

Jon H. Fiva og Torbjørn Hægeland

Resultatindikatorer i høyere utdanning

Studiepoeng, grader på normert tid og arbeidsmarkedsutfall

Rapporter I denne serien publiseres statistiske analyser, metode- og modellbeskrivelser fra de enkelte forsknings- og statistikkområder. Også resultater av ulike enkeltundersøkelser publiseres her, oftest med utfyllende kommentarer og analyser.

© Statistisk sentralbyrå, april 2008	Standardtegn i tabeller	Symbol
Ved bruk av materiale fra denne publikasjonen, skal Statistisk sentralbyrå oppgis som kilde.	Tall kan ikke forekomme	.
	Oppgave mangler	..
	Oppgave mangler foreløpig	...
ISBN 978-82-537-7367-4 Trykt versjon	Tall kan ikke offentligjøres	:
ISBN 978-82-537-7368-1Elektronisk versjon	Null	-
ISSN 0806-2056	Mindre enn 0,5 av den brukte enheten	0
	Mindre enn 0,05 av den brukte enheten	0,0
Emne	Foreløpige tall	*
04.02.40	Brudd i den loddrette serien	—
	Brudd i den vannrette serien	
Trykk: Statistisk sentralbyrå	Desimalskilletegn	,

Sammendrag

Formålet med denne rapporten er å drøfte ulike resultatindikatorer for høyere utdanningsinstitusjoner i Norge, som kan beregnes ut fra data som er tilgjengelige på institusjonsnivå. Spørsmålet er hvorvidt disse resultatindikatorene – eller mer presist forskjeller i resultatindikatorer mellom utdanningsinstitusjoner – sier noe om faktiske forskjeller mellom institusjonene i forhold til i hvilken grad de bidrar til studentenes læringsutbytte.

Det er særlig tre forhold som gjør at dette ikke nødvendigvis er tilfelle. For det første vil resultatforskjeller mellom utdanningsinstitusjoner kunne reflektere at de har ulike forutsetninger med hensyn til sammensetningen av studentmassen. Dersom resultater på individnivå er sterkt avhengig av studentenes forutsetninger og bakgrunn, og det er en systematisk ”sortering” av studenter med hensyn til disse kjennetegnene til ulike utdanningsinstitusjoner, vil resultatforskjeller mellom institusjonene delvis reflektere studentsammensetningen. For det andre er en del av resultatmålene som benyttes delvis under institusjonenes egen kontroll. For eksempel er karaktersetting, herunder beslutningen om å gi studenten bestått eller ikke bestått, for en stor del under institusjonenes egen kontroll. Etter innføringen av Kvalitetsreformen i høyere utdanning, har kravet om ekstern sensor ved eksamen blitt svekket. Dette kan medføre at samme prestasjon kan bli bedømt ulikt på tvers av institusjoner og over tid. I den grad det er systematiske forskjeller mellom utdanningsinstitusjonene i karakterpraksis, vil resultatmål som baserer seg på karaktersetting (merk at dette inkluderer indikatorer som studiepoeng pr. student og andel grader fullført på normert tid) ikke gi et korrekt bilde av forskjeller i kunnskapsproduksjonen mellom institusjonene. For det tredje er det slik at høyere utdanningsinstitusjoner i Norge er nokså heterogene med hensyn til hva slags utdanninger og fagkombinasjoner som tilbys. I den grad det er store forskjeller mellom fagområder knyttet til hvordan en gitt prestasjon slår ut i de resultatmålene man benytter, vil indikatorer som ikke tar hensyn til denne heterogeniteten kunne gi et misvisende bilde av resultatforskjeller mellom institusjonene.

Prosjektstøtte: Denne rapporten er en del av StatRes-prosjektet, finansiert av Fornyings- og administrasjonsdepartementet.

Innhold

Sammendrag	3
1. Innledning	7
2. Måling av resultater i høyere utdanning	8
2.1. Hva er de relevante resultatmålene?	8
3. Modell / Rammeverk	11
3.1. Generelt om indikatorer	11
3.2. Resultatindikatorer for høyere utdanningsinstitusjoner	13
4. Data	14
4.1. Resultatmål knyttet direkte til selve utdannelsen	14
4.2. Arbeidsmarkedsutfall	16
4.3. Kontrollvariable	17
5. Resultater	20
5.1. Indikatorer basert på produserte studiepoeng	21
5.2. Indikatorer basert på produksjon av grader på normert tid	29
5.3. Indikatorer basert på arbeidsmarkedsutfall	32
6. Konklusjoner	40
Referanser	43

Figurregister

3. Modell / Rammeverk

3.1. Illustrasjon av produksjonsløp	12
---	----

4. Data

4.1. Gjennomsnittlig avlagte studiepoeng etter antall år siden første gang registrert på høyere utdanning	19
---	----

5. Resultater

5.1. Sammenhengen mellom ukorrigerede og korrigerede studiepoengsindikatorer, basert på samtlige studenter som har meldt seg opp til eksamen i 2005 ¹	24
5.2. Sammenhengen mellom ukorrigerede og korrigerede studiepoengsindikatorer, basert på studenter som har meldt seg opp til eksamen i 2005, og det er tilgjengelig informasjon om vitnemål	26
5.3. Sammenhengen mellom ukorrigerede og korrigerede indikatorer, basert på studenter som avla bachelorgrad i 2005, i en av de ti største bachelorgradene	31
5.4. Sammenhengen mellom studiepoeng og lønnsforskjeller ett år etter avsluttet utdanning	34
5.5. Sammenhengen mellom studiepoeng og arbeidsledighet i løpet av fem første år etter avsluttet utdanning	35
5.6. Ingeniørutdanningen: Sammenhengen mellom studiepoeng og lønnsforskjell fem år etter avsluttet utdanning	39
5.7. Ingeniørutdanningen: Sammenhengen mellom studiepoeng og arbeidsledighet i løpet av fem første år etter avsluttet utdanning	39

Tabellregister

4. Data	
4.1. Avlagte grader inkludert i analysen av arbeidsmarkedsutfall (1995-2003).....	16
4.2. Foreldrenes utdanningsnivå da studenten var 16 år og studiepoengproduksjon i skoleåret 2004/2005	17
4.3. Vitnemål fra videregående skole og studiepoengproduksjon i skoleåret 2004/2005 (N=50893).....	18
4.4. Vitnemål fra videregående skole og institusjonstilhørighet i skoleåret 2004/2005	18
4.5. Fagfelt, nivå og studiepoengproduksjon i skoleåret 2004/2005	19
4.6. Gjennomsnittlig avlagte studiepoeng etter antall år siden første gang registrert på høyere utdanning.....	20
5. Resultater	
5.1. Forklaringskraft, kontrollvariable, basert på studenter som har meldt seg opp til eksamen i skoleåret 2004/2005 (N=168 883)	22
5.2. Forklaringskraft, kontrollvariable, basert på studenter som har meldt seg opp til eksamen i skoleåret 2004/2005 og med tilgjengelig informasjon om vitnemål fra videregående skole (N=50101).....	22
5.3. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, (N=168 827)	23
5.4. Korrelasjonsmatrise parametervektor presentert i 5.3	24
5.5. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen. Betrakter bare studenter hvor informasjon om vitnemål fra videregående skole er tilgjengelig (N= 50068)	25
5.6. Korrelasjonsmatrise til parametervektor presentert i 5.5.....	26
5.7. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, lavere grad (N=134 546)	27
5.8. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, lavere grad. Betrakter bare studenter hvor informasjon om vitnemål fra videregående skole er tilgjengelig. (N=42502).....	28
5.9. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, høyere grad (N=34 281).....	29
5.10. Grader produsert på normert tid, per studenter som avla bachelor grad studieåret 2004/2005, fordelt på de ti største bachelor-utdanninger, målt i antall produserte grader	30
5.11. Grader produsert på normert tid, per studenter som avla bachelor grad i 2005	30
5.12. Grader produsert på normert tid, per studenter som fullførte sykepleierutdanningen i 2005. 3077 observasjoner. Institusjoner med mindre enn 20 avlagte grader er utelatt fra analysen	31
5.13. Grader produsert på normert tid, per studenter som fullførte allmennlærer-utdanningen i 2005. Institusjoner med mindre enn 20 avlagte grader er utelatt fra analysen	32
5.14. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (HsId=5)	33
5.15. Korrelasjonsmatrise til parametervektor presentert i 5.14.....	33
5.16. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (HsId=5), kontrollert for næring og fylke	36
5.17. Korrelasjonsmatrise til parametervektor presentert i 5.16.....	36
5.18. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (HsId=5), sykepleierutdanningen.....	37
5.19. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (HsId=5), allmennlærerutdanningen	38
5.20. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (HsId=5), ingeniør-utdanningen.....	38

1. Innledning*

I de senere årene har det både i Norge og internasjonalt vært en økende oppmerksomhet knyttet til resultater og ressursbruk innenfor offentlig sektor. Flere land har utviklet kvalitetsvurderingssystemer som presenterer og sammenlikner informasjon om resultater og ressursbruk i institusjoner knyttet til offentlig sektor, særlig dem som leverer individrettede tjenester ("public sector accountability systems"). Generelt er det en utfordring å lage adekvate resultatindikatorer for virksomheter hvor produksjonen ikke fullt ut omsettes på et vanlig marked og derigjennom gis en verdsetting gjennom oppnådd pris. Ut fra virksomhetens formål er det heller ikke umiddelbart klart hvordan man kan lage operasjonaliserbare resultatindikatorer, dvs. en indikator som reflekterer hvordan virksomheten oppfyller de formålene den er opprettet for. Produktet, definert ut fra virksomhetens formål, har ikke alltid noe empirisk motstykke som det er enkelt å måle eller operasjonalisere.

Formålet med denne rapporten er å drøfte ulike resultatindikatorer for høyere utdanningsinstitusjoner i Norge, som kan beregnes ut fra data som er tilgjengelige på institusjonsnivå. Spørsmålet er hvorvidt disse resultatindikatorene – eller mer presist forskjeller i resultatindikatorer mellom utdanningsinstitusjoner – sier noe om faktiske forskjeller mellom institusjonene i forhold til i hvilken grad de bidrar til studentenes læringsutbytte.

Det er særlig tre forhold som gjør at dette ikke nødvendigvis er tilfelle. For det første vil resultatforskjeller mellom utdanningsinstitusjoner kunne reflektere at de har ulike forutsetninger med hensyn til sammensetningen av studentmassen. Dersom resultater på individnivå er sterkt avhengig av studentenes forutsetninger og bakgrunn, og det er en systematisk "sortering" av studenter med hensyn til disse kjennetegnene til ulike utdanningsinstitusjoner, vil resultatforskjeller mellom institusjonene delvis reflektere studentsammensetningen. For det andre er en del av resultatmålene som benyttes delvis under institusjonenes egen kontroll. For eksempel er karaktersetning, herunder beslutningen om å gi studenten bestått eller ikke bestått, for en stor del under institusjonenes egen kontroll. Etter innføringen av Kvalitetsreformen i høyere utdanning, har kravet om ekstern sensor ved eksamen blitt svekket. Dette kan medføre at samme prestasjon kan bli bedømt ulikt på tvers av institusjoner og over tid. I den grad det er systematiske forskjeller mellom utdanningsinstitusjonene i karakterpraksis, vil resultatmål som baserer seg på karaktersetning (merk at dette inkluderer indikatorer som studiepoeng pr. student og andel grader fullført på normert tid) ikke gi et korrekt bilde av forskjeller i kunnskapsproduksjonen mellom institusjonene. For det tredje er det slik at høyere utdanningsinstitusjoner i Norge er nokså heterogene med hensyn til hva slags utdanninger og fagkombinasjoner som tilbys. I den grad det er store forskjeller mellom fagområder knyttet til hvordan en gitt prestasjon slår ut i de resultatmålene man benytter, vil indikatorer som ikke tar hensyn til denne heterogeniteten kunne gi et misvisende bilde av resultatforskjeller mellom institusjonene.

Det finnes også andre måter å studere resultatforskjeller i offentlig tjenesteproduksjon. Kittelsen og Førstund (2001) gir en oversikt over metoder og empiriske resultater fra analyser som benytter den såkalte Data Envelopment Analysis (DEA)-metoden, som søker å identifisere effektivitetsforskjeller mellom ulike enheter relativt til en "beste praksis". Edvardsen og Førstund (2001) benytter DEA-metoden i en studie av norske høyere utdanningsinstitusjoner.

Vi finner at det er relativt stor grad av samsvar mellom ukorrigerede indikatorer og indikatorer som korrigerer for forskjeller i studentsammensetning (målt ved foreldres utdanning og/eller studentenes karaktergjennomsnitt fra videregående skole), når vi ser på resultatindikatorer basert på studiepoengsproduksjon og

* Takk til Sissel Ferstad og Alice Steinkellner for kommentarer.

fullføring av grader på normert tid. Resultatindikatorer som korrigerer for dette, gir altså stort sett det samme bildet som de ukorrigerede indikatorene, selv om korreksjonen er betydelig for enkelte institusjoner. Det er imidlertid systematiske forskjeller mellom fagfelt når det gjelder disse resultatmålene, og siden fagsammensetningen varierer betydelig på tvers av utdanningsinstitusjonene, er det viktig å korrigere for dette når man sammenligner institusjonenes resultater.

Dette er interessante resultat, men spørsmålet er hvorvidt disse resultatmålene sier noe om forskjeller mellom institusjonene når det gjelder å oppfylle det "egentlige" formålet med virksomheten. Dette spørsmålet kan deles i to: Gir det aktuelle resultatmålet en enhetlig måling av prestasjoner på tvers av institusjoner, og er disse prestasjonene det som faktisk verdsettes med utdanningen? Svaret på begge spørsmål må være ja, for at vi skal kunne si at indikatoren oppfyller dette "kravet".

Karakterer (som ligger bak begge våre resultatmål gjennom distinksjonen mellom bestått og ikke bestått) settes i stor grad av institusjonene selv. Det er i liten grad krav om ekstern sensur, og få mekanismer som sørger for en enhetlig karaktersetting. Det finnes klare indikasjoner på at kriteriene for karaktersetting som benyttes av ulike institusjoner for å vurdere studentenes prestasjoner tilpasses kvaliteten på studentmassen, se for eksempel Møen og Tjelta (2005).

Vi finner at det er betydelige forskjeller i arbeidsmarkedsutfall mellom studenter fra ulike utdanningsinstitusjoner. Hvis man godtar arbeidsmarkedet som en evalueringsinstans med hensyn til utdanningskvalitet, indikerer dette at det er betydelige forskjeller mellom utdanningsinstitusjoner. Det bør bemerkes at disse beregningene er av en viss "eksplorativ karakter". Indikatorene basert på arbeidsmarkedsutfall bare er svakt korrelert med indikatorene basert på studiepoengsproduksjon og grader fullført på normert tid. I mange tilfeller er det til og med en negativ korrelasjon. Hvis begge typer indikatorer delvis fanger opp "utdanningskvalitet", skulle man forvente at de var positivt korrelerte. Fraværet av positiv korrelasjon, er en indikasjon på at resultatindikatorer som baserer seg på data som har sitt utspring i karakterer satt av de enkelte utdanningsinstitusjonene er problematiske som "kvalitetsindikatorer". De inneholder mye informasjon som er relevant i mange sammenhenger, men man bør være forsiktig med å tolke dem som indikatorer som kvalitetsforskjeller mellom utdanningsinstitusjoner.

Resten av rapporten er disponert som følger. I neste kapittel gir vi en kort drøfting av enkelte utfordringer knyttet til måling av resultater på institusjonsnivå innenfor høyere utdanning. Kapittel 3 drøfter vårt empiriske rammeverk, mens datakilder og variable er beskrevet i kapittel 4. Resultater fra analysene presenteres i kapittel 5, mens siste kapittel oppsummerer og konkluderer.

2. Måling av resultater i høyere utdanning

2.1. Hva er de relevante resultatmålene?

Høyere utdanning har mange formål, hvorav flere, eksempelvis kulturbevaring og dannelse, ikke kan tillegges direkte økonomisk betydning. Ikke desto mindre er det bred enighet om at utdanning er en viktig kilde til økonomisk vekst. I et rent økonomisk perspektiv kan utdanning betraktes som en investering i kunnskapsskapital; den enkelte og samfunnet går i utdanningsfasen glipp av arbeidsinntekt, men dette tapet kan mer enn oppveies gjennom økt produktivitet i de yrkesaktive årene. Dette rent økonomiske perspektivet er mer relevant for høyere utdanning, dvs. utdanning utover videregående skole, enn for utdanning innenfor grunnskole og videregående skole. I den grad analyser av avkastning av kunnskapsinvesteringer benyttes som grunnlag for evalueringer av utdanningsinstitusjoner eller sektoren som helhet, taler dette for å ha økning i humankapital eller kunnskapsnivå som det "ideelle" produksjonsmål.

