Høgskulen på Vestlandet

Nærregion Sogn og Fjordane

Lærarkompetanse og skuleresultat – metoden bak resultata

Terje Myklebust og
Anne Norstein

N-NR 5/2017
Høgskulen på Vestlandet, Nærregion Sogn og Fjordane

Høgskulen på Vestlandet

NOTAT

Postboks 133, 6851 SOGNDAL telefon 57676000 telefaks 57676100

TITTEL	NOTATNR.	DATO
Lærarkompetanse og skuleresultat – metoden bak resultata	5/17	20.06.17
PROSJEKTTITTEL	**TILGJENGE**	**TAL SIDER**
Lærarkompetanse og skuleresultat	Open	35
FORFATTAR	**PROSJEKTLEIAR/-ANSVARLEG**	
Terje Myklebust og Anne Norstein	Anne Norstein	
OPPDRAGSGJEVAR	**EMNEORD**	
Høgskulen i Sogn og fjordane	Skuleresultat, Karakterar Kompetanse	

SAMANDRAG

I forskingsprosjektet Lærande regionar – Lærarkompetanse og skuleresultat såg vi på om klassestorleik samt lærarens erfaring og utdanning i matematikk har betydning for elevane sine prestasjonar i matematikk målt ved avgangseksamen i 10. klasse. Eit spørjeskjema vart sendt til alle skular i fylka Aust-Agder, Oppland, Nord-Trøndelag og Sogn og Fjordane. I tillegg har vi offentleg tilgjengelege data frå skoleporten.no som vi blant anna har nytta til å vurdere kvaliteten på dataa frå denne spørjeundersøkinga. Her gjer vi greie for detaljane i vurderingane og analysane som ligg bak konklusjonane som er publiserte i antologien *Skolens kvalitet skapes lokalt* (Langfeldt, 2015). Vi finn negativ effekt av klassestorleiken (store klassar er negativt) og relativt sterk effekt av lærarens studiepoeng i matematikk. Derimot ser vi ingen effekt av lærarens erfaring.

PRIS	ISSN	ANSVARLEG SIGNATUR
	0806-1696	20.6.2017 Osvald Lykkebø

Lærarkompetanse og skuleresultat – metoden bak resultata

Terje Myklebust og Anne Norstein

Forord

Dette notatet inneheld ein gjennomgang av metoden vi nytta i forskingsprosjektet Lærarkompetanse og skuleresultat. Prosjektet var ein del av eit større samarbeidsprosjekt, Lærande regionar, som var delfinansiert av Noregs Forskingsråd. Målet for forskingsprosjektet var å prøve å finne forklaringa på dei gode skuleprestasjonane i Sogn og Fjordane.

I delprosjektet Lærarkompetanse og skuleresultat ville vi undersøke om ei mogleg forklaring på dei gode elevprestasjonane i Sogn og Fjordane kunne ha samanheng med eit høgt utdanningsnivå hjå lærarane i fylket. Tidlegare studiar har samanlikna elevprestasjonar med lærarkompetanse generelt på skulen, men utan å kople kompetanse direkte mot den enkelte klasse.

Forskingsprosjektet byggjer på innsamla data over karakterar i matematikk ved avgangsprøva i 10. klasse, og kompetansen til lærarane som underviste elevane på ungdomssteget. Vi samla inn data frå skular i Sogn og Fjordane, Aust-Agder, Oppland og Nord-Trøndelag.

Resultata vart tidlegare publiserte i antologien Skolens kvalitet skapes lokalt (Langfeldt, 2015). I dette notatet går vi nærare inn på korleis vi analyserte datamaterialet, og dei val og avvegingar vi måtte gjera i dette analysearbeidet.

Sogndal, Juni 2017

Terje Myklebust og Anne Norstein

Innhaldsliste

Innleiing

I tidsrommet 2012-2015 deltok vi i det Nfr-finansierte prosjektet Lærande regionar. Forskingsprosjektet var eit samarbeid mellom Høgskulen i Sogn og Fjordane, Universitetet i Agder, Høgskulen i Hedmark og Høgskulen i Nord-Trøndelag. Målet med prosjektet var å undersøka kvifor elevar frå Sogn og Fjordane presterer så godt på Nasjonale prøvar, samanlikna med andre fylke som kan sjå ut til å ha dei same føresetnadene. Resultata vart publiserte i antologien *Skolens kvalitet skapes lokalt* (Langfeldt, 2015). I vårt bidrag til antologien tona vi ned den metodiske delen, og presenterte ei forenkla utgåve av metoden. I dette notatet gjer vi greie for og utdjupar dei statistiske vurderingane og analysane som vi har gjort i prosjektet, sjå Myklebust & Norstein (2015).

Vi ser på den gjennomsnittlege utdanninga i antal studiepoeng og erfaringa målt i antal undervisningsår hjå lærarane i åttande, niande og tiande klasse. Vi ser og på klassestorleik.

For å få data som koplar lærars utdanning til klassens resultat må vi samle data sjølve. Ingen registerdata har denne koplinga. Med registerdata kan ein i beste fall kople utdanningsnivået hjå lærarane på skulen mot skulens resultat. Dette er mykje dårlegare datasett for å observere ein eventuell samanheng mellom lærars utdanning og elevens prestasjon. Bakdelen med å samle data sjølv er blant anna at vi kan få ein låg svarprosent.

Om datasettet

I dette arbeidet har vi fem fleire observasjonar enn i Myklebust & Norstein (2015). Dette er observasjonar som vi fekk inn etter at arbeidet med Myklebust & Norstein (2015) var gjennomført. No har vi 138 observasjonar. Vi har også gått gjennom observasjonane og retta nokre feilrapporterte/feiltasta resultat i datasettet vårt. Samla sett har desse endringane svekka resultata litt, men hovudkonklusjonen er ikkje endra.

Gjennomsnittleg antal studiepoeng varierer frå 0 til 90 studiepoeng, i tillegg til eit datapunkt med 226 studiepoeng. I praksis vert sist nemnde datapunktet meiningslaust, sidan det ligg så langt unna majoriteten av datasettet. Avhengig av modellføresetnader kan dette punktet få uforholdsvis stor betydning, og vi har rimelegvis ikkje grunnlag for å vite noko om samanhengen der vi ikkje har data. Vi vel derfor å sette utdanningsnivået til denne respondenten til 100, alternativet er å ta han vekk. Vi gjer merksam på at karakteren til denne klassen er 2,9 som er under gjennomsnittet. I sju av klassane har det vore fleire lærarar involvert i same undervisingsår. I nokre av desse tilfella kan det sjå ut som at klassen har fått ny lærar i løpet av året, medan det i andre tilfelle kan sjå ut som det har vore fleire lærarar til stades i timane. I desse tilfella har vi valt å velje utdanninga og erfaringa til den av lærarane som har lengst utdanning. Tankegangen er at der det er to lærarar i timane, vil den best utdanna sette fagleg standard. Vi har likevel sett på konsekvensane av å heller velje kortast utdanning til dei to involverte lærarane i klassen. Det endrar ikkje resultatet vesentleg, men i den grad det vert endra, kjem det ut litt sterkare. For elleve av klassane manglar vi antal år læraren har undervist. I desse tilfella har vi forsøkt å sette erfaringa til gjennomsnittserfaringa i datamateriale, og vi har sett på kva resultatet vert om vi tar ut desse datapunkta frå materialet. Begge vala gir same resultat. Ingen effekt av lærarens erfaring.

I dette forsøket har vi svært god informasjon om karakterane på eksamenane i dei ulike fylka. Dette er informasjon som går langt utover vårt eige datasett. Det er relativt sjeldan at vi har så god informasjon utover innsamla data, og vi kan bruke informasjonen til å vurdere og justere våre data. I vår analyse brukar vi ein regresjonsmodell. Med så gode bakgrunnsdata trur vi blant anna at det er verdt å først forsøke å justere dataa for års- og fylkeseffekt. Alternativt kan vi bruke dummy-variablar for år og fylke i regresjonsmodellen.

La Y_{il} vere karakteren til ein tilfeldig elev l i år $i = 9, \dots, 13$. Vi antar at Y_{il} har ei miksa fordeling der vi tar omsyn til fylkeseffekten. Vi antar at Y_{il} har forventning μ_i og standardavvik $\sigma_{Y,i}$, og har tettleiksfunksjonen på forma

$$f\left(x; \mu_i, \sigma_{Y,i}^2\right) = \sum_{j \in B} w_j \phi\left(x; \mu_i + \mu_i^j, \sigma_i^2\right) \qquad (1)$$

der $\phi(\,\cdot\,; a, b)$ er tettleiksfunksjonen til normalfordelinga med forventning a og varians b. Parametrane μ_i^j er eit uttrykk for ulik forventning i dei ulike fylka, $j \in B$, i år $i = 9, \dots, 13$, der B er dei 19 fylka på skoleporten.no bortsett frå Svalbard. I utgangspunktet tenkjer vi at den fylkesvise forventninga kan variere frå år til år. Vi antar altså at variansen er lik i dei ulike fylka, men ikkje at han er lik frå år til år. Vi lar vektene w_j vere uavhengige av året, og vi lar vektene w_j vere lik den gjennomsnittlege andelen av populasjonen over åra 2009 til 2013, for fylke $j \in B$. Merk at talet på elevar som har vore oppe til eksamen i dei ulike fylka synes å vere litt skeivt i forhold til folketalet i dei ulike fylka, men vi kjenner ikkje til om det er skeivt i forhold til elevtalet.

For seinare bruk definerer vi delmengda $A = \{S, A, T, O\}$, der S, A, T og O står for høvesvis Sogn og Fjordane, Aust Agder, Nord Trøndelag og Oppland.

Vi brukar sannsynsmaksimeringsestimatorane

$$\hat{\mu}_i = \sum_{j \in B} w^j \frac{1}{N_i^j} \sum_{l=1}^{N_i^j} Y_{il}^j \,,$$

og den forventningsrette estimatoren

$$\hat{\sigma}_i^2 = \frac{1}{N_i - 19} \sum_{j \in B} \sum_{l=1}^{N_i^j} \left(Y_{il}^j - \bar{Y}_{i\cdot}^j\right)^2,$$

der N_i^j er talet på elevar som har vore oppe til eksamen i år i, og fylke $j \in B$, og $N_i = \sum_{j \in B} N_i^j$. Det vil seie at N_i er talet på elevar oppe til eksamen i år i.

La Y_{il}^j vere karakteren til ein elev l, i fylke j, i år i. Då kan vi skrive

$$Y_{il}^j = \mu_i + \mu_i^j + \varepsilon_{il}^j$$

der variablane ε_{il}^j er antatt å vere uavhengige $N(0, \sigma_i^2)$ fordelte, stor verdi av σ_i kan tolkast som at eksamen har god evne til å skilje elevane. Variabelen ε er eit uttrykk for tilfeldig variasjon mellom elevane på eksamen, gitt år og fylke. Vi skal sjå at det ikkje er grunn for å påstå at μ_i^j er avhengig av året $i = 9, \dots, 13$. Men vi antar likevel at $\mu_i^j = \mu^j \sigma_i$, det vil seie at vi reknar med at i år med stor variasjon mellom elevanes prestasjonar er fylkeseffekten også stor.

Føresetnaden at støyleddet ε_{il}^{j} er uavhengige variablar for $i, j\ og\ l$, held nok ikkje. Observasjonar innanfor same skule kan gjerne vere positivt korrelerte. Dessutan kan same lærar ha hatt fleire klassar same år eller i ulike år, og då er det rimeleg at vi har positiv korrelasjon. Problemet med korrelert støy er truleg ikkje veldig stort i denne delen av analysen, men meir aktuelt når vi seinare ser på våre eigne respondentar.

Talgrunnlaget er henta frå skuleporten						
Heile landet		2009	2010	2011	2012	2013
	N_i	$N_9 = 21059$	$N_{10} = 20488$	$N_{11} = 20900$	$N_{12} = 21561$	$N_{13} = 20701$
	$\hat{\mu}_i$	3,420	3,250	3,137	3,088	3,057
	$\hat{\sigma}_i$	1,254	1,184	1,229	1,223	1,267
Sogn og Fjordane	N_i^S	$N_9^S = 589$	$N_{10}^S = 521$	$N_{11}^S = 555$	$N_{12}^S = 517$	N_{13}^S=516
	$\bar{y}_{i\cdot}^S$	3,606	3,471	3,390	3,268	3,234
	Stdv, $\hat{\sigma}_i^S$	1,206	1,072	1,194	1,212	1,250
Aust-Agder	N_i^A	533	468	441	477	501
	$\bar{y}_{i\cdot}^A$	3,325	3,165	3,086	3,041	2,919
	Stdv, $\hat{\sigma}_i^A$	1,265	1,230	1,311	1,239	1,229
N.-Trøndelag	N_i^T	650	624	709	691	601
	$\bar{y}_{i\cdot}^T$	3,358	3,277	3,188	3,021	2,964
	Stdv, $\hat{\sigma}_i^T$	1,230	1,155	1,153	1,215	1,267
Oppland	N_i^O	727	724	805	855	769
	$\bar{y}_{i\cdot}^O$	3,377	3,247	3,07	3,082	2,935
	Stdv, $\hat{\sigma}_i^O$	1,271	1,241	1,244	1,254	1,312

Tabell 1. Merk at σ_i ikkje er standardavviket til Y_{il} som er $\sigma_{Y,i}$ og er marginalt større, jamfør (1).

