

Sigurd Melsom

Heterogene effekter av smågruppeundervisning på elevprestasjoner

En empirisk analyse basert på data fra norsk
barneskole

Masteroppgave i samfunnsøkonomi
Veileder: Hans Bonesrønning
Trondheim, mai 2018

Norges teknisk-naturvitenskapelige universitet
Fakultet for økonomi
Institutt for samfunnsøkonomi

Forord

Jeg vil rette en stor takk til min veileder, professor Hans Bonesrønning, for verdifulle tilbakemeldinger og interessante diskusjoner i forbindelse med oppgaven. Jeg ønsker også å takke Ingvild Eliassen og Eira Enodd for gode faglige diskusjoner, og min mor for hjelp med korrektur.

Trondheim, 29. mai 2018.

Sigurd Melsom

Sammendrag

Prestasjonene til norske elever i barneskolen er ikke på nivå med myndighetenes målsetting. Etter skuffende resultater i internasjonale undersøkelser i senere år, har oppmerksomheten rundt norske elevers matematikkferdigheter økt. I denne oppgaven benytter jeg data fra 1+1-prosjektet. Prosjektet tildeler ett ekstra lærerårsverk til tilfeldig utvalgte skoler som skal benyttes til smågruppeundervisning i matematikk. Totalt deltar 160 skoler fra 10 store norske kommuner i prosjektet, og datamaterialet inkluderer over 17 000 elever. Jeg finner at elever som deltar i smågruppeundervisning i gjennomsnitt presterer 0,184 standardavvik bedre enn elever i sammenligningsskolene. Effekten av smågruppeundervisning er særlig stor for middels presterende elever. En middels presterende elev som deltar i smågruppeundervisning presterer 0,233 standardavvik bedre enn lignende elever i sammenligningsskoler. Videre finner jeg høy variasjon i skolenes evne til å påvirke elevers testresultater. Denne variasjonen benyttes til å estimere positive korrelasjoner mellom hvor mange studiepoeng en lærer har og skolenes evne til å påvirke elevenes testresultater. Disse estimatene er imidlertid ikke nødvendigvis kausale sammenhenger.

Abstract

Norwegian students in elementary schools perform lower than the expectations set by the government. After disappointing results in comparative studies of educational achievement in recent years, more attention has been drawn to improvement of student performance. In this thesis, I use data from the 1+1 project which randomly assigns additional teacher resources to schools to be used in small group instruction in mathematics. The project is carried out in ten large Norwegian municipalities, and includes over 17,000 students from 160 schools. I find that students participating in small group instruction on average perform 0.184 standard deviations better than students in the comparison schools. The small group instruction has biggest impact on students located in the middle of the achievement distribution. These students have math improvements that are 0.233 standard deviations higher than similar students in the comparison schools. I also find significant variation in each school's ability to increase student test scores. This variation is exploited to estimate a positive correlation between teacher math education and each school's ability to increase student test scores, although these estimates are not necessarily causal in nature.

Innhold

1	Innledning	1
2	Beskrivelse av prosjektet	3
3	Tidligere litteratur	7
4	Teori	9
4.1	Produktfunksjon for utdanning	9
4.2	Kunnskapsproduksjon og økonometriske utfordringer	10
5	Deskriptiv statistikk	13
6	Kausal effekt av smågruppeundervisning	15
7	Heterogene programeffekter	19
7.1	Resultater	19
7.2	Oppsummerende kommentarer	22
8	Variasjon mellom tiltaksskoler	25
8.1	Antall uker smågruppeundervisning	25
8.2	Effekt av antall uker smågruppeundervisning	30
8.3	Skolespesifikke effekter	32
9	Konklusjon	43
	Referanser	45
	Appendiks A: Fordeling i kvintiler	i
	Appendiks B: Utvidelse av heterogene programeffekter	iii
	Appendiks C: Fordeling av antall uker smågruppeundervisning	v

1 Innledning

Et mye brukt og praktisk gjennomførbart politisk virkemiddel for å øke kvaliteten i grunnskolen, er å ansette flere lærere. Forskning på området gir imidlertid ikke noe klart svar på om dette er en gunstig tilnærming. Selv om det er bred enighet om at høy skolekvalitet er viktig, er det stor usikkerhet rundt hvordan man skal forbedre kvaliteten. Tradisjonell undervisning gir få muligheter til å individuelt tilpasse undervisningen for enkeltelever, noe som kan føre til svært ulik grad av måloppnåelse.

I 2017 økte regjeringen bevilgningen til økt lærertetthet på 1.–4. trinn med 460 millioner kroner. Med dette utgjør den samlede bevilgningen øremerket lærerårsverk 1,3 milliarder kroner. Økte bevilgninger for å sikre flere barnehagelærere og økt lærertetthet i barneskolen er en del av regjeringens strategi om «tidlig innsats», som går ut på å løfte alle elever så tidlig som mulig i utdanningsløpet og hindre at elever henger etter (Kunnskapsdepartementet, 2017b). Det er et viktig mål for myndighetene at alle som er i stand til det skal fullføre videregående opplæring, og svake prestasjoner fra grunnskolen er den viktigste årsaken til at elever ikke fullfører (Utdanningsdirektoratet, 2016).

For at færre elever skal henge etter i undervisningen foreslo regjeringen i juni 2017 en lovfestet plikt om å tilby intensiv opplæring til lavt presterende elever på 1.–4. trinn. Smågruppeundervisning nevnes som en del av denne intensive opplæringen (Kunnskapsdepartementet, 2017a). Smågrupper vil i denne oppgaven si undervisning i grupper på seks eller færre elever, med mindre annet er spesifisert.

Denne oppgavens bidrag til debatten er å undersøke om matematikkundervisning i smågrupper på småskoletrinnet kan ha positiv effekt på elevprestasjoner, og om smågruppeundervisning har størst effekt på prestasjonene til lavt presterende eller høyt presterende elever. Videre ønsker jeg å undersøke hvilke faktorer som kan forklare ulik effekt av smågruppeundervisning med utgangspunkt i hva som foregår på innsiden av skolene. Variasjon mellom skolene som deltar i prosjektet benyttes til å sette effekten av smågruppeundervisning i perspektiv, i tillegg til å utforske effekt av andre innsatsfaktorer i skolen. Det benyttes data fra 1+1-prosjektet. Prosjektet beskrives i kapittel 2.

Analyse av hvilke faktorer som har innvirkning på elevers prestasjoner i matematikk er viktig. Kvalifisert arbeidskraft blir stadig mer ettertraktet i arbeidsmarkedet, og etterspørselen vil fortsette å øke i fremtiden. Samtidig vil etterspørselen etter arbeidskraft med kun grunnskole reduseres. Siden svake prestasjoner fra grunnskolen er den viktigste årsaken til at elever ikke fullfører videregående opplæring, er et høyt prestasjonsnivå blant elever og høy skolekvalitet viktig for individet, men også for samfunnet som helhet.

En rekke studier konkluderer med at skolekvalitet, målt ved testresultater i realfag, har sterk påvirkning på fremtidige individuelle utfall og økonomisk vekst for et land (Hanushek, 2002). Hanushek & Woessmann (2012) benytter skolerresultater som et mål

på kognitive evner og finner en positiv sammenheng mellom disse evnene og økonomisk vekst. Resultatene er statistisk signifikante og av en størrelse som har klar økonomisk betydning.

Analysen av heterogene programeffekter er motivert av flere faktorer. Tall fra den store internasjonale undersøkelsen TIMSS i 2011 viser at Norge har lavest andel elever på 4. trinn som presterer på høyt og avansert nivå, og høyest andel elever som presterer på lavt eller under lavt nivå sammenlignet med nordiske land og fire referanseland (Grønmo et al., 2011). Kunnskap om hvilke elever som tjener mest på ulik undervisning er viktig for å løfte alle elever til et høyere nivå.

Basert på resultater fra TIMSS-undersøkelsen konkluderer Grønmo et al. (2011) med at målet om at elever i norsk skole skal levere gode resultater i matematikk er et godt stykke unna å bli oppfylt, både i barne- og ungdomsskolen. I PISA 2012 hadde norske 15-åringer de svakeste testresultatene i matematikk siden Norge ble med i undersøkelsen i 2000 (Kjærnsli & Olsen, 2013). Resultatene ble forbedret i PISA 2015, men ikke nok til å se noen positiv trend i norske 15-åringers ferdigheter i matematikk i perioden 2000-2015 (Kjærnsli & Jensen, 2016). Regjeringen har i mange år hatt et uttalt mål om å forbedre resultatene i realfag i den norske skolen (Kunnskapsdepartementet, 2015). Resultatene fra nevnte studier viser at prestasjonene ikke er blitt løftet til det nivået man skulle ønske.

Dette kapitlet har definert problemstillingen og motivert analysen. Oppgaven er videre disponert som følger. Kapittel 2 beskriver den delen av 1+1-prosjektet som er relevant for oppgaven, og kapittel 3 tar for seg relevant litteratur. I kapittel 4 presenteres produktfunksjonen for utdanning og noen økonometriske utfordringer i forbindelse med analyse av kunnskapsproduksjon. Deskriptiv statistikk blir gjennomgått i kapittel 5. I kapittel 6 undersøkes den kausale effekten av smågruppeundervisning, før jeg ser på ulik effekt på undergrupper av elever i kapittel 7. Den siste delen av analysen gjennomgås i kapittel 8. Her ser jeg nærmere på hvilke faktorer på innsiden av skolene som forårsaker ulik effekt av smågruppeundervisning estimert i kapittel 7, i tillegg til å estimere korrelasjon mellom mengde smågruppeundervisning og testresultater. Til slutt, i kapittel 8.3, benyttes variasjonen mellom skolene til å konstruere et mål på skolenes påvirkning på elevers testresultater, og det utforskes hva som kan forklare denne påvirkningsevnen. Kapittel 9 konkluderer.

2 Beskrivelse av prosjektet

1+1-prosjektet er en randomisert kontrollert studie (RCT) som strekker seg over fire år (2016–2020) og inkluderer 160 skoler i ti store norske kommuner. Formålet med prosjektet er å undersøke hvilken effekt undervisning i smågrupper har på elevers matematikkferdigheter. Målgruppen for prosjektet er elever på småskoletrinnet. De deltakende skolene deles likt i to grupper, hvor halvparten av skolene får tildelt ett ekstra lærerårsverk som brukes til å gjennomføre smågruppeundervisning i matematikk. Resten av skolene fortsetter som tidligere. Hvilke skoler som havner i hver av de to gruppene bestemmes ved tilfeldig parvis trekning i hver kommune, slik at alle kommuner som deltar i prosjektet har like mange skoler i hver av de to gruppene. Metoden sørger for at den eneste forskjellen mellom de to gruppene er ekstra lærerressurser.

