiebakgrunn. Det viser seg at sammenhengen mellom skolenes karaktergjennomsnitt og value added-indikatoren er vesentlig svakere for standpunktkarakterene enn for skriftlig eksamen. Korrelasjonen mellom value added-indikatoren basert på standpunktkarakterer og eksamenskarakterer er videre kun på 0,2. Tidligere analyser har vist at det er systematiske forskjeller mellom skoler i måten standpunktkarakterene settes på. Det ser ut til at skoler med faglig sterke elever har en strengere karakterpraksis enn skoler med ikke fullt så faglig sterke

elever. SSB advarer derfor mot å beregne value added-indikatorer basert på standpunktkarakterer. Variasjonen i karakterpraksis mellom skoler kan være betydelig og vanskelig å korrigere for.

Usikkerhet ved sammenligning av value added-indikatorer mellom skoler

En del forskjeller i value added-indikatorer mellom skoler er ikke reelle og skyldes tilfeldigheter. Hvordan skal vi vite om forskjeller i value added-indikatorerne mellom skolene er av betydning?

La oss tenke oss to skoler der skolebidragsindikatoren for skole A har høyere verdi enn for skole B. Dersom den nedre grensen i konfidensintervallet for A ligger høyere enn den høyeste for skolen B, da har vi en klar indikasjon på en signifikant forskjell mellom skolenes bidrag til elevenes læring.

Hvis det skulle bli aktuelt å implementere value added-indikatorer i NKVS bør brukerne få en forståelse av usikkerhet knyttet til sammenligning av value added-indikatorer mellom skoler. Det kan enkelt gjøres ved å oppgi to tall X og Y sammen med indikatorene. Hvis forskjellen mellom skolenes value added-indikatorer er mindre enn X, da er forskjellene mellom skolene ubetydelige, dvs ikke statistisk signifikante. Hvis forskjellen mellom skolenes indikatorer er større enn Y, da er forskjellene mellom skolenes bidrag til elevenes læring betydelig. Hvis forskjellen mellom skolenes indikatorer ligger et sted mellom X og Y, da er det usikkert hvorvidt forskjellen er statistisk signifikant eller ikke. Jo nærmere forskjellen er Y, desto større sjanse for at skolene bidrar ulikt.

Value added-indikator for videregående skoler

Rapporten er konsentrert om empiriske analyser av resultatmål for grunnskolen. Kapitlet om videregående skoler og value added-indikatorer inneholder derfor bare hovedfunnene fra Hægeland, Kirkebøen og Raaum (2010).

Utrekningen av value added-indikatorer ble, som vi har sett, noe mer komplisert da vi gikk fra mellomtrinnet til ungdomstrinnet. Beregningen av value added på videregående skoler er ytterligere mer komplisert enn på ungdomsskolen. Dette skyldes flere forhold. En utfordring er muligheten for frafall og manglende fullføring. Det er derfor ikke tilstrekkelig å se på skolens bidrag utelukkende ut fra elevenes karakterutvikling for de videregående skolene. En annen utfordring, gitt det vi nå vet om variasjon mellom skoler i måten standpunktkarakterene settes, er at elevene på videregående skole har få eksamener. Elevene er i tillegg inndelt i flere utdanningsprogram, noe som i en del tilfeller betyr at gruppestørrelsen kan bli vel knapp som grunnlag for meningsfulle statistiske analyser.

Hovedfunnet i analysene er at skolebidragsindikatorer for karakterer og skolebidragsindikator for gjennomstrømning *ikke* er sammenfallende, men ser ut til å fange opp ulike kvaliteter ved skolene. Det er følgelig ikke nok å beregne en skolebidragsindikator for videregående skoler. I tillegg finner SSB forskjeller mellom studieforberedende og yrkesfaglige utdanningsprogram.

Oppsummering

Value added-indikatorene forholder seg til elevenes faglige skåre på to tidspunkter, og kan tolkes som det resultatgjennomsnittet vi forventer at en skole ville hatt, dersom elevene på skolen var gjennomsnittlige i forhold til inkluderte elevkjenntegn. De kan identifisere skoler som bidrar mye eller lite til elevenes læring, men sier ikke noe om hva som må gjøres for å forbedre elevenes læringsutbytte.

De nasjonale prøvene kan benyttes i beregningene av value added-indikatorer for mellomtrinnet og ungdomstrinnet. Når man kontrollerer for elevenes tidligere resultater har det liten betydning hvorvidt vi i tillegg også kontrollerer for elevenes familiebakgrunn. Dette fører til en vesentlig forenkling av arbeidet med å beregne value added-indikatorer.

Måten value added-indikatorene beregnes påvirker i noen grad skolenes relative plassering, og størst utslag får skolene som er spesielt gode eller dårlige i å bidra til elevenes læring. For det store flertallet av skolene betyr value added-indikatorene mindre forskjeller mellom skolene. Det faktum at selve beregningsmåten i noen grad påvirker skolenes skåre på value added, tilsier i seg selv at indikatorene ikke kan brukes alene, men må ses i sammenheng med annen relevant informasjon om skolene.

SSBs anbefalinger i tilfellet value added-indikatorene skal implementeres i NKVS

- Anbefaler å beregne value added-indikatorer der elevutfall kun er justert for kjønn og tidligere resultater
- Informasjon om usikkerhet bør følge indikatoren
- For å redusere den statistiske usikkerheten bør indikatoren beregnes med bakgrunn i data fra flere årskull
- På grunn av mulig variasjon i karakterpraksis bør ikke indikatoren beregnes ut fra standpunktkarakterer
- Meget viktig at gjennomføringen av de nasjonale prøvene foregår mest mulig på samme måte mellom skolene