Når talet på observasjonar for heile landet er i storleiken 20 000 og standardavvika er i storleiken 1,25, vil feilen i estimatoren $\hat{\mu}_i$ vere innanfor $1.96 \cdot 1{,}25/\sqrt{20\,000} \approx 0{,}017$, dvs at vi har eit svært nøyaktig estimat der vi kan anta at feilen ligg innanfor hundredelar. Tilsvarande kan vi forvente at feilen i standardavvika er innanfor $1.96 \cdot \sqrt{2} \cdot 1{,}25\ /\sqrt{20\,000} \approx 0{,}0245$. Av dette kan vi seie at forskjellane i landsgjennomsnitta for dei ulike åra tyder på at eksamenane har hatt ulik forventing og varians. Det vil seie at eksamenane har hatt ulik vanskegrad, eller det kan tyde på nedgåande kunnskapsnivå hjå elevane (merk at data viser ein nedgåande trend). For vårt arbeid er det ikkje spesielt viktig, og vi har heller ikkje data til å vurdere om det skuldast eksamen eller nedgåande kunnskapsnivå, men i utgangspunktet er det nærliggande å forklare tala med ulik vanskegrad på eksamenane, spesielt sidan tidsperioden er så kort. Merk at dette er rimeleg. Vi kan vel forvente forskjell i vanskegraden på eksamen, og forskjell i varians frå år til år, noko anna ville vere merkeleg.

Vi har antatt at σ_i^2 er lik i alle fylka, stemmer det? Vi sjekkar det estimerte standardavviket i kvart fylke mot standardavviket basert på observasjonar frå heile landet i kvart av dei fem åra, med Fishers test. Den lægste p-verdien finn vi for Sogn og Fjordane i 2010, han er ca 0,06. Den nest lægste p-verdien er 0,12 i Nord Trøndelag i 2011. Dei andre p-verdiane er ikkje i nærleiken av å tyde på forskjellar i standardavvika. Vi har då gjort fem testar i fire fylke, det vil seie 20 testar. Det er med andre ord liten grunn til å tru at standardavviket er forskjellig i dei ulike fylka for nokon av åra.

Talet på observasjonar per fylke i dei fire fylka varierer frå 441 til 855. Det er som nemnt grunn til å tru at vanskegraden til eksamen varierer frå år til år. Men den fylkesvise effekten frå år til år er derimot særdeles stabil. Tabellen under viser differansen mellom fylkesgjennomsnittet og landsgjennomsnittet for dei ulike åra.

	$\bar{y}_{i\cdot}^{j} - \bar{y}_{i\cdot}$ $for\ fylke\ j \in A\ og$ år $i = 9, \ldots, 13$				
i	2009	2010	2011	2012	2013
Sogn og Fjordane	0,186	0,221	0,253	0,180	0,177
Aust-Agder	-0,095	-0,085	-0,051	-0,047	-0,138
N.-Trøndelag	-0,062	0,027	0,051	-0,067	-0,093
Oppland	-0,043	-0,003	-0,067	-0,006	-0,122

Tabell 2

Med data frå skuleporten.no kan vi utføre ein variansanalyse for å vurdere om dei observerte forskjellane på fylkeseffektane frå år til år er signifikante.

Vi startar med Sogn og Fjordane. La $\sigma_{S,i}^2 = \left(\sigma_i^S\right)^2$.

$$SST^S = \sum_{i=9}^{13} \sum_{l=1}^{N_i^S} \left(y_{il}^S - \bar{y}_{\cdot\cdot}^S\right)^2$$

$$SSG^S = \sum_{i=9}^{13} N_i^S \left(\bar{y}_{i\cdot}^S - \bar{y}_{\cdot\cdot}^S\right)^2 \approx 2{,}35$$

$$SSE^S = \sum_{i=9}^{13} \sum_{l=1}^{N_i^S} \left(y_{il}^S - \bar{y}_{i\cdot}^S\right)^2 = \sum_{i=9}^{13} \left(N_i^S - 1\right) \hat{\sigma}_{S,i}^2 = 3820$$

$$F_{(5-1),(N^S-5)} = \frac{\frac{SSG^S}{5-1}}{\frac{SSE^S}{N^S-5}} = \frac{2{,}35/4}{3820/(2698-5)} \approx 0{,}41$$

Der $N^S = N_9^S + \cdots + N_{13}^S = 589 + 521 + 555 + 517 + 516 = 2698$. Det gir ein p-verdi på ca. 0,80. Det vil seie at det er ingen grunn til å tru at Sogn og Fjordane-effekten er ulik frå år til år.

Aust-Agder:

$$SSG^A = \sum_{i=9}^{13} N_i^A \left(\bar{y}_{i\cdot}^A - \bar{y}_{\cdot\cdot}^A\right)^2 \approx 2{,}66$$

$$SSE^A = \sum_{i=9}^{13} \sum_{l=1}^{N_i^A} \left(y_{il}^A - \bar{y}_{i\cdot}^A\right)^2 = \sum_{i=9}^{13} \left(N_i^A - 1\right) \hat{\sigma}_{A,i}^2 \approx 3809$$

$$F_{(5-1),(N^A-5)} = \frac{\frac{SSG^A}{5-1}}{\frac{SSE^A}{N^A-5}} = \frac{2{,}66/4}{3809/(2420-5)} \approx 0{,}42$$

der $N^A = N_9^A + \cdots + N_{13}^A = 2420$

Dette gir ein p-verdi på ca 0,79. Det vil seie at det heller ikkje for Aust-Agder er grunn til å tru at den fylkesviseeffekten varierer frå år til år.

Nord-Trøndelag:

$$SSG^T = \sum_{i=9}^{13} N_i^T \left(\bar{y}_{i\cdot}^T - \bar{y}_{\cdot\cdot}^T\right)^2 \approx 2{,}66$$

$$SSE^T = \sum_{i=9}^{13} \sum_{l=1}^{N_i^T} \left(y_{il}^T - \bar{y}_{i\cdot}^T\right)^2 = \sum_{i=9}^{13} \left(N_i^T - 1\right) \hat{\sigma}_{T,i}^2 \approx 4747$$

$$F_{(5-1),(N^T-5)} = \frac{\frac{SSG^T}{5-1}}{\frac{SSE^T}{N^T-5}} = \frac{2{,}66/4}{4747/(3275-5)} \approx 1{,}83$$

Der $N^T = N_9^T + \cdots + N_{13}^T = 3275$. Dette gir ein p-verdi på ca 0,12. Heller ikkje for Nord-Trøndelag kan vi påstå at fylkeseffekten varierer frå år til år, sjølv om p-verdien her er noko mindre. Men at ein av p-verdiane skulle vere så låg som 0,12 blant dei fire fylke er på ingen måte uventa.

Oppland:

$$SSG^O = \sum_{i=9}^{13} N_i^O \left(\bar{y}_{i\cdot}^O - \bar{y}_{\cdot\cdot}^O\right)^2 \approx 7{,}5$$

$$SSE^O = \sum_{i=9}^{13} \sum_{l=1}^{N_i^O} \left(y_{il}^O - \bar{y}_{i\cdot}^O\right)^2 = \sum_{i=9}^{13} \left(N_i^O - 1\right) \hat{\sigma}_{O,i}^2 \approx 6203$$

$$F_{(5-1),(N^O-5)} = \frac{\frac{SSG^O}{5-1}}{\frac{SSE^O}{N^O-5}} = \frac{\frac{7{,}49}{4}}{\frac{6203}{3880-5}} \approx 1{,}17$$

Der $N^O = N_9^O + \cdots + N_{13}^O = 3880$. Dette gir ein p-verdi på ca 0,32.Det vil seie at Oppland-effekten heller ikkje ser ut til å variere frå år til år.

På bakgrunn av desse høge p-verdiane er det rimeleg å anta at den fylkesvise forventninga er uavhengig av året. Merk at med så mange observasjonar og berre fem grupper (år) er teststyrken rimeleg stor, derfor har vi god grunn tll å tru at fylkeseffekten ikkje varierer stort frå år til år.

Vi gjer likevel justeringa at $\mu_i^j = \mu^j \sigma_i$. Sannsynsmaksimeringsestimatoren (SME) er

$$\hat{\mu}^j = \frac{\sum_{i=9}^{13} \sum_{l=1}^{N_i^j} \frac{Y_{il}^j - \mu_i}{\sigma_i}}{\sum_{i=9}^{13} N_i^j} \quad \text{for j} \in \text{A},$$

og $\hat{\mu}_i^j = \hat{\mu}^j \sigma_i$ som gir litt andre estimat enn i tabell 1. Vi brukar $\hat{\mu}_i$ og $\hat{\sigma}_i^2$ frå tabell 1 for høvesvis σ_i^2 og μ_i, for $i = 9, \cdots, 13$, og ser på verdiane som kjende (gitt det store datamaterialet er ikkje det urimeleg). Vi får då følgjande punktestimat for den fylkesvise effekten og forventa karakter, $\mu_i + \mu_i^j = \mu_i + \mu^j \sigma_i$, for fylka i dei ulike åra.

	Estimat for $\mu_i^j = \mu^j \sigma_i$ for år i og fylke j.	Forventa karakter for fylket i 2009	Forventa karakter for fylket i 2010	Forventa karakter for fylket i 2011	Forventa karakter for fylket i 2012	Forventa karakter for fylket i 2013
Sogn og Fjordane	$0{,}165\sigma_i$	3,628	3,447	3,340	3,291	3,268
Aust-Agder	$-0{,}068\sigma_i$	3,335	3,169	3,054	2,970	3,106
N.-Trøndelag	$-0{,}021\sigma_i$	3,393	3,225	3,111	3,062	3,157
Oppland	$-0{,}038\sigma_i$	3,372	3,090	3,041	3,008	3,142

Tabell 3

I gjennomsnitt er standardavviket over desse åra: $\frac{(\sigma_9 + \cdots + \sigma_{13})}{5} \approx 1{,}231$, der vi brukar estimerte verdiar frå tabell 1. Det vil seie at fylkeseffektane er forventa å vere $0{,}165 \cdot 1{,}231 \approx 0{,}203$ for Sogn og Fjordane, $-0.068 \cdot 1{,}231 \approx -0{,}0837$ for Aust Agder, $-0{,}021 \cdot 1{,}231 \approx -0{,}0259$ for Nord Trøndelag og $-0{,}038 \cdot 1{,}231 \approx -0{,}0468$ for Oppland.

Er respondentane representative?

Med data frå skuleporten.no kan vi vurdere om respondentane ser ut til å vere tilfeldige med omsyn til karakterar. Vi lar n_i^j og m_i^j vere høvesvis talet på elevar i sample og utanfor sample, det vil seie at $m_i^j = N_i^j - n_i^j$ for år $i = 9, \ldots, 13$ og fylke $j \in A$. Vi brukar igjen estimert standardavvik med data frå heile landet som sikker verdi for σ_i for år $i = 9, \ldots, 13$. Talet på observasjonar er så stort at vi brukar normalfordelinga (Z-observatoren) i staden for t-fordelinga (t-observatoren). La $\bar{Y}_{i.}^j$ og $\bar{W}_i^j$ vere gjennomsnittskarakterane for elevar høvesvis blant respondentar og ikkje-respondentar i år $i = 9, \ldots, 13$ og fylke $j \in A$. Vi har at

$$\bar{W}_i^j = \frac{N_i^j E_i^j - n_i^j \bar{Y}_{i.}^j}{m_i^j}$$

der E_i^j er gjennomsnittskarakteren blant alle elevar som har avlagt eksamen i år $i = 9, \dots, 13$ og fylke $j \in A$, henta frå skuleporten.no.

Skilnaden mellom gjennomsnittskarakter hjå respondentar og ikkje-respondentar, tosidig test						
Sogn og Fjordane		2009	2010	2011	2012	2013
	n_i^S	14	98	282	424	177
	m_i^S	575	423	273	93	339
	$\bar{Y}_{i\cdot}^S - \bar{W}_i^S$	0,506	0,245	-0,142	0,345	0,298
	$\sigma_i \sqrt{\frac{1}{n_i^S} + \frac{1}{m_i^S}}$	0,341	0,134	0,105	0,141	0,119
	Z-observ.	1,48	1,83	-1,36	2,45	2,51
	p-verdi	0,14	0,07	0,17	0,01	0,01
Aust-Agder	n_i^A	9	121	129	178	178
	m_i^A	524	347	312	299	323
	$\bar{Y}_{i\cdot}^A - \bar{W}_i^A$	0,127	-0,115	-0,122	0,222	0,203
	$\sigma_i \sqrt{\frac{1}{n_i^A} + \frac{1}{m_i^A}}$	0,424	0,126	0,129	0,117	0,120
	Z-observ.	0,30	-0,91	-0,94	1,90	1,70
	p-verdi	0,76	0,36	0,35	0,06	0,09
Nord-Trøndelag	n_i^T	0	93	142	277	146
	m_i^T	650	531	567	414	455
	$\bar{Y}_{i\cdot}^T - \bar{W}_i^T$	-	0,039	-0,035	-0,018	0,100
	$\sigma_i \sqrt{\frac{1}{n_i^T} + \frac{1}{m_i^T}}$	-	0,134	0,116	0,096	0,122
	Z-observ.	-	0,29	-0,30	-0,19	0,82
	p-verdi	-	0,77	0,76	0,85	0,41
Oppland	n_i^O	0	41	74	64	0
	m_i^O	727	683	731	791	769
	$\bar{Y}_{i\cdot}^O - \bar{W}_i^O$	-	0,183	-0,099	0,041	-
	$\sigma_i \sqrt{\frac{1}{n_i^O} + \frac{1}{m_i^O}}$	-	0,192	0,150	0,160	-
	Z-observ.	-	0,96	-0,66	0,26	-
	p-verdi	-	0,34	0,51	0,80	-

Tabell 4a

Vi ser at p-verdiane for Nord Trøndelag og Oppland ikkje tyder på at vi har fått data frå verken spesielt gode eller dårlege skular. I Sogn og Fjordane ser det derimot ut som respondentane er skular med gode karakterar (bortsett frå i 2011, der respondentane er dårlegare enn ikkje-respondentane), p-verdiane for 2012 og 2013 er så låge som 0,01. For Aust-Agder har vi også noko låge p-verdiar for 2012 og 2013, høvesvis 0,06 og 0,09, dei andre åra ser ikkje ut til å vere skeive med omsyn til karakterar. Det er derimot gjennomgåande at vi har dei særleg låge p-verdiane for dei to siste åra.