I prosjektet benyttes altså randomisering på skolenivå, noe som krever et større utvalg enn randomisering på individnivå. Dette valget begrunnes med at randomisering på individnivå har noen ulemper. For det første kan det tenkes at skoleledelsen er motvillig til å delta i prosjektet hvis like elever ved samme skole behandles ulikt. For det andre vil det være vanskeligere å holde sammenligningsgruppen atskilt fra gruppen som tildeles ekstra lærerressurser hvis begge gruppene befinner seg i en og samme skole. For eksempel kan foreldre forsøke å få sine barn flyttet fra sammenligningsgruppen til gruppen med smågruppeundervisning, og skoleledelsen kan tildele sammenligningsskolene ekstra lærerressurser tiltenkt tiltaksgruppen. Denne type atferd vil være vanskelig å observere innenfor hver skole. Antallet skoler i kommunene i dette prosjektet er mer enn høyt nok for å gjennomføre randomisering på skolenivå, og denne metoden er derfor å foretrekke.

I randomiseringen blir det benyttet en parvis stratifisert utvelgning med tidligere resultater fra nasjonale prøver i regning som stratifiseringsvariabel for å øke styrken på den randomiserte kontrollerte studien. Resultatene fra nasjonale prøver brukes til å rangere skolene. Hver skole blir deretter delt inn i par med den skolen som befinner seg ved siden av i rangeringen. Til slutt blir én skole i hvert par randomisert til tiltaksgruppen, hvor den andre skolen havner i kontrollgruppen.

Randomiseringen gjennomføres blant skoler i kommuner som har meldt seg frivillig til å delta i prosjektet. Den frivillige deltakelsen, i kombinasjon med at tildeling av ekstra lærerressurser anses som en fordel for skolene, gjør at det trolig er liten risiko for at skoler trekker seg fra prosjektet.

For å følge elevenes utvikling i matematikk blir det gjennomført korte prøver i begynnelsen og på slutten av hvert skoleår. De to prøvene refereres til som pretest og posttest. Resultater fra pretesten gjør det mulig å dele inn elever i undergrupper basert på prestasjonsnivå i matematikk. Inndelingen legger grunnlaget for analysen av heterogene programeffekter som denne oppgaven i stor grad skal fokusere på, og basert på denne ana-

lysen ønsker jeg å undersøke om effekten av smågruppeundervisning i matematikk er ulik for lavt presterende elever og høyt presterende elever på småskoletrinnet. Denne oppgaven benytter data fra prosjektets første år, det vil si skoleåret 2016/2017. I løpet av det første året observeres 1. og 4. trinn mens 2. og 3. trinn tildeles et ekstra lærerårsverk.

Gjennomføringen av prosjektet overvåkes ved bruk av intervjuer, spørreundersøkelser og tett oppfølging av skoleledelsen i de aktuelle skolene. Lærerne i skolene hvor smågruppeundervisning skal benyttes, blir instruert i å undervise grupper med 4–6 elever i et annet klasserom enn der den ordinære undervisningen foregår. Dette gjør det mulig å tilpasse undervisningen for de små gruppene. Gruppene vil også overvåkes gjennom lærernes vurderinger av hvilke elever som skal delta i smågruppeundervisningen og fremgangen til disse elevene.

1+1-prosjektet foregår ikke i en boble, og vil blant annet bli påvirket av beslutningstakere i kommunen. Et potensielt problem ved prosjektet er fordelingen av ressurser blant skoler i kontrollgruppen. Enkelte kommuner har blitt tildelt ekstra ressurser utenfor dette prosjektet, og de kan tenkes å fortsatt motta ressurser i løpet av årene prosjektet foregår. Hvis kommunene tildeler mer av disse ressursene til sammenligningsskoler enn til tiltaksskoler, kan det gi skjevhet i estimerte effekter av smågruppeundervisning. Eventuelt uheldig fordeling av ressurser kan undersøkes med data fra Grunnskolens Informasjonssystem (GSI). I tillegg vil tett samarbeid med kommuner, skoleledelsen og lærere gi informasjon om i hvor stor grad ressurser fordeles skjevt.

Også beslutninger tatt på skolenivå kan medføre skjær i sjøen for validiteten av prosjektet. I de fleste tilfeller er skoleledelsen informert et år i forveien om tildeling av en ekstralærer. Et potensielt problem med dette er at skoleledelsen kan flytte allerede eksisterende assistenter til trinn som ikke berøres av prosjektet, slik at det tildelte lærerårsverket erstatter disse assistentene når prosjektet iverksettes. Dette vil føre til at lærertettheten varierer marginalt mellom sammenligningsskolene og tiltaksskolene, og estimerte effekter av smågruppeundervisning vil kun avhenge av forskjellen i kvaliteten på undervisningen mellom de to gruppene. En åpenbar løsning vil være å pålegge restriksjoner på skolene. Dette kan imidlertid potensielt redusere motivasjonen til de deltakende skolene, hvilket kan føre til mindre innrapportert datamateriale og dermed svekke resultatene i undersøkelsen. I stedet for anmodes skoleledere om å ikke endre spesialundervisning, bruk av assistentlærere eller annen ressursbruk ved skolen som følge av at skolen deltar i prosjektet.

For at testresultatene på starten og slutten av skoleåret skal kunne kobles sammen, kreves det samtykke fra foreldre. Mangel på samtykke fra foreldre kan gi skjevhet i estimatene, særlig hvis elever fra tiltaksskolene som blir unntatt deler flere av de samme karakteristikkene. Et lignende problem kan oppstå hvis like elever flytter med det formål å bytte skole fra en sammenligningsskole til en tiltaksskole i løpet av perioden. Dette er

trolig et lite problem, siden eleven må flytte til et annet skoledistrikt for å endre skole. På grunn av prosjektets varighet er det rimelig å anta at familiens kostnader ved å flytte er større enn gevinsten. Dessuten vil elevens testresultater fortsatt være tilgjengelig så lenge eleven fremdeles er bosatt i Norge. Eventuell skjevhet i resultatene vil i tillegg begrenses så lenge frafallet er tilfeldig når det gjelder hvilken av de to gruppene eleven tilhører.

3 Tidligere litteratur

Selv om tidligere litteratur ikke gir noe entydig svar på effekten av lærertetthet i barneskolen, viser flere studier utført på småskoletrinnet at reduksjon i klassestørrelse har størst effekt på elever med dårlige forutsetninger (Angrist & Lavy, 1999; Krueger & Whitmore, 2001). Denne tilleggseffekten på undergrupper av svakere elever gjelder også undervisning i smågrupper ifølge et begrenset antall studier av ungdomsskolen og småskoletrinnet (Fryer Jr, 2014; Cook et al., 2015). Flere av disse studiene baserer seg på STAR-prosjektet utført i Tennessee i 1985–1989. Konklusjonen i STAR er at mindre klasser har en statistisk signifikant positiv effekt på elevprestasjoner på småskoletrinnet (Krueger & Whitmore, 2001). Hvorvidt lavt presterende elever i STAR har større utbytte av små klasser enn høyt presterende elever, er særlig interessant for denne oppgaven. Nye et al. (2002) finner at effekten av mindre klasser i matematikkundervisning på småskoletrinnet er mindre for lavt presterende elever enn for høyt presterende elever. De finner motsatt resultat for leseferdigheter. Flere studier konkluderer i tillegg med at effekten av mindre klasser er størst på lave klassesertrinn (Finn et al., 2003).

Økt lærertetthet betyr ikke nødvendigvis mindre klasser. Flere lærerårsverk benyttes for eksempel i stor grad som assistenter i klasserommet. I skoleåret 2017/2018 utfører assistenter i Norge 7118 årsverk på 1.–7. trinn. Til tross for denne utstrakte bruken av assistenter i skolen har det vært lite forskning på området, og eksisterende studier gir ikke noe klart svar på hvilken effekt bruken har (Andersen et al., 2018). En britisk studie finner at bruk av assistenter i skolen har negativ effekt på elevers resultater (Webster et al., 2013). I tillegg konkluderes det med at dårlig kvalifiserte assistenter benyttes som et alternativ til ordinære lærere i undervisning av elever med dårlige forutsetninger, hvilket potensielt fører til at lavt presterende elever presterer enda dårligere.

Andersen et al. (2018) benytter samme randomiseringsstrategi som 1+1-prosjektet og finner statistisk signifikante effekter av assistenter i klasserommet på leseferdighetene til 13-åringene. De finner imidlertid ingen signifikant effekt på matematikkresultater. Dette forklares med at leseferdigheter i dansk skole prioriteres foran matematikk, noe som gjenspeiles i resultatene ved at assistentene deltok i 4,5 språktimer i uken, men kun i 3 matematikktimer. En annen årsak kan være at matematikkferdighetene til assistentene var lave.

Hvordan assistenter blir benyttet spiller tilsynelatende en større rolle enn hvor mange assistenter som er til stede i et klasserom. Finn et al. (2001) foreslår blant annet at assistenter bidrar mest til forbedring av elevresultater når elever med læringsvansker får hjelp i smågrupper. Når assistenter benyttes på en måte som fører til mer undervisning i smågrupper, gir dette mer effektiv utnyttelse av ekstra ressurser, i tillegg til at denne type undervisning viser seg å forbedre resultatene til lavt presterende elever (Webster et

al., 2013; Andersen et al., 2018).

Resultater fra andre studier tyder også på at bruk av ekstra lærere til undervisning i smågrupper gir positiv effekt på elevresultater. I en studie av «charter schools» i New York deles deltakende barneskoler i to grupper ved å benytte hver skoles gjennomsnittlige estimerte påvirkning på elevresultater. Skoler over medianen kategoriseres som høyt presterende, og skoler under medianen kategoriseres som lavt presterende. I studien finner Dobbie & Fryer Jr (2013) at 33 prosent av høyt presterende barneskoler tilbyr undervisning i smågrupper, mens bare 10 prosent av lavt presterende barneskoler tilbyr dette. Altså er det langt mer sannsynlig at en høyt presterende barneskole tilbyr undervisning i smågrupper enn at en lavt presterende barneskole gjør det. Videre finner studien at skoler som prioriterer undervisning i smågrupper har årlige forbedringer i matematikkresultater 0,044 standardavvik høyere enn andre skoler. I tillegg finner Angrist et al. (2013) en positiv sammenheng mellom et mål på undervisning i smågrupper og effekten charter-skoler har på matematikkresultater på 4. – 7. trinn.

I en omfattende randomisert kontrollert studie utført på elevnivå fra offentlige skoler i Chicago finner Cook et al. (2015) at smågruppeundervisning har signifikant positiv effekt på elevresultater i ungdomsskolen blant lavt presterende elever. Resultatene fra studien antyder at elever som deltar i undervisning i smågrupper bestående av tre elever lærer ett til to år mer matematikk enn elever i kontrollgruppen i løpet av et år. Videre argumenteres det for at denne type undervisning er kostnadseffektiv, blant annet fordi lærere ikke trenger utdanning i hvordan man håndterer større elevgrupper i et klasserom. En ytterligere konsekvens av å ikke kreve denne type pedagogisk utdanning i smågruppeundervisning er at man har et større utvalg av potensielle undervisere, mange med spesialkompetanse i enkeltemner.

