

7

Hvordan har likeverd i norsk skole endret seg de siste 20 årene?

TRUDE NILSEN, JULIUS KRISTJAN BJØRNSSON OG ROLF VEGAR OLSEN

SAMMENDRAG Likeverd er et viktig aspekt ved norsk skole, og forskjeller i prestasjoner knyttet til hjemmebakgrunn er små i Norge sett i et internasjonalt perspektiv. Likevel er det mye som tyder på at dette kanskje er i ferd med å endres. Denne studien ser på hvordan likeverd har endret seg i løpet av 20 år, og bruker data fra TIMSS (4. og 8. trinn) og PISA (10. trinn) fra 1995 til 2015. Prestasjonsforskjeller mellom skoler (variansen) er stigende gjennom perioden for to av de tre trinnene, men har en nedgang i siste runde i 2015. Samtidig viser det overordnede bildet at spredningen i prestasjoner synker, for så å øke i siste del av perioden. Dersom man sammenligner PISA og TIMSS når det gjelder sammenhengen mellom elevers hjemmebakgrunn og prestasjoner, er ikke mønsteret så tydelig. På 8. trinn i TIMSS har sammenhengen mellom hjemmebakgrunn og prestasjoner på skolenivå doblet seg. Den totale sammenhengen øker kun svakt i TIMSS i perioden, men er uendret i PISA.

SUMMARY Equity is an important aspect of the Norwegian school system, and differences in achievement related to home background are small in Norway seen in an international perspective. However, previous research indicates that this might be changing. This study examines how equity may have changed during the last 20 years, using data from both TIMSS (grade 4 and 8) and PISA (grade 10) for the whole period from 1995 to 2015. The differences in achievement (variance) at the school level increases throughout the period, in two of the three student groups included, but decreases in the last cycle of the studies. Concurrently the variance at the student level generally decreases, but increases between the last two cycles.

When comparing results between PISA and TIMSS, the relation between student home background and achievement does not appear to have a consistent tendency through-

hout the period. With the exception of grade 8 at the school level, where the strength of this relation has doubled between 1995 and 2015, the relationship between student home background and achievement increases slightly through the period for TIMSS while it is unchanged in PISA.

INNLEDNING

I de nordiske landene legges det betydelig politisk vekt på at utdanningen skal bidra til sosial utjamning, og det er rimelig å hevde at utdanningspolitikken i så måte har virket etter hensikten: Forskjellene mellom skoler har over tid vært små sett i et internasjonalt perspektiv (Blossing, Imsen & Moos, 2014; Hansen, Gustafsson & Rosén, 2011; OECD, 2016), og sammenhengen mellom hjemmebakgrunn og prestasjoner har vært relativt svak (Mullis, Martin, & Hooper, 2016; OECD, 2016). Imidlertid har flere studier i de siste årene indikert at forskjellene er økende i flere av de nordiske landene, inkludert Norge. (Se f.eks. Bakken & Elstad, 2012; Hansen, 2015; OECD, 2013; OECD, 2016.)

Det er likevel vanskelig å fastslå en slik utvikling – hovedsakelig av tre grunner. For det første må studier av endringer over tid være basert på mål eller indikatorer som er de samme over tid. I de internasjonale studiene har målene for sosioøkonomisk status (SES) gjerne endret seg litt fra én syklus til den neste. I analyser av nasjonale data brukes gjerne karakterer (standpunkt og/eller eksamen) som indikator for læringsutbytte, og dette er også mål som kan antas å endre seg over tid. For det andre reflekterer datasettene som studeres, kompliserte design hvor ikke alle elever svarer på de samme spørsmålene, noe som gjør at dataene må settes sammen på en bestemt måte. I tillegg har dataene et såkalt hierarkisk design med flere nivåer, for eksempel fordi de valgte elevene tilhører bestemte klasser, som igjen tilhører bestemte skoler, eller alle elever i en viss alder på valgte skoler blir inkludert i utvalget. Man må derfor benytte metoder som tar hensyn til både kompleksiteten i de prøvene som er brukt, og det hierarkiske designet. For det tredje har utjamning eller likeverd i utdanningen flere mulige både teoretiske og operasjonelle definisjoner, noe som kan føre til at man får litt ulike svar avhengig av hva man legger i begrepet, og hvordan man måler det.

Vel vitende om disse utfordringene ønsker vi i dette kapitlet å se på utviklingen av likeverd i Norge ved å bruke data fra alle syklusene som er gjennomført av de to studiene TIMSS og PISA. På denne måten søker vi å finne svar på spørsmålet: I hvilken grad har norsk skole blitt mer eller mindre utjammende i de siste 20 årene?

OM UTJAMNING OG LIKEVERD

Likeverd og utjamning er to begreper som ofte brukes for å karakterisere en kvalitet som de fleste utdanningssystemer har som mål å etterstrebe. Viktigheten av dette kommer blant annet fram i rapportene fra de internasjonale undersøkelsene, og temaet behandles ofte i egne dedikerte kapitler i hovedrapportene eller i egne delrapporter (f.eks. OECD, 2016). Kjernen i denne kvaliteten er relativt enkel å oppsummere: Alle elever skal få et utdanningstilbud som er likt, og som har samme (høye) kvalitet. Men når man begynner å grave litt mer i hvordan man skal kunne vurdere i hvilken grad et utdanningssystem virker utjammende, blir dette et for uklart utgangspunkt. Man må definere kjennetegnene eller dimensjonene ved elevene og skolene som man ønsker å ha utjamning langs, og man må etablere en operasjonell definisjon som angir hva det er som skal være likt eller utjevnet for ulike grupper elever. Dersom man til slutt skal trekke evaluerende konklusjoner, må man også ha en begrunnet formening om i hvilken grad det kan forventes likhet på tvers av disse kjennetegnene ved elevene.

I et internasjonalt perspektiv kompliseres dette ytterligere ved at en bestemt definisjon og operasjonalisering ikke nødvendigvis fungerer like godt på tvers av svært ulike land – både fordi målingene kan fungere ulikt, og ikke minst fordi de utdanningspolitiske målene kan være svært ulike i forskjellige deler av verden (Rutkowski, Rutkowski, & Plucker, 2012). I noen land kan ambisjonen være at *alle* barn skal få et skoletilbud. I andre land kan hovedfokuset være å redusere eksplisitt forskjellsbehandling, slik at målet er å gi underprivilegerte grupper (f.eks. jenter) et like godt utdanningstilbud som privilegerte grupper. I mange land er innsatsen rettet mot å redusere implisitt eller mer skjult diskriminering. En slik implisitt diskriminering kan for eksempel være at utbyttet av utdanningen er ulik for elever med ulik bakgrunn. I dette kapitlet bruker vi data fra de internasjonale undersøkelsene, men vi ser likevel kun på situasjonen i Norge. Utfordringene knyttet til sammenliknbarhet på tvers av land kan likevel være relevant også her fordi analysen inkluderer en sammenlikning av ulike kohorter over tid. Det kan eksempelvis tenkes at selv om variablene som analyseres i seg selv er uendret over tid, så kan variabelens funksjon som markør for sosioøkonomisk status endre seg over tid. For eksempel kan det tenkes at antall bøker hjemme, som har fungert som en indikator for sosioøkonomisk status, ikke lenger reflekterer sosioøkonomisk status fordi folk ikke lenger kjøper bøker i papirutgaver.

I den norske konteksten er likeverd en av grunntankene i utdanningssystemet og samfunnet forøvrig, og det er et uttalt politisk mål at alle elever skal ha lik rett til utdanning, uavhengig av for eksempel hjemmebakgrunn, kjønn, bosted og etnisitet. Det er også bred politisk enighet om at en av skolens viktigste oppgaver er å

bidra til å utjevne sosiale forskjeller mellom elever (Kunnskapsdepartementet, 2006). I praksis ser vi at en slik politikk gjennomføres blant annet ved hjelp av kompensatoriske tiltak, for eksempel ved å gi høyere lønn til lærere eller større tildelinger til skoler i regioner eller områder hvor man har elever med gjennomgående lavere SES, for eksempel fordelingsmodellen til skoler i Oslo hvor ressurser fordeles kompensatorisk,³⁸ eller ordningen for avskrivning av lån i Statens lånekasse for lærere ved skoler i Nord-Troms og Finnmark.³⁹

OM MÅLING AV LIKEVERD

Det finnes mange mål på likeverd, og de kan hovedsakelig deles i tre grupper: 1) spredningsmål, 2) sammenhenger mellom kjennetegn ved elevene (kjønn, etnisitet, foreldrenes utdanningsnivå osv.) og prestasjoner og 3) andel elever uten et minimum av elementære ferdigheter (Strietholt, 2014).