Kan det tyde på at rektorane hugsar resultata frå dei siste åra best og dermed unngår å rapportere om karakterane var dårlege?

P-verdiane i tabell 4a er for ein tosidig test. Viss vi derimot har som utgangspunkt at skular med dårlege resultat er meir tilbakehaldne med å rapportere, burde vi sjå på einsidige testar som resulterer i halvparten så store p-verdiar som for dei tosidige testane. Det vil seie at datasetta ser endå skeivare ut med einsidige testar. På den andre sida er det naturleg å observere ein og annan låg p-verdi når vi har mange testar. Vi kan sjå på p-verdiane for testane samla, seksten testar i alt. Vi nyttar Holm-Bonnferronis metode, som likevel gir at vi forkastar at respondentane er tilfeldige for Sogn og Fjordane i 2012 og 2013 når vi ser på einsidige testar på signifikansnivå 10%, men ikkje på 5% nivå. Ser vi derimot på tosidige testar gir metoden ikkje forkasting sjølv på 10% signifikansnivå. Sjølv om dette kan verka meir oppløftande, er likevel p-verdiane for Sogn og Fjordane gjennomgåande så små at vi ikkje føler oss heilt komfortable med å seie at datasettet ikkje er skeivt. Dersom vi til dømes ser på datasetta for dei fire fylka i år 2012, gir Holm-Bonnferronis metode at resultatet for Sogn og Fjordane ikkje er representativt med p-verdi ca. 3% for den tosidige testen. Vi er på grensa til å seie at Aust Agder heller ikkje er representativt med p-verdi ca. lik 0.12 for tosidig test og 0.06 for einsidig test.

Vi konkluderer med at datasettet ser skeivt ut med omsyn på karakterar hjå respondentar versus ikkje-respondentar i Sogn og Fjordane og kanskje også i Aust Agder. Likevel er ikkje datasettet nødvendigvis skeivt med omsyn til vårt formål. Vi studere samanhengen mellom lærarens utdanning og elevane sine karakterar, og om dette kan forklare noko av skilnaden mellom fylka. Dersom lærarane hjå ikkje-respondentane har relativt høg utdanning,. er skeivheita særleg alvorleg. Men dersom utdanningsnivået hjå ikkje-respondentane fordeler seg omtrent som hjå respondentane, treng ikkje skeivheita ha stor betydning for oss.

Er respondentane representative? Korrigert metode

Metoden ovanfor er brukt i Myklebust & Norstein (2015). Denne metoden er nok streng i den forstand at han i for stor grad konkluderer med at datasetta ikkje er representative med omsyn til karakterar. Grunnen er at metoden ikkje tar omsyn til at det er forskjell mellom klassar, og variansen vert dermed for liten. Her forsøkjer vi å ta omsyn til dette problemet, vi ser då berre på Sogn og Fjordane som i følgje førre metode gav grunn til å tru at utvala ikkje vare representative (grupper som ser representative ut med metoden over vil også vere det med denne metoden).

La Y_{ikl}^{j} vere karakteren til ein tilfeldig elev l i klasse k i år i i fylke $j \in A$. La M_{ik} vere talet på elevar i klasse *k, og* la K_i vere talet på aktuelle klassar i år i. Vi kan skrive

$$Y_{ikl}^{j} = L_{ik}^{j} + e_{ikl}^{j}$$

der variabelen L_{ik}^{j} er knytt til klassen og variabelen e_{ikl}^{j} er knytt til den enkelte elev. Vi antar at desse variablane er uavhengige. (Merk at notasjonen og variabelen L_{ik}^{j} berre er nytta i dette delkapitlet.)

La

$$Var(L_{ik}^{j}) = (\sigma_i^j)^2\, \alpha^j \quad og \quad Var(e_{ikl}^{j}) = (\sigma_i^j)^2(1 - \alpha^j)$$

der parameteren α^j må estimerast. Merk også at $\bar{Y}_{i\cdot\cdot}^{j} = \bar{Y}_{i\cdot}^{j}$.

$$\bar{Y}_{i\cdot\cdot}^{j} = \frac{\sum_{k=1}^{K_i}\sum_{l=1}^{M_{ik}}(L_{ik}^{j} + e_{ikl}^{j})}{\sum_{k=1}^{K_i} M_{ik}} = \frac{\sum_{k=1}^{K_i}(M_{ik}L_{ik}^{j} + \sum_{l=1}^{M_{ik}} e_{ikl}^{j})}{\sum_{k=1}^{K_i} M_{ik}}$$

Vi har at

$$Var(\bar{Y}_{i\cdot}^{j}) = Var(\bar{Y}_{i\cdot\cdot}^{j}) = (\sigma_i^j)^2 \left(\frac{\sum_{k=1}^{K_i}(M_{ik})^2}{\left(\sum_{k=1}^{K_i} M_{ik}\right)^2}\alpha^j + \frac{(1-\alpha^j)}{\sum_{k=1}^{K_i} M_{ik}} \right)$$

Tilsvarande får vi at utanfor sample er

$$Var(\bar{W}_i^{j}) = (\sigma_i^j)^2 \left(\frac{\sum_{k=1}^{\kappa_i}(\mathrm{K}_{ik})^2}{\left(\sum_{k=1}^{\kappa_i} \mathrm{K}_{ik}\right)^2}\alpha^j + \frac{(1-\alpha^j)}{\sum_{k=1}^{\kappa_i} \mathrm{K}_{ik}} \right)$$

der K_{ik} er talet på elevar i klasse *k, og* κ_i er talet på klassar, i år i utanfor sample. Vi estimerer α_i^j frå sample og brukar same punktestimat både i og utanfor sample. Av manglande tilgang på data brukar vi fast klassestorleik på 17 elevar her.

Sidan vi antar uavhengigheit, har vi at

$$Var(\bar{Y}_{i\cdot}^{j} - \bar{W}_{i\cdot}^{j}) = Var(\bar{Y}_{i\cdot}^{j}) + Var(\bar{W}_i^{j})$$

Vi estimerer α^S med tilsvarande metode som skildra seinare i artikkelen, det vil seie løysinga av likningssystemet (8a), og vi brukar data frå alle åra til å estimere denne storleiken. Vi finn at punktestimatet for α^S er 0,03. Merk at dette estimatet nok er relativt usikkert.

Sidan α^S er estimert kjenner vi ikkje fordelinga til observatoren

$$\frac{\bar{Y}_{i\cdot}^{S} - \bar{W}_i^{S}}{Stdav(\bar{Y}_{i\cdot}^{S} - \bar{W}_{i\cdot}^{S})},$$

men gitt det store talet på observasjonar, er det ikkje urimeleg å tru at den er tilnærma normalfordelt. Truleg har den ei fordeling med større varians enn standardnormalfordelinga, det vil seie at p-verdiane i alle fall ikkje vil vere mindre enn dei vi finn her. I denne situasjonen er det bra, vi håpar tross alt at det ikkje skal vere forskjell på sample og utanfor sample.

Skilnaden mellom gjennomsnittskarakter hjå respondentar og ikkje-respondentar i Sogn og Fjordane, tosidig test. Alternativ metode					
	2009	2010	2011	2012	2013
n_i^S	14	98	282	424	177
m_i^S	575	423	273	93	339
$\bar{Y}_{i\cdot}^{S} - \bar{W}_i^{S}$	0,506	0,245	-0,142	0,345	0,298
$Stdav(\bar{Y}_{i\cdot}^{j} - \bar{W}_{i\cdot}^{j})$	0,343	0,150	0,124	0,173	0,145
Z-observ.	1,47	1,63	-1,15	1,99	2,06
p-verdi	0,14	0,10	0,25	0,047	0,039

Tabell 4b

I tabell 4a var p-verdien 0,01 både i 2012 og 2013. Dei har i tabell 4b auka til høvesvis 0,047 og 0,039. Dei er altså framleis relativt låge, men likevel ikkje like urovekkande låge som dei var i tabell 4. Det må også seiast at punktestimatet for α^S er nokså lite, og at dess større α^S er, dess større p-verdi ville vi fått. Vår samla vurdering er at vi ikkje lenger er like urolege for at datasettet er skeivt, spesielt når vi tar i betraktning det høge talet på testar.

Er det forskjell på klassane innanfor kvart av fylka?

Tabell 5 gir sentrale storleikar i datasettet frå respondentane.

SAMPLE							
	År	2009	2010	2011	2012	2013	Sum
Sogn og Fjordane	n_i^S	$n_9^S = 14$	$n_{10}^S = 98$	$n_{11}^S = 282$	$n_{12}^S = 424$	n_{13}^S=177	$n^S = 995$
	Ant.kl.	1	6	18	21	11	M^S =57
	Gj.snitt	4,10	3,67	3,32	3,33	3,43	
	Stdv	-	0,480	0,356	0,415	0,389	
Aust-Agder	n_i^A	9	121	129	178	178	615
	Ant.kl.	1	5	7	8	8	M^A =29
	Gj.snitt	3,45	3,08	3,00	3,18	3,05	
	Stdv	-	0,319	0,417	0,456	0,300	
Nord-Trøndelag	n_i^T	0	93	200	306	197	796
	Ant.kl.	0	4	9	17	12	M^T =42
	Gj.snitt	-	3,31	3,19	2,99	2,97	
	Stdv	-	0,433	0,256	0,218	0,405	
Oppland	n_i^O	0	41	74	64	0	179
	Ant.kl.	0	2	4	4	0	M^O =10
	Gj.snitt	-	3,42	2,98	3,12	-	
	Stdv	-	0,145	0,441	0,325	-	

Tabell 5. Ant.kl. står for antal klassar som vi har fått data frå. Merk at gjennomsnittet er gjennomsnittleg karakter for elevane som inngår i alle klassane for dei ulike fylka og ulike åra. Det er altså ikkje gjennomsnittet av våre observasjonar. Merk at observasjonane er gjennomsnittskarakteren for kvar klasse som varierer i storleik. Stdv er derimot standardavviket av observasjonane og altså ikkje spesielt interessant.

Vi har observasjonar frå ulike år og dermed ulike eksamenar som har hatt ulik vanskegrad. Vi justerer derfor observasjonane etter år.

Vi forsøker å finne ut om lærarens utdanning har betydning for resultata til elevane. For å få så stort datasett som mogeleg, ser vi på alle fylka under eitt. Derfor justerer vi karakterane for forventa karakter i dei ulike fylka. Vi gjer desse justeringane:

$$\tilde{Z}_{il}^j = \frac{Y_{il}^j - (\mu_i + \mu_i^j)}{\sigma_i / \sigma_0} \qquad (2)$$

der vi brukar dei estimerte forventa karakterane i tabell 3 for $\mu_i + \mu_i^j$, for år $i = 9, \ldots, 13$ og fylke $j \in A$, og vi brukar dei estimerte verdiane i tabell 1 for σ_i for år $i = 9, \ldots, 13$. Vidare er konstanten $\sigma_0 = (\sigma_9 + \cdots + \sigma_{13})/5$, vi brukar estimata i tabell 1 og får at $\sigma_0 = 1{,}231$.

Når vi deler på σ_i/σ_0 betyr det at vi reknar med at ein kvar effekt varierer med standardavviket for året. Det vil seie at vi reknar med at ein eksamen der det er stor variasjon i elevprestasjonane, der vil også ein eventuell lærareffekt vere stor. Vi dividerer med σ_i/σ_0 i staden for σ_i for at observerte effektar skal vere i nokolunde rett storleiksorden. Effekten for år i finn vi då ved å multiplisere observert effekt med σ_i/σ_0, merk at σ_i/σ_0 er nokså nær 1 for alle år i.

Merk at

$$Var\left(\tilde{Z}_{il}^{j}\right) = \sigma_0^2$$

for alle mulege verdiar av $i, l\ og\ j$. I det justerte datasettet er altså variansen uavhengig av blant anna året. Og vi hugsar at vi i utgangspunktet har antatt at han er lik i alle fylka.