Hvordan måles produksjonen innenfor høyere utdanning i dag? Til tross for at det finnes rikelig med informasjon om bl.a. studenttall og offentlige utgifter innenfor ulike typer utdanning, er sektorens bidrag til den samlede økonomiske verdiskapingen, målt ved BNP, noe mangelfullt beskrevet. Tradisjonelt har man i inasjonalregnskapet ikke anslått produksjonen i noen deler av utdanningssektoren på grunnlag av informasjon om den produktivitets- og kompetansehevingen som utdanningen medfører. I stedet avleder man en indikator for produksjonen fra tall for de direkte produksjonskostnadene. Dette gjelder også produksjonen i de andre offentlige forvaltningssektorene. Denne metoden er selvsagt erkjent som svært mangelfull, og har kun vært benyttet i fravær av bedre alternativer i en situasjon der produksjonen ikke omsettes og registreres direkte. For høyere utdanning er en slik fremgangsmåte særlig misvisende og undervurderer ressursbruken betydelig. Det skyldes blant annet at studentenes tidsbruk ikke er inkludert i kostnadstallene. Den alternative lønnsinntekten som kunne vært høstet ved å arbeide i stedet for å studere, er alene større enn de samlede kostnadene som rapporteres i NR, se Ervik, Holmøy og Hægeland (2002). Det er heller ingenting som sikrer at produksjon målt ved samlet ressursbruk gir noe godt bilde på den faktiske verdiskapingen som skjer innenfor høyere utdanning.

I den senere tid har man implementert nye produksjonsmål i NR som i større grad reflekterer aktiviteten i blant annet høyere utdanning. Her benytter man for eksempel volumindekser basert på antall studenter og/eller antall produserte studiepoeng, jf Dam og Sørensen (2006).

Økningen i humankapital eller kunnskapsnivå som følge av utdanning, er uansett vanskelig å observere eller operasjonalisere direkte. Fullskala målinger av dette, med omfattende tester av kunnskapsnivå ved start og avslutning av en utdanning vil være svært kostbart, selv om man fant frem til relevante tester og målemetoder. Vi er derfor nødt til å bruke resultatmål og -indikatorer som i større eller mindre grad reflekterer den kunnskapsøkningen som skjer gjennom utdanningsløpet. For den enkelte student kan man tenke seg flere måter å måle resultatet av utdanningen.

Ett sett av resultatindikatorer er direkte knyttet til selve utdanningen, dvs. at de måles i direkte forbindelse med utdanningen.

- *Antall kurspoeng/vektall, oppnådde grader:* Et slikt produksjonsmål sier noe om at man har vært gjennom et definert pensum og at man gjennom eksamen eller annen evaluering har dokumentert kunnskapsnivå over et visst minstenivå (dvs. krav til bestått) innenfor det aktuelle feltet.
- *Karakterer:* I forhold til punktet ovenfor innebærer dette en finere gradering av oppnådd kunnskapsnivå etter et gjennomført studium.
- *Gjennomføringstid:* Dette målet sier ikke så mye om faktisk kunnskapsnivå ved avsluttet utdanning, men er et mål på "hvor raskt" studenten har nådd det nivået som kreves for å bestå eksamen eller få den oppnådde karakteren.

Bruk av disse målene til å sammenligne utdanningsinstitusjoner, fordrer at det er en stor grad av samsvar mellom institusjonene i hva som ligger i de ulike begrepene. Jo nærmere man i praksis er til å ha "nasjonale standarder" for kursinnhold, krav til pensum og resultater for at et studium skal føre fram til en bestemt grad og ikke minst kravene til bestått eksamen og bruk av karakterskalaen ellers, jo mer hensiktsmessig blir det å bruke slike mål til sammenligning mellom institusjoner. På den annen side, jo nærmere man i praksis er at institusjonene har sine egne kriterier, jo vanskeligere blir det å sammenligne dem i forhold til indikatorer *de selv har kontroll over*.

Det er derfor nærliggende å tenke på resultatmål som ikke settes av institusjonene selv, men av en ekstern "evaluator". Selv om ikke utdanningen selges direkte på et marked, vil utdanningen implisitt bli evaluert i arbeidsmarkedet og verdsettes gjennom sysselsetting og lønn, og eventuelt av andre utdanningsinstitusjoner ved

opptak til videre studier. Et sett med mulige resultatindikatorer knytter seg derfor til konsekvenser av utdanningen, dvs. hva som skjer etter at utdanningen er avsluttet. I forhold til de mer direkte resultatmålene ovenfor sier de i prinsippet mer om resultater i forhold til det opprinnelige målet med utdanningen (ikke en eksamen i seg selv, men å få kunnskaper og ferdigheter som er etterspurt og kan anvendes i verdiskapingen). Samtidig vil disse målene i større grad være påvirket av andre forhold som ikke kan tilskrives den utdanningen man har tatt, og dermed reflektere også andre faktorer enn selve utdanningen:

- *Videre utdanningskarrierer*: En rekke utdanninger, også innenfor høyere utdanning, kan påbygges med andre utdanninger. Overgangsrater til og resultater i videre utdanning, er mulige resultatindikatorer.
- *Sysselsetting*: Hvorvidt man får en jobb etter man er ferdig utdannet, er et direkte mål på arbeidsmarkedets vurdering av utdanningens kvalitet og relevans.
- *Lønn/livsinntekt*: I humankapitalteorien, hvor utdanning betraktes som en investering for den enkelte, er bruttoavkastningen av investeringen i utdanning den økningen i livsinntekt som følger av utdanningen. Den økte inntekten skyldes innenfor denne teorien at utdanningen har gjort den enkelte mer produktiv, slik at arbeidsgiver er villig til å betale en høyere lønn.

2.1.1. Resultatforskjeller mellom utdanningsinstitusjoner

Et viktig aspekt ved vurdering av resultatindikatorer for (høyere) utdanning, er spørsmålet om kausalitet. Den grunnleggende problemstillingen kan belyses ved et kjent eksempel: Kan inntektsforskjellene mellom personer med ulik utdanningslengde tilskrives utdanningen? Hvis det er faktorer som vi ikke kan observere (f.eks. evner) som påvirker både utdanningsvalg og lønn, kan man ikke uten videre si at inntektsforskjellene reflekterer en avkastning av utdanning. I vårt tilfelle, hvor vi skal sammenligne ulike utdanningsinstitusjoner, er ikke utfordringen først og fremst knyttet til valg av utdanningslengde, men valg av utdanningsinstitusjon. Sett at vi nå sammenligner to høyskoler som tilbyr de samme studieprogrammene. Vi antar at de er sammenlignbare hva gjelder innhold, og at karakterpraksis o.l. er lik på de to skolene, slik at resultater er fullt ut sammenlignbare. Den ene høyskolen skårer mye høyere enn den andre. I hvilken grad kan vi si at dette skyldes selve skolen? Vi kan aldri forvente å komme helt i mål på dette punktet, men det er likevel mulig å korrigere for en rekke forhold som påvirker resultatforskjeller mellom utdanningsinstitusjoner, men som ikke har noe med institusjonen å gjøre.

Det kan være nyttig å tenke at et gjennomsnittresultat på institusjonsnivå kan tilskrives tre typer faktorer:

- Studentenes forutsetninger
- Tilfeldig variasjon
- Institusjonens bidrag til resultatet

De to første faktorene er utenfor institusjonens kontroll, mens den siste representerer det den kan gjøre noe med, selv om handlingsrommet vil variere med rammene som settes av sentrale myndigheter. Skolene som skårer høyest, gir ikke nødvendigvis det beste bidraget til resultatet. Det kan skyldes fordelaktig sammensetning av studentmassen eller tilfeldigheter. På samme måte er det langt fra opplagt at skoler med svake resultater gir studentene et dårlig læringsutbytte.

Institusjonens bidrag til læringsutbytte, kan tilskrives mange forhold. Kvalitet på bygninger og utstyr, motivasjon og kunnskapsnivå hos undervisningspersonale, arbeidsmåter, organisering av undervisning/veiledning, samarbeidsklima mellom ansatte og kvalitet på ledelse, kan være med på å påvirke studentenes læringsutbytte. Hvorvidt resultatforskjeller mellom utdanningsinstitusjoner faktisk reflekterer forskjeller i hvordan de drives, avhenger kritisk av hvor viktige studentenes forutsetninger - og tilfeldig variasjon - er for resultatene.

2.1.2. Studentenes forutsetninger og bakgrunn

Mange undersøkelser fra Norge og andre land viser en sterk samvariasjon mellom elevenes familiebakgrunn og deres skoleprestasjoner. Dette gjelder på alle nivåer. Hægeland et al. (2004) undersøker sammenhengene mellom familiekjennetegn og skoleprestasjoner, målt ved karakterer for avgangskullene i 10.klasse i norske skoler i 2002 og 2003. Familiebakgrunn, målt ved et svært rikt sett av registerbaserte variable som reflekterer foreldrenes utdanning, inntekt, formue, arbeidsmarkedstilknytning, trygdeforhold, sivilstand, familiestørrelse osv., kan "forklare" omtrent en tredel av forskjellene i karakterer mellom enkeltelever. Det er imidlertid viktig å presisere at det ikke er snakk om et en-til-en forhold mellom familiebakgrunn og hva elever oppnår på skolen. Familiebakgrunn er den klart viktigste faktoren for å forklare skoleprestasjoner, men det er stort rom for andre faktorer, siden tross alt 70 prosent av variasjonen i karakterer skyldes forhold utenfor familien. Selv om barn av foreldre med høy utdanning og god økonomi – i gjennomsnitt - oppnår bedre resultater enn klassekamerater som har foreldre med kort skolegang og lav inntekt, finnes det mange *enkeltilfeller* hvor forholdet er motsatt. Elever med ulik familiebakgrunn fordeler seg ikke jevnt utover skoler. Det er tvert imot en klar tendens til at elever med relativt lik bakgrunn "klumper seg sammen" på samme skole, ved at like familier i stor grad velger tilsvarende boligområder og barn/ungdom typisk går på nabolagsskolen(e). Dermed vil karakterforskjeller mellom grunnskoler i betydelig grad være påvirket av elevsammensetningen.

Det samme fenomenet kan gjøre seg gjeldende også innenfor høyere utdanning. "Problemet" er trolig noe mindre rent empirisk sett, siden det er en viss seleksjon inn i høyere utdanning totalt sett basert på tidligere skoleprestasjoner. Det er likevel klart at elever med ulike forutsetninger ikke fordeler seg tilfeldig på høyere utdanningsinstitusjoner i Norge. En klar indikasjon på dette får man ved å se på opptakskravene til ulike studier. Her er det relativt store variasjoner, også mellom samme type studier ved ulike institusjoner. Det er all grunn til å tro at denne sorteringen av studenter påvirker gjennomsnittresultatene på institusjonsnivå. Dette tilsier at man bør kontrollere for sammensetningen av studentmassen når man sammenligner resultatene mellom utdanningsinstitusjonene.

2.1.3. Tilfeldig variasjon

Resultatforskjeller mellom utdanningsinstitusjoner vil også delvis skyldes tilfeldig variasjon. Denne usikkerheten skyldes både tilfeldigheter bak enkeltstudenters prestasjoner, og særskilte og 'uvanlige' hendelser på institusjonsnivå.

En viktig kilde til usikkerhet er knyttet til antall studenter ved institusjonen. Jo færre som danner grunnlaget for å regne ut et gjennomsnittresultat, desto større vil variasjonen i resultatet typisk være. Den tilfeldige variasjonen vil bidra sterkt til at toppen og bunnen av resultatfordelingen domineres av små institusjoner. Det viktig å ta hensyn til denne form for usikkerhet ved sammenlikninger av resultater mellom institusjoner. For de minste spiller tilfeldig variasjon en langt større rolle, og det er ofte umulig å legge vekt på gjennomsnittresultater for ett enkelt år. Det bør derfor ligge et minimum av antall observasjoner (studenter) bak en indikator.

3. Modell / Rammeverk

3.1. Generelt om indikatorer

Hva er en indikator? Som navnet sier, er det en størrelse som indikerer noe om et forhold man ønsker å belyse. En indikator måler ikke direkte det man egentlig er interessert i. Dette kan være vanskelig målbart, ved at det ikke finnes noe direkte empirisk motstykke til det man ser på. Det kan også være måleproblemer knyttet til at det man ønsker å måle først blir målbart på et senere tidspunkt enn når man ønsker informasjon. Endelig kan det være slik at aktiviteten eller resultatet til en

virksomhet er så mangfoldig at det ikke er meningsfylt å uttrykke det i en størrelse, men at ulike indikatorer til sammen gir et godt bilde.

Figur 3.1. Illustrasjon av produksjonsløp



Indikatorer kan beregnes på flere steder i produksjonsprosessen, fra ressursbruk og til resultater/effekter av produksjonen. Man kan ha målepunkter på flere steder. Jo lenger man kommer i kjeden fra ressursbruk til resultat, desto nærmere kommer man det som faktisk er virksomhetens formål. I en viss forstand vil man derfor ønske å måle resultater så langt til høyre i kjeden som mulig.

Samtidig er måling vanskeligere jo lenger ut i produksjonsforløpet man kommer. For det første blir det fenomenet man er ute etter, vanskeligere å måle. Samtidig blir det flere og flere forhold som ikke virksomheten selv har kontroll over, som spiller inn på resultatet. For at en indikator skal gi et uttrykk for resultatet av aktiviteten ved virksomheten, må man kontrollere for slike forhold, eller godtgjøre at slike forhold ikke påvirker den aktuelle indikatoren i vesentlig grad. Dette krever i mange tilfeller et omfattende datagrunnlag, og ulike sammenhenger må ofte estimeres. Dette fører til at indikatorene blir beheftet med en viss statistisk usikkerhet (som riktignok ofte kan kvantifiseres). I en del tilfeller eksisterer heller ikke det nødvendige datagrunnlaget. Kausale tolkninger av indikatorer vil ofte hvile på sterke og ofte urealistiske forutsetninger.

For å velge ut hva slags aktivitets- og resultatmål som skal beregnes, bør man ta utgangspunkt i hva som i vid forstand er de definerte formålene med den aktuelle virksomheten. Dette bør være retningsgivende for valget av indikatorer, både når det gjelder ressursbruk, aktivitet, produksjon og resultat.

Indikatorer bør samlet sett oppfylle følgende kriterier:

1: De må dekke hele institusjonens virksomhet (alle dens aktiviteter)

Mange institusjoner har mange typer aktiviteter. Disse kan være helt atskilte eller sterkt komplementære. For eksempel driver høyere utdanningsinstitusjoner med både undervisning og forskning. Å måle bare deler av virksomheten kan gi et feilaktig bilde, både når man sammenligner enheter på tvers og når man gjør sammenligninger over tid, fordi vektleggingen av de ulike aktivitetene kan endre seg over tid. Hvis indikatorene i tillegg benyttes som styringsinformasjon, dvs. at det eksplisitt eller implisitt er knyttet insentiver til dem, er det spesielt viktig at hele virksomheten måles, for å unngå uønskede vridninger av aktiviteten i retning av det som måles og belønnes.

2: De bør muliggjøre sammenligninger på tvers

Indikatorene bør være sammenlignbare på tvers, slik at man på en meningsfull måte kan sammenligne ulike enheter på samme nivå.

3: De bør muliggjøre sammenligninger over tid

Meningsfulle sammenligninger over tid fordrer at man kontrollerer for at teknologi og produkter endrer seg over tid.

4: De bør (ideelt sett) kunne aggregeres.

Samlemål er interessante i den grad de faktisk gir et meningsfullt uttrykk for den samlede virksomheten. Imidlertid er det ofte slik at det er vanskelig å veie sammen f.eks. to ulike aktiviteter ved en institusjon. Da vil det kunne være bedre ikke å aggregere.

3.2. Resultatindikatorer for høyere utdanningsinstitusjoner

Basert på drøftingen ovenfor, er det særlig tre spørsmål som må vurderes når man vurderer resultatindikatorer for høyere utdanning:

- I hvor stor grad reflekterer resultatindikatorerne forhold som institusjonen selv ikke kan påvirke?
- I hvor stor grad *bestemmes* resultatindikatorerne av institusjonene selv?
- I hvor stor grad reflekterer indikatorer som er beregnet på kort sikt, de mer langsiktige utfallene?

Når det gjelder det første spørsmålet, er en potensielt viktig faktor bak resultatforskjeller mellom ulike utdanningsinstitusjoner at de kan ha ulikt studentgrunnlag. Følgelig kan ikke variasjoner i enkle resultatindikatorer uten videre tilskrives utdanningsinstitusjonenes innsats. En av de største utfordringene for utarbeidelse av resultatindikatorer er å unngå at en utdanningsinstitusjon som følge av gunstig studentsammensetning kommer gunstig ut. Den ideelle resultatindikator kontrollerer derfor fullt ut for studentgrunnlag til institusjonene. Dette er ikke alltid mulig, fordi systematiske forskjeller i studentgrunnlaget ikke alltid er knyttet til lett observerbare kjennetegn ved studentene.