Spredningsmål eller variasjon i prestasjoner blir ofte brukt som et mål på likeverd (Mullis, Martin & Hooper, 2016; OECD, 2016). Dette målet knyttes av Strietholt (2014) til en egalitær ideologi hvor det ligger en forventning om små forskjeller i elevenes prestasjoner. I Norge og Norden forøvrig har denne ideologien hatt stor betydning for utviklingen av utdanningssystemet (Blossing et al., 2014).

Spredning eller variasjon kan angis på flere måter, for eksempel ved differansen mellom ulike skårprosentiler eller standardavvik i skår for ulike sentrale læringsutbytter. En interessant observasjon er at land med høye prestasjoner også har gjennomgående lave standardavvik (Gustafsson, Nilsen & Hansen, 2016; Kyriakides & Creemers, 2011; Mullis, Martin & Hooper, 2016; OECD, 2016). I TIMSS og PISA tilhører Norge gruppen av land med relativt sett lave standardavvik på begge trinn, men hvor prestasjonene i gjennomsnitt gjerne har vært omkring det internasjonale gjennomsnittet eller lavere. (Norge er sånn sett et unntak fra regelen om at små standardavvik er assosiert med høye prestasjoner.) Det er altså gjennomgående mindre forskjeller i elevenes prestasjoner i Norge enn i mange andre vest-europeiske land (Mullis, Martin & Hooper, 2016; OECD, 2016).

I tillegg kan spredningen mellom elever deles inn i (minst) to komponenter: Én del som er knyttet til at det er spredning i skolers gjennomsnitt, og én del som er

38. Noe som skaper reaksjoner fra skoler som er forfordelt, se <https://www.aftenposten.no/meninger/kronikk/i/9mlq2q/22-rektorer-pa-Oslo-vest-Grensen-for-skjevdeling-er-nadd-for-vare-skoler>

39. Se <https://www.regjeringen.no/no/aktuelt/satsar-pa-nye-lararar-forsking-og-samisk-i-nord-noreg/id2552838/>

knyttet til at det er forskjeller mellom elever på samme skole. Denne og andre former for dekomponering av spredningen i resultatene er et viktig analytisk grep for bedre å kunne identifisere om man bør iverksette universelle tiltak for alle elever, tiltak for bestemte grupper elever innen alle skoler eller tiltak som skal kompensere for ulikheter mellom skoler.

En annen og mer liberal definisjon av likeverd aksepterer at forskjeller mellom elever og skoler er unngåelige, men at slike forskjeller i så liten grad som mulig skal være knyttet til skole og elevkarakteristikker, som f.eks. kjønn, etnisitet, religion, geografi eller hjemmebakgrunn. Dette er opplagt en langt mer kompleks definisjon, som medfører at man må se på fordelingen av resultater langs et stort antall karakteristikker. Dette kapitlet er imidlertid kun begrenset til å se på sosioøkonomisk utjamning. I analysene legges det derfor vekt på sammenhengen mellom mål på elevenes sosioøkonomisk status (SES) og skolefaglige prestasjoner. Med et slikt perspektiv på utjamning er Norge gjennomgående på nivå med andre vestlige land som det er naturlig å sammenlikne seg med (Mullis, Martin, Foy & Hooper, 2016; OECD, 2013, 2016).

Et tredje mål på likeverd angis gjerne i form av andel elever under et laveste akseptabelt kompetansenivå. Dette er basert på ideen om at alle individer har en rett til å oppnå det minimumsnivået av helt elementære ferdigheter som er nødvendig for å klare seg i samfunnet. Dette kapitlet inneholder imidlertid ingen analyser som ser på utjamning fra et slikt perspektiv. Disse tre ulike måtene å studere likeverd på gir generelt ikke sammenfallende resultater (Strietholt, 2014, 2016).

HVA ER SES, OG HVORDAN MÅLES DET?

Sosioøkonomisk status (SES) er et begrep som inkluderer sosiale, kulturelle og økonomiske aspekter knyttet til hjemmet (Bourdieu, 1986). Det er vanlig å bruke indikatorer som foreldres yrke, utdanningsnivå, og økonomi, antall bøker hjemme og familiestørrelse (ibid.). I de fleste storskalaundersøkelser måles SES gjennom indikatorer knyttet til økonomiske og kulturelle karakteristika ved elevenes hjemmebakgrunn (Hansen & Munk, 2012; Martin, Foy, Mullis & O'Dwyer, 2013; Nilsen & Bergem, 2016; OECD, 2016). Bourdieu mfl. (1993) brukte blant annet antall bøker i hjemmet og foreldrenes utdanningsnivå som mål på kulturell kapital i sine studier. Marks, Cresswell, and Ainley (2006) fant at hjemmets kulturelle kapital hadde større betydning for elevenes prestasjoner i PISA enn indikatorer for hjemmets økonomiske status.

Antall bøker i hjemmet (som et uttrykk for kulturell kapital) har vist seg å ha spesiell sterk relasjon til elevers prestasjoner, og dette målet på SES har blitt brukt i

TIMSS og PISA helt siden oppstarten av studiene (Beaton et al., 1997; Mullis, Martin & Hooper, 2016; OECD, 2013; Strietholt, 2016). I PISA bruker man et stort antall variabler som sammen med antallet bøker i hjemmet kombineres til et felles mål for sosioøkonomisk hjemmebakgrunn, ESCS (Economic, Sosial and Cultural Status): foreldres utdanningsnivå og yrke og forekomsten av ulike gjenstander som reflekterer henholdsvis økonomisk, kulturell og sosial kapital (OECD, 2016; Olsen & Turmo, 2010). I TIMSS har man i de siste studiene rapportert en liknende indeks som er satt sammen av foreldres utdanningsnivå, antall bøker i hjemmet og hjemmets ressurser generelt (Mullis, Martin, Foy, et al., 2016). I tillegg til utfordringer med internasjonal sammenliknbarhet (Rutkowski & Rutkowski, 2013) er en hovedutfordring med disse sammensatte SES-indeksene at det ikke finnes en god teoretisk begrunnelse for vektingen av indikatorene som inngår i storskalaundersøkelser. I tillegg er det nesten umulig å finne gode indikatorer som kan selvrappeteres av elever.

I de internasjonale undersøkelsene brukes altså målene på SES for å kunne si noe om grad av likeverd i skolesystemer. I mer avanserte analyser av likeverd ønsker man også å undersøke hvilke faktorer (for eksempel læreres kompetanse) som kan bidra til å styrke eller svekke sammenhengen mellom SES og prestasjoner (moderasjon), eller som kan være med på å bidra til å øke/minke ulikheter mellom elever (mediering) (Gustafsson et al., 2016). Men SES-indeksene brukes også i sammenhenger hvor man hovedsakelig er interessert i å studere sammenhenger mellom to andre variabler, men hvor det er viktig å korrigere for (eller «viske bort») innvirkningen av SES på denne sammenhengen. I tillegg har målene for SES en viktig funksjon (sammen med mange andre bakgrunnsvariabler) ved at de inngår i prosedyrer for å beregne såkalte plausible verdier for elevenes skårer. (Se f.eks. Martin & Mullis, 2012.) Til sammen tilsier dette at det er svært viktig å ha gode mål for elevers hjemmebakgrunn i disse studiene.

Det viser seg at sammenhengen mellom SES og prestasjoner varierer mye fra studie til studie, og i en metaanalyse konkluderte Sirin (2005) med at denne variasjonen i betydelig grad skyldes måten SES ble målt på. Korrelasjonen mellom SES og elevprestasjoner ligger i de fleste land på 0,2 til 0,4 på elevnivå, og på klasse og skolenivå er denne korrelasjonen betydelig høyere, ofte fra 0,6 og høyere (Sirin, 2005; White, 1982).