Sidan vi no har data justert for årseffekten kan vi kutte indeksen i. Vidare har vi bruk for ein indeks k for klassenummer/gruppenummer k. La K_k^j vere talet på elevar i klasse k i fylke j, og M^j er talet på klassar i sample i fylke $j \in A$. Sidan talet på indeksar er det same, men tydinga er ulik, innfører vi no Z i staden for $\tilde{Z}$, for å unngå misforståingar. Vi antar at det justerte datasettet kan modellerast med

$$Z_{kl}^j = U_k^j + \varepsilon_{kl}^j \qquad (3)$$

der ε_{kl}^j er $N(0, \sigma^2)$ fordelte og antatt uavhengige med $\sigma^2 = \sigma_0^2(1-\alpha)$. Variabelen U_k^j er ein klasseeffekt, såkalla randomeffekt, og ε_{kl}^j er knytt til den enkelte elev. Konstanten α uttrykkjer kor stor del av variansen som ligg til U_k^j.

Vi antar at dei to komponentane er uavhengige. Då er

$$Var\left(Z_{kl}^j\right) = Var\left(U_k^j\right) + Var\left(\varepsilon_{kl}^j\right) = \sigma_0^2\left(\alpha + (1-\alpha)\right) \qquad (4)$$

Merk at for ein bestemt verdi av k har vi ein klasse i eit bestemt år i. Det vil seie at om vi går tilbake til ein årsavhengig modell er $U_{ik}^j = U_k^j \frac{\sigma_i}{\sigma_0}$.

Før vi studerer effekten av lærarens utdanning, ser vi om vi kan finne forskjell på klassane i datasettet vårt innanfor kvart fylke. Dersom vi ikkje finn nokon forskjell er det truleg også vanskeleg å sjå ein eventuell effekt av lærarens utdanning.

Vi har ikkje data på individnivå og har derfor ikkje data til å finne SSE, men vi har estimat for standardavviket som vi kan bruke til å finne tilnærma verdiar for SST (den totale variansen). Sidan Z_{kl}^j er justerte variablar er $Var\left(Z_{kl}^j\right) = \sigma_0^2$, og $\sigma_0 = 1{,}231$ er eit estimat for standardavviket uavhengig av fylke. Det vil seie at p-verdiane nedanfor berre er tilnærma riktige.

Det gir estimert

$$SST^S = \sum_{k=1}^{57} \sum_{l=1}^{K_k^S} (z_{kl} - \bar{z}_{..})^2 \quad \approx (n^S - 1)\sigma^2 \approx 994 \cdot \sigma_0^2 = 994 \cdot 1{,}231^2 \approx 1506$$

Der $n^S = 995$.

$$SSG^S = \sum_{k=1}^{57} K_k^S (\bar{z}_{k\cdot} - \bar{z}_{\cdot\cdot})^2 \approx 130$$

$$SSE^S = SST^S - SSG^S = 1506 - 130 = 1376$$

$$F_{(57-1),(n^S-57)} = \frac{\frac{SSG^S}{57-1}}{\frac{SSE^S}{n^S-57}} = \frac{130/57}{1506/(995-57)} \approx 1{,}58$$

(Merk at $SSE^S = SST^S - SSG^S$ er kji-kvadratfordelt med (995-57) fridomsgrader sidan SST^S og SSG^Ser kji-kvadratfordelt med høvesvis 995 og 57 fridomsgrader. Dette gir ein p-verdi på ca 0,005. Vi kan altså konkludere med at det er forskjell mellom klassane i Sogn og Fjordane som ikkje skuldast tilfeldigheiter. Spørsmålet er om noko av denne forskjellen kan forklarast med utdanningsnivået til læraren.

Aust-Agder:

$$SST^A = \sum_{k=1}^{29} \sum_{l=1}^{K_k^A} (z_{ik} - \bar{z}_{\cdot\cdot})^2 \quad \approx (n^A - 1)\sigma^2 \approx 614 \cdot \sigma^2 = 614 \cdot 1{,}231^2 \approx 932$$

Der $n^A = 615$.

$$SSG^A = \sum_{k=1}^{29} K_k^A (\bar{z}_{k\cdot} - \bar{z}_{\cdot\cdot})^2 \approx 96$$

$$SSE^A = SST^A - SSG^A = 932 - 96 = 836$$

$$F_{(29-1),(n^A-29)} = \frac{\frac{SSG^A}{29-1}}{\frac{SSE^A}{n^A-29}} = \frac{95/29}{836/(615-29)} \approx 2{,}41$$

Dette gir ein p-verdi på ca. 0,0001. Vi kan altså konkludere med at det er forskjell på klassane i Aust-Agder som ikkje skuldast tilfeldigheiter.

Nord-Trøndelag:

$$SST^T = \sum_{k=1}^{41} \sum_{l=1}^{K_k^T} (z_{ik} - \bar{z}_{\cdot\cdot})^2 \quad \approx (n^T - 1)\sigma^2 \approx 795 \cdot \sigma_0^2 = 795 \cdot 1{,}231^2 \approx 1205$$

Der $n^T = 796$.

$$SSG^T = \sum_{k=1}^{41} K_k^T (\bar{z}_{k\cdot} - \bar{z}_{\cdot\cdot})^2 \approx 52$$

$$SSE^T = SST^T - SSG^T = 1205 - 52 = 1152$$

$$F_{(41-1),(n^T-41)} = \frac{\frac{SSG^T}{41-1}}{\frac{SSE^T}{n^T-41}} = \frac{52/29}{1205/(795-41)} \approx 0{,}84$$

Dette gir ein p-verdi på ca. 0,75. Nord-Trøndelag gir altså eit veldig spesielt resultat, vi kan nemleg ikkje konkludere med at det er forskjell på klassane i Nord-Trøndelag som ikkje skuldast tilfeldigheiter. Vi konkluderer derfor med at våre data ikkje viser forskjellar mellom klassane i Nord-Trøndelag. Det spørst derfor om vi kan rekne med å finne at lærarens utdanning har påverknad på resultata til elevene i Nord-Trøndelag. Dersom vi skal finne ein slik effekt må i tilfelle utdanningseffekten vere oppheva av andre variablar som vi har kontroll over og derfor kan justere for.

Oppland:

Igjen har vi observasjonar frå ulike år, vi vel 1,31 som er konservativt i denne samanheng. Det gir estimert

$$SST^O = \sum_{k=1}^{10}\sum_{l=1}^{K_l^O}(z_{ik} - \bar{z}_{..})^2 \quad \approx (n^O - 1)\sigma^2 \approx 178 \cdot \sigma^2 = 178 \cdot 1{,}31^2 \approx 307$$

Der $n^O = 179$.

$$SSG^O = \sum_{k=1}^{10} K_k^O(\bar{z}_{k\cdot} - \bar{z}_{..})^2 \approx 25$$

$$SSE^O = SST^O - SSG^O = 307 - 25 = 271$$

$$F_{(10-1),(n^O-10)} = \frac{\frac{SSG^O}{10-1}}{\frac{SSE^O}{n^O-10}} = \frac{25/9}{271/(179-10)} \approx 1{,}94$$

Dette gir ein p-verdi på ca. 0,05. Det er rimeleg å konkludere med at det er forskjell på klassane i Oppland som ikkje skuldast tilfeldigheiter.

Konklusjonen er derfor at det er forskjellar mellom klassane i dei ulike fylka bortsett frå Nord-Trøndelag. Det tyder på at elevar i same klasse har noko til felles som påverkar resultata deira. Det kan til dømes vere læraren, klassekultur, at elevar i same klasse har foreldre med lik sosial bakgrunn og utdanning. Vi er som kjent interessert i om lærarens utdanning er ein slik forklaringsvariabel.

Resultat for Nord-Trøndelag er så vidt spesielt at vi finn grunn til å sjekke om dette også gjeld dei som ikkje har respondert, men då har vi berre høve til å sjekke forskjell mellom skulane. Forskjellen mellom respondentane i Sogn og Fjordane er stor nok til å observere han for så mange som 57 klassar, som antyder at forskjellen er betydeleg. I Aust Agder var han synleg for så mange som 29 klassar. Vi finn data på skoleporten.no. For enkelheits skuld ser vi på skular innan same år i Nord-Trøndelag og sjekkar om vi finn forskjell mellom desse skulane for kvart av åra 2013 til 2010. Talet på skular i desse åra som vi får tilgang til via skoleporten.no er høvesvis: 13, 15, 13 og 11. Då finn vi forskjell på skulane også i Nord-Trøndelag. P-verdiane er ca $2{,}7 \cdot 10^{-6}$, $1{,}1 \cdot 10^{-5}$, 0,06 og 0,002 for høvesvis åra 2013, 2012, 2011 og 2010. Dette betyr at vi kanskje ikkje har eit heilt representativt utval frå Nord-Trøndelag. Standardavvika i tabell 5 er eigentleg ikkje særleg informative sidan dei ikkje tar omsyn til at talet på elevar i klassane varierer mykje. Likevel er standardavviket i denne tabellen påfallande lågt for Nord-Trøndelag i 2012 (og 2011, men her er det berre 9 klassar) samanlikna med dei andre standardavvika i tabellen. Viss det er riktig at det er liten variasjon mellom desse klassane kan det kanskje vere med å skjule forskjellane mellom dei sidan vi har justert karakterane med estimert standardavvik som bygger på heile fylket.

Har lærarens utdanning betydning for resultatet til elevane?

Vi har observasjonar frå ulike år og dermed ulike eksamenar som har hatt ulik vanskegrad. Vi har derfor justert observasjonane våre etter år.

For å få så stort datasett som muleg, ser vi i første omgang på alle fylka under eitt. Og vi har justert datasettet som tidlegare forklart. Men merk at desse justeringane ikkje tar høgde for at fylkeseffekten kan skuldast ulik utdanning hjå lærarane i dei ulike fylka. Dersom det var tilfelle ville ei slik justering kunne gjere utdanningseffekten mindre synleg, og slik sett ville justeringa vere svært uheldig. Vi kjem tilbake til dette. Alternativet til å bruke justerte data er å innføre dummy-variablar i regresjonsmodellen. Det ville gi sju dummy-variablar! Tre for dei fire fylka og fire for dei fem åra 2009 til 2013.

Vi har antatt (3) som vi skriv opp att her

$$Z_{kl}^{j} = U_k^j + \varepsilon_{kl}^{j}$$

Vi har ikkje observasjonar på individnivå. Kvar observert karakter er eit gjennomsnitt av karakterane til ei gruppe/klasse elevar som varierer i storleik. Det vil seie at våre observasjonar er på forma

$$\bar{Z}_{k\cdot}^{j} = U_k^j + \bar{\varepsilon}_{k\cdot}^{j}$$

For å forenkle notasjonen skriv vi $Z_k^j = \bar{Z}_{k\cdot}^j \ og \ \epsilon_k^j = \bar{\varepsilon}_{k\cdot}^j$ og vi får modellen

$$Z_k^j = U_k^j + \epsilon_k^j \quad (5)$$

I store delar av det etterfølgjande har vi ikkje bruk får å skilje mellom fylka, derfor skriv vi

$$Z_k = U_k + \epsilon_k \qquad (6)$$

for $k = 1, \dots, M$, der $M = M^S + M^A + M^T + M^O = 138$. Heretter brukar vi notasjonen Z_k^j og Z_k ettersom det passar. Frå (4) og (5) får vi at

$$Var(Z_k) = \sigma_0^2 \left(\alpha + \frac{1-\alpha}{K_k}\right),$$

der K_k er talet på elevar i klasse $k = 1, \dots, M$. Det vil seie at

$$Var(U_k) = \sigma_0^2 \alpha \quad og \quad Var(\epsilon_k) = \frac{\sigma_0^2(1-\alpha)}{K_k} \qquad (7)$$

for $k = 1, \dots, M$, og vi hugsar frå (4) at

$$\sigma^2 = Var\left(\varepsilon_{kl}^j\right) = \sigma_0^2(1-\alpha).$$

Løysinga av følgjande likningssystem gir sannsynsmaksimeringsestimatoren for α (vi ser på σ_0 som kjent):

$$\sum_{k=1}^{M} \frac{K_k - 1}{\alpha(K_k - 1) + 1} - \sum_{k=1}^{M} \frac{(Z_k - \hat{\mu}_Z)^2\, K_k(K_k - 1)}{[\alpha(K_k - 1) + 1]^2 \sigma_0^2} = 0$$

(8a)

$$\hat{\mu}_Z = \sum_{k=1}^{M} \frac{Z_k K_k}{\alpha(K_k - 1) + 1} \left(\sum_{k=1}^{M} \frac{K_k}{\alpha(K_k - 1) + 1} \right)^{-1}$$

Vi klarer ikkje løyse dette likningssystemet når klassestorleiken K_k varierer for $k = 1, \dots, M$. Vi løyser likninga numerisk. Men merk at vi neppe er garantert verken eintydig løysing eller løysing i intervallet (0,1).

Vi estimerer α til å vere ca 0,024. Gitt at dette er eit korrekt estimat vil det seie at effekten av skulen varierer mellom $\pm 1{,}96 \cdot 1{,}231 \cdot \sqrt{0{,}024} \approx \pm 0{,}37$. Dette er ein relativt liten skuleeffekt i forhold til den Grøgaard (2012) finn. Han finn skuleeffekt tilsvarande α i storleiken 0,08-0,17. I internasjonal målestokk er også det lite, gjennomsnittet for OECD-landa er ca. 40% (Turmo & Lie, 2004). Merk likevel at vi finn skuleeffekten innanfor dei fire fylka, det er ikkje urimeleg å tru at vi hadde funne ein større skuleeffekt nasjonalt.