Hvor viktig det er å ta hensyn til ulik studentsammensetning når man beregner resultatindikatorer for norske utdanningsinstitusjoner, er et åpent spørsmål som kun kan avdekkes gjennom grundige empiriske undersøkelser. Til dette formål er registerdata fra SSB velegnet. Registerdata i SSB inneholder detaljert informasjon om studieforløp, familiebakgrunn og tidligere akademiske prestasjoner (vitnemål fra videregående skole). Dette gjør det mulig å estimere en regresjonsmodell av typen:

$$Y_{ij} = \alpha + \beta x_i + \eta S_{ij} + u_{ij} \quad (1)$$

Der Y_{ij} er et resultatmål for student i ved institusjon j . I denne analysen er Y basert på (i) produksjon av studiepoeng, (ii) fullføring av grader på normert tid og (iii) arbeidsmarkedsutfall. X_i er en vektor av bakgrunnsvariable for student i og S_{ij} er et sett av dummyvariable som er lik 1 dersom person i er student ved institusjon j og null ellers. Hovedinteressen knytter seg til vektoren η . η -vektoren fanger opp institusjonsspesifikke effekter, altså effekter som alle studentene ved institusjon j har felles. Disse kan tolkes som institusjonens bidrag til resultatet. For å oppnå forventningsrette estimat av η må restleddet i (1) være uavhengig og identisk fordelt. Denne antagelsen er problematisk dersom det er kjennetegn ved individene som ikke er inkludert i X -vektoren som korrelerer både med resultatmål og faste institusjonseffekter ($E(u_{ij} | X_i, S_{ij}) \neq 0$). Man står da ovenfor et utelatt variabel problem. Et utelatt variabelproblem kan i denne analysen oppstå dersom man utelater kjennetegn ved studentene som er korrelert med resultatet samtidig som studentene ikke fordeler seg tilfeldig på institusjonene. Man kan da komme til feilaktig å tilskrive effekter til institusjonene som egentlig skyldes andre faktorer.

Hægeland et al. (2005) beregner skolebidragsindikatorer for grunnskolen. Rammeverket og analyseopplegget vi benytter her har mange fellestrekk med beregningsopplegget for skolebidragsindikatorer. En tilsvarende analyse for høyere utdanning er mer krevende ettersom aktiviteten på høyskole- og universitetsnivå er betydelig mer heterogen enn på grunnskolen. For det første kan det være problematisk å behandle studiepoeng eller grader produsert på ulike fagfelt som ekvivalente. Dersom man estimerer ligning (1) antar man implisitt at utdanning er den samme på alle nivå og på alle typer utdanning. Denne heterogeniteten kan reduseres ved å inkludere høyresidevariable som tar ut effektene av fag(felt) og nivå. For det andre, er det grunn til å tro at det er store forskjeller på tvers av studentene i deres ambisjoner og intensjoner, og det er utfordrende å unngå at institusjoner med et høyt innslag av deltidsstudenter, ikke kommer dårlig ut om resultater måles som studiepoengproduksjon eller gradert produsert på normert tid.

Som diskutert i avsnitt 2.1, er et problem med resultatmål som studiepoeng og fullførte grader at de til en viss grad fastlegges av institusjonene selv. Det er rimelig å anta at studentene i noen grad vurderes relativt til hverandre innad i institusjonene heller enn i forhold til en objektiv skala. Dersom kravene det stilles til studentene varierer mye på tvers av institusjonene, kan faktiske kvalitetsforskjeller kamoufleres i resultatindikatorer basert på studiepoeng og fullførte grader. En nærliggende utvidelse er derfor å se på resultatmål som ikke bestemmes av institusjonene selv.

Å bruke arbeidsmarkedsutfall som basis for sammenligning av utdanningsinstitusjoner, innebærer implisitt at man rendyrker en dimensjon av kvalitetsaspektet ved en utdanning, nemlig hvordan den verdsettes i arbeidsmarkedet gjennom sysselsetting og lønn. Dette innebærer at flere relevante aspekter ved utdanningen ikke nødvendigvis fanges opp av slike resultatindikatorer. Dette er en imidlertid en innvending som vil gjelde alle typer indikatorer for institusjoner hvor aktivitetene har flere formål, og hvor det ikke er opplagt hvordan man skal måle og eventuelt veie sammen de ulike formålene. Dette innebærer at det ofte ikke finnes en enkeltindikator som på en udiskutabel måte fanger opp "resultatet", men at det kan være nødvendig med flere indikatorer, og at de indikatorene som eventuelt kan beregnes må vurderes i lys av at det finnes andre relevante aspekter knyttet til resultater og kvalitet som ikke så lett lar seg måle og beregne. Heller ikke disse resultatmålene vil perfekt reflektere i hvilken grad utdanningsinstitusjonene oppfyller de overordnede formålene med virksomheten, men de kan supplere det bildet som gis av resultatmål som institusjonene selv kontrollerer.

4. Data

I dette kapitlet gir vi en kort beskrivelse av datakilder og variable som vi benytter i analysen. Først diskuteres operasjonaliseringen av ulike resultatmål knyttet direkte til selve utdannelsen, deretter ser vi på ulike utfallsmål i arbeidsmarkedet. Til slutt beskriver vi ulike kontrollvariable som benyttes i analysene.

4.1. Resultatmål knyttet direkte til selve utdannelsen

Hovedkilden til dataene er Nasjonal utdanningsdatabase (NUDB), som er en individbasert database i Statistisk sentralbyrå. Denne forløpsdatabasen samler alt av SSBs individbaserte utdanningsstatistikk, fra avsluttet grunnskole til doktorgradsstudier tilbake til 1970, i én database. Hovedkilden til utdanningsopplysningene er utdanningsinstitusjonenes egne administrative systemer.

4.1.1. Produksjon av studiepoeng per student

Et naturlig produksjons- eller resultatmål for en utdanningsinstitusjon er avlagte studiepoeng, eventuelt beregnet pr. student. Indikatorer basert på studiepoeng tar utgangspunkt i resultatmål som er relevant for de individene som tjenesten er rettet mot (studentene), og sier i utgangspunktet noe om omfang (antall studiepoeng) av læring av en viss kvalitet (pga. krav om bestått) som har funnet sted. Som vi har drøftet over, fordrer bruk av studiepoeng pr. student som resultatmål at kravene for å få bestått er like på tvers av utdanningsinstitusjoner og over tid. Uten sentrale eksamener og med begrenset bruk av ekstern sensur er det en fare for at dette kravet ikke er oppfylt.

Hvordan skal man avgrense studentpopulasjonen? En utfordring når man skal sammenligne utdanningsinstitusjoner, er at mange studenter er deltidsstudenter. Det er stor variasjon mellom studenter når det gjelder hvor stor deltidsprosent de har, og andelen deltidsstudenter vil kunne variere mellom utdanningsinstitusjoner. I noen grad vil man kunne kontrollere for dette ved å kontrollere for forskjeller i

fagsammensetning mellom institusjonene. Dette vil i noen grad redusere ”problemet” knyttet til deltidsstudenter, siden det trolig er systematisk variasjon i omfanget av deltid mellom fag. Noen studier gir lite rom for annet enn heltidsløp, på andre studier er det store muligheter for tilpasning av studieintensitet.

Dersom man vil benytte studiepoeng per student som en resultatindikator innenfor høyere utdanning, vil det uansett være ønskelig å skille mellom fulltidsstudenter og deltidsstudenter. Den beste resultatindikatoren basert på studiepoeng ville være definert som studiepoeng produsert relativt til *planlagte* studiepoeng, dvs. hvor mange studiepoeng som blir realisert i forhold til hvor mye tid studenten brukte eller planla å bruke på studier i det aktuelle tidsrommet. Denne informasjonen er imidlertid ikke tilgjengelig. Gitt det datagrunnlaget som er tilgjengelig på det nåværende tidspunkt, kan man benytte tre tilnæringer: (i) inkludere alle registrerte studenter, (ii) inkludere alle studenter som har meldt seg opp til eksamen¹ eller (iii) inkludere alle studenter som har tatt studiepoeng. Ved å benytte definisjon (i) tar vi med alle passive studenter, dvs. som har registrert seg som studenter, men som ikke reelt sett studerer. I motsetning til under (ii) og (iii) får vi også med de studentene som faktisk studerer, men som ikke avlegger noen eksamen i det aktuelle året. Dette kan for eksempel være tilfelle med studenter som skriver hoved/masteroppgave e.l. Det er ikke åpenbart hvilken av disse tilnærmingene som er best. Vi benytter i denne analysen definisjon (ii).² Vi antar her at studenter som er registrert ved institusjonen, men som ikke har meldt seg opp til eksamen i løpet av et helt skoleår, er ’passive studenter’ som ikke bør regnes med når vi beregner gjennomsnittlig studiepoengproduksjon. Fremtidige analyser bør se nærmere på om dette er en rimelig antakelse. Dersom definisjon (iii) benyttes, er resultatene som presenteres i denne analysen, i all hovedsak de samme. Statistisk sentralbyrå har hittil produsert offisiell statistikk basert på definisjon (i) og (iii).

4.1.2. Fullførte grader på normert tid

Et alternativt resultatmål er fullførte grader på normert tid. Om en grad er fullført på normert tid, beregnes ut i fra første semesteret studenten var registrert på dette studiet³, og normert tid innebærer et krav om en gjennomsnittlig studieprogresjon tilsvarende 30 studiepoeng pr. semester. Utvalget er her alle studenter som registrerte seg på de studiene som er omfattet av analysen. Det gjør at alle deltidsstudenter er inkludert i datamaterialet. Hvis omfanget av deltidsstudenter innen hvert fagområde varierer mellom utdanningsinstitusjoner, vil det å ha med deltidsstudentene påvirke resultatene fordi restleddet i regresjonslikningen da vil være korrelert med indikatorvariablene for utdanningsinstitusjon. Det ideelle ville, som ovenfor, å ha informasjon om reelle deltidsstudenter og deres deltidsprosent, og å kontrollere for dette i analysen. Når dette ikke er tilgjengelig, er det uansett ikke helt opplagt om andelen deltidsstudenter utelukkende er et kjennetegn ved studentmassen og utenfor utdanningsinstitusjonens kontroll, eller om det også i noen grad kan påvirkes av utdanningsinstitusjonene.

I denne analysen har vi sett på de ti største kategoriene av bachelor-gradene basert på antall avlagte grader.⁴ Disse er (i) Førskolelærer, (ii) allmennlærer, (iii) ingeniør, (iv) sykepleier, (v) helsefagutdanning, (vi) humanistiske og estetiske fag, (vii) samfunnsfag og juridiske fag, (viii) økonomi og administrative fag, (ix) naturvitenskapelige fag, håndverksfag og tekniske fag og (x) helse-, sosial- og idrettsfag.

¹ Studenter som melder seg opp til eksamen, men som trekker seg, kommer ikke med.

² Enkeltobservasjoner på over 90 studiepoeng (mulige feilføringer) og studiepoeng som er tatt på nytt er ekskludert fra analysen.

³ Dette skiller seg fra statistikk som hittil er publisert av Statistisk sentralbyrå hvor det benyttes: ”Antall år fra studenten første gang var registrert i høyere utdanning”. Se <http://www.ssb.no/emner/04/02/40/hugjen/tab-2006-10-12-01.html>

⁴ Vi har ikke tilgjengelig informasjon om normert gjennomføringstid for samtlige studier. Dette gjelder blant annet praktisk-pedagogisk utdanning.

4.2. Arbeidsmarkedsutfall

Som nevnt tidligere i rapporten, er det i utgangspunktet problematisk å sammenligne institusjoner på grunnlag av resultatindikatorer som bygger på resultatmål som i stor grad er satt av institusjonene selv. Dette vil særlig kunne være problematisk hvis institusjonene har egeninteresse, eller økonomisk gevinst, knyttet til disse resultatmålene. Når det gjelder utdanning, vil arbeidsmarkedet i noen grad kunne fungere som en "ekstern evaluator", hvor man ser på hvorvidt det er systematiske forskjeller mellom studenter fra ulike utdanningsinstitusjoner i hvordan de klarer seg på arbeidsmarkedet, alt annet like.

Det er ikke opplagt hva slags arbeidsmarkedsutfall man skal benytte. Ut fra et ønske om indikatorer som har relevans for utdanningsinstitusjoner, slik de drives i dag, og ikke bare har historisk interesse, bør man tilstrebe å se på arbeidsmarkedsutfall relativt kort etter avsluttet utdanning. Det er først og fremst to argumenter som taler mot et slikt syn. For det første tar det ofte litt tid fra en student avslutter en utdanning til vedkommende er fullt ut etablert på arbeidsmarkedet. For det andre er det fra et teoretisk perspektiv riktig å se på en utdanning som en investering, som gir mulighet for en avkastning i form av en merinntekt på arbeidsmarkedet i den resterende delen av yrkesaktiv alder. Den riktige resultatindikatoren vil da være basert på sammenligninger av nåverdien av livsinntekten for utdanninger fra ulike institusjoner. Dette er for det første vanskelig å beregne gitt tilgjengelige data, samtidig som slike beregninger gir resultater som ikke er særlig relevante i forhold til å gi informasjon om utdanningsinstitusjoner slik de drives i dag. Man må da finne et kompromiss, og velge utfallsmål som lar seg beregne innen rimelig tid etter avsluttet utdanning, og som forhåpentligvis samvarierer sterkt med det ideelle resultatmålet.

Tabell 4.1. Avlagte grader inkludert i analysen av arbeidsmarkedsutfall (1995-2003)

	Antall	Andel
Allmennlærer	13,291	11,63
Førskolelærer	12,414	10,86
Ingeniør	12,146	10,63
Sykepleier	20,384	17,84
Helsefag	9,457	8,28
Siviløkonom	4,283	3,75
Cand.philol	5,615	4,91
Can.polit	6,258	5,48
Cand.scient	5,997	5,25
Cand.jur	6,964	6,09
Cand.med	3,109	2,72
Cand.agric	1,859	1,63
Cand.oecon	326	0,29
Cand.psychol	1,387	1,21
Cand.pharm	474	0,41
Cand.odont.	756	0,66
Cand. Ed. /cand.paed	493	0,43
Sivilingeniør	9,052	7,92
Totalt	114,265	100

Vi betrakter et sett av produserte grader avlagt mellom 1995 og 2003 (og som ikke har tatt mer utdanning etterpå). Hvilke typer grader som er inkludert er illustrert i Tabell 4.1.

Vi betrakter videre to typer resultatmål:

1. opptjent pensjonsgivende inntekt i år t for en lønnstaker (se under) som tok sin utdanning i t-1, t-3, t-5 eller t-7. Dvs. at vi ser på lønnsinntekten henholdsvis ett, tre, fem og sju år etter avsluttet utdanning
2. ledighetsinsidens (registrert ledig i minst en måned) i løpet av fem første år etter avsluttet utdanning.

For at datamaterialet skal være så velegnet som mulig for analyseformålet begrenses utvalget under 1) til fulltidsarbeidende lønnstakere (forventet arbeidstid

på minst 30 timer per uke) mellom 20 og 64 år. Alle som har biarbeidsforhold og som har vært registrert som ledige utelates, det samme gjør de som begynte eller sluttet i en jobb i løpet av året. Videre fjernes personer med spesielt lav eller høy inntekt (under 100 000 NOK eller over 2 000 000 NOK i faste 1998-kr). Personer som arbeider i primærnæringene utelates også.

4.3. Kontrollvariable

Som vi har drøftet tidligere i dette notatet, er det potensielt mange grunner til at resultater varierer mellom utdanningsinstitusjoner, uten at dette nødvendigvis kan tilskrives forskjeller i institusjonenes bidrag til dette resultatet. For å kontrollere for i alle fall deler av resultatvariasjonen som skyldes ”andre faktorer” inkluderer vi en rekke kontrollvariable i regresjonsmodellene som vi benytter for å beregne resultatindikatorene. Nedenfor redegjør vi mer spesifikt for hvilke variable vi har inkludert og hvordan de er operasjonalisert.

4.3.1. Elev- og familiekarakteristika

Foreldrenes utdanningsnivå: Vi har informasjon om høyeste fullførte utdanning for hver av foreldrene da studenten var 16 år. Vi skiller mellom fem utdanningsnivåer – grunnskole, videregående, mellomnivået, kort høyere utdanning (inntil fire år) og lang høyere utdanning – i tillegg til en kategori for manglende utdanningsopplysninger. Dette gir seks dummyvariable (variable som antar verdien 1 hvis ”sann”, og null ellers) for hver forelder. Vi har også eksperimentert med et mer detaljert sett av dummyvariable for hver mulige kombinasjon av mor og fars utdanning (36 dummyvariable), men dette endrer ingen resultater. Tabell 4.2 viser at avlagte studiepoeng stiger, som forventet, med foreldrenes utdanningsnivå. Studenter som har mødre med lang høyere utdanning, avlegger i gjennomsnitt fire studiepoeng pr. år mer enn studenter som har mødre med bare grunnskoleutdanning. Tilsvarende forskjeller finnes for fars utdanning.