De fleste studier har konkludert med at elever som går på skoler med lav SES, presterer dårligere enn elever som går på skoler med høy SES, selv etter å ha kontrollert for elevenes hjemmebakgrunn (Gustafsson et al., 2016; Liu, Van Damme, Gielen & Van Den Noortgate, 2015; Palardy, 2013; Schmidt, Burroughs, Zoido & Houang, 2015; Van Ewijk & Sleegers, 2010; Willms, 2010). Når man måler sko-

lens SES, og i hvilken grad skolen eventuelt påvirker gapet mellom elever, er det to måter å gjøre dette på. Skolens rektorer kan få spørsmål som dreier seg om skolens SES, for eksempel om hvorvidt skolen ligger i et ressurssterkt nabolag. Den andre måten å gjøre dette på er å beregne såkalte komposisjonseffekter (på engelsk: *compositional effects*), hvor man beregner effekten av SES aggregert til skolenivå korrigert for effekten av SES på elevnivå innen skolen (Gustafsson et al., 2016; Willms, 2010). Man kan argumentere med at disse også er to forskjellige teoretiske konstrukter, men i tillegg skiller disse to målene seg fra hverandre ved at det ene er aggregert og det andre ikke.

Studier som benytter slike metoder, har vist at det er grunnlag for å anta at forskjeller i læringsutbytte på forskjellige skoler delvis skyldes sosiale og institusjonelle faktorer som er relatert til skolens SES, i tillegg til de individuelle elevenes SES. Med andre ord kan skolene bidra til å redusere eller øke forskjeller mellom elever (Schmidt et al., 2015; Willms, 2010).

UTVIKLING AV LIKEVERD I NORSK SKOLE

Det finnes flere studier av betydningen av SES også i norsk skole, men det er få studier som ser på utviklingen over tid. Bakken (2004) brukte data fra undersøkelsen «Ung i Norge», og fant en svak tendens til at betydningen av SES økte fra 1992 til 2002. I en nyere rapport viser Bakken og Elstad (2012) gjennom analyser av registerdata i perioden 2002–2011 at sammenhengen mellom SES og skoleprestasjoner var stabil i første del av 2000-tallet, men at det er et knekkpunkt med økende ulikhet som i tid sammenfaller med innføringen av Kunnskapsløftet i 2006. Videre rapporterer denne studien også resultater fra en flernivåanalyse. Denne viser at den sosiale segregeringen i skoler mellom elever ikke har endret seg betydelig. I en videre analyse av disse dataene finner de at en elev med gjennomsnittlig SES har en fordel av å gå på en skole med høyere SES. I tillegg finner de støtte for en hypotese om at skoler med en positiv utvikling i gjennomsnittlige resultater i perioden også har en reduksjon i betydningen av SES.

Også Bonesrønning, Falch og Strøm (2005) peker på at elever ved skoler med høy SES (altså skoler med mange elever med høy SES) presterer bedre, og viser at norske skoler er med på å øke forskjellen mellom elevene. Dette knyttes til at skoler med en relativt stor andel elever med høy SES også har de best kvalifiserte lærerne (Bonesrønning et al., 2005). Dette vil bety at de mest ressurssterke elevene får tilgang til de beste lærerne, noe som fører til at gapet mellom elevene kan bli ytterligere forstørret. Dette støttes av annen forskning fra andre land (Gustafsson et al., 2016; Schmidt et al., 2015).

Mullis, Martin og Hooper (2016) sammenstiller utviklingen av likeverd i TIMSS-studiene gjennom de siste 20 årene. Formålet med denne rapporten er altså sammenfallende med formålet for dette kapitlet, men likeverd er her definert som forskjeller i spredning, og mer spesifikt som differansen i skår mellom høyt presterende 4. trinns elever i matematikk (de som skårer over 90-prosentilene) og lavt presterende (de som skårer under 10-prosentilene). Forskjellene mellom disse gruppene er generelt store i alle land, men i Norge ble gapet redusert med 12 poeng fra 1995 til 2015 i matematikk på 4. trinn. Dette kan hovedsakelig knyttes til en økning i prestasjonene til de lavt presterende elevene. For naturfag ble gapet redusert med hele 48 prosentpoeng i løpet av de 20 årene, men dette kan hovedsakelig knyttes til at prestasjonen til den tidelen av elevene med høyeste skårer var betydelig lavere i 2015 enn i 1995. Tall for 8. trinn ble ikke analysert for hele perioden, men rapporten viser at gapet i naturfag på 8. trinn økte mellom 2011 og 2015, mens det ikke var noen signifikant endring i matematikk.

I tillegg må det nevnes at de dataene som analyseres i dette kapitlet, også har blitt analysert og presentert i flere av de nasjonale og internasjonale rapportene som utgis etter hver studie. I noen grad presenterer disse rapportene også endringer over tid både for spredningsmål og sammenhenger mellom prestasjoner og ulike kjennetegn ved elevene. Men disse endringsanalysene er gjerne i form av sammenlikninger mellom to påfølgende studier. Dette kapitlet har som hovedformål å oppsummere slike resultater over langt flere studier og over et større tidsrom. Vi legger derfor mindre vekt på endringer fra én studie til den neste.

FORSKNINGSSPØRSMÅL

Med unntak av studien til Bakken og Elstad (2012) kjenner vi ikke til studier av hvordan forskjeller mellom *skoler* har utviklet seg de siste tiårene i Norge, og i hvilken grad disse forskjellene kan knyttes til SES. I dette kapitlet presenteres derfor resultater fra såkalte flernivåanalyser som oppsummerer hvordan resultatforskjellene (i form av varians) fordeler seg mellom skoler og mellom elevene på samme skole. Videre viser vi i hvilken grad disse varianskomponentene over tid kan knyttes til SES på skole- og elevnivå.

Analysene som presenteres i dette kapitlet, søker å finne svar på følgende forskningsspørsmål:

1. Hvordan har forskjeller i prestasjoner (målt ved varians) endret seg over tid, både totalt og dekomponert i andeler knyttet til henholdsvis skoler og elever?

2. Hvordan har andelen varians – forklart av SES – endret seg over tid, både totalt og dekomponert i andeler knyttet til henholdsvis skoler og elever?
3. Hvordan har størrelsen eller gradienten til relasjonen mellom SES og prestasjoner endret seg over tid, både mellom skoler og mellom elever?

DATA OG METODE

I denne studien var utgangspunktet å inkludere data fra alle sykluser i PISA og TIMSS i analysene. TIMSS ble gjennomført i 1995, 1999, 2003, 2007, 2011 og 2015. PISA ble gjennomført i 2000, 2003, 2006, 2009, 2012, og 2015. Resultater blir kun presentert for matematikk. PISA 2000 er imidlertid tatt ut av analysene, siden den faglige prøven i matematikk kun inneholdt oppgaver fra to av de fire tematiske områdene i rammeverket for prøven. Sammenlikninger med resultater fra de senere studiene er dermed ikke gyldige. Videre deltok ikke Norge i TIMSS 1999.

PISA måler lesing, naturfag og matematikk for 15-åringer. Alle undersøkelsene har inkludert faglige prøver for alle de tre faglige områdene, men ett av fagområdene vektlegges hver gang med en rulling mellom de tre (lesing i 2000 og 2009, matematikk i 2003 og 2012 og naturfag i 2006 og 2015). TIMSS måler elevers kompetanse i matematikk og naturfag på 4. og 8. trinn. Måten TIMSS trakk sitt utvalg på i 1995 var annerledes enn alle de andre syklusene. Da deltok elever fra 2. og 3. trinn på barnetrinnet. På ungdomstrinnet deltok elever fra 6. og 7. trinn. Data fra denne undersøkelsen er likevel inkludert, men det anbefales å være forsiktig med å legge for mye vekt på punkttestimatene fra 1995 i tolkningen av trender. Spredningsmål vil eksempelvis være noe større dette året fordi resultater fra to klasses trinn er inkludert. Tabell 7.1 og 7.2 viser antall elever, skoler og klasser samt gjennomsnittlig matematikkskår for alle sykluser i PISA og TIMSS som er inkludert i analysene.