Ovanfor såg vi på σ_0 som kjent. Ser vi på σ_0^2 som ukjent, gir løysinga av følgjande likningssystem sannsynsmaksimeringsestimatorane for $\alpha\ og\ \sigma_0^2$:

$$\hat{\sigma}_0^2 = \frac{1}{M} \sum_{k=1}^{M} \frac{\left(Z_k - \hat{\mu}_Z\right)^2 K_k}{\alpha(K_k - 1) + 1}$$

$$\hat{\mu}_Z = \sum_{k=1}^{M} \frac{Z_k K_k}{\alpha(K_k - 1) + 1} \left(\sum_{k=1}^{M} \frac{K_k}{\alpha(K_k - 1) + 1} \right)^{-1} \qquad (8b)$$

$$\sum_{k=1}^{M} \frac{K_k - 1}{\alpha(K_k - 1) + 1} - \sum_{k=1}^{M} \frac{(Z_k - \hat{\mu}_Z)^2 K_k(K_k - 1)}{[\alpha(K_k - 1) + 1]^2 \hat{\sigma}_0^2} = 0$$

Vi estimerer då α så liten som 0,005, og $\hat{\sigma}_0 \approx 1{,}43$ og effekten av skulen vil variere mellom $\pm 1{,}96 \cdot 1{,}231 \cdot \sqrt{0{,}005} \approx \pm 0{,}20$. Men vi har større tiltru til å bruke kjent sigma her.

I alle tilfelle er alfa liten, estimert til ca. 0,024 og då er estimatet relativt sett svært usikkert og blant anna svært sensitivt overfor rett verdi av sigma. Det er til dømes relativt sett stor skilnad på 0,01 og 0,03, sjølv om skilanden i talverdi er liten. Merk også at vi ikkje er garantert verken eintydig løysing eller løysing av dette systemet i intervallet (0,1).

Anta at randomeffekten U_k kan skrivast som

$$U_k = \beta_0 + \beta_1 X_{1k} + \cdots + \beta_{\tau k} + \xi_k \ .$$

Dessverre observerer vi ikkje U_k, men Z_k (riktig nok er heller ikkje Z_k direkte observert). Når vi ser på alle fylka under eitt, har vi derfor

$$\begin{aligned} Z_k &= \beta_0 + \beta_1 X_{1k} + \cdots + \beta_{\tau k} + \xi_k + \epsilon_k \\ &= \beta_0 + \beta_1 X_{1k} + \cdots + \beta_{\tau k} + \eta_k \ , \end{aligned}$$

der $\eta_k = \xi_k + \epsilon_k$. Variabelen ξ_k er då ein randomeffekt for klasse k. Vi får då at η_k er uavhengige $N(0, \sigma^2_{\eta,k})$ for $k = 1, \ldots, M$. Når vi ser på kvart fylke, får vi tilsvarande notasjon,

$$Z_k^j = \beta_0^j + \beta_1^j X_{1k}^j + \cdots \beta_\tau^j X_{\tau k}^j + \eta_k^j.$$

Variansen

$$\sigma^2_{\eta,k} = \sigma^2 \left(\delta + \frac{1}{K_k} \right) = \sigma_0^2 (1 - \alpha) \left(\delta + \frac{1}{K_k} \right), \qquad (9)$$

der α er som i (4). Det vil seie at $Var(\xi_k) = \sigma^2 \delta = \sigma_0^2 (1 - \alpha) \delta$.

Parameteren δ er eit uttrykk for varians som skuldast variasjon mellom klassane, men som regresjonsmodellen ikkje forklarar. Dette kan til dømes vere varians mellom lærarar som ikkje kan forklarast av utdanninga til lærarane. Det kan også skuldast eventuelle skule- og kommuneeffektar.

Det er vanskeleg å anslå denne variansen, men den er liten i forhold til variansen mellom dei ulike elevane i klassen. Vi finn eit anslag for δ ved følgjande prosedyre:

- Vi utfører ein førebels regresjonsanalyse og brukar residualane $\hat{\eta}_k$ i neste trinn.
- Sannsynsmaksimeringsestimatoren for δ er løysinga av likningssystemet

$$\sum_{k=1}^{M} \frac{K_k}{\delta K_k + 1} - \sum_{k=1}^{M} \frac{\left(\hat{\eta}_k - \hat{\mu}_\eta \right)^2 K_k^2}{[\delta K_k + 1]^2 \sigma^2} = 0$$

$$\hat{\mu}_\eta = \sum_{k=1}^{M} \frac{Z_k K_k}{\delta K_k + 1} \left(\sum_{k=1}^{M} \frac{K_k}{\delta K_k + 1} \right)^{-1} \qquad (10)$$

  Vi klarer ikkje å løyse dette likningssystemet når klassestorleiken K_k varierer med k. (Dessutan er vi vel ikkje garantert ei eintydig løysing). Vi løyser likninga numerisk.

- No brukar vi estimert δ, og gjennomfører ein ny regresjonsanalysen med denne verdien.
- Vi gjentek denne prosedyren inntil δ stabiliserer seg.

Det gir estimert δ ca lik 0,018. Det vil seie at vi kan rekne med at $\pm 1{,}96\sqrt{0{,}018} \cdot 1{,}231 \approx \pm 0{,}33$ av elevens forventa karakter skuldast andre faktorar enn eleven og er ikkje forklart av modellen. Det betyr med andre ord at dersom kvar observasjon bygde på uendeleg mange elevar som alle var tenkt å komme frå same kommune, undervist av ein lærar med like lang utdanning og gå i ein klasse identisk med den observerte, ville standardavviket til dei justerte observasjonane vere ca. $\sqrt{0{,}018} \cdot 1{,}23 \approx 0{,}17$.

Vi brukar ein vekta regresjonsmodell med vektene $\left(\delta + \frac{1}{K_k}\right)^{-1}$, der δ er estimert til ca 0,016. Merk også at dess større verdi vi vel for δ, dess svakare resultat kan vi i utgangspunktet rekne med.

Føresetnaden at η_k^j er uavhengige for alle k og $j \in A$, held truleg ikkje. Dette problemet er større i den følgjande analysen enn tidlegar. Det kjem av at vi no har relativt sett fleire observasjonar per kommune og per skule, og truleg er fleire av klassane undervist av same lærar. Vi forsøker med ein primitiv sjekk av korrelasjonen innanfor same skular som antyder korrelasjon på ca. 0,3. Ein tilsvarande primitiv korrelasjonssjekk mellom tilfeldig klassar frå ulike skular antyder korrelasjonen 0,09, altså tilnærma null som vi ville forvente. Den relativt høge korrelasjonen 0,3 innanfor same skular antyder at støyledda ikkje er uavhengige. I første omgang ser vi likevel på kva modellen seier om vi antar at støyledda er ukorrelerte. Deretter forsøker vi å justere datasettet for å redusere problemet med korrelert støy.

Plott av utdanningsnivået til lærarane viser at det er opphoping av lærarar som har 0, 30 og 60 studiepoeng, det er relativt få observasjonar mellom desse. Forklaringa er nok at eit typisk matematikkurs i lærarutdanningane er på 30 studiepoeng. Dersom vi grupperer lærarens utdanningsnivå i intervalla $[0{,}30)$, $[30{,}60), [60, \infty)$ får vi at [30,60) studiepoeng gir ein estimert forventa effekt på høvesvis 0,142 og 0,215 for utdanningsintervalla $[30{,}60)\ og\ [60, \infty)$. Det antyder ein kvadratrotfunksjon. Det betyr at det ser ut som effekten av lærarens utdanning aukar mest for dei første studiepoenga, og det er kanskje ikkje overraskande. Dette har leda oss til å bruke følgjande regresjonsmodell:

$$Z_k = \beta_0 + \beta_1 \sqrt{X_{1k}/30} + \cdots + \beta_\tau X_{\tau k} + \eta_k$$

eller når vi ser på kvart fylke for seg

$$Z_k^j = \beta_0^j + \beta_1^j \sqrt{X_{1k}^j/30} + \cdots + \beta_\tau^j X_{\tau k}^j + \eta_k^j$$

Der X_{ik} og X_{1k}^j er gjennomsnittleg utdanning til lærarane til klassen på 8., 9. og 10. klassesteg. Vi deler med 30 for å lette tolkinga av koeffisienten sidan eit matematikkurs i lærarutdanninga typisk er 30 studiepoeng. Når koeffisienten er ca. 0,18, som i tabell 6, tyder det derfor at 30 studiepoeng gir ein forventa effekt 0,18, medan til dømes 60 studiepoeng gir forventa effekt lik $\sqrt{2} \cdot 0{,}18 \approx 1{,}4 \cdot 0{,}18 \approx 0{,}24$ effekt. Merk at denne regresjonsmodellen gir ein p-verdi for utdanningseffekten som er omtrent halvparten av det den lineære modellen gir. Det er likevel ikkje prov på at effektauken ikkje er lineær.

I tillegg til den uavhengige variabelen utdanning har vi høve til å bruke forklaringsvariablane:

- Kvadratrota av utdanninga til læraren i 10. klasse
- Talet på elevar i klassen
- Erfaringa til læraren som har undervist klassen, vi brukar gjennomsnittleg erfaring til lærarane som har undervist klassen på 8., 9. og 10. klassesteg
- Læraren har allmennlærarutdanning eller fagutdanning
- Prosentdelen kvinner i kommunen som har vidaregåande skule
- Prosentdelen kvinner i kommunen som har ei kort høgare utdanning
- Prosentdelen kvinner i kommunen som har minst vidaregåande skule
- Prosentdelen kvinner i kommunen som har minst ei kort høgare utdanning
- Prosentdelen kvinner i kommunen som har ei lang høgare utdanning
- Prosentdelen menn i kommunen som har vidaregåande skule
- Prosentdelen menn i kommunen som har ei kort høgare utdanning
- Prosentdelen menn i kommunen som har minst vidaregåande skule
- Prosentdelen menn i kommunen som har minst ei kort høgare utdanning
- Prosentdelen menn i kommunen som har ei lang høgare utdanning

I tillegg har vi dummy-variablar for fylke og år. Men gitt den gode bakgrunnsinformasjonen meiner vi det justerte datasettet kan vere å føretrekkje framfor dummy-variablar for fylke og år.

Vi har mange variablar som angår utdanningsnivået til kvinner og menn i kommunane. Her har vi sjølvsagt multikolineæritet. Det ser ut som variabelen kvinner med minst vidaregåande skule er den som forklarar datasettet best, vi endar derfor opp med å redusere dei ti variablane som har med utdanningsnivået i kommunen til ein variabel. Merk at dette ikkje betyr at vi leitar i datasettet etter signifikante resultat, noko som er uakseptabelt. Hypotesane våre går framleis på lærarens utdanning og erfaring, samt klassestorleik. Dei andre variablane er berre meint for å justere modellen og redusere variansen og vi ønskjer rimelegvis ikkje å bruke for mange variablar.

Datasettet tyder på at det ofte er same læraren som har undervist klassen både på 8., 9. og 10. klassesteg. Derfor er det rimelegvis multikolineæritet mellom gjennomsnittleg utdanning til lærarane i desse tre åra og utdanninga til læraren i 10. klasse. Vi kan derfor ikkje bruke begge desse forklaringsvariablane, men det er interessant at det gjennomsnittlege utdanningsnivået over dei siste tre åra ser ut til å ha betre prediksjonsevne enn utdanningsnivået til læraren i 10. klasse. Det tyder på at ein godt utdanna lærar i avgangsåret ikkje rettar opp eit dårleg utgangspunkt. Dette er sjølvsagt ikkje uventa, men det er likevel interessant at det kanskje kjem til syne også i datasettet. Men merk at vi ikkje har grunnlag for å trekkje denne konklusjonen.

Det er argument for å bruke logaritmen til elevtalet, det vil seie at ein tenkjer at det er den prosentvise forskjellen i klassestorleik en bør sjå etter. Tanken er at den relative forskjellen mellom fem og ti elevar er mykje større enn mellom 25 og 30 elevar, sjølv om differansen er fem elevar i begge tilfella. Dersom vi brukar logaritmen av elevtalet, vil derimot forskjellen mellom fem og ti elevar vere lik forskjellen mellom 15 og 30 elevar. Sjølv om argumenta for logaritmen kan virke tiltalande er det argument mot å bruke denne transformasjonen. Til dømes kan vi tenkje oss at fem elevar ekstra tar like mykje av dei andre elevane si tid enten ein går frå fem til ti elevar, eller ein går frå 25 til 30 elevar. Dessutan er det ikkje urimeleg å tenkje seg at det er verre å gå frå 15 til 30 elevar enn det er å går frå 5 til 10 elevar, sjølv om begge tilfella representerer ei dobling. Vi finn ikkje at logaritmen til elevtalet passar betre enn elevtalet, og vi vel å bruke elevtalet direkte.