Tabell 4.2. Foreldrenes utdanningsnivå da studenten var 16 år og studiepoengproduksjon i skoleåret 2004/2005

	Antall observasjoner	Gjennomsnittlig studiepoeng
Mors høyeste utdanning		
Ingen informasjon	1597	36
Grunnskole	38863	38
Videregående	72911	39
Mellomnivået	4055	41
Høyere utdanning, lavere grad	48056	43
Høyere utdanning, høyere grad	6350	43
Fars høyeste utdanning		
Ingen informasjon	1415	38
Grunnskole	29057	38
Videregående	71400	40
Mellomnivået	6555	41
Høyere utdanning, lavere grad	36562	42
Høyere utdanning, høyere grad	24415	42

Vitnemål fra videregående skole: Vitnemålsopplysninger for videregående opplæring fra Nasjonal vitnemålsdatabase (NVB) innhentes fra og med skoleåret 2000/2001. Ut fra dette beregner vi gjennomsnittskarakter fra videregående skole, som et uveid snitt av karakterene i alle fag. Tabell 4.3 viser at studenter med høyere karaktergjennomsnitt fra videregående skole, i gjennomsnitt tar flere studiepoeng enn studenter med lavere karaktergjennomsnitt fra videregående. Vi har her delt inn studentpopulasjonen i høyere utdanning i ti grupper (desiler) etter karaktergjennomsnittet fra videregående. Vi ser at jo høyere karakterer, jo flere avlagte studiepoeng. Sammenligner vi den øverste desilen med den nederste, er forskjellen på hele 18 studiepoeng. I den grad studenter ikke fordeler seg tilfeldig på studiesteder med hensyn til karakterer, indikerer dette klart at det å kontrollere

for studentenes karakterer fra videregående skole, er svært relevant når man sammenligner resultater mellom høyere utdanningsinstitusjoner.

Tabell 4.3. Vitnemål fra videregående skole og studiepoengproduksjon i skoleåret 2004/2005 (N=50893)

Persentiler i fordelinga av karakterer fra videregående skole	Gjennomsnittlig studiepoeng
0-0,1	37
0,1-0,2	42
0,2-0,3	45
0,3-0,4	46
0,4-0,5	48
0,5-0,6	49
0,6-0,7	51
0,7-0,8	52
0,8-0,9	54
0,9-1,0	55
Totalt	48

Tabell 4.4. Vitnemål fra videregående skole og institusjonstilhørighet i skoleåret 2004/2005

Hslid	Gjennomsnittlige karakterer fra videregående
6	3,75
17	3,78
20	3,80
13	3,82
14	3,83
1	3,84
29	3,86
31	3,87
15	3,88
19	3,90
35	3,94
37	3,96
28	3,97
7	3,98
4	3,99
38	4,00
10	4,01
32	4,01
36	4,05
2	4,06
9	4,11
16	4,19
34	4,20
5	4,21
39	4,24
25	4,28
24	4,33
12	4,34
18	4,35
30	4,42
26	4,43
23	4,48
21	4,49
27	4,55
22	4,61
11	4,65
8	4,75
33	4,99
3	5,17
Total	4,26

At det er klare forskjeller karakternivået til studenter som rekrutteres til ulike institusjoner, kommer klart fram i tabell 4.4. (I denne og følgende tabeller er navnet på utdanningsinstitusjonene erstattet med en anonymisert institusjonskode, "HsId".) Snittkarakter fra videregående varierer mellom 3.75 og 5.17 for de ulike institusjonene. En del av denne variasjonen kan trolig henføres til at de ulike institusjonene tilbyr ulike studier som rekrutterer forskjellig med hensyn til karakterer, men dette forklarer langt fra alle forskjeller.

4.3.2. Studiekarakteristika

Fagfelt: Norsk standard for utdanningsgruppering (NUS2000) gir informasjon på detaljert nivå om utdanningens fagfelt. Tabell 4.5 viser at det er store forskjeller på tvers av fagfelt når det gjelder studiepoengproduksjon. I den grad dette er st strukturelt trekk ved fagene og ikke ved selve institusjonene, er det rimelig å kontrollere for fagfelt i analysene. Derfor inkluderes et sett med dummyvariable for fagfelt. Dersom en student har studiepoeng fra mer enn ett fagfelt benyttes det fagfeltet hvor studenten har tatt flest studiepoeng.

Universitets- og høyskolegruppering (uhgruppe): NUS 2000 klassifiserer også studier på et mer detaljert nivå enn fagfelt. Siden klassifiseringen følger gradsbenevningen fanges lavere og høyere grads studium opp. Vi har også benyttet dummyvariable for UH-gruppe i analysene

Tabell 4.5. Fagfelt, nivå og studiepoengproduksjon i skoleåret 2004/2005

Fagfelt	Fagfelt kode	Antall observasjoner	Gjennomsnittlig studiepoeng
Humanistiske og estetiske fag	1	24098	34
Lærerutdanninger og utdanninger i pedagogikk	2	26722	45
Samfunnsfag og juridiske fag	3	27541	38
Økonomiske og administrative fag	4	18507	35
Naturvitenskapelige fag, håndverksfag og tekniske fag	5	31102	40
Helse-, sosial og idrettsfag	6	37389	45
Primærnæringsfag	7	1049	46
Samferdsels- og sikkerhetsfag og andre servicefag	8	1997	42
Uoppgitt fagfelt	9	420	13
Nivå			
Lavere grad	6	134525	40
Høyere grad	7	34300	43

Figur 4.1. Gjennomsnittlig avlagte studiepoeng etter antall år siden første gang registrert på høyere utdanning



4.3.3. Heltid eller deltid?

For å forsøke å fange opp heltids-deltids-dimensjonen, kontrollerer vi for antall år siden studenten første gang ble registrert på høyere utdanning i form av 12 dummy variable: 0 år siden (samme vinter), 1 år siden, 2 år siden, ..., 10 år siden og '11 år siden eller mer'. Ideen er at heltids-deltids-dimensjonen til en viss grad vil gå langs studentenes ansiennitet på høyere utdanning.

Figur 4.1 viser at blant fortsatt registrerte studenter, er det de som første gang ble registrert på høyere utdanning for tre år siden, som i gjennomsnitt avlegger flest studiepoeng (53). Studenter som registrerte seg for første gang for mer enn 10 år siden avlegger i gjennomsnitt ca. halvparten (26). Det er interessant å merke seg at nesten en av fem studenter som meldte seg opp til eksamen i 2005, første gang registrerte seg på høyere utdanning for mer enn 10 år siden.

Tabell 4.6. Gjennomsnittlig avlagte studiepoeng etter antall år siden første gang registrert på høyere utdanning

År siden første gang registrert	Andel av studentene	Gjennomsnittlig studiepoeng
0*	0,02	12
1	0,16	42
2	0,13	48
3	0,13	51
4	0,10	47
5	0,08	44
6	0,06	39
7	0,05	36
8	0,04	33
9	0,03	31
10	0,03	29
>10	0,18	29

* samme vinter

4.3.4. Kontrollvariable arbeidsmarked

Når vi estimerer (1) med arbeidsmarkedsutfall som resultatmål, er det relevant å kontrollere for variable som en omfattende empirisk litteratur har funnet påvirker slike utfall. I analysene kontrollerer vi for personens kjønn og han/hennes alder når graden ble avsluttet. I tillegg kontrollerer vi for type grad (uhgruppe) og fagfelt. Man kan også argumentere for at vi bør kontrollere for arbeidssted siden det er betydelige regionale lønnsforskjeller i Norge. I den grad studenter fra ulike institusjoner fordeler seg ujevnt på ulike regioner, vil dette kunne influere estimerte institusjonseffekter. Tilsvarende dersom studenter fra institusjoner fordeler seg ujevnt på ulike typer næringer. Vi mener likevel at det ikke er hensiktsmessig å kontrollere for bosted og næring fordi dette i en viss forstand er å betrakte som et utfall av utdanningen. Vi rapporterer likevel resultater der vi kontrollerer for næring og arbeidsfylke.

5. Resultater

I dette kapitlet presenterer vi resultater fra beregninger og estimeringer av varianter av modellen vi har presentert og drøftet tidligere i rapporten. I tråd med formålet med denne rapporten, fokuserer vi presentasjonen rundt følgende punkter:

- Hvor mye av variasjonen i ulike resultatmål kan forklares av ulike typer variable?
- I hvor stor grad endrer observerte resultatforskjeller mellom institusjoner seg når man kontrollerer for ulike studentkjennetegn?
- I hvor stor grad er det samsvar mellom indikatorer som er beregnet ut fra ulike resultatmål?
- Er det samsvar mellom indikatorer beregnet på bakgrunn av resultatmål satt av institusjonene selv og indikatorer basert på arbeidsmarkedsutfall?

Som nevnt innledningsvis, er formålet med denne rapporten ikke først og fremst å produsere nye indikatorer, men mer å drøfte egenskaper ved eksisterende

indikatorer i lys av andre indikatorer som baserer seg på et rikere datagrunnlag og kontrollerer for flere forhold. Det er altså de kvalitative aspektene ved indikatorene som er de sentrale, ikke selve tallene for hver enkelt institusjon. Vi presenterer derfor ikke navn på utdanningsinstitusjonene i denne rapporten.

Vi går først gjennom resultatene basert på studiepoeng og grader fullført på normert tid, deretter ser vi på indikatorer basert på arbeidsmarkedsutfall. Til slutt drøfter vi sammenhengen mellom dem.

5.1. Indikatorer basert på produserte studiepoeng

5.1.1. Hvor mye av den individuelle variasjonen forklares av ulike typer variable?

I modellen presentert i kapittel 3, representerer de ulike parametrene sammenhengen mellom den tilhørende aktuelle variabelen og det resultatmålet vi ser på, dersom man holder alle de andre variablene konstante. Tolkningen av parameteren for hver enkelt utdanningsinstitusjon kan dermed tolkes som et justert gjennomsnitt, dvs. hva gjennomsnittresultatet ville vært dersom studentsammensetningen var lik gjennomsnittet i datamaterialet.

Størrelsen på de justerte gjennomsnittene, og dermed tolkningen av dem, er selvsagt avhengig av hvilke kjennetegn som er inkludert i modellen. Vi er her ikke først og fremst opptatt av den direkte sammenhengen mellom studentkjennetegn og resultater (størrelsen på a-parametrene), selv om dette er interessant i seg selv. (Det er også verd å merke seg at de estimerte sammenhengene ikke nødvendigvis reflekterer direkte årsakssammenhenger). Det vi ønsker er å "fjerne" den delen av utdanningsinstitusjonenes resultatgjennomsnitt som kan tilskrives ulik studentsammensetning. For å få til dette må modellen inkludere alle variable som (i) varierer på institusjonsnivå og som samtidig (ii) har en egen effekt på resultater. I den grad modellen ikke fanger opp alle slike variable, vil vi i økonometriske termer ha et 'utelatt variabel problem'. Legg merke til at både (i) og (ii) må gjelde dersom dette skal være tilfelle. Variable som har en egen effekt på resultatene, men som ikke systematisk varierer på institusjonsnivå er ikke problematisk å utelate. Modellen vil da gi mindre presise, men forventingsrette estimat. Videre er det uproblematisk å inkludere variable som ikke har en egen effekt på resultatene, men som varierer på institusjonsnivå. Den eneste konsekvensen blir at modellen blir mer omfattende enn nødvendig. I den empiriske implementering av modellen legger vi (i) og (ii) til grunn i valget av kontrollvariable. Videre er vi begrenset til å benytte variable hvor informasjonen som ligger til grunn er tilgjengelig for hele studentmassen, siden vi ønsker å lage indikatorer for alle utdanningsinstitusjoner.

Det er også verd å merke seg at enkelte variable kan ha liten betydning i forhold til å "forklare" den totale variasjonen i resultater, men stor betydning for resultatene på enkeltinstitusjoner. Forklaringskraft er derfor ikke det eneste kriteriet som bør ligge til grunn når man velger hvilke variable som bør inkluderes.

Selv om vi ikke primært er interessert i den faktiske sammenhengen mellom ulike variable og ulike skolerresultater, er det likevel nyttig å se nærmere på hvor mye av den individuelle variasjonen i studiepoeng som kan "forklares" av ulike studentkjennetegn. Dette gjøres ved hjelp av regresjonsanalyse. Tabell 5.1 og 5.2 viser hvor mye av den totale variasjonen i studiepoeng som forklares av ulike sett med variable. Kolonnene presenterer tre ulike mål på forklart variasjon. "Total" angir andelen forklart variasjon av en modell som inkluderer variablene til og med den aktuelle raden. "Marginal" angir hvor mye ekstra forklaringskraft man får ved å inkludere det aktuelle settet med variable *i tillegg til* de variablene som var inkludert i modellen i raden over. "Partiell" angir andel forklart variasjon i en modell hvor bare det aktuelle variabelsettet er inkludert. Legg merke til at partiell forklart variasjon ofte er betydelig høyere enn den marginale. Dette skyldes at det ofte er betydelig samvariasjon mellom de ulike forklaringsvariablene. Tabell 5.1

viser resultatene for hele studentpopulasjonen, mens tabell 5.2 er basert på delpopulasjonen av studenter hvor vitnemål fra videregående skole er tilgjengelig.

Tabell 5.1. Forklaringskraft, kontrollvariable, basert på studenter som har meldt seg opp til eksamen i skoleåret 2004/2005 (N=168 883)

	Total R ²	Marginal R ²	Partiell R ²
Antall år på høyere utdanning	0,14		0,14
Fagfelt	0,17	0,03	0,03
Uhgruppe	0,24	0,07	0,17
Foreldrenes utdanning	0,24	0,00*	0,01
Institusjonseffekter	0,25	0,01	0,02

* = 0.004

Tabell 5.2 Forklaringskraft, kontrollvariable, basert på studenter som har meldt seg opp til eksamen i skoleåret 2004/2005 og med tilgjengelig informasjon om vitnemål fra videregående skole (N=50101)

	Total R ²	Marginal R ²	Partiell R ²
Antall år på høyere utdanning	0,09		0,09
Fagfelt	0,12	0,03	0,04
Uhgruppe	0,15	0,03	0,11
Foreldrenes utdanning	0,16	0,01	0,00
Snittkarakter fra vgs (lineært)	0,21	0,05	0,05
Institusjonseffekter	0,22	0,01	0,03

En sentral forklaringsfaktor er 'antall år siden studenten første gang ble registrert på høyere utdanning'. Et sett av dummyvariable som fanger opp denne egenskapen ved studentene forklarer 14 prosent av den totale variasjonen i avlagte studiepoeng. Variasjoner på tvers av fagfelt og særlig universitets- og høyskolegruppering forklarer også en betydelig del av variasjonen i avlagte studiepoeng.

Selv om studenter som har foreldre med høyere utdanning, kontrollert for antall år på høyere utdanning, fagfelt og uh-gruppe, produserer flere studiepoeng, er denne effekten svak. Kontrollert for andre faktorer er forskjellen mellom det å ha en mor med høyere utdanning i forhold til å ha en mor med grunnskole i gjennomsnitt 3 studiepoeng. Tilsvarende forskjell for far er 2 studiepoeng. Effektene av både mor og fars utdanning er sterkt statistisk signifikante, men kan forklare lite av variasjonen i studiepoengproduksjonen. Forskjellen til analyser av skoleprestasjoner i grunnskolen er stor. Mens foreldrenes utdanning forklarer 18 prosent av variasjonen i grunnskolepoeng (Hægeland et al. 2005), er det tilsvarende tallet for variasjonen i studiepoengproduksjon under 1 prosent. En viktig grunn til dette er sannsynligvis at vi her ser på et langt mer selektert utvalg av individer enn når vi ser på avgangselever fra grunnskolen, hvor hele fødselskullene i prinsippet er med. Gjennom videregående skole og valg av høyere utdanning skjer det et systematisk frafall som reduserer variasjonen i foreldreutdanning blant de "gjenværende" individene. Dette kan forklare at de til dels betydelige forskjellene etter foreldres utdanning som vi så i tabell 4.2, går sammen med en lav andel forklart variasjon. I tillegg bruker vi med avlagte studiepoeng et grovere resultatmål (bestått/ikke bestått) som fokuserer på den nedre enden av prestasjonsskalaen, enn når vi bruker hele karakterskalaen. Dette kan også bidra til at andelen forklart variasjon er lavere.

Den marginale forklaringskraften til et sett av felles institusjonseffekter (dummyvariable for hver institusjon) er i overkant av 1 prosent når andre faktorer er kontrollert for. Dette illustrerer at en relativt beskjedne del av studiepoengproduksjon kan forklares av faktorer som studenter ved samme institusjon har felles. Hægeland et al. (2005) finner at tilsvarende variasjon (2 prosent) i grunnskolepoeng kan forklares av skolefaste effekter.