Utvalgsdesignet til begge studiene er likt ved at man først trekker et utvalg av skoler, men designene skiller seg ved at PISA trekker et tilfeldig utvalg elever fra skolene på tvers av klassene (gjennomsnittlig 23 elever per skole), mens TIMSS trekker ut hele klasser innen hver av skolene (i snitt 29 elever per skole). Hovedregelen i TIMSS er at det skal trekkes to klasser fra hver skole når den har flere klasser på trinnet (Foy, Arora & Stanco, 2013; Martin, Mullis & Hooper, 2016). Gjennomsnittlig antall elever per skole er derfor noe høyere i TIMSS enn i PISA.

Norge skiller seg fra mange andre land ved å ha mange små skoler. De fleste av disse små skolene har bare én klasse per trinn, hvor antall elever også kan være svært lavt. For å få et effektivt utvalg har derfor store skoler større sannsynlighet

for å bli trukket ut. Dette fører til at små skoler er underrepresentert i utvalget, men det fører samtidig til at alle elever i Norge har omtrent lik sannsynlighet for å bli trukket ut, uavhengig av størrelsen på skolen. I analysene brukes vektorer for å korrigere for slike ulike utvalgssannsynligheter for henholdsvis elev- og skolenivå.

Tabell 7.2 viser at antall klasser per skole varierer noe fra år til år i TIMSS, og antall elever per skole varierer også noe. Til sammen gir dette noen begrensninger knyttet til tolkning av endringer over tid, spesielt knyttet til dekomponeringen av varians mellom og innen skoler.

TABELL 7.1 Gjennomsnittlige prestasjoner i matematikk med standardfeil i parentes i alle sykluser for PISA

	2003	2006	2009	2012	2015
Matematikkskår	495 (2,4)	490 (2,6)	498 (2,4)	489 (2,7)	502 (2,2)
Antall elever	4064	4692	4660	4686	5456
Antall skoler	182	203	197	197	228

TABELL 7.2 Gjennomsnittlige prestasjoner i matematikk med standardfeil i parentes samt antall elever og skoler i alle sykluser for TIMSS 4. trinn og 8. trinn

		1995	2003	2007	2011	2015
4. trinn	Matematikkskår	476 (3,0)	451 (2,2)	473 (2,6)	495 (2,8)	493 (2,3)
	Antall elever	4227	4167	4009	3033	4230
	Antall klasser	275	228	290	197	219
	Antall skoler	140	139	145	119	140
	Antall klasser per skole	2,0	1,6	2,0	1,7	1,6
8. trinn	Matematikkskår	498 (2,2)	461 (2,5)	469 (2,0)	475 (2,5)	487 (2,0)
	Antall elever	5642	4053	4561	3821	4634
	Antall klasser	287	179	231	170	216
	Antall skoler	249	138	139	134	143
	Antall klasser per skole	1,2	1,3	1,7	1,3	1,5

Antallet bøker i hjemmet er det eneste målet på SES som er uendret siden undersøkelsene ble satt i gang, og spørsmålet som er stilt til elevene, er så og si identisk i de

to studiene. Den lille forskjellen består i at den øvre kategorien i TIMSS er «mer enn 200 bøker», mens PISA har en kategori for «201–500 bøker» og en for «mer enn 500 bøker». I analysene er derfor dataene fra PISA rekodet ved å slå sammen svarene i de to høyeste kategoriene. De siste tiårene har vi sett en rivende digital utvikling, og man kan derfor forvente at det er en nedgang i antallet bøker i elevenes hjem. Dette ser vi også tydelig i dataene fra undersøkelsene hvor gjennomsnittlig antall bøker i hjemmet faller med hele 40 % i TIMSS-undersøkelsene fra 1995 til 2015. I perioden 2003 til 2015, hvor alle de tre undersøkelsene kan sammenliknes, er nedgangen i antallet bøker 16–17 % i alle de tre populasjonene som er inkludert. I tillegg kan det faktum at vi nå leser mye mer på digitale medier ha ført til at variabelen over tid gradvis har endret betydning som markør for hjemmets sosiale og kulturelle ressurser. Likevel viser empirien at antall bøker i hjemmet fortsatt fungerer som en sterk indikator for kulturell kapital, og sammenhengen med prestasjoner er fortsatt betydelig (Gustafsson, Hansen & Rosèn, 2013; Hansen & Munk, 2012; Strietholt, 2014).

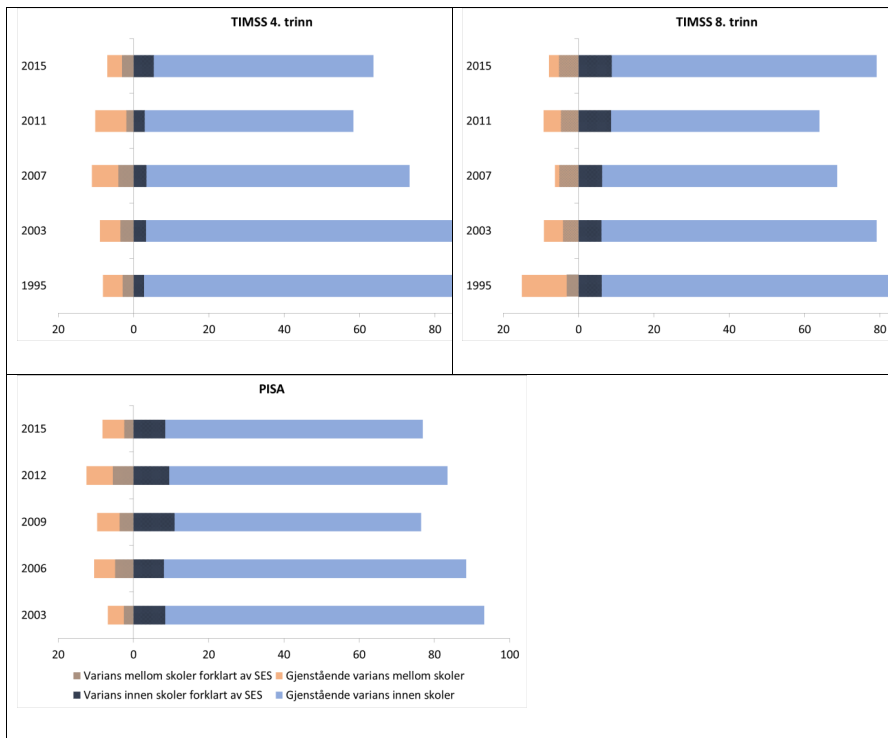
Et annet mål på SES som er inkludert i begge studiene, er spørsmål om foreldres utdanningsnivå. Dette spørsmålet varierer litt fra syklus til syklus i både PISA og TIMSS, både når det gjelder ordlyd og antall svaralternativer. I tillegg er det en stor andel elever som ikke svarer, og enkelte år har også kategorien «vet ikke» vært inkludert, noe som fører til at det ikke finnes data for mer enn 40–50 % av elevene enkelte år. Til sammen har dette gjort det vanskelig å inkludere denne variabelen i analysene. I dette kapitlet er derfor elevenes svar på antall bøker i hjemmet brukt som den eneste indikatoren for SES.

Analysene ble gjort på elev- og skolenivå i såkalte flernivå regresjonsanalyser i Mplus versjon 8.0 (Muthén & Muthén, 1998–2017). En slik analyse tar utgangspunkt i en dekomponering av variansen i to komponenter: én knyttet til forskjeller mellom skoler og én knyttet til forskjeller mellom elever innen samme skole. For enkelhets skyld referer vi videre ofte til disse to delene av løsningen som henholdsvis skole- og elevnivået. I en slik regresjonsanalyse modelleres elevenes matematikkprestasjoner som en funksjon av både elevens og skolens gjennomsnittlige SES. I tillegg til at programvaren Mplus gir estimater for regresjonskoeffisienter for begge nivåer, ivaretar det også en korrekt bruk av de fem såkalte plausible verdiene for hver elevs matematikkskårer (Rutkowski, von Davier & Rutkowski, 2013). Det gir gode muligheter for å håndtere at enkelte elever har manglende verdier for én eller flere variabler (MLR, maximum robust likelihood), og programmet gir samtidig estimater for varians og indekser for hvor godt modellen er tilpasset dataene (Hox & Roberts, 2011).