Til tross for generelt oppløftende resultater er det få studier som har intervenert i skolen for å undersøke effekten smågruppeundervisning har på elevresultater. Fryer Jr (2014) undersøker effekt av ulike tiltak i lavt presterende «charter schools» i Texas ved bruk av en tilsvarende randomisering som i 1+1-prosjektet. Han finner at ungdomsskoleelever som deltok i smågruppeundervisning presterte 200 prosent bedre enn elever i kontrollgruppen. Effektene var ikke like overbevisende i barneskolen. Dette kan skyldes at gruppene var større i barneskolen eller at elevene her ikke hang like mye etter som elevene i ungdomsskolen, og dermed ikke fikk like stort utbytte av smågruppeundervisningen.

4 Teori

Dette kapitlet er en gjennomgang av teori som kan benyttes til å forstå effekter av smågruppeundervisning. Det tas utgangspunkt i produktfunksjonen for utdanning og hvordan smågruppeundervisning inngår i denne. Videre diskuteres det i kapittel 4.2 vanlige økonomiske utfordringer i estimering av kunnskapsproduksjon.

4.1 Produktfunksjon for utdanning

Produktfunksjonen for utdanning inkluderer alle typer faktorer som kan tenkes å påvirke skoleprestasjonene til et individ. En vanlig spesifisering av funksjonen er gitt på følgende form (se for eksempel Todd & Wolpin (2003)):

$$O_{it} = f(S_{it}, F_{it}, \alpha_i), \quad (1)$$

hvor O_{it} er prestasjonsnivået til individ i på tidspunkt t , S_{it} er skoleinnsatsfaktorer for elev i tidspunkt t , F_{it} er familiebakgrunnen til elev i tidspunkt t og α_i er et mål på evnene til elev i . Disse antas gjerne å være medfødte egenskaper som ikke varierer over tid.

I denne oppgaven er det i hovedsak tre skoleinnsatsfaktorer som er viktige for å forstå effekten smågruppeundervisning har på elevprestasjoner: Lærertetthet, elevsammensetning og lærerkvalitet. Alle disse faktorene endres av smågruppeundervisning. Den første og mest åpenbare effekten er at alle elever i tiltaksskoler eksponeres for høyere lærertetthet enn elevene i sammenligningsskoler.

Elevsammensetningen som innsatsfaktor i produktfunksjonen for utdanning er i dette tilfellet noe mer komplisert. Noen elever får en mer gunstig elevsammensetning i både storgruppen og smågruppen som følge av tiltaket. Dette er trolig «gode» medelever som blir plassert i smågrupper sammen med andre elever med lignende atferd, og dermed oppnår en gunstig elevsammensetning i smågruppen. «Gode» medelever kjennetegnes blant annet ved at de inspirerer andre elever og ikke forstyrrer undervisningen (Chris, 2017). Disse elevene vil også oppnå en gunstig elevsammensetning i storgruppen fordi de resterende elevene med problematferd får undervisning i smågrupper når de «gode» medelevene er i storgrupper. Avhengig av hvordan smågruppene settes sammen på hvert trinn kan andre elever eksponeres for mer eller mindre gunstige elevsammensetninger i de to gruppene.

For elever som opplever en kombinasjon av høyere lærertetthet og gunstig elevsammensetning forventes en positiv effekt av smågruppeundervisning, alt annet likt. For elever som opplever en ugunstig elevsammensetning som følge av smågruppeundervisningen, trenger ikke effekten å være utelukkende positiv. Hvorvidt smågruppeundervisning fører til bedre prestasjoner for denne gruppen avhenger av hvor sterke effektene av elevsam-

mensetning er relativt til effektene av høyere lærertetthet.

Hva har dette å si for estimert forskjellig effekt av smågruppeundervisning for høyt og lavt presterende elever? Skoler som setter sammen prestasjonsmessig homogene grupper av elever i smågrupper kan oppleve svak positiv eller negativ effekt av smågruppeundervisning for lavt presterende elever hvis prestasjonsnivå er negativt korrelert med problematferd. De høyt presterende elevene som oppnår gunstige elevsammensetninger i både smågrupper og storgrupper vil ha høyere effekt av smågruppeundervisning.

Det rekrutteres en ny lærer på hvert trinn i forbindelse med prosjektet. Denne læreren representerer en endring i lærerkvalitet for alle elever i tiltaksskoler. Kvaliteten på denne læreren kan naturligvis enten være høyere eller lavere enn den allerede eksisterende læreren, og disse endringene kan påvirke utfallet. Lærerkvalitet er en viktig innsatsfaktor i skolen, og har stor innvirkning på testresultater og elevers sosiale status senere i livet (Chetty et al., 2014a,b). I kapittel 8.3 undersøkes det om ulike egenskaper hos læreren kan forklare skolenes evne til å påvirke elevers testresultater.

4.2 Kunnskapsproduksjon og økonometriske utfordringer

Et problem som ofte oppstår i estimering av produktfunksjonen for utdanning er endogenitet som følge av mangelfull data eller valg gjort av foreldre, lærere og skoler (Todd & Wolpin, 2003). For å redusere denne endogeniteten bidrar randomisering til eksogen variasjon for noen av disse variablene, slik som i dette prosjektet. Den randomiserte kontrollerte studien sørger for robuste resultater for estimert effekt av smågruppeundervisning. Likevel er det vanskelig å gå fra en analyse av *om* tiltaket fungerer til *hvorfor* tiltaket fungerer. Dette er fordi lærerne i skolene har mulighet til å gjøre mange valg. De kan bestemme størrelse og varighet på smågruppeundervisningen innenfor budsjettammer og retningslinjer gitt av prosjektet. Resultater i kapittel 8.1 viser at lærerne foretar mange valg som har innvirkning på hvordan undergrupper av elever deltar i prosjektet. I tillegg kan skoleledelsen ansette den smågruppelæreren de selv ønsker, noe som kan gi ulik lærerkvalitet mellom skoler. Dette diskuteres i kapittel 8.3. Alle disse valgene involverer uobserverbare vurderinger og atferd som kompliserer analysen.

Analysene som gjennomføres i denne oppgaven har som primært formål å undersøke om effektene av tiltaket varierer mellom undergrupper av elever. Avdekking av kausale årsaker til en eventuell slik variasjon begrenses av omfanget av denne oppgaven og datamaterialet på dette stadiet i prosjektet. Det er heller ikke et formål for analysen å vurdere ytterligere kausale sammenhenger utover det som gjennomføres. Likevel utforskes ulike korrelasjoner innad i tiltaksskolene i kapittel 8. I kapittel 8.1 avdekkes i hvilken grad elever behandles ulikt basert på prestasjonsnivå. Dette er direkte relevant for diskusjon av heterogene programeffekter analysert i kapittel 7. Videre undersøkes det i kapittel 8.2

om læringsutbyttet korrelerer med mengde smågruppeundervisning. Variasjonen mellom tiltaksskoler benyttes i kapittel 8.3 til å undersøke om læreres formelle kvalifikasjoner korrelerer med skolens effekt på elevresultater.

5 Deskriptiv statistikk

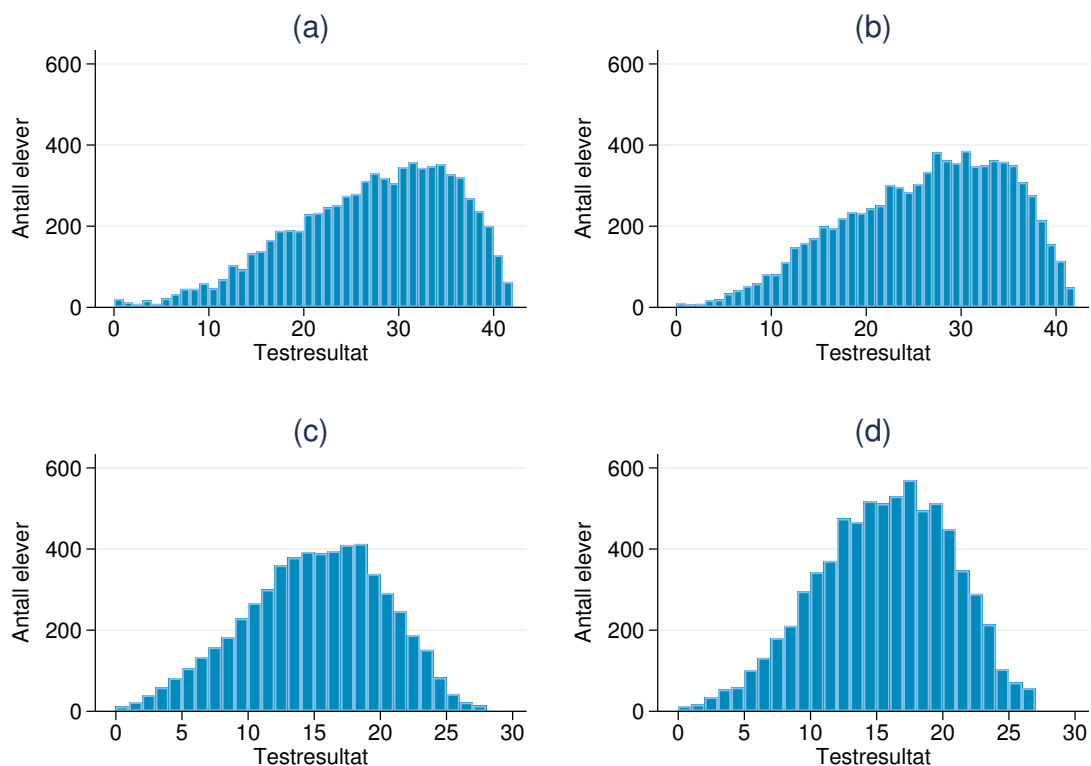
En stor del av datasettet benyttet i denne oppgaven baserer seg på innrapportering fra lærere. For å begrense skjevhet i estimerte effekter er noen observasjoner ekskludert fra datasettet. Det benyttede datamaterialet inkluderer alle elever som er rapportert å ha deltatt i smågruppeundervisning i 18 uker eller færre hvis ikke annet er spesifisert. Både tidsbegrensninger og budsjettmessige forhold tilsier at smågruppeundervisning i over 18 uker for enkeltelever trolig ikke er gjennomførbart, og disse verdiene skyldes antakeligvis feilrapportering. I tillegg er elever i tiltaksskoler som ikke mottok smågruppeundervisning ikke medregnet i store deler av oppgaven. Dette er elever som av ulike grunner fikk fritak for tiltaket, noe som gjør dem lite interessante for denne analysen. Jeg kommer tilbake til dette i kapittel 6.

Av de gjenværende 17 138 elevene i datasettet deltok 16 150 på pretesten, mens 13 138 elever deltok på posttesten. Sammenligningsgruppen er marginalt større enn tiltaksgruppen, men som det fremgår av Tabell 1 er skolene i tiltaksgruppen ikke overraskende noe mer ivrig i innrapporteringen av testresultater.