Det er alltid eit problem å avgjere kva for uavhengige variablar som bør vere med i modellen. Analysane i dette tilfellet peikar mot at vi bør redusere dei uavhengige variablane, inkludert dummy-variablane, til

- prosentdelen kvinner i kommunen med minst vidaregåande skule
- talet på elevar i klassen
- kvadratrota av lærarens utdanning (i studiepoeng dividert med 30)

Vi får følgjande modell når vi ser på heile datasettet:

Model Summary[b,c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,349[a]	,122	,102	1,246320

a. Predictors: (Constant), RotUtD30, vgPk, Elevtal

b. Dependent Variable: KarMiForvDsigmaG1.24

c. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients		
		B	Std. Error	Beta	t	Sig.
1	(Constant)	-1,612	,542		-2,976	,003
	vgPk	,025	,008	,269	3,283	,001
	Elevtal	-,013	,005	-,214	-2,601	,010
	RotUtD30	,182	,078	,194	2,352	,020

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

Når vi ser på fylka kvar for seg får vi desse modellane

Sogn og Fjordane:

Coefficients[a,b,c]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1,120	,775		-1,445	,154
	vgPk	,018	,011	,217	1,663	,102
	Elevtal	-,014	,008	-,216	-1,665	,102
	RotUtD30	,202	,111	,237	1,818	,075

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

c. Selecting only cases for which Fylke = S

Aust-Agder

Coefficients[a,b,c]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-4,684	1,287		-3,639	,001
	vgPk	,075	,019	,606	3,975	,001
	Elevtal	-,026	,012	-,326	-2,101	,046
	RotUtD30	,123	,173	,110	,709	,485

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

c. Selecting only cases for which Fylke = AA

Oppland:

Coefficients[a,b,c]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-6,504	3,420		-1,901	,106
	vgPk	,078	,041	,853	1,885	,108
	Elevtal	,028	,030	,422	,959	,375
	RotUtD30	,496	,313	,496	1,585	,164

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

c. Selecting only cases for which Fylke = O

Nord-Trøndelag:

Coefficients[a,b,c]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	,316	1,159		,273	,786
	vgPk	-,004	,016	-,043	-,242	,810
	Elevtal	-,001	,008	-,019	-,106	,916
	RotUtD30	-,022	,170	-,023	-,132	,896

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

c. Selecting only cases for which Fylke = NT

Vi har lite data til å sjå på kvart fylke kvar for seg, men likevel er det interessant å merke seg at Nord-Trøndelag skil seg ut med at ingen av variablane ser ut til å ha betydning. Dette er konsistent med det vi fann tidlegare, nemleg at det ikkje ser ut til å vere forskjell mellom dei observerte klassane i Nord-Trøndelag. Men dette er som nemnt tidlegare noko merkeleg.

Merk at storleiken R^2 sjølvsagt er ubetydeleg i alle desse testane. Han er i omtrent 0,12 for heile datasettet. Det kjem av at storparten av variansen er mellom elevar innanfor klassane og derfor er det nærast berre ein ubetydeleg del av variansen i vårt datasett som kan la seg forklare. (Merk at varians mellom enkeltelevar moglegvis også kan forklarast, men rimelegvis ikkje i vårt datasett der vi ikkje har observasjonar på individnivå). Vi antar at ingen av forklaringsvariablane (regresjonsvariablane) kan forklare $var(\varepsilon_{kl}^{j})$ (eller tilsvarande $Var(\epsilon_k)$ i (7)) som er variasjon mellom elevane i ein klasse. Denne variansen utgjer som nemnt storparten av observert varians. Derimot er det interessant å sjå på kor mykje modellen forklarar av den delen av variansen som kan la seg forklare. Med andre ord av variansen mellom klassar. Vi forsøker å estimerer denne andelen til

$$\frac{var(U_k) - var(\xi_k)}{var(U_k)} = 1 - \frac{var(\xi_k)}{var(U_k)} = 1 - \frac{\sigma_0^2(1-\alpha)\delta}{\sigma_0^2\alpha} \approx 1 - \frac{(1-\hat{\alpha})\hat{\delta}}{\hat{\alpha}} \approx 1 - \frac{0{,}976 \cdot 0{,}016}{0{,}024} \approx 0{,}35. \quad (11)$$

Modellen forklarar altså ca. 35% av variansen mellom klassar. Det vil seie av variansen som ikkje er mellom individ i same klasse, eller vi kan seie av variansen som skuldast faktorar som påverkar heile klassen. Det betyr at modellen forklarar relativt mykje av den delen av variansen som kan forklarast. Men det må likevel seiast at dette talet er svært usikkert, sidan estimata av α og δ er svært usikre. Vi kan derfor ikkje feste lit til dette talet 0,35 slik vi vanlegvis gjer til R^2 i ein regresjonsmodell.

Vi justerer for korrelasjonen i støyleddet

Vi gjer følgjande korreksjonar for å justere for korrelerte støyledd: Dersom to klassar innanfor same skule har hatt lærarar med like mykje utdanning, slår vi saman desse observasjonane slik at vi ser på gjennomsnittskarakteren til elevane i dei klassane det gjeld. Elevtalet lar vi vere gjennomsnittleg tal på elevar i klassane, denne variabelen vert derfor ikkje heilt riktig i desse tilfella. Vektene i den vekta regresjonsmodellen vert då ein funksjon av summen av elevar i dei klassane det gjeld.

I dei tilfella at to eller fleire klassar på same skule har hatt lærarar med lik utdanning, er det ofte snakk om same lærar og då vil denne korreksjonen vere særs viktig. Spesielt på mindre skular vil det ofte vere tilfelle at vi har fleire observasjonar med same lærar. I desse tilfella korrigerer denne metoden for lærareffekt (som ikkje skuldast utdanning). Denne effekten er truleg stor og er derfor ein viktig grunn til avhengigheit i støyledda.

Dersom det ikkje er snakk om same lærar, korrigerer metoden berre for ein eventuell skuleeffekt og kommuneeffekt. Denne effekten er truleg mindre enn lærareffekten, men kan likevel vere stor nok til å skape uheldig avhengigheit i støyledda.

Bakdelen med desse justeringane er rimelegvis at vi får færre observasjonar, sjølv om kvar observasjon vert litt sterkare. Det kjem av at vi får fleire elevar bak kvar observasjon, og vi reduserer variansen som ikkje er forklart av modellen. For alle fire fylka går vi frå 138 observasjonar til 93 observasjonar. I utgangspunktet medfører dette at parameterestimata vert meir usikre og hypotesetestane mister styrke. Derfor kan vi i utgangspunktet forvente høgare p-verdiar etter desse justeringane. Men modellføresetnadane er truleg meir korrekte og derfor har vi høgare tillit til desse p-verdiane, og vi meiner at dei er å føretrekkje.

I regresjonsanalysen brukar vi vektfunksjonen

$$\left(\delta_k + \frac{1}{K_k}\right)^{-1}$$

der $\delta_k = \delta(\gamma + \frac{1-\gamma}{S_k})$, der S_k er talet på klassar involvert i justeringa og $\delta = 0{,}018$ som før. Og i dei tilfella to eller fleire klassar er slått saman, er K_k samla elevtal i desse klassane. Vi vel $\gamma = \frac{2}{3}$. Desse justeringane er litt tilfeldige, men ideen er at noko av variansen mellom klassar som modellen ikkje kan forklare (randomeffekt) vert redusert med talet på klassar vi studerer. Vi har lite data til å estimere denne storleiken. Denne verdien er neppe veldig viktig og vi set han til ein tredel, det vil seie $\gamma = \frac{2}{3}$.

Model Summary

Model	R UavhStoy = 0 (Selected)	R Square	Adjusted R Square	Std. Error of the Estimate
1	,388[a]	,151	,122	1,2427488

a. Predictors: (Constant), RotUtD30, vgPk, NoKorElevtal

Coefficients[a,b,c]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-1,758	,558		-3,154	,002
	vgPk	,026	,008	,336	3,315	,001
	NoKorElevtal	-,012	,005	-,234	-2,305	,023
	RotUtD30	,181	,082	,220	2,215	,029

a. Dependent Variable: NoKorKarakter

b. Weighted Least Squares Regression - Weighted by NoKorVektAlph

c. Selecting only cases for which UavhStoy = 0

Tabell 6

Vi ser at p-verdiane vert noko større som forventa. Elles er estimata nokså like. Effekten av elevtalet er svekka mest, dette er også som forventa sidan elevtalet ikkje lenger er heilt korrekt. For utdanninga går estimert verdi frå 0,182 til 0,181 (tilnærma identiske estimat) og p-verdien går frå ca. 0,02 til 0,029.

Kva vert resultatet om vi vel andre verdiar der observasjonane er tvitydige?

Som nemnt i innleiinga er sju av observasjonane uklare når det gjeld lærarens utdanning. I utgangspunktet valde vi å notere utdanninga til den læraren som hadde lengst utdanning i desse klassane, no ser vi på kva som skjer om vi vel å notere utdanninga til den av lærarane som hadde minst utdanning. Hovudkonklusjonen er at dette ikkje betyr mykje, men i den grad det betyr noko, vert konklusjonane litt styrka.

Coefficients[a,b]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	-1,650	,537		-3,075	,003
	vgPk	,025	,007	,275	3,379	,001
	Elevtal	-,014	,005	-,234	-2,827	,005
	MMRotUtD30	,207	,075	,229	2,746	,007

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

Kva vert resultatet om vi inkluderer fylkesvariablar og årsvariablar?

På bakgrunn av all tilgjengeleg informasjonen om eksamensresultata, justerte vi observasjonane våre for årseffekt og fylkeseffekt. Dette er ikkje nødvendigvis optimalt. Her ser vi på effekten av å inkludere dummy-variablar for fylkeseffekten, tre variablar, og årseffekten, fire variablar. Til saman sju dummy-variablar. Resultata kjem styrka ut. Men vurdert ut frå justert R^2, seier modellen likevel at vi ikkje skal inkludere desse variablane, sjølv om estimata og p-verdiane altså kjem styrka ut. Med andre ord kan det kanskje tyde på ei overtilpassing til datasettet. Estimert effekt av utdanning aukar til 0,227 og p-verdien er så låg som 0,007 når vi inkluderer alle dei sju dummy-variablane. Estimata for mors utdanning og antal elevar i klassen er uforandra på høvesvis 0,025 og -0,014, men p-verdiane er litt svekka.

Coefficients[a,b]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-1,612	,542		-2,976	,003
	vgPk	,025	,008	,269	3,283	,001
	Elevtal	-,013	,005	-,214	-2,601	,010
	RotUtD30	,182	,078	,194	2,352	,020
2	(Constant)	-1,465	,613		-2,391	,018
	vgPk	,024	,008	,261	3,172	,002
	Elevtal	-,015	,005	-,239	-2,778	,006
	RotUtD30	,215	,081	,229	2,661	,009
	ti	,001	,303	,001	,004	,997
	elleve	-,171	,293	-,217	-,585	,560
	tolv	-,110	,294	-,153	-,375	,708
	tretten	-,079	,298	-,092	-,265	,791
3	(Constant)	-1,547	,624		-2,481	,014
	vgPk	,025	,008	,270	3,227	,002
	Elevtal	-,014	,005	-,226	-2,569	,011
	RotUtD30	,227	,082	,242	2,764	,007
	ti	,020	,306	,020	,067	,947
	elleve	-,151	,296	-,191	-,510	,611
	tolv	-,089	,296	-,123	-,299	,765
	tretten	-,053	,300	-,063	-,178	,859
	AA	-,005	,077	-,006	-,065	,948
	NT	-,086	,071	-,113	-1,202	,231
	O	-,019	,120	-,014	-,157	,876

a. Dependent Variable: KarMiForvDsigmaG1.24

b. Weighted Least Squares Regression - Weighted by Vekt1Alfa0026

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,349[a]	,122	,102	2,786856
2	,380[b]	,144	,098	2,792966
3	,395[c]	,156	,089	2,807023

a. Predictors: (Constant), RotUtD30, vgPk, Elevtal

b. Predictors: (Constant), RotUtD30, vgPk, Elevtal, tolv, ti, tretten, elleve

c. Predictors: (Constant), RotUtD30, vgPk, Elevtal, tolv, ti, tretten, elleve, O, AA, NT

Her, som I tilfellet utan dummy-variablar, er justert-R^2 ubetydeleg, faktisk så lågt som 0,098. Vi forsøker på same måten som ovanfor å estimere kor stor del av variansen modellen forklarar, av den delen av variansen som modellen I teorien kan forklare. Vi hugsar at mesteparten av variansen er mellom elevar innanfor same klasse og kan ikkje forklarast av modellen. Vi går tilbake til (6)

$$Z_k = U_k + \epsilon_k \qquad (12)$$

Vi antar at randomeffekten U_k kan skrivast som

$$U_k = \beta_0 + \beta_1 \sqrt{\frac{X_{1k}}{30}} + \beta_2 X_{2k} + \beta_3 X_{3k} + \beta_4 X_{4k} \ldots + \beta_{10k} + \xi_k \,.$$

Der $\beta_4, \cdots, \beta_{10}$ er dummy-variablar for fylke og år. Igjen forsøker vi å estimere andelen av variansen som modellen forklarer, av den andelen som modellen kan forklare (bortsett frå variansen forklart av dummy-variablane). Det vil seie at vi lar

$$U_k^* = U_k - (\beta_4 X_{4k} \ldots + \beta_{10k} X_{10k}).$$

Der $\beta_4, \ldots, \beta_{10k}$ er estimert til høvesvis -0,05, -0,086, og -0,02 (for Aust-Agder, Nord-Trøndelag og Oppland) og 0,023, -0,15, -0,88 og -0,053 (for høvesvis år 10 til 13).