Det samme mønsteret gjør seg gjeldende dersom man betrakter bare studenter hvor vitnemål fra videregående skole er tilgjengelig, se Tabell 5.2. I hovedsak innebærer dette studenter som gikk ut fra videregående tidligst våren 2001. Gjennomsnittskarakter fra vitnemål kan forklare omkring 5 prosent av variasjonen i studiepoengproduksjon. Estimaten antyder at dersom gjennomsnittskarakterer fra videregående

skole øker med en enhet (for eksempel fra fire til fem), øker antall avlagte studiepoeng med 9,4.⁵

5.1.2. Institusjonens bidrag

Gjennomsnittlig studiepoengproduksjon for studenter som har meldt seg opp til eksamen varierer relativt mye på tvers av institusjonene. Blant de institusjonene vi har med i denne analysen, varierer gjennomsnittet mellom 34 og 59 studiepoeng i året.⁶

Tabell 5.3. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, (N=168 827)

Hslid	1	2	3	4	5
12	34	35	32	32	32
19	36	37	38	40	41
14	36	36	39	39	39
29	36	38	38	39	40
21	36	37	37	38	37
8	37	36	39	30	29
18	37	40	38	39	39
30	38	38	37	37	36
33	38	40	44	43	43
37	39	38	39	40	41
13	40	40	40	41	41
10	40	42	43	43	43
23	40	39	37	35	35
34	41	42	44	47	48
17	41	40	41	41	42
28	41	40	41	41	42
38	41	41	42	42	43
35	41	41	40	40	41
16	41	40	41	40	40
25	42	41	43	42	42
36	42	42	43	44	44
2	42	41	42	42	42
6	42	44	42	42	42
9	43	43	42	43	43
7	43	40	42	41	42
15	43	42	43	43	43
4	44	44	44	45	45
32	44	42	43	45	45
1	44	44	44	44	44
24	44	43	43	43	43
20	45	44	44	44	45
31	45	43	43	42	43
5	45	44	44	43	43
39	46	44	44	43	43
3	47	46	36	56	55
26	48	45	51	46	46
11	50	49	44	47	46
27	53	52	55	52	52
22	59	56	60	56	56
ÅR	-	JA	JA	JA	JA
FAGFELT	-	-	JA	JA	JA
UHGRUPPE	-	-	-	JA	JA
FORELDRENES UTDANNINGSNIVÅ	-	-	-	-	JA

Tabell 5.3 viser gjennomsnittlig studiepoengproduksjon for fire ulike spesifikasjoner for studieåret 2004/2005. Spesifikasjon 1 viser ukorrigerte gjennomsnitt, mens spesifikasjon 2, 3, 4 og 5 kontrollerer for ulike andre kjennetegn som er angitt i tabellen. Tolkningen av tallene i disse kolonnene er hva som ville vært gjennomsnittlig antall studiepoeng ved institusjonen dersom den hadde en sammensetning av studentmassen som var lik gjennomsnittet for utvalget av studenter som er med i analysen for de variablene som er inkludert i den aktuelle modellen. Vi ser av tabellen at det gjennomgående er høy grad av samsvar

⁵ Forklaringskraften økes ikke dersom man introduserer en ikke-lineær sammenheng i form av et andreordensledd eller et sett av dummy variable basert på desiler i fordelinga av vitnemål fra videregående skole.

⁶ En F-test av felles signifikans av skole-effektene er også statistisk signifikant på 1% nivå, med en F-verdi på 239.

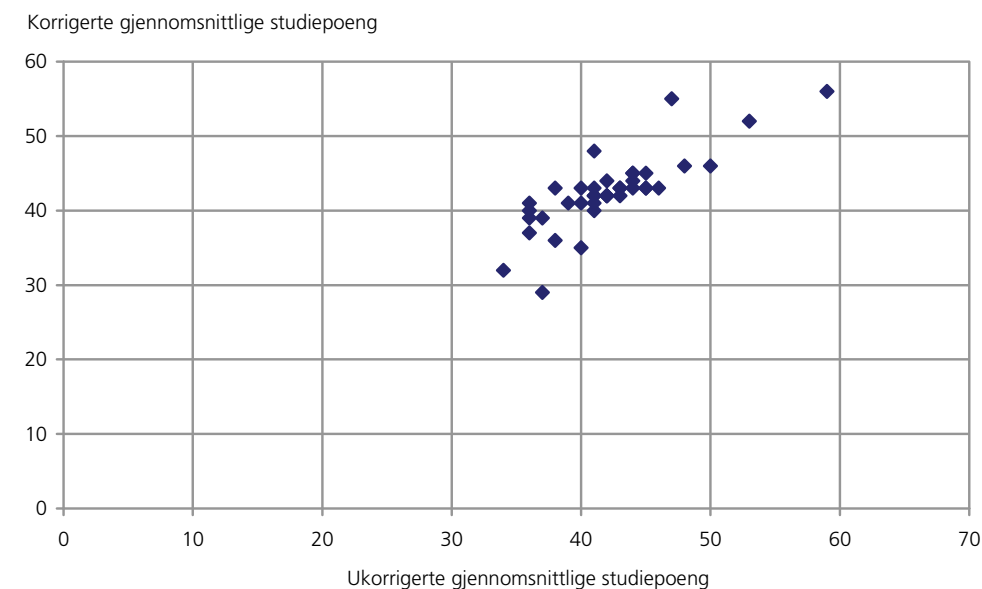
mellom ”ukorrigerede gjennomsnitt” og ulike estimerte korrigerede gjennomsnitt. Institusjoner med høy ukorrigeret gjennomsnittlig studiepoengproduksjon fremstår i stor grad også som ”effektive” dersom man kontrollerer for ’antall år siden studenten første gang registrerte seg på høyere utdanning’, fagfelt, uhgruppe eller foreldrenes utdanningsnivå. Som rapportert i Tabell 5.4, er korrelasjonen mellom den ukorrigerede indikatoren og mer sofistikerte indikatorer er alltid over 0,8.⁷ Figur 5.1 illustrerer den nære sammenhengen mellom den ukorrigerede indikatoren (spesifikasjon 1) og den indikatoren med rikset sett av kontrollvariable (spesifikasjon 5).

Interessant er det også å merke seg at det er ulikheter på tvers av fagfelt og uhgruppe som gjør at korrelasjonen av parametervektorer endres. Når det er kontrollert for fagspesifikke kjennetegn og antall år studenten har vært på høyere utdanning, har det liten betydning å kontrollere for foreldrenes utdanningsbakgrunn. Det er likevel verd å merke seg at selv om korrelasjonen mellom ukorrigerede resultater og estimerte indikatorer gjennomgående er høy, så har det å kontrollere for studentkjennetegn relativt stor betydning for enkeltinstitusjoner. Hvis korreksjonene var uten betydning, ville punktene i figur 5.1 ligge langs en rett 45-graders linje i diagrammet. For en del institusjoner er dette tilnærmet tilfellet, mens for andre innebærer det å kontrollere for sammensetningen av studentmassen en ikke ubetydelig korreksjon.

Tabell 5.4. Korrelasjonsmatrise parametervektor presentert i tabell 5.3

	Spes 1	Spes 2	Spes 3	Spes 4	Spes 5
Spes 1	1,00				
Spes 2	0,97	1,00			
Spes 3	0,85	0,85	1,00		
Spes 4	0,84	0,87	0,70	1,00	
Spes 5	0,82	0,86	0,70	1,00	1,00

Figur 5.1. Sammenhengen mellom ukorrigerede og korrigerede studiepoengsindikatorer, basert på samtlige studenter som har meldt seg opp til eksamen i 2005¹



¹ Spesifikasjon 1 vs. Spesifikasjon 5, tabell 5.3.

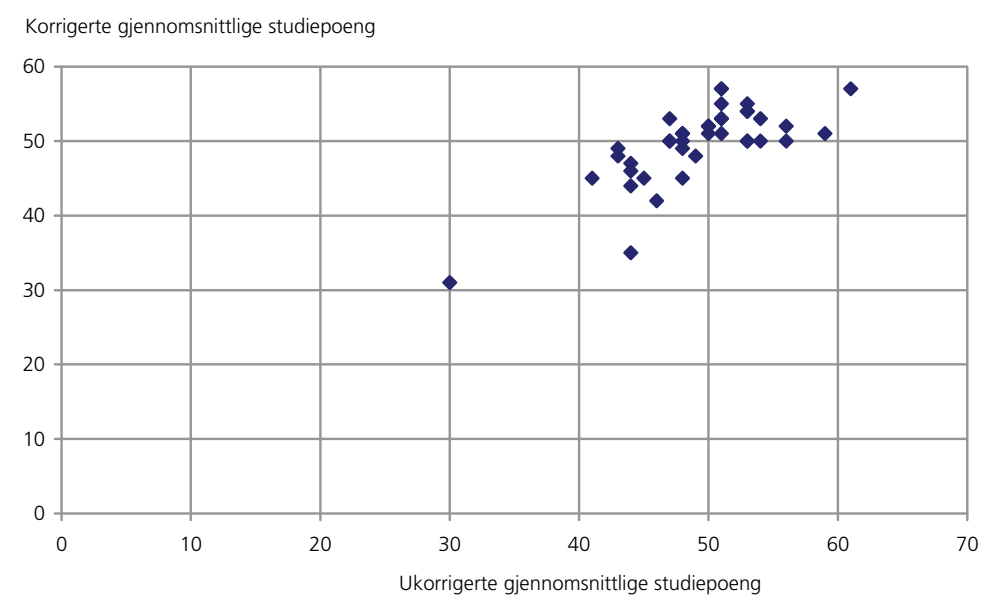
⁷ Korrelasjonen er også over 0,8 dersom institusjoner med mindre enn hhv. 1000, 2000 eller 3000 studenter utelates, man bare betrakter universiteter eller man ser på korrelasjonen mellom tilsvarende indikatorer i studieåret 2003/2004.

Tabell 5.5. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen. Betrakter bare studenter hvor informasjon om vitnemål fra videregående skole er tilgjengelig (N= 50068)

Hslid	1	2	3	4	5	6	7
26	30	28	30	32	32	29	31
18	41	44	43	45	45	40	45
14	43	44	44	45	45	47	48
29	43	47	46	46	47	46	49
9	44	46	45	46	46	46	46
17	44	43	43	43	43	49	47
8	44	44	49	36	36	39	35
21	44	45	45	46	46	41	44
30	45	44	45	45	45	43	45
12	46	46	44	43	43	45	42
7	47	47	47	48	48	50	50
4	47	49	50	52	52	50	53
37	47	46	46	48	48	50	50
38	48	49	49	49	49	51	51
35	48	48	48	48	48	52	49
2	48	48	49	49	49	50	50
6	48	47	47	46	47	53	51
23	48	47	47	45	45	45	45
16	49	48	49	48	48	50	48
19	50	49	48	49	49	53	51
15	50	49	50	50	50	53	52
10	50	50	50	50	51	53	52
34	51	51	53	56	58	51	57
20	51	51	51	51	51	55	55
28	51	50	50	50	51	54	51
13	51	50	50	50	50	55	53
1	51	51	51	51	51	55	53
3	51	50	43	64	63	42	57
25	53	52	51	51	51	52	50
36	53	53	53	53	53	55	54
31	53	52	51	51	51	56	54
32	53	53	53	53	53	56	55
5	53	52	52	51	51	54	50
39	54	53	52	52	52	54	50
24	54	54	53	53	53	53	53
22	56	55	60	56	55	52	52
11	56	55	53	51	50	52	50
33	59	58	59	58	57	52	51
27	61	60	65	61	60	58	57
ÅR	-	JA	JA	JA	JA	-	JA
FAGFELT	-	-	JA	JA	JA	-	JA
UHGRUPPE	-	-	-	JA	JA	-	JA
FORELDRENES UTDANNINGSNIVÅ	-	-	-	-	JA	-	JA
KARAKTERER FRA VGS	-	-	-	-	-	JA	JA

Tabell 5.5 presenterer tilsvarende resultater for delpopulasjonen av studenter hvor informasjon om vitnemål fra videregående karakterer er tilgjengelig. Dette reduserer utvalget fra 168 883 observasjoner til 50068 observasjoner. Det samme mønsteret som er presentert tidligere gjør seg gjeldende (grafisk illustrert i figur 5.2). Ukorrigerte studiepoengsindikatorer reflekterer i høy grad mer sofistikerte indikatorer. Å kontrollere for foreldrenes utdanningsnivå, har relativt liten betydning for å forklare individuell variasjon i studiepoeng. Prestasjoner på videregående skole kan imidlertid i høyere grad forklare variasjoner i studiepoengproduksjon. Når alle andre kontrollvariable er inkludert, endrer det å inkludere snittkarakterer fra videregående skole parametervektorene noe (korrelasjon på 0,91 mellom spesifisering 5 og spesifisering 7). Institusjoner som rekrutterer studenter med lavt snitt fra videregående, justeres opp, mens institusjonene som rekrutterer studenter med høyt snitt fra videregående, justeres ned.

Figur 5.2. Sammenhengen mellom ukorrigerde og korrigerte studiepoengsindikatorer, basert på studenter som har meldt seg opp til eksamen i 2005, og det er tilgjengelig informasjon om vitnemål¹



¹ Spesifikasjon 1 vs. Spesifikasjon 7, tabell 5.5.

Tabell 5.6. Korrelasjonsmatrise til parametervektor presentert i tabell 5.5

	Spes 1	Spes 2	Spes 3	Spes 4	Spes 5	Spes 6	Spes 7
Spes 1	1,00						
Spes 2	0,97	1,00					
Spes 3	0,92	0,93	1,00				
Spes 4	0,83	0,84	0,72	1,00			
Spes 5	0,82	0,83	0,72	1,00	1,00		
Spes 6	0,82	0,83	0,77	0,64	0,67	1,00	
Spes 7	0,74	0,78	0,65	0,89	0,91	0,82	1,00

5.1.3. Lavere og høyere grad?

Som vi har drøftet flere steder, er utdanningsinstitusjonene til dels svært forskjellige med hensyn til hva slags utdanninger de tilbyr. I den grad denne heterogeniteten delvis driver resultatforskjellene mellom institusjonene, pga. at utdanningene i seg selv er forskjellige i forhold til hvor lett det er å oppnå et gitt resultat målt for eksempel ved studiepoeng, er det viktig å ta høyde for dette i beregningene. For å kontrollere for heterogenitet på tvers av fagfelt og nivå, har strategien ovenfor vært å inkludere variable som fanger opp disse dimensjonene som kontrollvariable. En annen mulig strategi er å analysere ulike typer utdanninger separat. Her viser vi et eksempel på dette, hvor vi deler inn studentene i høyere- og laveregradsstudenter. Resultatene er presentert i tabellene 13, 14 og 15. Kvalitativt sett er resultatene tilsvarende de vi presenterte ovenfor. Når vi ser på forskjeller i studiepoengsproduksjon pr. student mellom utdanningsinstitusjoner, er bildet i hovedsak det samme om vi ser på ukorrigerde resultater eller om vi kontrollerer for en rekke studentkarakteristika. Imidlertid er det enkelte institusjoner hvor forskjellen på ukorrigerde og korrigerte resultatmål er betydelig. Det er imidlertid interessant å merke seg at korrelasjonen mellom parametervektorer på tvers av lavere og høyere grad er nær null. Det vil si at studiepoeng per student på lavere grad er en svært dårlig indikator på hvor mange poeng per student som produseres på høyere grad ved samme institusjon.

Tabell 5.7. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, lavere grad (N=134 546)

Hslid	1	2	3
18	30	36	36
23	33	34	34
21	33	32	32
30	33	38	37
12	34	29	29
19	36	37	37
14	36	38	38
29	36	39	39
28	40	39	39
37	40	42	41
33	40	40	41
34	41	41	42
17	41	39	40
38	41	39	40
35	41	41	42
25	41	46	47
13	41	39	40
16	41	40	40
10	41	40	40
9	42	41	41
2	42	33	32
6	42	40	41
8	42	41	42
7	43	40	40
24	43	41	41
36	44	44	44
4	44	42	43
15	44	44	44
32	44	44	45
1	44	43	43
20	45	41	42
31	45	42	42
5	45	43	44
39	46	42	41
26	48	47	47
22	57	52	52
27	58	54	53
3	64	64	63
ÅR	-	JA	JA
FAGFELT	-	JA	JA
UHGRUPPE	-	JA	JA
FORELDRENESUTDANNINGSNIVÅ	-	-	JA
KARAKTERER FRA VGS	-	-	-

Tabell 5.8. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, lavere grad. Betrakter bare studenter hvor informasjon om vitnemål fra videregående skole er tilgjengelig. (N=42502)

Hslid	1	2	3	4
26	30	32	32	31
18	34	42	42	42
30	42	44	44	43
14	43	45	46	48
29	43	43	43	44
23	43	44	44	47
21	43	46	46	44
9	44	42	43	46
17	44	45	45	45
8	44	34	34	34
12	46	42	41	41
7	47	51	51	52
4	47	47	47	49
37	47	47	47	49
38	48	48	48	50
35	48	46	46	49
2	48	48	48	49
6	48	47	47	48
16	49	47	47	47
19	50	48	48	50
15	50	49	49	51
10	50	49	50	51
34	51	49	49	52
20	51	55	57	56
28	51	50	50	54
13	51	49	50	50
1	51	50	50	52
25	53	51	51	50
36	53	50	51	53
31	53	50	50	48
32	53	52	52	53
5	53	50	50	49
24	53	52	52	54
39	54	51	51	49
22	56	55	54	51
33	59	57	56	50
27	61	60	59	57
3	69	63	62	56
ÅR	-	JA	JA	JA
FAGFELT	-	JA	JA	JA
UHGRUPPE	-	JA	JA	JA
FORELDRENES UTDANNINGSNIVÅ	-	-	JA	JA
KARAKTERER FRA VGS	-	-	-	JA

Tabell 5.9. Studiepoengproduksjon i høyere utdanning, per student som har meldt seg opp til eksamen i skoleåret 2004/2005, høyere grad (N=34 281)

Hsld	1	2	3
15	17	28	28
8	21	23	23
13	22	39	39
36	24	33	33
20	32	42	42
32	32	37	38
33	32	39	39
1	33	42	43
5	35	45	45
31	36	42	42
10	37	49	49
37	37	41	41
6	37	42	43
19	38	51	52
38	39	41	41
27	39	44	45
21	39	45	45
16	41	43	44
28	42	45	46
4	42	43	43
12	42	51	51
14	45	56	56
2	46	44	44
3	46	27	27
23	46	42	42
30	46	49	49
24	46	47	48
18	48	47	47
39	50	52	52
11	50	58	58
25	51	58	58
9	52	66	66
22	66	74	74
ÅR	JÅ	JÅ	JÅ
FAGFELT	-	JÅ	JÅ
UHGRUPPE	-	JÅ	JÅ
FORELDRENES UTDANNINGSNIVÅ	-	-	JÅ
KARAKTERER FRA VGS	-	-	-

5.2. Indikatorer basert på produksjon av grader på normert tid

Vi ser nå på et annet resultatmål, nemlig andelen av studentene som gjennomfører studiet på normert tid. Vi ser på de ti hyppigst avlagte bachelor-gradene. Det er tydelige forskjeller med hensyn til gjennomføringstid. Tabell 5.10 viser at mens bare 34 prosent av studentene som avla en bachelorgrad innenfor samfunnsvitenskapelige og juridiske fag gjorde det på de normerte tre årene, gjennomførte 67 prosent av studentene på helsefaglig utdanning (unntatt sykepleier) på tre år eller mindre.