Figur 1 viser at elevene samlet sett presterer veldig bra på pretesten. Det kan være flere årsaker til dette, for eksempel at lærere hjalp elevene under testen. En annen forklaring er at pretesten trolig var noe enkel. Dette fører til en venstreskjev fordeling av testresultatene i begge gruppene, hvilket vil si at testen fungerer dårligere til å sortere de middels presterende og høyt presterende elevene. Posttesten ble forbedret sammenlignet med pretesten ved inkludering av vanskeligere oppgaver. Flere lærere uttrykte bekymring for at oppgavene skulle være for vanskelige og medføre en belastning for elevene. Et potensielt problem i forbindelse med dette er at bekymrede lærere kan hjelpe elever med de vanskeligste oppgavene, hvilket ville gjort testen til en uriktig representasjon av elevenes ferdigheter. Fordelingen av elevene i panel (c) og (d) viser imidlertid at resultatene er gjennomgående svakere i begge grupper. Dette gjør at posttesten sorterer de middels presterende og høyt presterende elevene bedre, og tyder på at intervensjon fra lærerne ikke er noe stort problem i dette tilfellet.

En enkel grafisk analyse av testresultatene fra posttesten antyder at elevene i sammenligningsskolene presterer bedre enn elevene i tiltaksskolene. Resultater fra Tabell 1 underbygger dette, og viser samtidig at gjennomsnittsskåren er høyere for sammenligningsskolene i pretesten. Selv om begge gruppene presterer dårligere på posttesten, er reduksjonen i gjennomsnittlig testskår lavere for tiltaksskolen. Den gjennomsnittlige testskåren i sammenligningsskolene reduseres med 12,166 poeng mot 10,727 poeng for tiltaksskolene. Denne overfladiske analysen antyder at elevene har en gevinst ved å delta i smågruppeundervisning.

Figur 1: Fordeling av testresultater



Panel (a) viser testresultater på pretest for elever i sammenligningsskolene, panel (b) viser testresultater på pretest for elever i tiltaksskolene, panel (c) viser testresultater på posttest for elever i sammenligningsskolene og panel (d) viser testresultater på posttest for elever i tiltaksskolene.

Tabell 1: Deskriptiv statistikk

	Før intervensering		Etter intervensering	
	(1)	(2)	(3)	(4)
	Tiltak	Sammenligning	Tiltak	Sammenligning
Gjennomsnittsskår	25.780	26.691	15.053	14.525
(Standardavvik)	(8.499)	(8.629)	(5.269)	(5.093)
Min	0	0	0	0
Maks	42	42	27	28
Skoler	82	81	81	81
Elever	8 475	7 675	7 426	5 712

Tabellen viser deskriptiv statistikk for tiltaksgruppen og sammenligningsgruppen før og etter smågruppeundervisning.

6 Kausal effekt av smågruppeundervisning

For å besvare spørsmålet om effekten av smågruppeundervisning er ulik for undergrupper av elever, undersøker jeg først om smågruppeundervisning har noen effekt på elevers testresultater. Den kausale effekten for et individ av å delta i smågruppeundervisning er forskjellen mellom to ulike utfall. Det første utfallet er testresultater hvis eleven deltar i smågruppeundervisning, og det andre utfallet er testresultater hvis eleven ikke deltar i smågruppeundervisning. Det er umulig å observere begge utfallene for et enkelt individ, fordi det ene utfallet utelukker det andre. Det er imidlertid mulig å observere det gjennomsnittlige utfallet for en gruppe som deltar i smågruppeundervisning og en gruppe som ikke deltar. Den beste måten å dele inn i en tiltaks- og sammenligningsgruppe er gjennom en kontrollert randomisering. Målet for randomiseringen er at de to gruppene skal være så like som mulig, eller i det minste ikke ha noen systematiske forskjeller (Weiss et al., 2014).

For å undersøke effekten av smågruppeundervisning i tiltaksskolene estimeres en regresjonsmodell med testskår på posttest som avhengig variabel på en indikator for tiltaksskoler. Det kontrolleres også for elevenes resultater på pretesten. Modellen er gitt på følgende form, og estimert med minste kvadraters metode (MKM):

$$A_{1,is} = \beta_0 + \beta_1 Tiltaksskole_s + \beta_2 A_{0,is} + \varepsilon_{is}, \quad (2)$$

hvor $A_{1,is}$ er verdien til testskår etter intervensjonen for elev i ved skole s , standardisert til å ha gjennomsnittsverdi null og standardavvik lik én. Indikatorvariabelen $Tiltaksskole_s$ tar verdien én hvis skole s ble randomisert til gruppen med smågruppeundervisning, og null hvis skole s er en sammenligningsskole. $A_{0,is}$ er verdien til testskår for elev i , skole s før intervensjonen, standardisert til å ha gjennomsnittsverdi null og standardavvik lik én. ε_{is} er et restledd.

De standardiserte verdiene for testresultatene er generert basert på testresultatene for hvert av de to trinnene. Da de to trinnene ikke gjennomførte de samme testene, er standardisering av variablene nødvendig for å kunne aggregere resultatene. Videre er skalaen for pre- og posttesten ulik, slik at standardisering gjør det lettere å estimere tolkbare resultater. Standardiseringen utføres ved å først beregne trinnsesifikke gjennomsnittsverdier. Dette gir to gjennomsnittsverdier for hvert trinn, én for hver av testene. Den tilhørende gjennomsnittsverdien trekkes fra hver observerte verdi på testresultatet, for så å deles på standardavviket til hver test på hvert trinn.

Parameteren β_1 viser gjennomsnittlig forskjell i testresultat mellom tiltaksskolene og sammenligningsskolene, og tolkes dermed som *intention to treat*-effekten. Randomisering sørger for identifisering ($E[\varepsilon_{is}|Tiltaksskole_s] = 0$), hvilket gjør at MKM-parameteren

Tabell 2: Effekt av smågruppeundervisning

	Posttest		
	(1) Hele utvalget	(2) Tiltakselever	(3) Anbefalt lengde
Pretest	0.661*** (0.00686)	0.660*** (0.00691)	0.661*** (0.00699)
Tiltaksskole	0.184*** (0.0136)	0.189*** (0.0136)	0.189*** (0.0138)
Konstant	-0.121*** (0.0103)	-0.121*** (0.0103)	-0.121*** (0.0103)
Observasjoner	12 795	12 649	12 359
R^2	0.422	0.421	0.421

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser tre separate MKM-regresjoner med standardiserte verdier på testresultater. Kolonne (1) benytter hele utvalget, kolonne (2) inkluderer ikke elever ved tiltaksskoler som ikke deltok i smågruppeundervisning og kolonne (3) ekskluderer i tillegg elever ved tiltaksskoler som deltok i smågruppeundervisning i kortere eller lengre periode enn anbefalt (4-12 uker). Standardfeil i parentes.

foran indikatorvariabelen kan tolkes som estimert effekt av å delta i smågruppeundervisning. Ved å inkludere elevenes resultater fra pretesten kontrolleres det for sammenhengen mellom elevenes resultater på de to testene.

Tabell 2 viser resultater fra spesifikasjonen i ligning (2). Begge variablene for testresultater er standardisert til å ha gjennomsnittsverdi lik null og standardavvik lik én. Dette gjør at verdiene kan tolkes som effektstørrelser. De tre kolonnene viser resultater for tre ulike utvalg. Alle resultater i denne estimeringen er statistisk signifikante for alle fornuftige nivåer på signifikansnivå.

Kolonne (1) viser resultater for hele utvalget. En estimert effektstørrelse på 0,184 betyr at elever ved tiltaksskolene forbedret matematikkresultatene sine med 18,4 prosent av et standardavvik sammenlignet med sammenligningsgruppen. Sagt på en annen måte er skåren til gjennomsnittspersonen i tiltaksgruppen 0,184 standardavvik høyere enn skåren til gjennomsnittspersonen i sammenligningsgruppen.

I kolonne (2) er alle elever ved tiltaksskoler som ikke deltok i smågruppeundervisning utelatt. Dette utgjør en reduksjon på 146 elever. Elevene i tiltaksskoler som ikke deltok i smågruppeundervisning er fanget opp gjennom innrapportering fra lærere ved tiltaksskolene om hvor mange uker hver enkelt elev har deltatt i smågruppeundervisning. Alle disse elevene er i kolonne (1) beholdt i gruppen for tiltaksskoler, da det kan være ulike

smitteeffekter av de ekstra tildelte lærerressursene i tiltaksskoler, også for elever som ikke deltar i smågruppeundervisning. Disse effektene er vanskelig å observere, og gjør at elever som faller innenfor denne kategorien ikke er blitt inkludert i sammenligningsgruppen i denne delen av oppgaven.

Det er flere potensielle årsaker til at elever ved tiltaksskoler ikke deltok i smågruppeundervisning. Disse elevene er i stor grad elever som mottok ulike former for spesialundervisning eller var funksjonshemmet. En siste årsak kan være elever som endret skole eller byttet klasse innad på et trinn. Dette gjelder svært få elever, fordi de fleste elever som ikke kunne delta i smågruppeundervisning heller ikke tok testene. Ekskludering av den nevnte undergruppen av elever endrer resultatene i liten grad. Resultatene er fortsatt statistisk signifikante for utvalget i kolonne (2), og effekten er av nærmest identisk størrelse. Elever som deltok i smågruppeundervisning presterte i gjennomsnitt 0,189 standardavvik bedre enn elever i kontrollskolen.

I kolonne (3) rapporteres resultater for elever som deltok i smågruppeundervisning med varighet på 4-12 uker. Tiltaksskolene ble instruert om å gjennomføre smågruppeundervisning i minimum fire uker per gruppe. I kolonne (3) er alle elever ved tiltaksskoler med færre enn fire uker smågruppeundervisning ekskludert. I tillegg er alle elever som deltok i smågruppeundervisning i mer enn 12 uker ekskludert. Disse observasjonene kan skyldes feilrapportering (noen lærere rapporterte svært lang varighet på smågruppeundervisningen) eller at enkelte skoler misforstod prosjektet. Uansett årsak ser ikke ekskludering av disse elevene ut til å endre resultatene, da effekten av smågruppeundervisning er den samme for elever som kun deltok i smågruppeundervisning i 4-12 uker som for resten av tiltakselevne.

Resultatene i Tabell 2 viser en klart signifikant positiv effekt av smågruppeundervisning, og effekten er lik for alle tre utvalgene. Sammenlignet med resultater fra tidligere studier av klassestørrelse, er størrelsen på effekten forholdsvis lik. For eksempel finner Krueger (1999) i STAR-prosjektet at en reduksjon i klassestørrelse på 8 elever fører til en forbedring i elevprestasjoner med 0,22 standardavvik i andre klasse og 0,19 standardavvik i tredje klasse. Dette er marginalt høyere enn mine estimerte effekter av smågruppeundervisning. Angrist & Lavy (1999) finner noe svakere effekt enn Krueger (1999) av en tilsvarende reduksjon i klassestørrelse.