RESULTATER

Presentasjonen av resultatene følger samme rekkefølge som forskningsspørsmålene: Først presenteres resultater for variasjonen mellom elever, mellom skoler, og totalt (figur 7.1), deretter oppgis andelen varians på begge nivå som kan knyttes til skolenes og elevenes SES (figur 7.1, 7.2 og 7.3), og til sist gis resultatene fra regresjonsanalysen for sammenhengen mellom SES og prestasjoner (figur 7.4).

Figur 7.1 viser variansen i matematikkskår for alle undersøkelsene som er gjennomført for TIMSS (4. og 8. trinn) og PISA. Figurene er komplekse og krever en relativt grundig leseveiledning: Siden studiene har et lenket design, er alle skårene (og dermed også variansene) direkte sammenliknbare over alle syklusene. De originale varianskomponentene er imidlertid uttrykt på skalaer som ikke er umiddelbart enkle å tolke. I figur 7.1 er derfor variansene reskalert ved å sette variansen det første året for hver av de tre studiene lik 100. Alle varianser for påfølgende studier er uttrykt relativt til denne verdien.



FIGUR 7.1 Relativ fordeling av varians mellom og innen skoler, og andeler av varians forklart av SES. Se tekst for ytterligere forklaring.

Hele søylen for hvert år svarer til den totale variansen. I tillegg er variansen dekomponert på følgende måte: Søylen som peker mot venstre fra den vertikale akse, viser variansen som kan knyttes til at det er forskjeller i skolenes gjennomsnittlige prestasjoner, mens søylene som peker mot høyre, viser variansen som er knyttet til forskjeller i matematikkskår for elever innen skoler. I tillegg angir figuren den andelen av variansen som kan knyttes til elevens sosioøkonomiske bakgrunn som de mørkere/skraverte feltene. Disse andelen av variansen som kan forklares av SES mellom og innen skoler, er også vist mer lesbart i henholdsvis figur 7.2 og 7.3

Til sammen gir dermed de grafiske framstillingene i figur 7.1 oss muligheten til å se på:

- Hvordan forskjeller mellom elever totalt sett har utviklet seg over tid, gitt her som utvikling over tid for den totale variansen.
- Hvordan forskjeller mellom skoler har utviklet seg over tid, gitt her som utviklingen i mellom skolevarians over tid.
- Hvordan andelen av variansen mellom skoler og innen skoler som kan forklares ved SES, har utviklet seg over tid, her gitt som utviklingen i den totale bredden på de to mørke feltene for hver av søylene

Det er for øvrig verdt å legge merke til at vi ikke indikerer andelen varians mellom skoler ved hjelp av de såkalte intraklassekorrelasjonene (ICC), noe man vanligvis gjør for flernivåanalyser. Årsaken til dette er at ICC beregnes relativt til spredningen man har i de enkelte år. ICC for disse studiene over tid er dermed ikke sammenliknbare fordi den totale variansen også er ulik over tid. Av samme grunn bruker vi heller ikke kvadrert forklart varians (R^2) for å angi andelen varians som er forklart av SES.

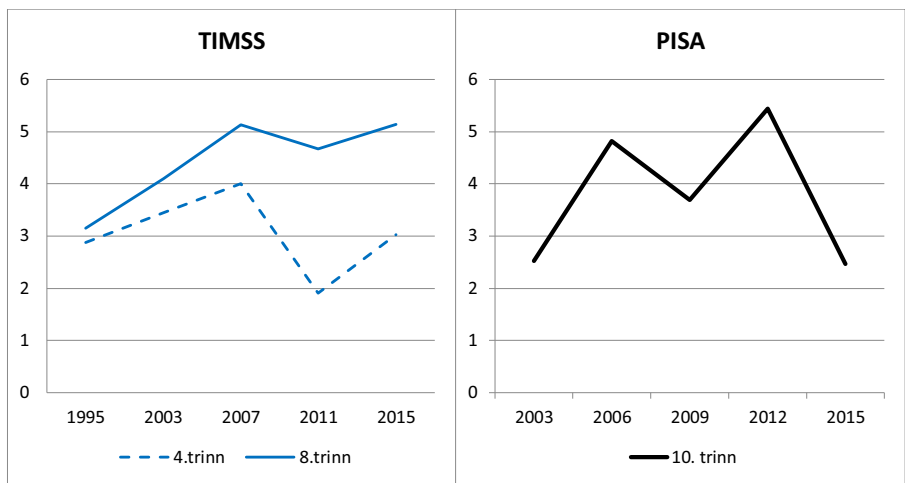
Det er ikke helt enkelt å oppsummere noen klare trender i utviklingen av varianskomponentene over tid. Den totale variansen har sin maksimalverdi i det første året for alle undersøkelsene for deretter å ha en overveiende synkende tendens over tid. Denne tendensen er tydeligst for 4. trinn i TIMSS. Det må imidlertid gjentas at variansen for 1995 kan antas å være større fordi elever fra to klassetrinn var inkludert. Ser vi helt bort fra søylen for dette året, er det likevel en svak tendens til synkende total varians. For PISA er den totale variansen over tid mer stabil, uten en tydelig utvikling. Her bør det nevnes at matematikk var fokusområde i 2003 og 2012. Det vil si at det var kun i disse årene at fagområdet ble inkludert i hele sin bredde, og dette kan ha virket inn på variansestimaterne. Fra 2003 til 2012 er det en liten nedgang på 5 %.

Når det gjelder variansen (og alle andre estimater) for skolenivået, er det viktig å huske på at disse er estimert med relativt stor standardfeil. Dette betyr at analysene knyttet til skolenivået ikke har styrke til å slå fast om små eller moderate endringer er statistisk signifikante. Hvis vi ser nærmere på forskjellene mellom skoler i figur 7.1, altså den delen av søylene som peker mot venstre, er det litt ulik utvikling for de ulike studiene. I PISA er det en signifikant økning fra cirka 6 % i 2003 til cirka 12 % i 2012, men med et brudd i utviklingen for 2015, hvor forskjellene mellom skoler er mindre igjen (8 %). For TIMSS 4. trinn er det en svak økning fram til 2007, med en tilsvarende nedgang frem til 2015. Utviklingen for 8. trinn er først og fremst preget av en veldig stor nedgang fra 1995 til 2003, mens utviklingen deretter er ustabil med ikke-signifikante endringer fra ett år til det neste i begge retninger.

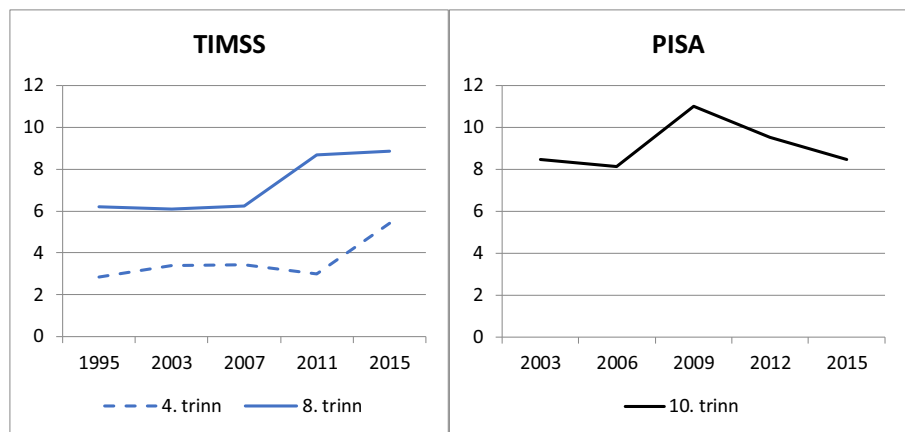
De mørkere områdene i figur 7.1 viser delen av variansen som kan knyttes til elevenes SES. Helt overordnet er tendensen at disse har vokst over tid for ungdomstrinnet, men ikke dramatisk og ikke helt konsistent. Utviklingen for total varians forklart av SES er tydeligst for 8. trinn i TIMSS, hvor utviklingen er jevnt stigende fra 9 % av totalvariansen i 1995 til 14 prosentpoeng på den samme variansskalaen i 2015. I PISA (10. trinn) steg den totale variansen forklart av SES fra 11 % i 2003 til 14 % i 2012, men med en liten nedgang fra 2012 til 2015. For TIMSS 4. trinn, har den totale variansen forklart av SES ligget mellom 6 % og 8 % i perioden.⁴⁰

Figur 7.2 framhever tydeligere varianskomponentene på skolenivå som kan knyttes til SES. Den venstre halvdel av figuren viser utvikling for TIMSS. På 8. trinn (heltrukket linje) har andelen varians på skolenivå forklart av SES steget fra cirka 3 % i 1995 til cirka 5 % i 2015. Dette er ingen dramatisk økning, og den er ikke statistisk signifikant. Utviklingen for 4. trinn (stiplet linje) er ustabil over tid. Den høyre halvdel av figur 7.2 viser resultatene for PISA. Her er det heller ikke et klart bilde. Variansen på skolenivå forklart av SES stiger og faller for hver syklus, men starter i 2003 og ender i 2015 på samme relativt lave verdi (omtrent 2.5 %). For PISA er heller ingen av endringene fra én syklus til den neste signifikante. Oppsummert på tvers av de tre populasjonene viser figuren at det jevnt over er noe større andel varians som er forklart av SES på ungdomstrinnet.