Det er også betydelige forskjeller på tvers av institusjoner, rapport i tabell 5.11. Andelen som fullførte på normert tid, varierer fra 0,06 til 0,89. Dersom vi, gjennom en lineær sannsynlighetsmodell (MKM), tar hensyn til at ulike institusjoner tilbyr ulike kombinasjoner av bachelor-programmer, reduseres forskjellen mellom institusjonene noe. Typisk vil de institusjonene hvor store deler av de produserte gradene er innenfor samfunnsvitenskapelige og juridiske fag og humanistiske og estetiske fag, hvor det nasjonale snittet er lavt, få oppjustert sine resultater når vi korrigerer for gradssammensetning. Tilsvarende justeres en del institusjoner med overvekt av bachelorgrader hvor en høy nasjonal andel gjennomfører på normert tid, for eksempel institusjoner som har stor andel sykepleier- og helsefagstudenter. Selv om korreksjonen har stor betydning for enkeltinstitusjoner, er korrelasjonen

mellom ukorrigerede og korrigerede andeler som gjennomfører på normert tid relativt høy, 0,86. Figur 5.3 illustrerer dette grafisk.

Tabell 5.10. Grader produsert på normert tid, per studenter som avla bachelor grad studieåret 2004/2005, fordelt på de ti største bachelor-utdanninger, målt i antall produserte grader

Bachelorutdanning	Uh-gruppe	Andelen på normert	Antall semester, normert	Observasjoner
Førskolelærer	9/9B	0,57	6	1071
Allmennlærer	8/8B	0,44	8	1872
Ingeniør	11/11B	0,63	6	1725
Sykepleier	12/12B	0,65	6	3097
Helsefagutdanning (ikke sykepleier)	13/13B	0,67	6	1632
Humanistisk og estetiske fag	26B	0,41	6	1189
Samfunnsfag og juridiske fag	28B	0,34	6	763
Økonomi og administrative fag	29B	0,53	6	1258
Naturvitenskapelige fag, håndverksfag og tekniske fag	30B	0,49	6	592
Helse-, sosial- og idrettsfag	96B	0,58	6	1577
Totalt		0,56		14776

Tabell 5.11. Grader produsert på normert tid, per studenter som avla bachelor grad i 2005¹

Hslid	Ukorrigert snitt	Snitt korrigeret for sammensetning av grader (MKM)	Antall grader avlagt per institusjon
35	0,06	0,10	146
18	0,10	0,27	69
21	0,36	0,51	595
17	0,37	0,27	189
38	0,38	0,38	383
30	0,38	0,56	216
20	0,41	0,43	250
23	0,41	0,57	361
19	0,47	0,44	448
36	0,50	0,60	332
13	0,51	0,42	419
4	0,51	0,65	578
37	0,51	0,65	255
27	0,51	0,51	56
22	0,51	0,51	117
15	0,52	0,52	454
6	0,53	0,43	292
5	0,56	0,53	1989
16	0,59	0,61	890
29	0,60	0,61	469
1	0,60	0,57	604
14	0,61	0,56	169
31	0,61	0,56	358
39	0,63	0,59	1108
28	0,64	0,56	125
10	0,64	0,60	358
9	0,65	0,62	397
25	0,65	0,61	1396
7	0,65	0,58	235
2	0,66	0,70	758
32	0,66	0,66	338
12	0,79	0,83	295
24	0,80	0,89	96
Totalt	0,56	0,56	14745

¹ Begrenset til bachelor grad av typen nevnt i tabell 5.10 (N=14745).

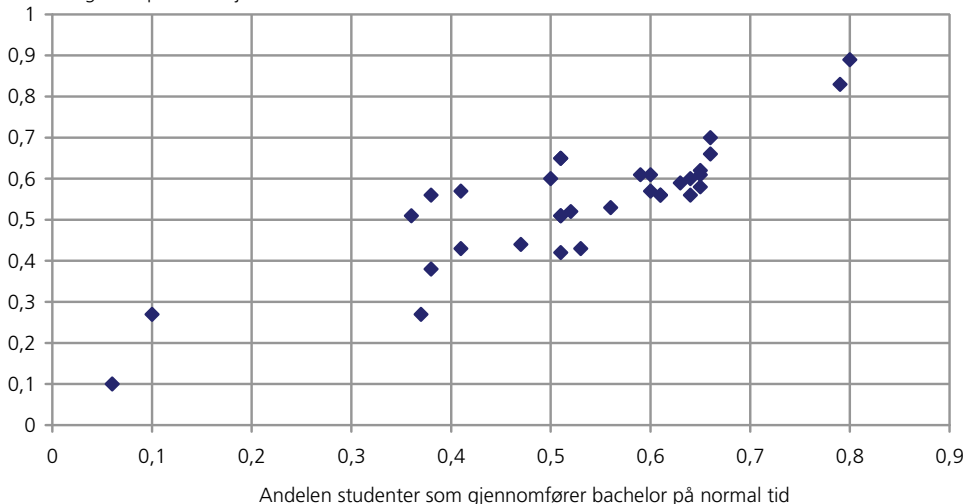
Det er også store forskjeller på tvers av institusjoner for hver av de større utdanningsgruppene. For å illustrere dette presenterer tabell 5.12 og tabell 5.13 grader produsert på normert tid for to store grupper, henholdsvis sykepleier og allmennlærerutdanningen.

På samme måte som i forrige avsnitt, har vi også gjort beregninger av indikatorer som også tar hensyn til at institusjonene er forskjellige med hensyn til sammen-

setning av studentmassen når det gjelder foreldrenes utdanning. Disse resultatene rapporteres ikke her, men hovedbildet er igjen det samme: Foreldres utdanning forklarer relativt lite av den totale individuelle variasjonen i studenters fullføring av studiet på normert tid. Korrelasjonen mellom ukorrigerede og korrigerede indikatorer er relativt høy, men enkelte institusjoner får justert sine resultater i betydelig grad.

Figur 5.3. Sammenhengen mellom ukorrigerede og korrigerede indikatorer, basert på studenter som avla bachelorgrad i 2005, i en av de ti største bachelorgradene

Andelen studenter som gjennomfører bachelor på normal tid, korrigeret for sammensetning av bachelorgrader på institusjonsnivå



Tabell 5.12. Grader produsert på normert tid, per studenter som fullførte sykepleie-utdanningen i 2005. 3077 observasjoner. Institusjoner med mindre enn 20 avlagte grader er utelatt fra analysen

Hslid	Ukorrigert snitt	Antall grader avlagt per institusjon
15	0,39	118
37	0,41	170
13	0,46	112
17	0,51	51
38	0,56	160
19	0,58	184
9	0,61	119
39	0,62	157
10	0,63	131
14	0,64	92
6	0,64	116
5	0,66	424
1	0,67	135
7	0,68	94
4	0,73	71
29	0,74	135
2	0,76	197
31	0,79	183
20	0,84	42
28	0,86	38
16	0,87	163
25	0,88	185
Totalt	0,65	3077

Tabell 5.13. Grader produsert på normert tid, per studenter som fullførte allmennlærer-utdanningen i 2005. Institusjoner med mindre enn 20 avlagte grader er utelatt fra analysen

Hslid	Ukorrigert snitt	Antall grader avlagt per institusjon
35	0,08	89
38	0,14	78
13	0,17	77
20	0,25	47
36	0,32	85
4	0,33	77
19	0,37	95
1	0,40	63
15	0,46	113
16	0,47	116
31	0,49	70
39	0,50	184
5	0,50	202
25	0,53	177
29	0,56	164
9	0,61	115
2	0,66	103
TOTALT	0,44	1855

5.2.1. Oppsummert om studiepoeng og gradsproduksjon

Familiebakgrunn har i utallige studier vist seg som en av de mest sentrale forklaringsfaktorene for akademiske prestasjoner. I denne analysen finner vi imidlertid relativt svake effekter av familiebakgrunn. Antageligvis skyldes dette blant annet at studenter på høyere utdanning er en sterkt selektert gruppe.

Sentralt for denne analysen er imidlertid at effekten av institusjonenes bidrag endres relativt lite dersom vi kontrollerer for familiebakgrunn. Det er noe større effekter av tidligere akademiske prestasjoner på studiepoeng, men å kontrollere for disse fører heller ikke til en dramatisk endring i institusjonseffektene.

Det er uklart hvordan det å kontrollere for tidligere akademiske prestasjoner vil slå ut på produksjon av grader siden data på vitnemål fra videregående skole ikke er tilgjengelig tilstrekkelig langt tilbake i tid. Videre er det relativt store forskjeller på tvers av såkalte uhgrupper, men heller ikke det gir en dramatisk endring i institusjonseffektene.

Det kan imidlertid være grunn til å stille spørsmålstegn ved resultatmål basert på studiepoeng og gradsproduksjon. For det første siden disse er noe som fastlegges institusjonene selv. For det andre er de sensitive til andelen deltids- og heltidsstudenter, som det per dags dato ikke finnes gode data på.

5.3. Indikatorer basert på arbeidsmarked utfall

Som drøftet tidligere, vil arbeidsmarkedet i i noen grad kunne fungere som en ”ekstern evaluator”, hvor man kan se på hvorvidt det er systematiske forskjeller mellom studenter fra ulike utdanningsinstitusjoner i hvordan de klarer seg på arbeidsmarkedet, alt annet like. I den grad man aksepterer at arbeidsmarkedet gir en relevant evaluering av ”utdanningskvalitet”, vil det være ønskelig at de indikatorene man kan beregne på kort sikt, basert på for eksempel studiepoengsproduksjon, viser positiv samvariasjon med indikatorer basert på arbeidsmarked utfall beregnet på litt lengre sikt. Hvis de ikke gjør det, er det mer problematisk å tolke de kortsiktige indikatorene som indikatorer på kvalitet, i alle fall den kvaliteten som implisitt måles og verdsettes i arbeidsmarkedet. Det er selvsagt ikke dermed sagt at de er irrelevante og uten verdi som resultatindikatorer, men de måler da eventuelt kvalitet og resultater langs en annen dimensjon enn hva indikatorer basert på arbeidsmarked utfall gjør.

Tabell 5.14. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (Hsld=5)

	Lønn (1. år)	Lønn (3. år)	Lønn (5. år)	Lønn (7. år)	Ledighet (i løpet av 5 år)	Studie- poeng 2003	Grad 2005
Spes	1	2	3	4	5	6	7
Hsld							
24	-0,167	-0,205	-0,232	-0,221	-0,011	41	0,89
14	-0,078	-0,056	-0,063	-0,069	0,066	40	0,56
7	-0,063	-0,026	-0,035	-0,031	-0,024	47	0,58
28	-0,056	-0,066	-0,108	-0,115	0,076	45	0,56
23	-0,049	-0,047	-0,044	-0,012	-0,025	46	0,57
25	-0,048	-0,047	-0,057	-0,051	0,034	54	0,62
18	-0,045	-0,049	-0,038	-0,024	-0,028	49	0,27
2	-0,042	-0,043	-0,051	-0,061	0,076	50	0,66
19	-0,038	-0,040	-0,070	-0,090	0,094	50	0,44
36	-0,036	-0,027	-0,037	-0,035	0,023	44	0,60
30	-0,035	-0,028	-0,006	-0,003	-0,006	44	0,56
39	-0,033	-0,029	-0,033	-0,016	0,027	52	0,59
17	-0,029	-0,030	-0,026	-0,032	0,002	45	0,42
37	-0,029	-0,024	-0,027	-0,034	0,009	39	0,27
31	-0,026	-0,033	-0,036	-0,028	-0,002	42	0,57
4	-0,026	-0,024	-0,047	-0,045	0,028	46	0,51
1	-0,026	-0,034	-0,025	-0,024	-0,001	51	0,56
21	-0,023	-0,015	0,007	0,033	-0,043	41	0,51
9	-0,020	-0,019	-0,032	-0,044	0,018	47	0,38
38	-0,020	-0,020	-0,026	-0,026	0,006	53	0,61
13	-0,019	-0,031	-0,041	-0,047	-0,006	46	0,51
29	-0,019	-0,019	-0,035	-0,030	-0,014	42	0,61
15	-0,014	0,001	-0,017	-0,018	0,008	46	0,52
32	-0,011	-0,028	-0,035	-0,035	-0,005	53	0,70
10	-0,010	-0,010	-0,015	-0,033	-0,007	50	0,60
35	-0,009	-0,022	-0,035	-0,033	0,025	44	0,10
20	-0,006	-0,008	-0,011	-0,011	0,081	50	0,43
5	0,000	0,000	0,000	0,000	0,000	48	0,53
16	0,017	0,021	0,011	0,008	-0,014	46	0,61
6	0,019	0,008	-0,002	0,004	-0,030	42	0,43
12	0,058	0,029	0,043	0,019	-0,053	43	0,83
33	0,068	0,116	0,130	0,130	0,000	44	,
Laveste verdi	-0,167	-0,205	-0,232	-0,221	-0,053	39	0,10
Høyeste verdi	0,068	0,116	0,130	0,130	0,094	54	0,89
Differanse mellom laveste og høyeste verdi	0,235	0,322	0,361	0,351	0,147	15	0,79
Observasjoner	42182	46284	33012	19731	61216		
Antall kull	9	7	5	3	5		
Forklaringskraft i modell uten institusjonseffekter	0,475	0,443	0,441	0,44	0,071		
Forklaringskraft i modell med institusjonseffekter	0,483	0,452	0,452	0,45	0,077		
Differanse i r ² -adj	0,008	0,009	0,011	0,01	0,006		

Kontrollvariable er inkludert i alle spesifikasjoner, som rapportert i teksten. Forklaringskraft i ledighetsspesifikasjonen referer til pseudo R².

Tabell 5.15. Korrelasjonsmatrise til parametervektor presentert i tabell 5.14

	Spes 1	Spes 2	Spes 3	Spes 4	Spes 5	Spes 6	Spes 7
Spes 1	1,00						
Spes 2	0,94	1,00					
Spes 3	0,91	0,97	1,00				
Spes 4	0,85	0,93	0,98	1,00			
Spes 5	-0,23	-0,16	-0,29	-0,36	1,00		
Spes 6	0,07	0,05	0,02	0,01	0,27	1,00	
Spes 7	-0,21	-0,28	-0,25	-0,26	-0,18	0,09	1,00

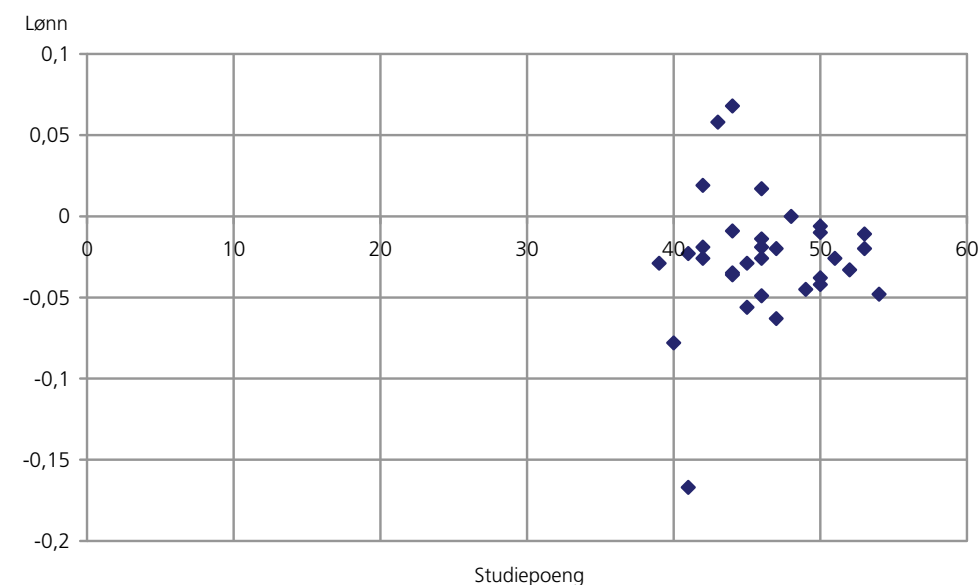
Som diskutert i forrige kapittel, beregner vi indikatorer basert på en modell hvor vi søker å forklare individuelle variasjoner i (logaritmen til) pensjonsgivende inntekt med variable som kjønn, alder ved avlagt grad, år graden ble avlagt og fagfelt og uh-gruppe graden tilhører. I tillegg kontrollerer vi for hvilken utdanningsinstitusjon

studenten ble uteksaminert fra, og det er som før, disse estimerte forskjellene mellom utdanningsinstitusjoner som er våre resultatindikatorer. Resultatene i tabell 5.14 er basert på arbeidsmarkedsutfall for studenter som avsluttet sin høyere utdanning mellom 1995 og 2003. Å inkludere faste institusjonseffekter, øker modellens forklaringskraft med rundt 1 prosentpoeng både for lønn etter 1 år, 3 år, 5 år og 7 år etter eksamensåret. Institusjonseffektene er samlet sett statistisk signifikante, og flere av dem er også signifikant forskjellige fra hverandre.