7 Heterogene programeffekter

I denne delen av oppgaven ønsker jeg å undersøke problemstillingen om smågruppeundervisning har størst effekt på lavt presterende elever eller høyt presterende elever. For å avdekke eventuell ulik effekt av behandling mellom ulike undergrupper av elever benytter jeg flere ulike modellspesifikasjoner. Først deles elevene i to grupper basert på elevenes resultater på pretesten, slik at elevenes prestasjonsnivå blir grunnlaget for inndelingen. Deretter deles elevene inn i fem kvantiler rangert etter resultater på pretesten.¹ Avhengig variabel er som tidligere testskår på posttesten, som måler elevenes prestasjoner etter perioden med smågruppeundervisning.

7.1 Resultater

Tabell 3 viser fire ulike modeller for å estimere effekt av smågruppeundervisning på to undergrupper av elever, elever som hadde testskår over medianen på pretesten og elever som hadde testskår under medianen på pretesten. Kolonne (1) inkluderer en dummyvariabel, Lavtpresterende, som er en indikatorvariabel for om eleven er høyt eller lavt presterende. Variabelen tar verdien én hvis eleven presterte lavere enn medianskåren for hele utvalget på pretesten og null ellers. Videre inkluderes det i kolonne (1) et interaksjonsledd mellom indikatorvariabelen og variabelen for om eleven er plassert i en tiltaksskole. Dette gjøres for å avdekke ulik effekt av behandling for de to elevgruppene. Kolonne (2) inkluderer et interaksjonsledd mellom indikatorvariabelen for om eleven går på en tiltaksskole og testskår på pretesten.

Kolonne (1) viser at elever som har testskår under medianen presterer 0,0755 standardavvik dårligere enn elever som har testskår over medianen. Dette betyr at de lavt presterende elevene har lavere skår på posttesten enn elever som presterte over medianen på pretesten, uavhengig av om eleven går på en tiltaksskole eller en sammenligningsskole. Det er ingen signifikant ulik estimert effekt av interaksjonsleddet i kolonne (1), som viser om effekt av smågruppeundervisning er avhengig av hvorvidt en elev presterte over eller under medianen på pretesten. Denne inndelingen er trolig for grov til å avdekke noen signifikant ulik effekt. Interaksjonsleddet i kolonne (2) viser imidlertid at effekten av smågruppeundervisning varierer avhengig av hvordan elevene presterte på pretesten. Jeg finner en signifikant negativ sammenheng mellom hvor godt en elev presterte på pretesten og effekten av smågruppeundervisning. Dette betyr ikke at høyt presterende elever taper på å delta i smågrupper, men at gevinsten er mindre enn for lavt presterende elever.

Kolonne (3) og (4) viser to separate regresjoner for to forskjellige utvalg, elever som presterte over medianen og elever som presterte under medianen på pretesten. Disse re-

¹Se Appendiks A for fordelingen av elever i kvantiler.

Tabell 3: Effekt av smågruppeundervisning for undergrupper

	Posttest			
	(1) Hele utvalget	(2) Hele utvalget	(3) Over median	(4) Under median
Pretest	0.634*** (0.0123)	0.672*** (0.0107)	0.759*** (0.0203)	0.575*** (0.0158)
Tiltaksskole	0.175*** (0.0187)	0.191*** (0.0137)	0.179*** (0.0175)	0.202*** (0.0213)
Lavtpresterende	-0.0755*** (0.0287)			
Tiltaksskole x Lavtpresterende	0.0302 (0.0274)			
Tiltaksskole x Pretest	-0.0237* (0.0140)			
Konstant	-0.0855*** (0.0171)	-0.123*** (0.0104)	-0.188*** (0.0211)	-0.207*** (0.0207)
Observasjoner	12 643	12 643	6 660	5 983
R^2	0.421	0.421	0.181	0.188

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser MKM-regresjon med standardiserte verdier på testresultater. Kolonne (1) inkluderer interaksjonsledd mellom indikatorvariabel for tiltaksskole og testskår fra pretest. Kolonne (2) inkluderer interaksjonsledd mellom indikatorvariabel for henholdsvis tiltaksskole og om eleven hadde testskår lavere enn medianverdien for hele utvalget på pretesten. Standardfeil i parentes.

gresjonene tillater ulik sammenheng mellom elevenes resultater på de to testene, som er svakere for de lavt presterende elevene. Punkttestimatene for hver av gruppene er 0,759 og 0,575, og forskjellen er signifikant ulik på under én prosent signifikansnivå (p -verdi=0,000). I likhet med resultatene i kolonne (2) finner jeg også her en sterkere effekt av smågruppeundervisning for lavt presterende elever. Den estimerte effekten av smågruppeundervisning på elevprestasjoner er høyere for elevene som presterte under medianskåren på pretesten. En elev som presterer over medianskåren i en tiltaksskole presterer 0,179 standardavvik bedre enn en tilsvarende elev i en sammenligningsskole. For en elev som presterer under medianskåren er forskjellen 0,202 standardavvik. Differansen mellom dem er -0,0230. Estimert ulik effekt for de to undergruppene er av en størrelse som er nærmest identisk som estimert effekt i kolonne (2), da en økning i testresultat på pretesten på ett standardavvik gir 0,0237 standardavvik lavere forbedring i testresultat av å delta i smågruppeundervisning. Dette tyder på robuste estimater.

Både elever som presterer over og under medianen gjør det bedre i tiltaksskolene enn i sammenligningsskolene. Dette fremgår tydelig av en signifikant positiv effekt av indikatorvariabelen for om eleven er i en tiltaksskole. Effekten er klart positiv i alle de fire ulike modellspesifikasjonene i Tabell 3, men de lavt presterende elevene presterer ikke overraskende svakere på posttesten enn de høyt presterende elevene i begge gruppene av skoler. Sammenhengen mellom resultat på pretest og posttest er også svakere for de lavt presterende elevene enn de høyt presterende elevene, som man ser i kolonne (3) og (4). Det er en signifikant forskjell i effekt av smågruppeundervisning avhengig av prestasjoner på pretesten. Elever får større gevinst av å delta i smågruppeundervisning jo lavere prestasjonsnivå de ligger på, slik det fremgår av kolonne (2). Det er også ulik effekt av smågruppeundervisning mellom elever som presterer over medianverdien og elever som presterer under medianverdien når det tillates ulik sammenheng mellom pretest og posttest for de to gruppene elever.

Alle elever som presterer dårligere enn medianen kan imidlertid ikke kategoriseres som lavt presterende. I tillegg ønsker jeg å undersøke om effekten av smågruppeundervisning er ulik for elever som presterer svært dårlig eller svært godt. For å avdekke heterogene programeffekter mellom elever med ulikt prestasjonsnivå, er det derfor nødvendig med ytterligere inndelinger basert på testresultater. I tidligere litteratur er det benyttet ulike kategoriseringer. Her velger jeg å dele elevene inn i fem kvantiler med utgangspunkt i testskår på pretesten.

Tabell 4 viser fem separate regresjoner, én for hvert av kvintilene. Effekten av smågruppeundervisning avtar jo lenger ut i halene av fordelingen over testresultater eleven befinner seg. En elev i det midterste kvintilet kan sies å være en elev med middels prestasjonsnivå. Gjennomsnittseleven i denne gruppen i en tiltaksskole skårer 0,233 standardavvik høyere enn en tilsvarende elev i en sammenligningsskole. Denne effekten er langt større

Tabell 4: Effekt av smågruppeundervisning for ulike kvantiler

	Posttest				
	(1) 1. kvantil	(2) 2. kvantil	(3) 3. kvantil	(4) 4. kvantil	(5) 5. kvantil
Pretest	0.456*** (0.0359)	0.652*** (0.0769)	0.671*** (0.0854)	0.628*** (0.0925)	0.850*** (0.0655)
Tiltaksskole	0.188*** (0.0341)	0.208*** (0.0321)	0.233*** (0.0301)	0.193*** (0.0285)	0.125*** (0.0270)
Konstant	-0.388*** (0.0583)	-0.171*** (0.0435)	-0.187*** (0.0268)	-0.113* (0.0686)	-0.271*** (0.0867)
Observasjoner	2 548	2 443	2 664	2 610	2 378
R^2	0.069	0.043	0.044	0.035	0.072

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser MKM-regresjon med standardiserte verdier på testresultater. Hver kolonne er en separat regresjon for hver av kvantilene. Standardfeil i parentes.

enn effekten for en svært høyt presterende elev, som forbedrer prestasjonene med 0,125 standardavvik. Forskjellen mellom punkttestimatene er statistisk signifikante på én prosent signifikansnivå (p -verdi=0,007). De resterende forskjellene mellom kvintilene i effekt av smågruppeundervisning er ikke statistisk signifikante.

Effekten er også lavere for de svært lavt presterende elevene. For denne gruppen presterer gjennomsnittseleven i en tiltaksskole 0,188 standardavvik bedre enn gjennomsnittseleven i en sammenligningsskole. Resultatene i Tabell 4 antyder altså en forskjell i effekt mellom de svært høyt og svært lavt presterende elevene, selv om forskjellen er for liten til å utgjøre noen statistisk signifikans. Generelt er effekten av smågruppeundervisning tilsynelatende noe høyere for lavt presterende enn høyt presterende elever, og forskjellen øker jo større forskjell det er mellom elevenes prestasjonsnivå. Dette er konsistent med tidligere resultater fra Tabell 3.

7.2 Oppsummerende kommentarer

Tabell 4 gir en rik beskrivelse av ulik effekt av smågruppeundervisning, og viser at effekten avtar jo lenger ut i halene av fordelingen over resultater på pretesten en elev befinner seg. I tillegg indikerer estimatene i Tabell 4 at lavt presterende elever har noe høyere utbytte av smågruppeundervisning enn høyt presterende elever. Ulike modellspesifikasjoner i Tabell

3 underbygger disse resultatene.²

Tabell 4 viser i tillegg en forskjell i sammenheng mellom de to testene avhengig av hvilket kvintil elevene befinner seg i. Forskjellen mellom de to punkttestimatene for 1. kvintil og 5. kvintil er statistisk signifikant på under én prosent signifikansnivå (p-verdi=0,000). For elevene i det øverste kvintilet er sammenhengen svært høy, noe som vil si at en høyt presterende elevs evner er veldig avgjørende for testskåren. Disse elevene har ikke like mye å gå på som elever på et lavere nivå, og presterer høyt uansett hvilken undervisningsform de deltar i. Dette kan være med på å forklare den noe lavere effekten smågruppeundervisning har på disse elevene.

Det er nærliggende å sammenligne resultatene med det tidligere nevnte randomiserte prosjektet STAR gjennomført på småskoletrinn i Tennessee. Nye et al. (2002) antyder at høyt presterende elever har høyere utbytte av mindre klasser når det gjelder leseferdigheter, både når lavt presterende elever kategoriseres som elever under medianen og elever i det nederste kvartilet i fordelingen av testresultater. Effekten er motsatt for matteferdigheter på den måten at høyt presterende elever har mer utbytte av små klasser.

Konstantopoulos (2008) undersøker om gapet mellom høyt presterende og lavt presterende elever kan tettes ved å endre klassestørrelser. Også denne analysen baserer seg på STAR-prosjektet. Forfatteren finner indikasjoner på at høyt presterende elever har mer utbytte av mindre klasser enn lavt presterende elever, slik at manipulering av klassestørrelser ikke er egnet til å tette prestasjonsgapet.