Vi har at $Var(U_k^*) = Var(U_k)$. I (12) byter vi U_k med U_k^* slik at

$$Var(U_k^*) = \alpha\sigma_0^2 \quad og \quad Var(\epsilon_k) = \frac{\sigma_0^2(1-\alpha)}{K_k}$$

og som før er

$$Var(\varepsilon_{kl}) = \sigma_0^2(1-\alpha).$$

Vi estimerer denne andelen, α, med same prosedyre som ovanfor (8a), der vi igjen ser vi på σ_0 som kjent lik 1,231.

$$\frac{var(U_k^*) - var(\xi_k)}{var(U_k^*)} = 1 - \frac{var(\xi_k)}{var(U_k^*)} = 1 - \frac{\sigma_0^2(1-\alpha)\delta}{\sigma_0^2\alpha} \approx 1 - \frac{(1-\hat{\alpha})\hat{\delta}}{\hat{\alpha}} \approx 1 - \frac{0{,}975 \cdot 0{,}016}{0{,}025} \approx 0{,}38.$$

Alternativ metode:
Dersom vi også estimerer σ_0 får vi punktestimatet 1,28 og punktestimatet 0,019 for α (løysing av (8b). Og i staden for å estimere $\sigma^2 = Var(\varepsilon_{kl})$ via α der vi bruker samanhengen $Var(\varepsilon_{kl}) = \sigma^2 = \sigma_0^2(1-\alpha)$, kan vi også estimere σ^2 direkte ved å inkludere

$$\hat{\sigma}^2 = \frac{1}{M}\sum_{k=1}^{M}\frac{\left(Z_k - \hat{\mu}_Z\right)^2 K_k}{\hat{\delta}K + 1}$$

i likningssystemet (10). Då får vi punktestimatet for σ^2 til 1,28 og punktestimatet for $\hat{\delta}$ er uforandra på 0,016. Forklart del av varians som kan forklarast vert då

$$\frac{var(U_k^*) - var(\xi_k)}{var(U_k^*)} = 1 - \frac{var(\xi_k)}{var(U_k^*)} = 1 - \frac{\sigma_0^2(1-\alpha)\delta}{\sigma_0^2\alpha} \approx 1 - \frac{(1-\hat{\alpha})\hat{\delta}}{\hat{\alpha}} \approx 1 - \frac{(1-0{,}19)\cdot 0{,}016}{0{,}019} \approx 0{,}17.$$

Men i staden for å estimere $\sigma^2 = Var(\varepsilon_{kl})$ via α der vi bruker samanhengen $Var(\varepsilon_{kl}) = \sigma^2 = \sigma_0^2(1-\alpha)$ kan vi bruke punktestimatet for σ^2 (1,46), og vi får då at andelen av variansen som kan forklarast av modellen er

$$\frac{var(U_k^*) - var(\xi_k)}{var(U_k^*)} = 1 - \frac{var(\xi_k)}{var(U_k^*)} = 1 - \frac{\sigma^2\delta}{\sigma_0^2\alpha} \approx 1 - \frac{\sigma^2\hat{\delta}}{\hat{\sigma}_0^2\hat{\alpha}} \approx 1 - \frac{1{,}46\cdot 0{,}016}{1{,}28^2\cdot 0{,}019} \approx 0{,}25.$$

Både 0,17 og 0,25 er vesentleg mindre enn 0,38 som vi fann i (11), og dette seier noko om at usikkerheita i desse estimata er relativt stor.

Kva seier desse testane

Det viser seg at estimata for $\hat{\beta}_1^j$ varierer sterkt frå fylke til fylke og at koeffisienten ikkje er signifikant større enn null. P-verdien i tilfellet Sogn og Fjordane er rett nok på grensa til å vere signifikant, han er 0,075.

Nord Trøndelag er spesielt, her er ingen av variablane i nærleiken av å vere signifikante. Men dette er kanskje ikkje overraskande, for tidlegare fann vi at dei ulike klassane i Nord Trøndelag ikkje presterte signifikant forskjellig frå kvarandre. Som vi peika på då, må det då vere andre variablar som trekkjer i motsett retning av utdanning, dersom vi skal kunne sjå ein mogleg effekt av lærarens utdanning. Ei muleg forklaring på fenomenet kunne tenkjast å vere at lærarane i Nord Trøndelag hadde omtrent like mykje utdanning slik at utdanningseffekten er usynleg. Vi ser av plotta over lærarens utdanning i dei ulike fylka at det er relativt få som har mindre enn 30 studiepoeng i Nord Trøndelag, og når resultatet ser ut til å vere ein funksjon av kvadratrota av antal studiepoeng er forskjellen relativt liten mellom 30 og 60 studiepoeng. Slik sett kunne vel det tenkjast å vere ei mogleg forklaring, men forklaring har nok ikkje belegg i data.

Vi må også peike på at Nord Trøndelag er det einaste fylket der det ser ut som om respondentane er tilfeldige med omsyn til karakterar, respondentane i dei andre fylka har som nemnt betre karakterar enn vi skulle forvente om dei var eit tilfeldig utval. Kan det vere at Nord Trøndelag viser sanninga, og at sanninga då er ingen effekt av utdanning? Det er som nemnt truleg skeivheit i datasettet frå dei andre fylka. Men som nemnt treng ikkje denne skeivheita påverke samanhengen mellom utdanning og elevprestasjonar. Men det er mistenkjeleg at det fylket der vi ikkje ser nokon samanheng mellom lærarens utdanning og elevens prestasjon, er det einaste fylket som ikkje har ei openbar skeivheit i respondentanes karakternivå.

I gjennomsnitt har elevar (blant respondentane) i dei ulike fylka lærarar med følgjande utdanning.

I snitt er elevane underviste av ein lærar med antal studiepoeng:	Sogn og Fjordane	Aust-Agder og Oppland	Nord-Trøndelag
	40,9	47,0	47,6

Sidan vi brukar ein ikkje-lineær modell for samanhengen mellom elevens resultat og lærarens utdanning er det ikkje nok å sjå på gjennomsnitta slik dei kjem fram i tabellen over. Vi lar modellen predikere resultatet for det gitte antalet elevar i kvart fylke og for kvart år undervist av lærarar med den gitte utdanninga, men der elevane er like med omsyn på variablane vgPk (mors utdanning) og elevtal i klassen. Det vil seie at vi lar

$$\hat{z}_{ikl}^{j} = 0{,}182\sqrt{x_{1,ikl}^{j}/30}$$

der $X_{1,ikl}^{j}$ er lærarens utdanning, i studiepoeng utover matematikk 1.

Då får vi følgjande predikerte effekt med det faktisk observerte utdanningsnivået:

Predikert effekt av lærarens utdanning i dei ulike fylka					
	Alle år	2010	2011	2012	2013
S	0,180	0,141	0,186	0,192	0,160
A og O	0,195	0,182	0,201	0,200	0,92
T	0,200	0,183	0,216	0,195	0,201

Av tabellen ser vi at utdanningsnivået til lærarane i dei ulike fylka ikkje forklarar forskjellen mellom fylka. Faktisk predikerer modellen at utdanningsnivået til lærarane talar mot Sogn og Fjordane, men likevel berre i storleiken to hundredelar som sjølvsagt er ubetydeleg og neppe signifikant. Vi ser også at sjølv om elevane i Aust Agder og Oppland i gjennomsnitt er undervist av lærarar med mest utdanning, kjem likevel Nord Trøndelag litt betre ut i følgje modellen.

Dersom vi tilsvarande ser på kva resultat klassestorleiken predikerer, får vi tala -0,285, -0,312 og -0,307 for høvesvis Sogn og Fjordane, Aust-Agder saman med Oppland og Nord-Trøndelag. Det vil seie at klassestorleiken heller ikkje forklarar forskjellen mellom fylka.

Ser vi tilsvarande på andelen kvinner i dei ulike fylka som har vidaregåande skule eller meir utdanning får vi tala 1,752, 1,724 og 1,762 for høvesvis Sogn og Fjordane, Aust-Agder saman med Oppland og Nord-Trøndelag. Heller ikkje denne variabelen ser ut til å forklare forskjellen.

Talet på elevar i klassen og andelen kvinner med minst vidaregåande utdanning ser ut til å slå positivt ut for Sogn og Fjordane. Samla sett predikerer modellen at Sogn og Fjordane skulle gjere det omtrent 5,5 hundredelar betre enn Aust-Agder med Oppland på desse to variablane i lag. Dette er nok ikkje signifikant, men på den andre side er fem hundredelar ca. 25% av Sogn og Fjordan-effekten som er estimert til ca. 0,23. Og det er denne effekten vi forsøker å forklare. Med andre ord er det i denne samanheng snakk om å forklare ein tross alt relativt beskjeden effekt! Dessutan ville det vere utruleg om vi faktisk hadde identifisert dei variablane som forklarte heile denne effekten.

Vi finn også at mange elevar i klassen har negativ effekt. Det kan vere nærliggjande å tenkje at effekten ikkje er lineær, men at det er særleg negativt med store klassar kanskje over 25 elevar. Vi har forsøkt med ein kvadratfunksjon, $ax^2 + bx + c$, men det synes som om berre det lineære leddet har effekt.

Det er kjent at spesielt kvinners utdanning har betydning for elevars prestasjonar (Turmo & Lie, 2004). Også vi finn denne samanhengen. Effekten er forholdsvis sterk med standardisert koeffisient ca 0,29. Han er likevel i storleiksorden med tidlegare funn, faktisk nokså lik den Turmo & Li (2004, s39) finn. Dei finn at mors og fars utdanningsnivå gir til saman ein standardisert koeffisient lik 0,11+0,16=0,27. (Merk at sidan vi har multikolinaritet mellom mors og fars utdanning, er det berre summen av desse estimata det kan festast lit til og effekten av summen er truleg nokså samanliknbar med berre ein av variablane.) På bakgrunn av at vi berre har andelen kvinner i kommunen med minst vidaregåande utdanning og ikkje direkte kopling mellom mor og barn er effekten likevel kanskje større enn vi skulle forvente. I utgangspunktet brukar vi denne variabelen til å korrigere observasjonane. Rimelegvis er det multikolinaritet mellom utdanningspredikatorane som har med utdanning for kvinner og menn å gjere. Vi har variablar for andelen av befolkninga i kommunane som har høvesvis vidaregåande skule, kort høgare utdanning og lang utdanning. Sidan vi har multikolinaritet mellom desse variablane er estimata av dei tilhøyrande koeffisientane usikre. Men for oss er det ikkje viktig å identifisere kor mykje dei enkelte av desse variablane har å seie. Derimot er det viktig for vårt spørsmål at det ikkje er multikolinaritet mellom utdanningsnivået i kommunen og lærarens utdanning, noko det ikkje ser ut til å vere. Det er signifikant samanhengen mellom andelen i kommunen som har minst vidaregåande skule og prestasjonane til elevane. Merk at vi ikkje har mors utdanning til den enkelte elev, men kvinners utdanning i kommunen. Det er nærliggande å tru at det er mors utdanning til den observerte elev som har betyding, men det kan også tenkjast at kvinners utdanning i nærmiljøet har positiv effekt sjølv på elevar som ikkje har mor med utdanning. Slike fenomen er ikkje ukjente, blant anna finn Valijarvi & Malin (2003) at elevar har fordel av å gå på ein skule med eit høgare sosioøkonomisk nivå enn sitt eige. Vi har ikkje data til å avgjere dette spørsmålet, men vi peikar på kva konklusjonar vi ikkje kan trekkje.

Som nemnt meiner vi at effekten av utdanning er relativt stor. Dei standardiserte koeffisientane er

- 0,269 for andelen kvinner med minst vidaregåande skule
- $-0{,}214$ for antal elevar i klassen
- $0{,}194$ for kvadratrota av lærarens utdanning

Med dei standardiserte koeffisientane kan vi samanlikne desse effektane direkte. Vi ser at betydinga av lærarens utdanning og antal elevar i klassen ikkje ligg veldig langt etter kvinners utdanningsnivå som tross alt er kjent som ein av dei mest betydingsfulle variablar på elevars prestasjonar (Turmo & Lie, 2004). Med andre ord ser det ut til at både lærarens utdanningsnivå og klassestorleiken er to særs viktige variablar for skuleresultata. Merk likevel at vi ikkje kan legge for stor vekt på estimata, dei er svært usikre, men vi har rimeleg grunn til å tru at desse variablane har effekt. I praksis er effekten av klassestorleiken likevel avgrensa. Om vi ser på ein klasse med 15 elevar versus 25 elevar, er effekten forventa å vere ca. 0,13 karakterpoeng, eller ca. 10% av eit standardavvik. Det er elles verdt å merke seg at Bonesrønning & Iversen (2010, s. 64) finn at store skular gjer det betre enn små skular i Sogn og Fjordane. Vi har ikkje skulestorleiken med i vår analyse, men det er rimeleg å anta at skulestorleiken er positivt korrelert med klassestorleiken.