Parametervektorene i spesifikasjon 1, 2, 3 og 4 kan tolkes som prosentvis endring i lønn relativt til referanseinstitusjonen. Det er betydelige forskjeller på tvers av institusjonene. På institusjonen med den høyeste indikatoren får studentene, alt annet likt, 7 prosent høyere enn lønn enn personer fra referanseinstitusjonen. Tilsvarende oppnår personer med avlagt grad fra institusjonen med lavest indikator 17 prosent lavere lønn. Forskjellene på tvers av institusjonene avhenger lite om man ser på lønn 1, 3, 5 eller 7 år etter eksamen (se tabell 5.15).

Korrelasjonen mellom studiepoengsproduksjon i 2003 og forskjeller i lønn ett år etter avsluttet utdanning (1995-2003) er imidlertid meget lav (0,07).⁸ Et høyt nivå på gjennomsnittlig studiepoeng ser altså ikke ut til å være noen god indikator på hvor mye studenter fra disse institusjonene får i lønn, når andre faktorer er kontrollert for. Figur 5.4 illustrerer dette grafisk. Denne konklusjonen baserer seg imidlertid på en antakelse om konstante institusjonseffekter over tid, siden de er beregnet på grunnlag av forskjellige kull med studenter.

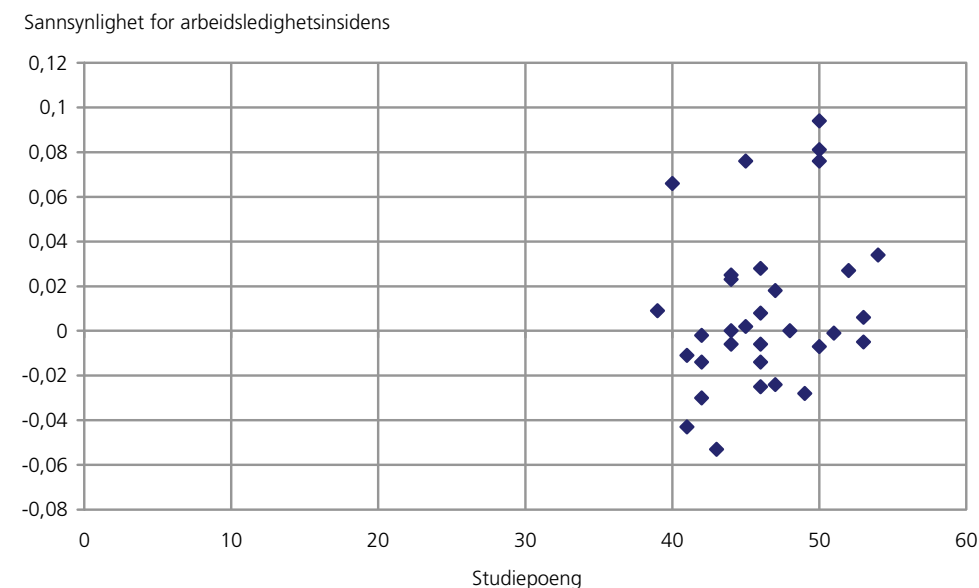
Figur 5.4. Sammenhengen mellom studiepoeng og lønnsforskjeller ett år etter avsluttet utdanning



Grader produsert på normert tid er heller ingen god indikator på hvor mye studenter fra disse institusjonene får i lønn. Her er korrelasjonen faktisk negativ. En høy andel kandidater på normert tid trekker i retning av lavere lønn. Denne effekten drives imidlertid i noen grad av ekstremobservasjonen som har høyest andel av kandidatene på normert tid, men lavest lønn. Når denne observasjonen utelates snur korrelasjonen mellom lønn og avlagte grader til å bli svakt positiv.

⁸ Studiepoengsindikatoren er basert på samme utvalget av uhgrupper som lønnsindikatoren.

Figur 5.5. Sammenhengen mellom studiepoeng og arbeidsledighet i løpet av fem første år etter avsluttet utdanning



Hva så med ledighet? Spesifikasjon 5 i tabell 5.14 rapporterer resultater hvor den avhengige variabelen er ledighetsinsidens i løpet av de fem første årene etter avsluttet utdanning. Estimaten er basert på en ikke-lineær sannsynlighetsmodell (probit) der vi igjen kontrollerer for kjønn, alder ved avlagt grad, år graden ble avlagt og fagfelt/uh-gruppe graden tilhører. I gjennomsnitt opplever 13,8 prosent av de studentene som er analysert her, arbeidsledighet en eller annen gang i løpet av de fem første årene etter avlagt eksamen. Også her er det betydelig variasjon på tvers av institusjonene. Differansen mellom laveste og høyeste verdi er hele 15 prosentpoeng.

Korrelasjonene mellom indikatorer basert på lønn og ledighet er negativ som forventet (rundt -0,25). Korrelasjonen mellom ledighet og avlagte studiepoeng er positiv (0,27). Dette er motsatt av hva man skulle forvente. Institusjoner som i gjennomsnitt produserer mange studiepoeng har, alt annet likt, studenter som i høyere grad opplever arbeidsledighet i løpet av sine fem første år på arbeidsmarkedet. Figur 5.5 illustrerer dette grafisk.

Korrelasjonen mellom produserte grader på normert tid og sannsynlighet for ledighet er imidlertid negativ (-0,18).

Når vi kontrollerer for næring og regionale lønnsforskjeller, reduseres forskjellene mellom institusjonene noe, men hovedbildet er det samme.

Vi har også sett nærmere på korrelasjonen mellom indikatorer basert på studiepoengsproduksjon og fullføring med indikatorer basert på arbeidsmarkedsutfall separat for spesifikke utdanningstyper, nemlig sykepleier-, almennlærer- og ingeniøruddanningen.

Tabell 5.18 viser resultatene for sykepleierutdanningen. Det er store variasjoner både i lønn og arbeidsledighetsinsidens. Lønna varierer med 9 prosentpoeng mellom ”beste” og ”verste” institusjon. Sannsynligheten for å ha opplevd ledighet varierer med 13 prosentpoeng mellom beste og verste institusjon. Korrelasjonen mellom lønn og ledighet er som forventet også her negativ (-0,20), mens korrelasjonen med studiepoeng (0,12) og produserte grader (-0,04) er nær null. Korrelasjonen mellom både henholdsvis ledighet og studiepoeng og ledighet og grader er relativt svake og har ’feil fortegn’ (0,08 og 0,18).

Tabell 5.16. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (Hslid=5), kontrollert for næring og fylke

	Lønn (5. år)	Lønn (5. år)	Lønn (5. år)	Lønn (5. år)
		Kontroll for fylke	Kontroll for næring	Kontroll for næring/fylke
Spes.	1	2	3	4
Hslid				
24	-0,230	-0,199	-0,136	-0,107
28	-0,107	-0,062	-0,051	-0,015
19	-0,070	-0,005	-0,053	0,005
14	-0,063	0,003	-0,052	0,006
25	-0,057	0,008	-0,047	0,008
2	-0,051	-0,004	-0,032	0,015
4	-0,046	-0,013	-0,033	-0,003
23	-0,043	-0,017	-0,017	0,010
13	-0,040	0,017	-0,027	0,027
36	-0,037	0,025	-0,029	0,025
1	-0,036	-0,009	-0,008	0,016
18	-0,036	0,010	-0,022	0,015
35	-0,035	0,006	-0,024	0,005
7	-0,035	0,022	-0,022	0,023
32	-0,035	0,028	-0,027	0,027
29	-0,035	0,015	-0,010	0,031
38	-0,032	0,009	-0,029	0,023
39	-0,032	0,026	-0,022	0,023
17	-0,027	0,032	-0,005	0,039
9	-0,026	0,034	-0,010	0,036
37	-0,026	0,010	-0,006	0,024
31	-0,025	-0,011	-0,023	0,019
15	-0,017	0,016	-0,021	0,007
10	-0,015	0,012	0,012	0,034
20	-0,011	0,041	-0,001	0,032
30	-0,004	-0,001	-0,003	0,019
6	-0,001	-0,002	0,012	0,013
5	0,000	0,000	0,000	0,000
21	0,009	-0,014	0,019	0,005
16	0,011	0,003	-0,003	0,027
12	0,045	0,031	0,121	0,117
33	0,130	0,167	0,114	0,148
Laveste verdi	-0,230	-0,199	-0,136	-0,107
Høyeste verdi	0,130	0,167	0,121	0,148
Differanse mellom laveste og høyeste verdi	0,360	0,366	0,257	0,255
Observasjoner	33039	33039	33039	33039
Antall kull	5	5	5	5
Forklaringskraft i modell uten institusjonseffekter	0,441	0,465	0,527	0,541
Forklaringskraft i modell med institusjonseffekter	0,451	0,471	0,532	0,544
Differanse i forklaringskraft	0,010	0,006	0,005	0,003

Kontrollvariable er inkludert i alle spesifikasjoner, som rapportert i teksten. Forklaringskraft i ledighetsspesifikasjonen referer til pseudo R².

Tabell 5.17. Korrelasjonsmatrise til parametervektor presentert i tabell 5.16

	Spes 1	Spes 2	Spes 3	Spes 4
Spes 1	1,00			
Spes 2	0,89	1,00		
Spes 3	0,91	0,75	1,00	
Spes 4	0,91	0,89	0,92	1,00

Tabell 19 viser resultatene for allmennlærerutdanningen. Variasjonene i lønn på tvers av institusjonene er betydelig mindre enn for sykepleierutdanningen og bidrar bare til en marginal økning i forklaringskraften på 0,3 prosent. Det er derfor liten grunn til å legge vekt på disse forskjellene. Det er imidlertid ganske store (og sterkt statistisk signifikante) forskjeller i ledighetsinsidens. Sannsynlighet for ledighet varierer med 11 prosentpoeng mellom 'beste' og 'verste' institusjon. Samvariasjonen mellom ledighet og studiepoeng er lav (-0,07), men har 'riktig

fortegn'. Samvariasjonen mellom ledighet og grader er noe høyere og har 'feil
fortegn' (0,17).

Tabell 5.18. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (Hslid=5), sykepleierutdanningen

Hslid	Lønn (femte året)	Ledighet (i løpet av fem år)	Studie- poeng (2005)	Grad (2005)
14	-0,087	0,036	43	0.64
19	-0,084	0,029	50	0.58
7	-0,078	-0,037	47	0.68
25	-0,077	-0,01	56	0.88
2	-0,075	0,062	50	0.76
39	-0,072	-0,016	53	0.62
17	-0,068	0,013	48	0.51
31	-0,068	0,01	59	0.79
13	-0,067	-0,03	37	0.46
16	-0,065	-0,007	50	0.87
4	-0,062	0,093	44	0.73
38	-0,06	-0,016	45	0.56
9	-0,052	-0,024	51	0.61
28	-0,047	-0,015	54	0.86
15	-0,047	-0,013	51	0.39
10	-0,047	0,04	47	0.63
29	-0,045	-0,02	32	0.74
37	-0,041	-0,004	49	0.41
1	-0,041	-0,014	52	0.67
20	-0,028	0,021	67	0.84
6	-0,028	-0,022	46	0.64
5	0	0	49	0.66
Laveste verdi	-0,087	-0,037	32	0.39
Høyeste verdi	0	0,093	67	0.88
Differanse mellom laveste og høyeste verdi	0,087	0,13	35	0,49
Observasjoner	3476	10272	11791	3097,00
Forklaringskraft i modell uten institusjonseffekter	0,156	0,011		
Forklaringskraft i modell med institusjonseffekter	0,167	0,044		

Kontrollvariable er inkludert i alle spesifikasjoner, som rapportert i teksten. Forklaringskraft i ledighetsspesifikasjonen referer til pseudo R².

Tabell 5.19. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (Hslid=5), allmennlærerutdanningen

Hslid	Lønn (femte året)	Ledighet (i løpet av fem år)	Studiepoeng (2005)	Grad (2005)
19	-0,052	0,086	52	0.37
2	-0,033	0,042	48	0.66
36	-0,029	0,038	54	0.32
13	-0,025	-0,011	48	0.17
16	-0,025	-0,005	53	0.47
25	-0,024	0,055	50	0.53
39	-0,022	0,034	54	0.5
35	-0,02	0,005	48	0.08
9	-0,019	0,054	51	0.61
15	-0,018	-0,012	58	0.46
31	-0,017	0,011	55	0.49
29	-0,016	0,002	51	0.56
4	-0,01	0,025	51	0.33
1	-0,008	-0,022	53	0.4
20	-0,003	0,048	59	0.25
38	-0,001	0,021	58	0.14
5	0	0	55	0.5
Laveste verdi	-0,052	-0,022	48	0.08
Høyeste verdi	0	0,086	59	0.66
Differanse mellom laveste og høyeste verdi	0.052	0,108	11	0,58
Observasjoner	4030	6594	9668	1855,00
Forklaringskraft i modell uten institusjonseffekter	0.305	0,008		
Forklaringskraft i modell med institusjonseffekter	0.308	0,023		

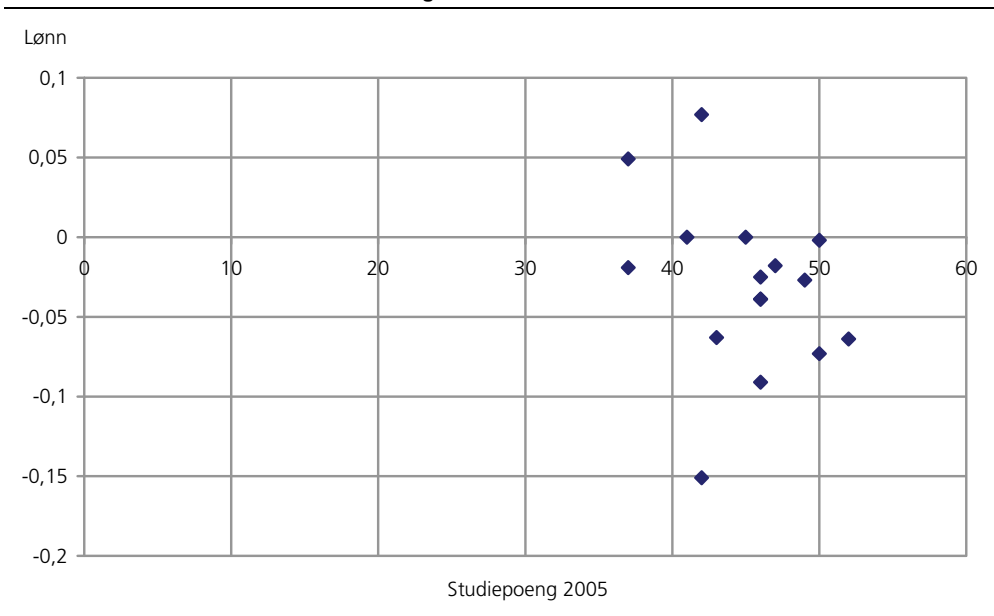
Kontrollvariable er inkludert i alle spesifikasjoner, som rapportert i teksten. Forklaringskraft i ledighetsspesifikasjonen referer til pseudo R².