²I Appendiks B foretas det ytterligere robusthetssjekker.

8 Variasjon mellom tiltaksskoler

Resultatene i kapittel 6 tyder på at smågruppeundervisning har effekt på elevresultater. Kapittel 7 konkluderer med at denne effekten er noe høyere for lavt presterende elever enn for høyt presterende elever, mens det tilsynelatende er de middels presterende elevene som har størst utbytte av smågruppeundervisning. I denne delen av analysen skal jeg se nærmere på tiltaksskolene og undersøke ulike faktorer som kan påvirke estimert ulik effekt mellom lavt presterende og høyt presterende elever. Her skal jeg undersøke potensielle årsaker til at jeg ikke finner en mer tydelig ulik effekt, men også drøfte potensielle sammenhenger som kan drive en estimert signifikant ulik effekt mellom lavt presterende og høyt presterende elever.

Variasjon i antall uker smågruppeundervisning kan være en viktig faktor for å forklare ulik effekt av tiltaket på undergrupper av elever. For å avgjøre hvor robuste resultatene i kapittel 7 er, undersøkes denne variasjonen nøye i kapittel 8.1.

Forskjellen i hvordan ulike elevgrupper har deltatt i prosjektet påvirker ikke bare estimert ulik effekt av smågruppeundervisning, men kan også i seg selv benyttes som et mål på tiltaket. Noen elever har deltatt mer i smågruppeundervisning enn andre elever. Ulikt antall uker smågruppeundervisning for undergrupper av elever benyttes kapittel 8.2 til å undersøke en sammenheng mellom antall uker en elev deltok i smågruppeundervisning og elevens testresultater.

I tillegg utforskes ytterligere variasjon mellom tiltaksskoler ved å estimere skolespesifikke effekter i kapittel 8.3. Spesielt undersøkes hvordan variasjon mellom tiltaksskoler kan benyttes til å forklare elevenes prestasjoner. Er det mulig å identifisere årsaker til variasjon mellom skoler som forklarer testresultater? For å svare på dette spørsmålet benyttes variasjon mellom skoler som et mål på skolens evne til å forbedre testresultater. Et mål på skolens evne til å forbedre resultater kan blant annet brukes til å si noe mer presist om størrelsen på effekten av smågruppeundervisning. For å benytte variasjonen mellom skoler ytterligere, estimeres effekten som lærernes egenskaper har på dette målet for skolens effektivitet.

8.1 Antall uker smågruppeundervisning

Den første faktoren som kan drive observert ulik effekt mellom høyt og lavt presterende elever kan være at de svake elevene fikk tettere oppfølging av lærere ved at de tilbrakte mer tid i smågrupper. En annen faktor kan være at disse elevene opptok en større andel av lærerens tid i løpet av undervisningen. I denne delen av oppgaven benyttes innrapporterte verdier på hvor mange uker hver elev deltok i smågruppeundervisning til å avdekke ulik behandling av lavt presterende og høyt presterende elever i prosjektet. Ulikt antall

Tabell 5: Deskriptiv statistikk for antall uker

	Gjennomsnitt	(Standardavvik)	Min	Maks	Elever
Hele utvalget	7.610	(2.351)	1	18	6401
Lavtpresterende	8.045	(2.311)	1	18	3205
Høytpresterende	7.172	(2.310)	1	17	3196
1. kvintil	8.265	(2.296)	1	18	1447
2. kvintil	7.920	(2.295)	1	17	1245
3. kvintil	7.598	(2.271)	1	18	1344
4. kvintil	7.160	(2.268)	1	17	1207
5. kvintil	6.996	(2.318)	1	15	1056

Tabellen viser deskriptiv statistikk for antall uker ulike elevgrupper deltok i smågruppeundervisning. Inkluderte elever deltok i smågruppeundervisning minst 1 og maksimalt 18 uker.

uker smågruppeundervisning kan forklare hvorfor effekten av tiltaket varierer mellom undergrupper. Diskusjonen er viktig for å analysere hvor robuste resultatene i kapittel 7 er. Antall uker elever ved tiltaksskolene deltok i smågruppeundervisning ble innrapportert av skolene.³ I tråd med tidligere diskusjon er ekstremverdier utelatt fra analysen for å skape et mest mulig troverdig bilde av antall uker elever deltok i smågruppeundervisning.

Resultatene i denne delen av analysen er ikke nødvendigvis kausale sammenhenger, da det kan forekomme uobserverte faktorer som driver korrelasjonene mellom forklaringsvariabler og resultater på posttesten. Den største årsaken til dette er at lærerens preferanser og ferdigheter påvirker både antall uker smågruppeundervisning og elevenes testresultater.

En annen av disse faktorene kan være evnene til rektorer eller skoleledelsen. Foreldres mulighet til å påvirke undervisningen kan også være en årsak til korrelasjon. Press fra ressurssterke foreldre kan påvirke antall uker en elev deltar i smågruppeundervisning. Dette kan svekke resultatene, da disse elevene trolig deler mange av de samme karakteristikene.

Elever og foreldre kan til en viss grad selektare seg til ulike skoler. Elevsammensetningen varierer mye mellom ulike skoler i Norge, og særlig skoler hvor elevene har foreldre med høy utdanning presterer høyt på nasjonale prøver (Steffensen et al., 2017). At elever og foreldre selekterer seg til skoler kan føre til ulik deltakelse i smågruppeundervisning avhengig av hvor belastet område skolen befinner seg i.

For å undersøke ulik deltakelse i smågruppeundervisning presenteres deskriptiv statistikk i Tabell 5. Tabellen viser at elever som presterte under medianskåren på pretesten deltok i gjennomsnitt 0,9 uker mer i smågruppeundervisning enn de elevene som presterte over medianskåren (8,045 mot 7,172). Forskjellen mellom lavt presterende og høyt preste-

³Se appendiks C for fordeling av antall uker smågruppeundervisning.

Tabell 6: Sammenheng mellom testresultater og antall uker i smågrupper

	Antall uker		
	(1)	(2)	(3)
Pretest	-0.0544*** (0.00336)		
Lavtpresterende		0.874*** (0.0578)	
Kvintiler			-0.340*** (0.0206)
Konstant	9.033*** (0.0914)	7.172*** (0.0409)	8.598*** (0.0658)
Observasjoner	6 299	6 401	6 299
R^2	0.040	0.035	0.041

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser MKM-regresjon for elever som mottok minimum én uke og maksimum 18 uker smågruppeundervisning. Standardfeil i parentes.

rende elever gjør seg enda mer gjeldende hvis man tar utgangspunkt i resultatene fra hver ende i fordelingen av resultater på pretesten. Inndelingen i kvintiler viser at forskjellen mellom de aller svakeste og aller sterkeste elevene var 1,3 uker smågruppeundervisning i gjennomsnitt.

Den deskriptive statistikken fra tiltaksskolene tyder på en sammenheng mellom elevenes prestasjonsnivå og hvor mange uker eleven deltok i smågruppeundervisning. For å undersøke denne sammenhengen mer formelt kjøres tre regresjoner på formen:

$$Uker_{is} = \sigma + \tau S_{0,is} + e_{is}, \quad (3)$$

hvor $Uker_{is}$ er antall uker elev i ved skole s deltok i smågruppeundervisning og $S_{0,is}$ er tre ulike måter å kategorisere elev i ved skole s sitt prestasjonsnivå på. Tabell 6 viser resultatene.

I kolonne (1) viser τ økningen i antall uker smågruppeundervisning hvis eleven forbedrer testskåren sin med ett poeng. Effekten av å prestere bedre på pretesten er signifikant negativ på antallet uker en elev deltar i smågruppeundervisning. Estimater viser at en forbedring av resultatet på pretesten med 10 poeng ville ført til en halv uke, eller 2,5 dager, kortere smågruppeundervisning.

Kolonne (2) viser forskjellen i antall uker smågruppeundervisning mellom en elev som

presterte under og over medianen på pretesten. Elever som presterte under medianen på pretesten deltok i smågruppeundervisning i 0,9 uker mer enn elever som presterte over medianen. Dette er identisk med resultatene fra den deskriptive analysen.

Kolonne (3) viser sammenheng mellom hvilket kvintil eleven befant seg i basert på resultater fra pretesten og antall uker smågruppeundervisning. For hvert kvintil forbedring i testresultater mottok eleven 0,34 uker mindre smågruppeundervisning. Alle de tre ulike kategoriseringene av elevenes ferdigheter gir robuste resultater på den måten at effekten av en forbedring av testresultater på antall uker smågruppeundervisning er av tilnærmet samme størrelsesorden i hver av spesifikasjonene. Resultatene i Tabell 6 tyder på at lavt presterende elever mottok mer smågruppeundervisning enn høyt presterende elever. Alle fortegnene er som forventet, og resultatene er statistisk signifikante. Som diskutert i kapittel 4 øker lærertettheten for alle elever i en tiltaksskole, men den øker altså minst for de høyt presterende elevene og mest for de lavt presterende elevene.

At de svakeste elevene gjennomgående blir plassert i smågrupper oftere enn de sterkeste elevene tyder på at skolene ønsker å løfte flest mulig av de lavt presterende elevene til et høyere nivå. Dette er et eksempel på subjektive vurderinger foretatt av lærere som diskutert i kapittel 4.2. Lærernes subjektive vurdering av hvordan ulike elever skal delta i prosjektet kan ha innvirkning på estimerte effekter i kapittel 7.

Plasseringen av elever i smågrupper tyder på kompenserende atferd for lavt presterende elever fra lærerne. I så fall kan det også tenkes at lærerne kompenserer de lavest presterende elevene ytterligere ved å vie dem mer oppmerksomhet i selve undervisningen av smågruppene. Denne ekstra lærerinnsatsen rettet mot de svakeste elevene kan bidra til skjevhet i estimerte effekter av smågruppeundervisning basert på resultater fra pretesten, men dette er trolig ikke noe problem her. Hvis lavt presterende elever blir viet mer oppmerksomhet i smågruppeundervisningen, vil de samme elevene sannsynligvis også få tettere oppfølging i den ordinære undervisningen. Dette gjør at kontrasten ved å delta i smågruppeundervisning ikke blir større for lavt presterende enn for høyt presterende elever, og den høyere estimerte effekten av smågruppeundervisning i kapittel 7 for de lavt presterende elevene blir mer troverdig. I tillegg er smågruppene trolig satt sammen av elever med et tilnærmet likt prestasjonsnivå for lettere å kunne tilpasse undervisningen, slik at variasjonen mellom elever innad i hver gruppe er lav.

Det kan også tenkes at lærerne og skoleledelsen ved de ulike skolene i prosjektet har et ønske om at deres elever skal fremstå på best mulig måte i studien. Siden lærerne overvåkes og alle resultater rapporteres, kan lærerne kompensere svake elever med mer undervisning i smågrupper for å begrense mengden av veldig svake testresultater.