Det har vore gjort mange forskingsarbeid på effekten av klassestorleiken. Konklusjonane er ikkje eintydige, men mange viser liten til ingen effekt. Verken Chingos (2012) med data frå Florida, eller Hoxby (2000) med data frå Connecticut fann klare bevis for effekt av klassestorleiken, andre studiar konkluderer med ein liten effekt (Cho, Glewwe, & Whitler, 2012) (Rivkin, Hanushek, & Kain, 2005). Dei nemnde undersøkingane gjeld amerikanske forhold og kan ikkje utan vidare overførast til norske tilhøve. Wössmann & West (2006) med data frå TIMSS finn berre minimal til ubetydeleg effekt av klassestorleiken i åtte av dei elleve landa i undersøkinga, og betydeleg (negativ) effekt av klassestorleiken i Hellas og Island. (Noreg er ikkje med i undersøkinga.) Dei hevdar at betydinga av klassestorleiken kjem an på skulesystemet, og at effekten berre er til stades i land der sitat: «the average capability of the teaching force appears to be low» (s. 727). Bonesrønning (2003) finn også at klassestorleiken har ulik effekt for ulike elevgrupper. Jakubowski & Sakowski (2006) meiner at å auke klassestorleiken i rurale område i Polen, vil auke gapet i eksamensresultat mellom urbane og rurale område ytterlegare. Angrist & Lavy (1999) finn også klassestorleikseffekt i Israel, men også her spriker resultata noko, til dømes mellom klassetrinna. Iversen & Bonesrønning (2013) finn at norske elevar i klassar med 15 elevar gjer det 22% av eit standardavvik betre på Nasjonale prøver på 4. trinn enn når det er 23 elevar i klassen. Bonesrønning & Iversen (2010) finn at denne forskjellen berre er på 3% på Nasjonale prøver på 8. trinn. Både Iversen & Bonesrønning (2013) og Bonesrønning & Iversen (2013) tyder på at det er elevar med lågt utdanna foreldre som har mest nytte av små klassar. Leuven mfl. (2008) finn ikkje nokon effekt av klassestorleiken på eksamenskarakterar i Norge.

Vi finn ikkje forskjell mellom allmennlærarar og fagutdanna lærarar. Det må seiast at vi har lite data til å sjå ein slik eventuell forskjell.

Vi finn ingen effekt av lærarens erfaring verken i heile datasettet eller innanfor kvart fylke, ei heller om vi ser på kvart fylke for kvart år. P-verdiane for lærarens erfaring er faktisk ikkje i nærleiken av å vere signifikante. Dette er litt overraskande, spesielt sidan datasettet er godt nok til å vise effekt av utdanning. Vi har også forsøkt å sjå om effekten eventuelt kan vere ikkje-lineær slik at han ikkje er synleg i ein lineær modell. Til dømes at det er ein forskjell mellom dei som har svært lite erfaring, til dømes mindre enn 5 år og dei som har meir. Men heller ikkje då er vi nær ved å finne nokon effekt av lærarens erfaring. Merk at vi berre har registrert undervisningsår generelt og ikkje antal år som matematikklærar. Det kunne sjølvsagt gitt ein forskjell, men på den andre sida er desse truleg korrelerte, og vi skulle derfor kunna sjå effekten også i vårt datasett.

Oppsummering og konklusjon

Det kan kanskje sjå ut til at respondentane frå Sogn og Fjordane ikkje er eit tilfeldig utval frå fylket. Meir presist, det synest som respondentane har betre karakterar enn det ein skulle forvente om dei var eit tilfeldig utval.

Om vi likevel studerer datasettet tyder det på at lærarens utdanning har betydning for elevane sine prestasjonar. Det synes å vere ein ikkje-lineær samanheng. Datasettet peikar mot ei god tilpassing mellom kvadratrota av lærarens studiepoeng (utover matematikk 1) og resultata til elevane. Det vil seie at dei første studiepoenga har sterkast påverknad på elevens prestasjon. Nærare bestemt estimerer modellen at 30stp gir ein forventa effekt på ca. 0,18 og 60 studiepoeng gir ein forventa effekt på $0{,}18\sqrt{2} \approx 0{,}24$. Det gir høvesvis ein effekt på 16% og 22% av standardavviket. Dette må vurderast som ein stor effekt av lærarens utdanning.

Vi kjenner ikkje til at det har vore dokumentert liknande sterk effekt mellom lærarens utdanning og resultata til elevane. Til dømes rapporterer Goldhaber og Brewer (2000) om ein samanheng på ca. 8% av standardavviket for testen. Falch & Naper (2008) rapporterer at andelen lektorar i matematikk på skulen har positiv effekt, men dei ser ingen effekt av lågare grads utdanning. Dei rapporterer effekten til 0,08 multiplisert med andelen lektorar. Merk at dette resultatet ikkje er så lett å tolke. Det er derimot lett å feiltolke, det kan nemleg vere nærliggande å tolke det som at om ein klasse vart undervist av ein lektor ville forventa forbetring vere 0,8. Men det er ei feiltolking. Grøgaard (2012, s. 80) finn ingen effekt av andelen lærarar på skulen med godkjent lærarutdanning, ei heller finn han at lærartettleiken har effekt. Det må nemnast at verken Falch & Naper eller Grøgaard ser ut til å ha observasjonar som knyter lærars utdanning til klasse. Dei har berre data på skulenivå. Vi har med andre ord mykje betre data for å sjå effekt av lærars utdanning, og det kan vere ei plausibel forklaringa på at effekten er mykje tydlegare hjå oss.

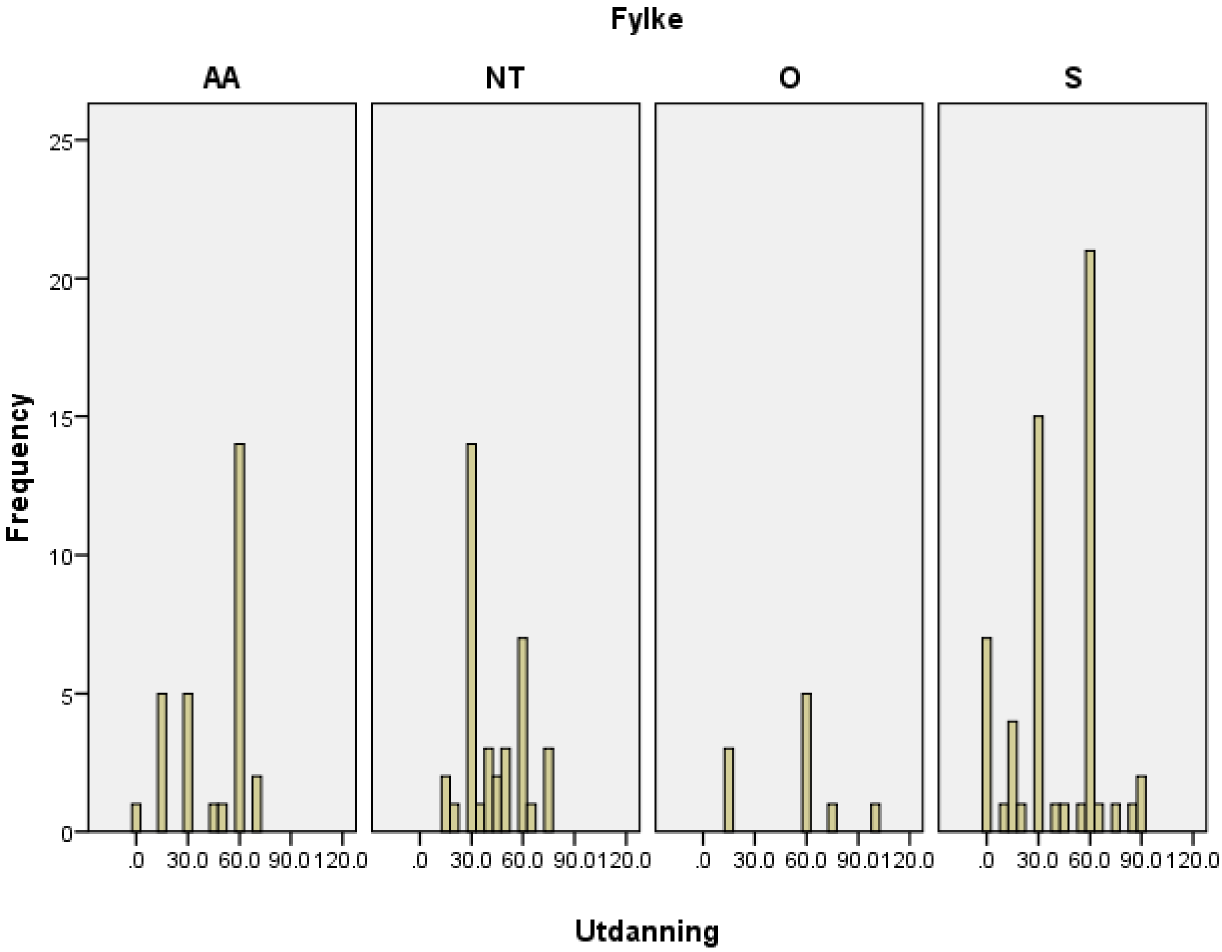
Kanskje unødig å seie, men ein samanheng mellom lærarens utdanning og prestasjonane til elevane, treng sjølvsagt ikkje vere eit resultat av utdanninga. Det kan like gjerne skuldast at personar som i utgangspunktet ville vere gode matematikklærar har ein tendens til å studere meir matematikk.

Vi finn også at mange elevar i klassen har negativ effekt. Det synes dessutan å vere ein lineær effekt.

Det er kjent at spesielt kvinners utdanning har betydning for elevars prestasjonar. Også vi finn denne samanhengen. Det vil seie at vi finn samanheng mellom andelen kvinner i kommunen med minst vidaregåande utdanning og resultatet på eksamen.

Det er likevel slik at ingen av desse variablane forklarar forskjellen mellom fylka.

I motsetnad til det vi kanskje skulle tru, finn vi ingen effekt av erfaringa til læraren målt i antal år som lærar. Vi er ikkje nær ved å sjå effekt av erfaringa verken i heile datasettet eller innanfor kvart fylke, ei heller om vi ser på kvart fylke for kvart år. P-verdiane for lærarens erfaring er faktisk ikkje i nærleiken av å vere signifikante.


Fylke
AA
NT
O
S
Frequency
25
20
15
10
5
0
.0
30.0
60.0
90.0
120.0
Utdanning

Referansar

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

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

Bonesrønning, H., & Iversen, J. V. (2010). *Presteasjonsforskjeller melom skoler og kommuner: Analyse av nasjonale prøver 2008.* (SØF-rapport nr. 01/10). Trondheim: Senter for økonomisk forskning AS.

Chingos, M. M. (2012). The impact of a universal class-size reduction policy: Evidence from Florida's statewide mandate. *Economics of Education Review, 2012*(31), pp. 543-562.

Cho, H., Glewwe, P., & Whitler, M. (2012). Do reductions in class size raise students' test scores? Evidence from population variation in Minnesota's elementary schools. *Economics of Education Review, 33*(3), pp. 77-95.

Falch, T., & Naper, L. (2008). *Lærerkompetanse og elevresultater i ungdomsskolen.* (SØF-rapport nr. 01/08). Trondheim: Senter for økonomisk forskning AS.

Goldhaber, D., & Brewer, D. (2000). Does Teacher Certification Matter? High School Teacher Certification Status and Student Achievement. *Educational Evaluation and Policy Analysis, 22*(2), pp. 129-145.

Grøgaard, J. B. (2012). *Hva kjennetegner barneskoler som oppnår høy skår på nasjonale prøver? Delrapport 5 fra prosjektet 'Ressurser og resultater i grunnopplæringen'.* (Rapport 38/2012). Oslo: NIFU.

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of Economics, 115*(4), pp. 1239-1285.

Iversen, J. V., & Bonesrønning, H. (2013). Disadvantaged students in the early grades: will smaller classes help them? *Education Economics, 21*(4), pp. 305-324.

Jakubowski, M., & Sakowski, P. (2006). Quasi-experimental estimates of class size effect in primary schools in Poland. *Educational Research, 45*(3), pp. 202-215.

Langfeldt, G. (2015). *Skolens kvalitet skapes lokalt. presentasjon av funn fra forskningsprosjektet "Lærende regioner".* Bergen: Fagbokforlaget.

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi-experimental estimates of the effect of class size on achievement in Norway. *The Scandinavian Journal of Economics, 110*(4), pp. 663-693.

Myklebust, T., & Norstein, A. (2015). Lærarkompetanse og skuleresultat. In G. Langfeldt (red.), *Skolens kvalitet skapes lokalt. Presentasjon av funn fra forskningsprosjektet "Lærende regioner".* (pp. 193-216). Bergen: Fagbokforlaget.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica, 73*(2), pp. 417-458.

Turmo, A., & Lie, S. (2004). Hva kjennetegner norske skoler som skårer høyt i PISA 2000? *Acta Didactica, 2004*(1), pp. 1-65.

Välijärvi, J., & Malin, A. (2003). The two-level effect of socioeconomic background. In S. Lie, P. Linnakylä, & A. Roe (Eds.), *Northern Lights in PISA. Unity and Diversity in the Nordic Countries in PISA 2000.* Department of Teacher Education and School Development, University of Oslo.

Wössmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variations in TIMSS. *European Economic Review, 50*(3), pp. 695-736.