Tabell 5.20. Arbeidsmarkedsutfall, relativt til referanseinstitusjonen (Hslid=5), ingeniørutdanningen

Hslid	Lønn (femte året)	Ledighet (i løpet av fem år)	Studiepoeng (2005)	Grad (2005)
19	-0,151	0,147	42	0.65
28	-0,091	0,106	46	0.53
1	-0,073	0,071	50	0.8
4	-0,064	0,05	52	0.63
2	-0,063	0,055	43	0.65
25	-0,039	0,07	46	0.72
13	-0,039	0,15	46	0.53
7	-0,027	-0,011	49	0.61
37	-0,025	0,009	46	0.64
9	-0,019	0,035	37	0.38
39	-0,018	0,02	47	0.66
10	-0,002	-0,048	50	0.68
15	0	-0,001	41	0.54
5	0	0	45	0.61
31	0,049	-0,052	37	0.26
16	0,077	-0,006	42	0.55
Laveste verdi	-0,151	-0,052	37	0.26
Høyeste verdi	0,077	0,150	52	0.80
Differanse mellom laveste og høyeste verdi	0,228	0,202	15	0,55
Observasjoner	4253	6337	7986	1725
Forklaringskraft i modell uten institusjonseffekter	0,086	0,023		
Forklaringskraft i modell med institusjonseffekter	0,113	0,027		

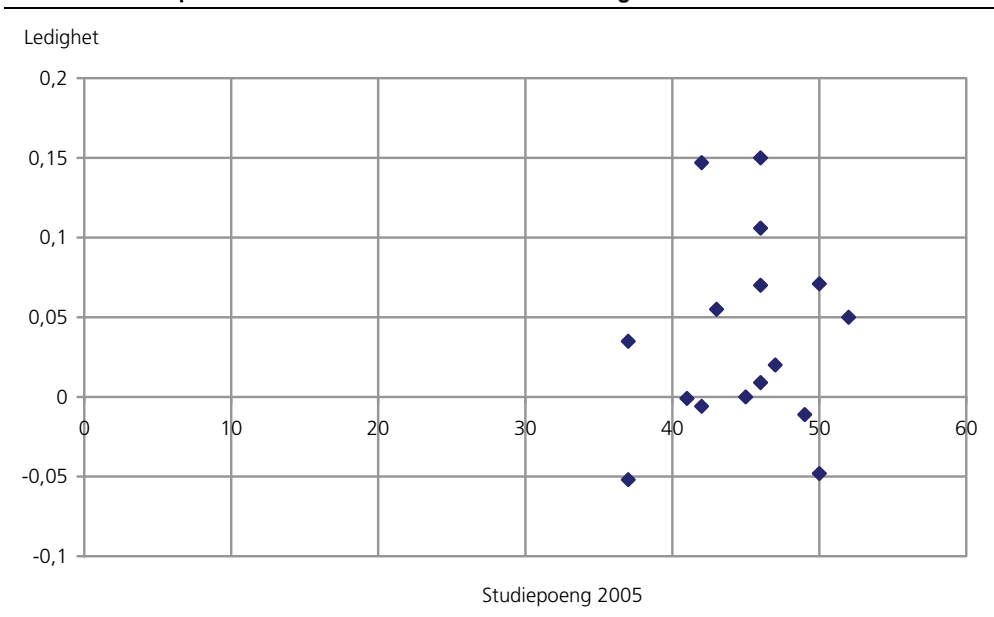
Kontrollvariable er inkludert i alle spesifikasjoner, som rapportert i teksten. Forklaringskraft i ledighetsspesifikasjonen referer til pseudo R².

Figur 5.6. Ingeniørutdanningen: Sammenhengen mellom studiepoeng og lønnsforskjell fem år etter avsluttet utdanning



Tabell 5.20 viser resultatene for ingeniørutdanningen. Alle ingeniørutdanninger på lavere grad er inkludert (sivilingeniører er altså ikke inkludert). Institusjonseffektene forklarer betydelig mer av variasjonen i lønn for ingeniører enn for både lærere og sykepleiere. Dette er ikke overraskende siden ingeniører i større grad enn sykepleiere og lærere har innslag av individuell lønnsfastsettelse, og i større grad arbeider i privat sektor hvor det er større forskjeller i lønnsnivå mellom arbeidsgivere. Videre viser indikatorer basert på lønn og ledighet en stor grad av samvariasjon (-0,77). Men gunstige arbeidsmarkedsutfall går ikke sammen med verken høy grad av studiepoengsproduksjon eller høy grad av gjennomstrømning. Tvert i mot ser det ut til å være slik at høy gjennomsnittlig studiepoengsproduksjon går sammen med lavere lønn og høyere sannsynlighet for ledighet. Dette er illustrert i figur 5.6 og figur 5.7.

Figur 5.7. Ingeniørutdanningen: Sammenhengen mellom studiepoeng og arbeidsledighet i løpet av fem første år etter avsluttet utdanning



6. Konklusjoner

Formålet med denne rapporten har vært å drøfte ulike resultatindikatorer for høyere utdanningsinstitusjoner i Norge, som kan beregnes ut fra data som er tilgjengelige på institusjonsnivå. Spørsmålet er hvorvidt disse resultatindikatorerne – eller mer presist forskjeller i dem mellom utdanningsinstitusjoner – sier noe om faktiske forskjeller mellom institusjonene i forhold til i hvilken grad de bidrar til studentenes læringsutbytte. Vi har begrenset oss til å se på to slike indikatorer, studiepoeng pr. student, og andelen av grader som fullføres på normert tid.

Når man skal vurdere dette spørsmålet, er det særlig to forhold som må vurderes, og som vi har sett spesielt på i denne rapporten. Det første spørsmålet er hvorvidt forskjeller i indikatorerne mellom institusjoner *skyldes* forhold som institusjonene selv er herre over, eller om de i stor grad påvirkes av forhold som ikke har noe med hvordan institusjonene faktisk drives. Det andre spørsmålet er hvorvidt resultatindikatoren faktisk sier noe om resultatet av utdanningen i henhold til formålet.

For å belyse det første spørsmålet, har vi sett nærmere på i hvor stor grad ”ukorrigerede” resultatindikatorer samvarierer med indikatorer basert på de samme resultatmålene som tar hensyn til at utdanningsinstitusjoner rekrutterer studenter med ulik bakgrunn og forutsetninger. Som mål på studentenes bakgrunn og forutsetninger, benytter vi informasjon om foreldrenes utdanningsnivå og/eller studentenes gjennomsnittskarakterer fra videregående skole.

Det kan diskuteres hvorvidt studentsammensetningen langs disse dimensjonene er utenfor institusjonenes kontroll og hvorvidt man dermed bør korrigere for disse når man er ute etter ”kvalitetsforskjeller” mellom institusjonene. Studenter søker seg til ulike skoler. Det er ikke urimelig å anta at de til en viss grad er klar over skolens faktiske kvalitet og at dette påvirker deres søkeatferd. Dette kan skape en positiv korrelasjon mellom skolens faktiske kvalitet og sammensetningen av studentmassen. Når man da korrigerer skolens resultater for forskjeller i student-sammensetning, risikerer man dermed å foreta en overkorreksjon, hvis sorteringen av studenter er så sterk at ikke alle typer studenter er representert på alle skoler. Da vil deler av resultatforskjellene som egentlig skal tilordnes institusjonenes bidrag, feilaktig bli tilordnet sammensetningen av studentmassen. I tillegg kommer det faktum at det kan finnes studentkjennetegn som vi ikke observerer som kan påvirke både studentenes valg av skole og deres resultater. Så lenge koblingen mellom student og utdanningsinstitusjon foregår ved seleksjon etter søknad fra studentenes side, er det vanskelig fullt ut å separere de to kildene til resultatvariasjon. For med større sikkerhet å kunne gi estimerte institusjonseffekter tolkningen av kvalitetsforskjeller, trenger man en form for ”eksogen variasjon” i fordelingen av studenter mellom institusjoner, hvor det er en viss grad av tilfeldig allokering av studenter til ulike utdanningsinstitusjoner. Slik variasjon vil man kunne finne, men analyser av denne typen har mer karakter av forskning, og er ikke særlig aktuell for produksjon av statistikk og indikatorer. Se Hægeland, Raaum og Salvanes (2007) for en diskusjon av dette og et eksempel på en slik analyse når det gjelder sammenhengen mellom ressursbruk og resultater i ungdomsskolen. Selv om man skal være varsom med å gi de indikatorerne vi har beregnet og presentert i denne rapporten en kausal tolkning, vil de likevel kunne gi verdifull informasjon og en større forståelse av hvilke faktorer som har betydning når det gjelder å forklare forskjeller i resultater mellom utdanningsinstitusjoner

Vi finner at det er relativt stor grad av samsvar mellom ukorrigerede indikatorer og indikatorer som korrigerer for forskjeller i studentsammensetning (målt ved foreldres utdanning og/eller studentenes karaktergjennomsnitt fra videregående skole), når vi ser på resultatindikatorer basert på studiepoengsproduksjon og fullføring av grader på normert tid. Resultatindikatorer som korrigerer for dette, gir altså stort sett det samme bildet som de ukorrigerede indikatorerne, selv om

korreksjonen er betydelig for enkelte institusjoner. Det er systematiske forskjeller mellom fagfelt når det gjelder disse resultatmålene, og siden fagsammensetningen varierer betydelig på tvers av utdanningsinstitusjonene, er det viktig å korrigere for dette når man sammenligner institusjonenes resultater. Et problem man har med dagens datasituasjon er at man ikke har noen fullgod måte å identifisere deltidsstudenter, dvs. studenter som *planlegger* å ha en lavere studieprogresjon. I den grad deltidsstudenter fordeler seg ujevnt på institusjonene, vil dette kunne påvirke resultatene. Vi har forsøkt å kontrollere for forekomsten av deltidsstudenter ved å inkludere ”antall år siden studenten første gang ble registrert i høyere utdanning” som en kontrollvariabel i analysene. Begrunnelsen for dette er at antall avlagte studiepoeng pr. år samvarierer sterkt med denne variabelen (færre studiepoeng jo lenger studenten har vært i høyere utdanning), og at dette trolig fanger opp distinksjonen mellom heltid og deltid i noen grad. Den er imidlertid ikke en fullgod kontrollvariabel. Alt annet like vil studenter som er forsinket i forhold til normert tid, ha hatt flere år i høyere utdanning enn andre studenter, selv om de er heltidsstudenter. I en analyse av resultatforskjeller i forhold til fullføring på normert tid risikerer man dermed å korrigere for heltidsstudenter som er forsinket, og ikke bare for deltidsstudenter. Dette gir en overkorreksjon som bidrar til å undervurdere faktiske resultatforskjeller. Det er vanskelig å skille mellom en heltidsstudent som er litt forsinket, og en deltidsstudent som bruker store deler av (men ikke hele sin tid på studier) med de dataene vi har tilgjengelig. Det er ikke opplagt hvordan modellen bør spesifiseres på dette punktet. I fravær av pålitelig informasjon om ”studieintensitet” er det selvsagt mulig å prøve ut andre indirekte mål, for eksempel gjennom å se på studentenes deltakelse i arbeidsmarkedet under studiene. Heller ikke dette vil imidlertid kunne gi noen fullgod inndeling i heltids- og deltidsstudenter.

Til tross for at a) familiebakgrunn og karakterer fra videregående skole forklarer en del av den individuelle variasjonen i studiepoeng og b) det er en systematisk sortering av studenter på tvers av institusjoner, synes det altså ikke å ha stor betydning å kontrollere for familiebakgrunn når faste institusjonseffekter beregnes.

Dette er et interessant resultat, men det bringer oss over til det andre spørsmålet vi har drøftet i denne rapporten, nemlig i hvilken grad disse resultatmålene sier noe om forskjeller mellom institusjonene når det gjelder å oppfylle det ”egentlige” formålet med virksomheten. Dette spørsmålet kan deles i to: Gir det aktuelle resultatmålet en enhetlig måling av prestasjoner på tvers av institusjoner, og er disse prestasjonene det som faktisk verdsettes med utdanningen? Svaret på begge spørsmål må være ja, for at vi skal kunne si at indikatoren oppfyller dette ”kravet”.

Karakterer (som ligger bak begge våre resultatmål gjennom distinksjonen mellom bestått og ikke bestått) settes i stor grad av institusjonene selv. Det er i liten grad krav om ekstern sensur, og få mekanismer som sørger for en enhetlig karaktersetting. I tillegg har institusjonene interesser i at flest mulig skal bestå og gjennomføre på normert tid, siden dette er med på å påvirke bevilgningene. Dette *kan* føre til at karakterpraksisen blir ulik, og at karakterer ikke er sammenlignbare på tvers av institusjoner, selv om de fungerer i forhold til å rangere studenter på samme institusjon. Tilsvarende kan man tenke seg at det finnes ulik praksis på hva som skal til for å stå på eksamen. For eksemplets skyld kan vi tenke oss en situasjon der alle institusjonene har som tommelfingerregel at 10 % av studentene skal stryke på hver eksamen. Da vil det typisk være slik at selv om familiebakgrunn forklarer en del av variasjonen i studiepoeng, har det ingen betydning de estimerte institusjonseffektene om disse er inkludert eller ikke. I den grad dette er tilfelle, er resultatmål som baserer seg på karakterer vanskelige å benytte til sammenligning av institusjoner, uavhengig om man korrigerer for elevsammensetning eller ikke.

Det finnes klare indikasjoner på at kriteriene for karaktersetting som benyttes av ulike institusjoner for å vurdere studentenes prestasjoner, tilpasses kvaliteten på studentmassen. En interessant studie i så henseende er Møen og Tjelta (2005). De

benytter data fra Norges handelshøyskole, hvor det er mulig å søke opptak til tredje året ved siviløkonomstudiet på bakgrunn av to års studier i økonomisk-administrative fag ved andre høyskoler. Opptaksgrunnlaget er karakterene fra disse studiene. Når de kontrollerer for opptakspoeng, finner de systematiske forskjeller mellom ”opprinnelsesskoler” i studentenes prestasjoner på siviløkonomstudiet. Dette tyder på ulik karakterpraksis ved de ulike institusjonene. Hvis det var en enhetlig karakterpraksis på tvers av skoler, slik at den samme prestasjonen ble belønnet med den samme karakteren, ville man ikke finne slike systematiske forskjeller. Gitt at dette resultatet har gyldighet også for andre typer av høyere utdanning, er indikatorer basert på karakterer ikke velegnet til å studere kvalitetsforskjeller mellom norske høyere utdanningsinstitusjoner.

Strategien som er benyttet i Møen og Tjelta (2005) har mye til felles med den strategien vi benytter når vi ser på forskjeller i arbeidsmarkedsutfall mellom studenter fra ulike utdanningsinstitusjoner. Ideen er å utsette studenter fra ulike institusjoner for en ”felles eksamen” med enhetlig sensur (det ideelle fra vårt isolerte perspektiv ville selvsagt vært nasjonale eksamener for ulike studier, men slike eksisterer ikke). I Møen og Tjeltas studie var den felles vurderingen eksamener på siviløkonomstudiet ved NHH, i vårt tilfelle benytter vi arbeidsmarkedet som vårt vurderingsinstrument. Vi finner at det er betydelige forskjeller i arbeidsmarkedsutfall mellom studenter fra ulike utdanningsinstitusjoner. Hvis man godtar arbeidsmarkedet som en evalueringsinstans med hensyn til utdanningskvalitet, indikerer dette at det er betydelige forskjeller mellom utdanningsinstitusjoner. Det bør bemerkes at disse beregningene er av en viss ”eksplorativ karakter”.

Den viktigste lærdommen vi kan trekke fra disse beregningene er imidlertid ikke de faktisk estimerte forskjellene mellom utdanningsinstitusjonene, men at indikatorene basert på arbeidsmarkedsutfall bare er svakt korrelert med indikatorene basert på studiepoengsproduksjon og grader fullført på normert tid. I mange tilfeller er det til og med en negativ korrelasjon. Hvis begge typer indikatorer delvis fanger opp ”utdanningskvalitet”, skulle man forvente at de var positivt korrelerte. Fraværet av positiv korrelasjon, i tillegg til de forholdene vi har diskutert ovenfor, er ytterligere en indikasjon på at resultatindikatorer som baserer seg på data som har sitt utspring i karakterer satt av de enkelte utdanningsinstitusjonene er problematiske som ”kvalitetsindikatorer”. De inneholder mye informasjon som er relevant i mange sammenhenger, men man bør være forsiktig med å tolke dem som indikatorer som kvalitetsforskjeller mellom utdanningsinstitusjoner. For at indikatorer skal kunne gis en slik tolkning, er et minstekrav at resultatmålene de bygger på er fastsatt på en måte som er konsistent på tvers av de enhetene som sammenlignes. Dette synes ikke å være tilfelle med studiepoeng og grader fullført på normert tid.

Referanser

Dam, K.W og K.Ø. Sørensen (2006) : Volume measures of education in the Norwegian National Accounts, paper presentert på konferansen "Measurement of non-market output in education and health", London, oktober 2006.

Edvardsen, D.F og F.R Førstund (2001): De statlige høyskolene som produsenter: Ressursbruk og resultater 1994-1999. Rapport 3/2001, Frischsenteret.

Ervik, A.O., E. Holmøy og T. Hægeland (2002): A Theory-Based Measure of the Output of the Education Sector. Discussion Papers 353, Statistisk sentralbyrå.

Hægeland, T., L.J. Kirkebøen, O. Raaum og K.G. Salvanes (2004): Karakterforskjeller mellom norske ungdomsskoler: Hva kan forklares med forskjeller i elevsammensetning og ressursbruk. Rapporter 2004/11, Statistisk sentralbyrå.

Hægeland, T., L.J. Kirkebøen, O. Raaum og K.G. Salvanes (2007): Skolebidragsindikatorer. Beregnet for avgangskarakterer fra grunnskolen for skoleårene 2002-2003 og 2003-2004. Rapporter 2005/33, Statistisk sentralbyrå.

Hægeland, T., O. Raaum og K.G. Salvanes (2005): Pennies from heaven. Using exogenous tax variation to identify effects of school resources on pupil achievement, Discussion Papers 508, Statistisk sentralbyrå.

Kittelsen, S.A.C. og F.R Førstund (2001) Empiriske forskningsresultater om effektivitet i offentlig tjenesteproduksjon, *Økonomisk Forum* nr. 6, 22-29.

Moen, J. og M. Tjelta (2005): Bruker ulike høyskoler karakterskalaen ulikt? En analyse av sammenhengen mellom skolebakgrunn og faglig suksess ved NHH. *Økonomisk Forum* nr. 6 2005, 33-45.