Hvis lavt presterende elever tilbys mer intensiv undervisning gjennom hele prosjektets varighet, blir det vanskeligere å fange opp en potensiell ekstragevinst sterke elever får ved å delta i smågruppeundervisning. Dette bidrar til å dempe estimert effekt av

smågruppeundervisning for høyt presterende elever samtidig som effekten for lavt presterende elever kan bli overestimert. Til tross for dette finner jeg høy signifikant effekt av smågruppeundervisning på elever i de to øverste kvintilene.

Blant de svært lavt presterende elevene vil det trolig være en høy andel elever som sliter veldig med matematikk selv om de deltar mye i smågruppeundervisning. Disse elevene vil ha liten fremgang til tross for sterk medisin i form av en høy dose smågruppeundervisning, og estimert effekt av smågruppeundervisning vil følgelig bli lavere for denne undergruppen av elever. Likevel finner jeg en sterk effekt av smågruppeundervisning for elevene i det nederste kvintilet. Dette tyder på at smågruppeundervisning er egnet til å løfte mange av de elevene som presterer på et svært lavt nivå.

Det vil også være en stor andel blant de svært høyt presterende elevene det blir vanskelig å måle noen fremgang for da de allerede presterer på et svært høyt nivå. Som nevnt i kapittel 5 var særlig pretesten noe enkel, hvilket kompliserer estimerte effekter for denne undergruppen av elever, da elevene er inndelt etter resultater på pretesten. Bedre sortering av svært høyt presterende elever enn det som lar seg gjøre med dette datamaterialet vil være mulig på et senere tidspunkt i prosjektet. For å forbedre målbarheten av effekter for svært høyt presterende elever, kreves andre former for undervisning rettet spesielt mot denne elevgruppen i tillegg til undervisning av et annet pensum og vanskeligere tester.

Som forklart i kapittel 4 vil høyt presterende elever i enkelte skoler oppleve en mer gunstig elevsammensetning i både storgruppen og smågruppen som følge av prosjektet. Dette vil isolert sett bidra til sterkere estimert effekt av smågruppeundervisning for disse elevene, men at denne gruppen deltar mindre i smågruppeundervisning kan være med på å dempe effektene. Likevel har disse elevene såpass lite forbedringspotensial innen det fastsatte pensum at økt mengde smågruppeundervisning trolig ikke ville spilt en stor rolle.

Den gruppen som står igjen som vinnerne i prosjektet er de middels presterende elevene. Denne gruppen består av elever som stort sett mottar en anbefalt mengde smågruppeundervisning. Disse elevene har trolig noe ulikt evnenivå, og resultatene deres kan ha sammenheng med innsats og undervisningsform. Flere av elevene med middels prestasjonsnivå kan ved hjelp av riktig undervisningsform forbedre resultatene sine markant. Smågruppeundervisning kan for denne gruppen bidra til økt motivasjon, slik at elever med et høyt evnenivå men lav innsats forbedrer resultatene. I tillegg kan smågruppeundervisning ikke bare påvirke innsatsen, men også ha signifikant positiv effekt på hvor mye en middels presterende elev tar til seg av pensum.

Resonnementet er med på å forklare den sterke effekten smågruppeundervisning har på middels presterende elever rapportert i Tabell 4. Diskusjonen i denne delen av oppgaven sett under ett tyder på at resultatene i kapittel 7 virker å være robuste, til tross for at ulik behandling av elever kan påvirke resultatene til en viss grad.

Hva med elevsammensetningen for de middels presterende elevene? I likhet med de høyt presterende elevene kan middels presterende elever oppleve mer gunstig elevsammensetning hvis elever med problematferd ikke er til stede i deres smågrupper, eller hvis elever med problematferd plasseres sammen i smågrupper når de resterende elevene er i storgrupper. Disse effektene undersøkes ikke spesifikt i denne oppgaven, men er del av et større bilde som forklarer elevprestasjoner diskutert i kapittel 4.

Selv om jeg finner signifikant ulik effekt av smågruppeundervisning avhengig av prestasjonsnivået til elevene, og særlig stor effekt på middels presterende elever, kan det være faktorer som demper sammenhengene. En forklaring på at forskjellene ikke er tydeligere kan være at de lavt presterende elevene allerede har deltatt i smågruppeundervisning før tiltaket ble igangsatt, da smågruppeundervisning ikke er en ukjent undervisningsform i norsk skole. Kontrasten mellom å delta og ikke delta i smågruppeundervisning vil i så fall bli mindre for de lavt presterende elevene enn de høyt presterende elevene. Dette kan bidra til å redusere den estimerte forskjellen i effekt av smågruppeundervisning mellom elevgruppene diskutert i kapittel 7.

En annen faktor som kan dempe estimert ulik effekt er diskutert i kapittel 4.1. Skoler som setter sammen prestasjonsmessig homogene grupper vil trolig oppleve en sterkere effekt av smågruppeundervisning for høyt presterende elever. Siden disse elevene tilsynelatende har noe svakere effekt av smågruppeundervisning, vil disse effektene dempe estimert forskjell mellom undergrupper. Siden det er opp til lærere og skolene i dette prosjektet å sette sammen smågruppene basert på subjektive vurderinger, er det vanskelig å måle disse effektene presist. I en utvidelse av denne analysen kunne man undersøkt om det er en positiv korrelasjon mellom antall uker lavt presterende elever deltar i smågrupper og prestasjoner til høyt presterende elever.

For å undersøke graden av ekstraundervisning for de lavt presterende elevene måtte studien vært gjennomført noe annerledes. For eksempel kunne et av årskullene i tiltaksskolene blitt holdt likt sammenligningsskolene det første året for å undersøke ressursallokeringen i skolen når skoleledelsen er klar over at de vil få et ekstra lærerårsverk neste år. Dette ville også redusert et annet problem med ulik kontrast mellom tiltaksskoler og sammenligningsskoler som oppstår hvis det ekstra tildelte lærerårsverket benyttes som en erstatning for allerede eksisterende assistenter.

8.2 Effekt av antall uker smågruppeundervisning

Antall uker smågruppeundervisning har et standardavvik på 2,351. Siden det er mye variasjon i antall uker elever deltar i smågruppeundervisning, er det interessant å undersøke om den variasjonen kan benyttes til å se effekter av antall uker smågruppeundervisning på testresultater. Tabell 7 viser sammenheng mellom antall uker smågruppeundervis-

Tabell 7: Sammenheng mellom antall uker smågruppeundervisning og posttest

	Posttest		
	(1)	(2)	(3)
Antall uker	-0.163*** (0.0128)	-0.0227** (0.0102)	-0.0232** (0.0102)
Pretest		0.632*** (0.0101)	0.634*** (0.0102)
Antall uker x Pretest			-0.0241** (0.00999)
Konstant	0.0789*** (0.0126)	0.0940*** (0.00974)	0.0888*** (0.00997)
Observasjoner	5 849	5 812	5 812
R^2	0.027	0.418	0.418

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser MKM-regresjon for elever som mottok minimum 1 uke og maksimum 18 uker smågruppeundervisning med standardiserte verdier på testresultater og antall uker. Standardfeil i parentes.

ning og posttest. Hvis antall uker smågruppeundervisning var tilfeldig bestemt for alle elever i tiltaksskolene, ville resultatene vært kausale sammenhenger. Uten noen kontrollvariabler er det en signifikant negativ sammenheng mellom antall uker en elev deltok i smågruppeundervisning og hvordan eleven presterte på posttesten. Årsaken til dette er at lavt presterende elever deltok mer i smågruppeundervisning enn høyt presterende elever, som diskutert tidligere. Ved å kontrollere for pretesten endres resultatene noe, men jeg finner fortsatt en signifikant negativ sammenheng.

Antall uker en elev deltar i smågruppeundervisning har trolig ikke negativ effekt på elevresultater. Gjennomgående signifikant positiv effekt av smågruppeundervisning som diskutert tidligere i oppgaven bekrefter denne mistanken. Likevel finner jeg i tiltaksskolene en negativ sammenheng mellom hvor mange uker en elev deltar i smågruppeundervisning og elevenes resultater på posttesten. Dette skyldes trolig toveis kausalitet i sammenhengen mellom antall uker en elev deltar i smågruppeundervisning og testskår på posttest. Antall uker en elev deltar i smågruppeundervisning påvirker testresultater, men testresultatene påvirker også antall uker en elev deltar i smågruppeundervisning. Estimert effekt av antall uker smågruppeundervisning på testresultater i Tabell 7 viser dermed også den negative sammenhengen mellom testresultater og antall uker smågruppeundervisning.

Hvis alle faktorer som bidrar til toveis kausalitet fanges opp av resultatene på pretes-

ten, vil en løsning på problemet være å inkludere denne variabelen i estimeringen. Testskår på pretesten inkluderes i kolonne (2). Som det fremgår av kolonne (2) i Tabell 7 reduseres den negative sammenhengen mellom antall uker smågruppeundervisning og testskår på posttesten kraftig ved å kontrollere for elevenes resultater på pretesten. Likevel finner jeg fortsatt en signifikant negativ effekt av antall uker smågruppeundervisning på elevenes resultater på posttesten. Det er antakeligvis gjenværende toveis kausalitet som bidrar til skjevhet i estimerte effekter og som ikke fanges opp av testskår på pretest. Lærerne har mye informasjon om elevene som bidrar til en subjektiv vurdering av hvor mange uker hver elev skal delta i smågruppeundervisning. Som nevnt i kapittel 4.2 skaper dette en utfordring i avdekking av kausale sammenhenger. Disse faktorene som tas med i lærens vurdering er mer omfattende enn det som fanges opp av resultater på pretesten, og forklarer dermed den gjenværende omvendte kausaliteten.

En signifikant negativ effekt av interaksjonsleddet i kolonne (3) i Tabell 7 tilsier at en økning i antall uker smågruppeundervisning har desto mer negativ effekt på posttesten jo bedre eleven presterte på pretesten. Effekten kan imidlertid like gjerne skyldes at den positive sammenhengen mellom posttest og pretest blir svakere jo flere uker eleven deltok i smågruppeundervisning. Denne sammenhengen er i seg selv uinteressant, og skyldes den omvendte kausaliteten diskutert tidligere. Siden elever deltar mer i smågruppeundervisning jo svakere de presterer, er den siste effekten konsistent med tidligere resultater, som viser en svakere sammenheng mellom de to testene for lavt presterende elever enn for høyt presterende elever.