40. Vi minner om at alle disse prosentandelene refererer til den totale variansen i det første året studien ble gjennomført.



FIGUR 7.2 Andel varians på skolenivå forklart av SES i TIMSS og PISA.

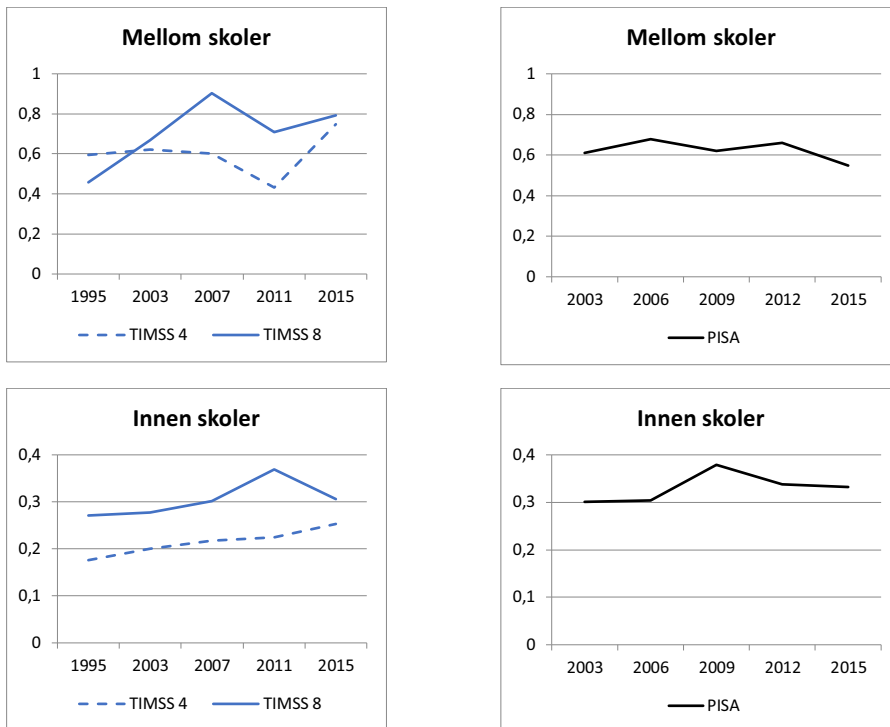


FIGUR 7.3 Andel varians innen skoler forklart av SES i TIMSS og PISA.

Figur 7.3 viser hvordan variansen innen skoler som kan knyttes til SES, utvikler seg over tid. Nok en gang er det tydelig at SES er sterkere relatert til prestasjoner for økende klassetrinn. Det er imidlertid ikke et konsistent bilde for utviklingen over tid på tvers av studiene. Det er en liten økning mot slutten av perioden for TIMSS-resultatene, og det er et maksimum midt i perioden for PISA.

Figur 7.4 viser de standardiserte regresjonskoeffisientene når prestasjoner er modellert som en funksjon av SES. Grafene øverst viser de standardiserte regre-

sjonskoeffisientene på skolenivå; det vil si i hvilken grad elevers matematikkprestasjoner kan predikeres av skolens SES. Grafene i nedre del av figuren viser tilsvarende koeffisienter på elevnivå. For begge nivåer er det tydelig at SES er en sterkere prediktor for 8. trinn enn for 4. trinn i TIMSS. Det er en overordnet tendens til at sammenhengen mellom SES og prestasjoner på skolenivået har økt merkbart i første del av perioden for 8. trinn for så å flate ut. For 4. trinn er det en markert økning mellom de to siste årene, men over hele perioden er det ingen tydelig tendens. I PISA er bildet heller at sammenhengen er stabil over tid for skolenivået. For elevnivået, korrigert for skoleforskjellene, er det en svakt økende tendens, og for elevnivået er feilmarginene så små at disse trendene er statistisk signifikante. Denne tendensen er tydeligst for 8. trinn i TIMSS, mens tendensen er noe mer ustabil for PISA og 4. trinn i TIMSS.



FIGUR 7.4 Sammenhengen mellom SES og matematikkprestasjoner i TIMSS og PISA på skolenivå (mellom skoler) og elevnivå (innen skoler).

OPPSUMMERING OG KONKLUSJON

Tabell 7.3 gir en kompakt oppsummering av funn knyttet til de tre forsknings-spørsmålene. Selv om PISA og TIMSS måler forskjellige typer kompetanser for ulike alderstrinn, er det meningsfullt å oppsummere det totale bildet på tvers. I tabellen er dette gjort i den nederste raden som en hovedtendens på tvers av de tre studiene. I oppsummeringene tar vi hensyn til at feilmarginene er betydelig større for alle resultater på skolenivået.

TABELL 7.3 Oppsummering av funn fra første til siste syklus av PISA og TIMSS.

	Forskningsspørsmål 1		Forskningsspørsmål 2		Forskningsspørsmål 3	
	Varians på skolenivå	Total varians og varians mellom elever	Varians på skolenivå forklart av SES	Varians på elevnivå forklart av SES	Sammenheng mellom SES og prestasjoner på skolenivå	Sammenheng mellom SES og prestasjoner på elevnivå
TIMSS 4. trinn	Svakt økende til 2007, nedgang i 2015	Synker	Ingen tydelig tendens.	Ingen tydelig tendens	Ingen tydelig tendens. Økning fra 2011 til 2015	Svak og jevn økning
TIMSS 8. trinn	Ingen tydelig tendens, men stor nedgang fra 1995 til 2007	Synker (men øker fra 2011 til 2015)	Stigende til 2007, deretter uendret	Et hopp fra 2007 til 2011	Økende	Svak økning fram til 2011
PISA	Tydelig økning til 2012, stor nedgang i 2015	Ustabilt synkende	Ingen tydelig tendens	Et maksimum for 2009, men ingen tydelig tendens	Ingen tydelig tendens	Ingen tydelig tendens
Hovedtrend	Stigende tendens i to av tre studier, men et brudd i siste punkt med nedgang	Synkende	Ingen konsistent tendens	Ingen konsistent tendens	Ingen konsistent tendens	Svak økning i to av tre studier

En helt overordnet konklusjon for det første forskningsspørsmålet er at norsk skole, ut fra kriteriet om minimal spredning i resultater, har beveget seg i retning av mer utjæmning i perioden. Forskjellene mellom skoler endrer seg ikke vesentlig i perioden, men estimatene er ustabile og preget av at feilmarginene for skolenivået er relativt store. Her kan det tilføyes at vi også har undersøkt om det er en tendens til økt sosial segregering. Dette kan indikeres ved om elever på samme skoler blir relativt mer lik hverandre over tid når det gjelder SES. Dette uttrykkes i andelen varians i variabelen «bøker i hjemmet» som er knyttet til hvilken skole elevene går på. For PISA ligger denne andelen stabilt på 5 % over hele perioden med noen små variasjoner. For TIMSS øker imidlertid denne andelen gradvis fra

om lag 3 % i 1995 til 8 % i 2015. Det er imidlertid ikke gitt at endringer i slike spredningsmål i seg selv er gode indikatorer for utjamning *som en kvalitet* i norsk skole. Spredningen reduseres jo både gjennom å løfte elevene nederst i fordelingen og gjennom at færre elever har høye skårer. Et løft i bunnen er opplagt en kvalitativ god utvikling, men at det blir færre elever som presterer godt, er ikke ønskelig. Oversikten over utviklingen for den 10. og 90. prosentilen i elevenes prestasjoner gitt i kapittel 1 viser at begge deler i noen grad har skjedd i denne perioden.

Forskningsspørsmål 2 og 3 angår begge hvorvidt sammenhengene mellom elevenes hjemmebakgrunn og prestasjoner endrer seg i perioden. For dette spørsmålet er den mest rimelige konklusjonen at norsk skole ikke viser tegn til å ha blitt mer utjammende over tid. Andelen av varians som kan knyttes til elevenes SES, angir styrken for denne sammenhengen (forskningsspørsmål 2). Selv om det finnes noen variasjoner i perioden, er det likevel mest rimelig å konkludere med at sammenhengen mellom elevenes prestasjoner og henholdsvis elevenes og skolenes SES er relativt uendret i perioden. Regresjonskoeffisientene angir størrelsen eller gradienten for denne sammenhengen. For skolenivå er det noen relativt store endringer i perioden, men likevel ingen entydig utviklingstrend. På elevnivået er det imidlertid en svakt økende gradient i perioden for TIMSS-undersøkelsene. I PISA er imidlertid estimatene både for skole- og elevnivå stabile over tid.

Tabell 7.3 kan imidlertid også oppsummeres på tvers av de tre forskningsspørsmålene for de tre studiene, altså med et fokus på de overordnede trekkene for de tre ulike kohortene eller klassetrinnene som er inkludert. I en slik lesing av tabellen er det spesielt resultatene for 8. trinn i TIMSS som står fram. For dette trinnet og denne studien er det en relativt konsistent tendens til en svak økning i ulikhet mellom elever med ulik SES.

DISKUSJON OG VEIEN VIDERE

I likhet med de andre kapitlene i denne boka har denne analysen tatt utgangspunkt i at de gjentatte undersøkelsene over tid har et stort potensiale for å gi informasjon om utviklingstrekk i norsk skole for et tidsrom som har vært preget av relativt store endringer i utdanningspolitikken, med to store utdanningsreformer. (Se f.eks. Olsen, Hopfenbeck & Lillejord, 2013.) En unik styrke ved de internasjonale studiene er at skalaene for elevenes prestasjoner er lenket over tid, og de gir dermed et fast punkt som tilfører robust empiri om endringer i eksempelvis gjennomsnitt og spredning i elevenes prestasjoner, eller sammenhengene mellom prestasjonene og elevenes hjemmebakgrunn. Dersom gjennomsnitt og fordeling i

matematikkprestasjonene endrer seg i perioden, kan man derfor være rimelig trygg på å konkludere med at elevenes dyktighet i matematikk har endret seg i perioden – slik dette er definert i undersøkelsene (men opplagt uten at man kan slå fast hvorfor). Det finnes ingen andre nasjonale data som kan brukes for å studere hvordan gjennomsnittlig dyktighet og fordelingen av prestasjoner i sentrale skolefaglige områder har endret seg i et så langt tidsrom.

Analysene våre gir et litt sprikende bilde. For det første er det et sprik når vi ser på ulike operasjonaliseringer av utjamning. Spredningen i norske elevers prestasjoner har en minkende tendens, noe som med en slik definisjon tilsier at norsk skole blir mer utjamnende. Når vi ser på hvordan SES er relatert til elevenes prestasjoner, og hvordan denne endrer seg over tid, har vi konkludert med at det ikke er noe i vårt datamateriale som tilsier at sammenhengen har blitt svekket. Snarere er det slik at i de tilfellene hvor en tendens framstår, peker de i retning av en svak forsterking av denne sammenhengen. Ulike definisjoner gir altså ulike svar på et tilforlatelig likt spørsmål, noe som for så vidt er i tråd med det Strietholt (2014) finner i sine analyser av data fra de internasjonale undersøkelsene. For det andre viser analysene et sprik mellom TIMSS og PISA. PISA gir i de fleste tilfellene et bilde av stabilitet over tid. I TIMSS, og da spesielt for undersøkelsen på 8. trinn, er det en svak tendens til økning. Dette kan også være sammenfallende med annen forskning som tilsier at utjamning kan være ulik fra ett skolefaglig område til et annet (Bakken & Elstad, 2012; Mullis, Martin & Hooper, 2016). For å belyse dette nærmere vil vi i en senere publikasjon presentere tilsvarende analyser også for de andre fagområdene i studiene, inkludert også lesing i PIRLS. Et sammenfall på tvers av fagområder for samme klassetrinn vil kunne bidra til at konklusjoner om utjamning i norsk skole kan generaliseres utover et bestemt faglig område.

I analysene i dette kapitlet har vi studert skolers utjamning over tid. Både PISA og TIMSS har inkludert ganske mange variabler som indikatorer for elevers sosio-økonomiske status (SES). Flere av variablene er imidlertid litt ulike på tvers av studiene, og innen de enkelte studiene har også noen av variablene endret seg over tid. Et eksempel er spørsmål om foreldrenes utdanningsnivå, som regnes som en indikator på SES, og som er inkludert i begge studiene. Dette spørsmålet varierer litt fra syklus til syklus i både PISA og TIMSS, både når det gjelder ordlyd og antall svaralternativer.

Vi har derfor kun hatt muligheten til å studere denne sammenhengen gjennom et enkelt selvrapportert mål om antall bøker i hjemmet. Til tross for at dette fortsatt er den enkeltvariabelen som fanger inn mest av variansen i SES, og til tross for at den foreløpig har vist seg å være en rimelig robust indikator (Gustafsson et al., 2013; Hansen & Munk, 2012; Strietholt, 2014), er det grunn til å anta at dette er

en variabel som med stor sannsynlighet endrer sin betydning som markør for hjemmets sosiale og kulturelle kapital i løpet av de neste 20 årene. En ideell (men vanskelig gjennomførbar) løsning hadde vært å kunne kople data fra registre til de internasjonale undersøkelsene. I de nordiske landene finnes svært gode registre som inkluderer både foreldrenes utdanningslengde, yrkeskarriere, innvandringsbakgrunn og inntekter/formue. Slike data ville kunne gitt et langt bredere og reliabelt mål for elevers bakgrunn.

Videre illustrerer denne analysen at undersøkelsen design er gode for å få stabile estimater for sammenhenger mellom variabler på elevnivået. Imidlertid viser også våre analyser at undersøkelsene ikke er følsomme for relativt små, men like fullt potensielt meningsfulle endringer over tid for skolenivået. Dette kan knyttes til to litt ulike problemer i studiene. En side av problemet er variasjonen over tid for antallet klasser som er inkludert fra de deltakende skolene i TIMSS (se tabell 7.2). Denne variasjonen i utvalgsdesign over tid kan være en mulig grunn til de relativt store variasjonene som man ser fra TIMSS-undersøkelsen for effekter knyttet til skolenivået. Til tross for at dette er undersøkelser som inkluderer flere tusen elever, er likevel antallet skoler ganske lavt (150–200 skoler). For hver enkelt studie er for så vidt antallet skoler så stort at mange estimater knyttet til skolenivået er statistisk signifikante. Selv om enkeltskoler kan endre seg raskt, representerer skolene et større skolesystem som endrer seg sakte og gradvis. Endringene i effektene på skolenivå over tid kan derfor forventes å være så små at de er vanskelige å estimere med statistisk sikkerhet. En direkte løsning på denne utfordringen ville være å øke antallet deltakende skoler i de internasjonale undersøkelsene, men dette ville ha medført kostnader både i form av penger og belastning for skolene.

For å studere utviklingen over tid for skolenivået vil vi anbefale at man legger vekt på analyser av registerdata. Så langt har imidlertid studier som bruker nasjonale registre, manglet gode mål for elevenes faglige prestasjoner. Karakterer både i form av standpunkt og eksamen har problemer knyttet til manglende lenking over tid, og for standpunkt er det også problemer knyttet til at lærere og skoler tolker kriterier for karakterer i fag ulikt, og/eller at de foretar en lokal normering innenfor sitt eget klasserom og skole. I den perioden som vi har studert her, gir derfor heller ikke analyser av nasjonale registerdata problemfrie estimater. Imidlertid er de nasjonale prøvene for 5. og 8. trinn nå lenket over tid, og når vi skal gjøre en tilsvarende oppsummering 20 år fra nå, vil disse gi stabile mål for elevers prestasjoner over tid. Kombinert med registerdata for elevers og foreldres bakgrunn vil man ha et godt verktøy for å studere hvordan utjamning i norsk skole eventuelt endrer seg over tid, og ikke minst hvordan skolenes bidrag til utjamning

endrer seg. De internasjonale undersøkelsene gir imidlertid et viktig tilskudd til slike analyser ved å tilby et relevant sammenlikningsgrunnlag for å kunne vurdere om ulikhetene i norsk skole er store eller små.

REFERANSER

- Bakken, A. (2004). Økt sosial ulikhet i skolen? *Tidsskrift for ungdomsforskning*, 4(1).
- Bakken, A., & Elstad, J. I. (2012). *For store forventninger?: Kunnskapsløftet og ulikhetene i grunnskolekarakterer*. Oslo, Norsk institutt for forskning om oppvekst, velferd og aldring. 7/2012.
- Beaton, A. E., Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1997). *The Mathematics Achievement in the Middle School Years*. Boston Collage: TIMSS and PIRLS International Study Centre.
- Blossing, U., Imsen, G., & Moos, L. (2014). Nordic schools in a time of change. In *The Nordic Education Model: «A School for All» Encounters Neo-Liberal Policy* (pp. 1–14). Springer Netherlands.
- Bonesrønning, H., Falch, T., & Strøm, B. (2005). Teacher sorting, teacher quality, and student composition. *European Economic Review*, 49(2), 457–483.
- Bourdieu, P. (1986). *The forms of capital* In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258): New York: Greenwood.
- Bourdieu, P., Wacquant, L. J., & Kvalsvik, B. N. (1993). *Den kritiske ettertanke: grunnlag for samfunnsanalyse*. Norske samlaget.
- Foy, P., Arora, A., & Stanco, G. M. (2013). *TIMSS 2011 User Guide for the International Database*. Lynch School of Education, Boston College: TIMSS & PIRLS International Study Center and International Association for the Evaluation of Educational Achievement (IEA).
- Gustafsson, J.-E., Hansen, K. Y., & Rosén, M. (2013). *Effects of Home Background on Student Achievement in Reading, Mathematics, and Science at the Fourth Grade*. In M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS and PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement at the Fourth Grade—Implications for Early Learning*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Gustafsson, J.-E., Nilsen, T., & Hansen, K. Y. (2016). School characteristics moderating the relation between student socio-economic status and mathematics achievement in grade 8. Evidence from 50 countries in TIMSS 2011. *Studies in Educational Evaluation*. doi:<https://doi.org/10.1016/j.stueduc.2016.09.004>
- Hansen, K. Y. (2015). *Trend in Educational Inequality in the Nordic countries between 2000 and 2009: Evidence from OECD PISA studies*. Paper presented at the ECER, Budapest.
- Hansen, K. Y., Gustafsson, J.-E., & Rosén, M. (2011). *School performance differences and policy variations in Finland, Norway and Sweden*. In N. C. o. Ministers (Ed.), *Northern Lights on TIMSS and PIRLS* (pp. 25–48).
- Hansen, K. Y., & Munk, I. (2012). Exploring the measurement profiles of socioeconomic background indicators and their differences in reading achievement: A two-level latent class analysis. *IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments*, 5, 49–77.

- Hox, J., & Roberts, J. K. (2011). *Handbook of advanced multilevel analysis*: Psychology Press.
- Kunnskapsdepartementet. (2006). ... og ingen sto igjen. Tidlig innsats for livslang læring. Retrieved from <https://www.regjeringen.no/no/dokumenter/stmeld-nr-16-2006-2007-/id441395/>
- Kyriakides, L., & Creemers, B. P. (2011). Can schools achieve both quality and equity? Investigating the two dimensions of educational effectiveness. *Journal of Education for Students Placed at Risk (JESPAR)*, 16(4), 237–254.
- Liu, H., Van Damme, J., Gielen, S., & Van Den Noortgate, W. (2015). School processes mediate school compositional effects: model specification and estimation. *British Educational Research Journal*, 41(3), 423–447.
- Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation*, 12(02), 105–128.
- Martin, M. O., Foy, P., Mullis, I. V. S., & O'Dwyer, L. M. (2013). *Effective Schools in Reading, Mathematics, and Science at Fourth Grade*. In M. O. Martin & I. V. S. Mullis (Eds.), TIMSS and PIRLS 2011: Relationships among reading, mathematics, and science achievement at the fourth grade—Implications for early learning (pp. 109–178). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., & Mullis, I. V. S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (2016). *Methods and Procedures in TIMSS 2015*. In. Retrieved from <https://timssandpirls.bc.edu/publications/timss/2015-methods.html>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS 2015 International Results in Mathematics*. In T. P. I. S. C. Boston College (Ed.).
- Mullis, I. V. S., Martin, M. O., & Hooper, M. (2016). *20 years of TIMSS. International Trends in Mathematics and Science Achievement, Curriculum, and Instruction*. Boston College: TIMSS & PIRLS International Study Cente, and IEA.
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (Eight ed.). Los Angeles, CA: Muthén & Muthén.
- Nilsen, T., & Bergem, O. K. (2016). *Hjemmebakgrunn. En likeverdig skole*. In O. K. Bergem, H. Kaarstein, & T. Nilsen (Eds.), Vi kan lykkes i realfag. Resultater og analyser fra TIMSS 2015. Oslo: Universitetsforlaget.
- OECD. (2013). *PISA 2012 results: excellence through equity: giving every student the chance to succeed* (volume II): OECD, Paris, France.
- OECD. (2016). *PISA 2015 Results* (Volume I). Excellence and Equity in Education: OECD Publishing.
- Olsen, R. V., Hopfenbeck, T. N., & Lillejord, S. (2013). Elevenes læringssituasjon etter Kunnskapsløftet. *Norsk Pedagogisk Tidsskrift*, 97(6), 355–369.
- Olsen, R. V., & Turmo, A. (2010). *Et likeverdig skoletilbud? In På rett spor – Norske elevers kompetanse i lesing, matematikk og naturfag i PISA 2009* (pp. 207–225): Universitetsforlaget.
- Palardy, G. J. (2013). High school socioeconomic segregation and student attainment. *American Educational Research Journal*, 50(4), 714–754.

- Rutkowski, D., & Rutkowski, L. (2013). Measuring Socioeconomic Background in PISA: One Size Might not Fit all. *Research in Comparative and International Education*, 8(3), 259–278. doi:10.2304/rcie.2013.8.3.259
- Rutkowski, D., Rutkowski, L., & Plucker, J. A. (2012). Trends in education excellence gaps: A 12-year international perspective via the multilevel model for change. *High Ability Studies*, 23(2), 143–166.
- Rutkowski, L., von Davier, M., & Rutkowski, D. (2013). *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis*. Boca Raton (FL): CRC Press.
- Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The Role of Schooling in Perpetuating Educational Inequality An International Perspective. *Educational Researcher*, 44(7), 371–386.
- Sirin, S. R. (2005). Socioeconomic Status and Academic Achievement: A Meta-Analytic Review of Research. *Review of Educational Research*, 75(3), 417–453. doi:10.3102/00346543075003417
- Strietholt, R. (2014). Studying Educational Inequality: Reintroducing Normative Notions. In R. Strietholt, W. Bos, J. E. Gustafsson, & M. Rosén (Eds.), *Educational Policy Evaluation* (pp. 51). Münster, New York: Waxman.
- Strietholt, R. (2016). *How is inequality measured in international studies on educational achievement?* Paper presented at the AERA, Washington.
- Van Ewijk, R., & Slegers, P. (2010). The effect of peer socioeconomic status on student achievement: A meta-analysis. *Educational Research Review*, 5(2), 134–150.
- White, K. R. (1982). The relation between socioeconomic status and academic achievement. *Psychological Bulletin*, 91(3), 461.
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *The Teachers College Record*, 112(4), 3–4.