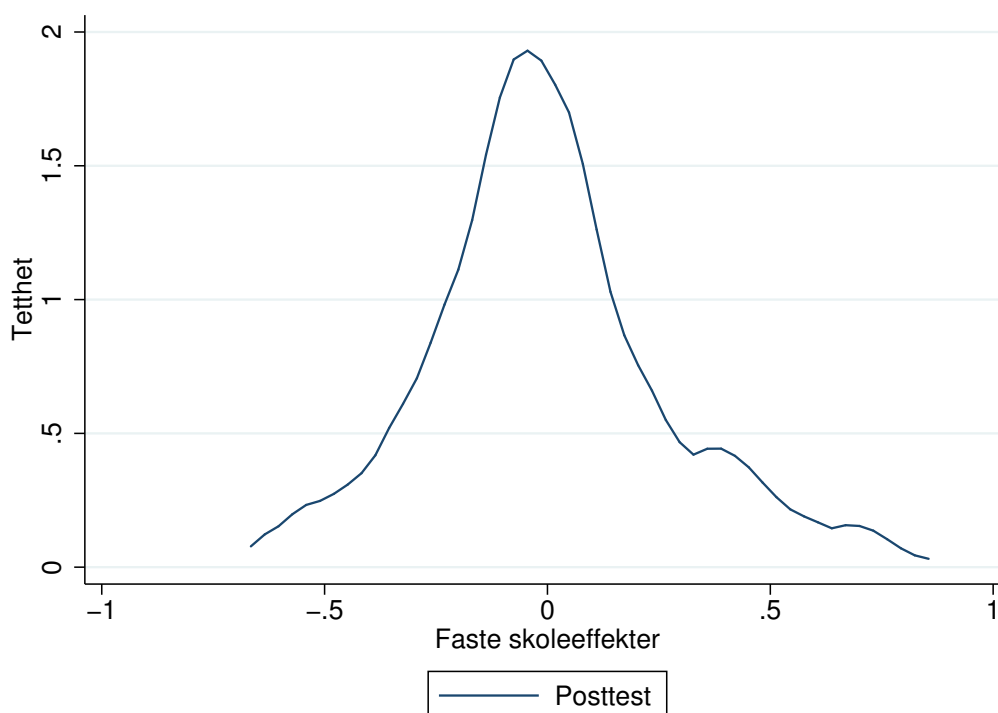
For å avdekke en kausalitet mellom antall uker en elev deltar i smågruppeundervisning og elevens prestasjoner, kreves analyser som går utenfor denne oppgavens problemstilling og som ikke kan gjennomføres på grunn av begrensninger i datamaterialet. En mulig fremgangsmåte for å løse problemet med toveis kausalitet er å randomisere antall uker en elev mottar smågruppeundervisning. Antall uker vil da bidra til eksogen variasjon istedenfor å være endogent bestemt av prestasjonsnivået til eleven.

8.3 Skolespesifikke effekter

For å avdekke ytterligere variasjon mellom tiltaksskolene, i tillegg til ulikt antall uker i smågrupper, benyttes faste skoleeffekter. Disse skolespesifikke effektene fanger opp alle faktorer som kun varierer mellom tiltaksskolene og påvirker elevenes resultat på posttesten. Det kontrolleres også for elevenes resultat på pretesten. De faste effektene estimeres ved hjelp av:

$$A_{1,is} = \alpha + FE_s + \theta A_{0,is} + \xi_{is}, \quad (4)$$

Figur 2: Empirisk fordeling av skolespesifikke effekter



Figuren viser kernel-fordeling av skolespesifikke effekter for 80 tiltaksskoler med båndbredde 0,07. Ett standardavvik er 0,263.

hvor FE er et mål på hvor mye hver tiltaksskole bidrar til å forbedre elevenes testresultater. Ligning (4) estimeres med skolespesifikke effekter. Konstantleddet, α , er gjennomsnittsverdien til de skolespesifikke effektene. I denne delen av analysen benyttes de skolespesifikke effektene til å estimere ulike korrelasjoner mellom hva som foregår på innsiden av tiltaksskolene og skolens evne til å forbedre elevens resultater. På grunn av uobserverte faktorer som kan drive korrelasjonene, er effektene ikke nødvendigvis direkte kausale sammenhenger.

De skolespesifikke effektene estimert i ligning (4) har et standardavvik på 0,263, noe som indikerer stor variasjon i effekten hver tiltaksskole har på testresultater. Tettheten til den empiriske fordelingen av de estimerte skolespesifikke effektene vises i Figur 2. Forskjellen mellom en elev ved en tiltaksskole ett standardavvik under og ett standardavvik over gjennomsnittsskolen er lik 0,526 standardavvik på posttesten. Denne forskjellen er 2,8 ganger større enn den gjennomsnittlige forskjellen mellom en tiltaksskole og en sammenligningsskole, som i Tabell 2 er estimert til å forbedre testskåren med 0,189 standardavvik.

13,2 prosent av variansen i det sammensatte restleddet skyldes varians i skolespesifikke effekter. Dette er beregnet på følgende måte:

$$\rho = \frac{\sigma_{FE}^2}{\sigma_{FE}^2 + \sigma_{\xi}^2}, \quad (5)$$

hvor σ_{FE}^2 er variansen til de skolespesifikke effektene og σ_{ξ}^2 er variansen til et idiosynkratisk restledd som varierer mellom både skoler og elever. ρ forklarer dermed hvor stor andel av den totale variansen i faktorer som påvirker elevenes posttestresultat (og ikke forklares av pretestresultat) som skyldes variasjon mellom skoler.

Hva skyldes denne store forskjellen i hvordan skoler kan påvirke elevenes testresultater? I litteraturen blir det foreslått mange forklaringer. Chetty et al. (2014a) benytter et *value added*-mål for å undersøke lærerens kausale effekt på en elevs testresultater. De finner at lærere med høy *value added* forbedrer elevenes testresultater signifikant. I Chetty et al. (2014b) finner forfatterne videre at dette målet er en god indikator for kvaliteten på læreren, hvor lærerkvalitet blir målt ved blant annet elevens fremtidige inntekt, opp- tak til høyere utdanning og antall barn i tenårene. De konkluderer med at forbedring av lærerkvalitet har store økonomiske og sosiale fordeler.

Roland G. Fryer (2017) nevner flere årsaker til store forskjeller mellom skoler. Mangel på skolevalg og liten konkurranse mellom skoler, forskjeller mellom familier og foreldres inntekt, foreldres ulike grader av engasjement, påvirkning fra medelever og venner, kulturelle forskjeller og ulike strategier fra skoleledelsen, er alle faktorer som varierer mellom skoler og påvirker elevresultater. Selv om smågruppeundervisning har en stor signifikant effekt, er det tydelig at forbedring av testresultater er et mer sammensatt problem som krever ulike politiske virkemidler.

Hvilken tiltaksskole en elev er plassert i spiller tilsynelatende en større rolle enn om eleven er plassert i en sammenligningsskole eller tiltaksskole. For å undersøke hvilke egenskaper ved lærere som forklarer variasjonen i tiltaksskolene benyttes de skolespesifikke effektene fra ligning (4) som et mål på en skoles effektivitet i å forbedre elevens testresultater. Lærerkvalitet nevnes i kapittel 4 som en viktig innsatsfaktor i produktfunksjonen for utdanning. I denne delen av analysen diskuteres lærernes kompetanse målt ved antall års erfaring som matematikklærer og antall studiepoeng i matematikk.

Jeg har separate observasjoner for storgruppelærer og smågruppelærer for hver elev. Storgruppelæreren er den ansvarlige matematikklæreren i klassen og står for den ordinære undervisningen, mens smågruppelæreren ble rekruttert i forbindelse med prosjektet. I konstruksjon av variabler som inkluderer både storgruppelærer og smågruppelærer, for eksempel summert antall studiepoeng, er alle manglende observasjoner ekskludert. Hvis det mangler observasjon for en av lærerne, er begge lærerne ekskludert. Det vil si at disse variablene kun inkluderer lærere som innrapporterte egenskapene til både storgruppelæreren og smågruppelæreren. Dersom det er flere lærere som underviser en av gruppene, er den observerte verdien gjennomsnittet mellom dem.

Tabell 8: Egenskaper ved tiltaksskoler

	Storgruppe		Smågruppe		Totalt	
	Over median	Under median	Over median	Under median	Over median	Under median
Studiepoeng	38.38632 (27.40945)	34.3841 (20.78521)	64.86208 (50.47527)	46.02653 (21.42383)	83.53888 (64.47168)	51.66955 (36.06818)
Observasjoner	2 452	2 308	2 168	1 583	2 810	2 946
Erfaring	13.14825 (8.23877)	12.91729 (8.90906)	10.62385 (8.152433)	12.98705 (9.728943)	23.76941 (9.946095)	30.70674 (12.94884)
Observasjoner	2 421	2 404	2 168	1 583	1 829	1 256

Tabellen viser deskriptiv statistikk for ulike egenskaper hos lærere ved tiltaksskoler. Standardavvik i parentes.

Tabell 8 viser deskriptiv statistikk for antall studiepoeng i matematikk og antall års erfaring som matematikklærer for lærerne. Utvalget deles inn i mer effektive skoler og mindre effektive skoler basert på estimerte verdier av de skolespesifikke effektene. Utvalget deles ved medianverdien for den estimerte påvirkningen hver skole har på elevenes testresultater. Jeg finner at høyt presterende skoler rekrutterer smågruppelærere med et mye høyere antall studiepoeng enn lavt presterende skoler. Gjennomsnittlig antall studiepoeng for smågruppelærere er 64,9 i høyt presterende skoler mot 46,0 for smågruppelærerne i en lavt presterende skole. Det er altså langt mer sannsynlig at en smågruppelærer med høyt antall studiepoeng er ansatt i en høyt presterende skole enn i en lavt presterende skole.

Forskjellen mellom høyt og lavt presterende skoler er mye mindre for storgruppelærere. Storgruppelærerne i en høyt presterende skole har i gjennomsnitt 38,4 studiepoeng, mens storgruppelærere i en lavt presterende skole har 34,4 studiepoeng. Det er også stor forskjell i totalt antall studiepoeng for storgruppelæreren og smågruppelæreren avhengig av om skolen er høyt presterende eller lavt presterende. Gjennomsnittet av summen av studiepoengene til storgruppelæreren og smågruppelæreren er 83,5 i en høyt presterende skole og 51,7 i en lavt presterende skole.

Når det gjelder antall års erfaring er forskjellene mindre. Antall års erfaring for storgruppelærere er nærmest identisk i høyt presterende og lavt presterende skoler. Smågruppelærerne har noe lenger erfaring i de lavt presterende skolene, mens det totalt er flere års matematikkerfaring i lavt presterende skoler enn i høyt presterende skoler. Gjennomsnittet av totalt antall års erfaring for storgruppelæreren og smågruppelæreren er 23,8 i

en høyt presterende skole og 30,7 i en lavt presterende skole.

Den deskriptive statistikken indikerer en sammenheng mellom ulike egenskaper ved lærerne og skolens effektivitet. For å undersøke denne sammenhengen nærmere, presenterer jeg her ulike korrelasjoner mellom egenskaper ved lærere og en skoles effektivitet. Spesifikt estimeres ligninger på formen:

$$FE_s = \textit{konstant} + \phi E_{is} + \zeta_{is}, \quad (6)$$

hvor FE_s er estimert effekt av tiltaksskole s på testresultater og E_{is} er ulike egenskaper ved lærerne. Her benyttes antall studiepoeng i matematikk en lærer har, og antall års erfaring som lærer i matematikk.

Tabell 9 viser separate regresjoner for korrelasjoner mellom studiepoeng hos lærere og skolens effektivitet. I kolonne (1) og (2) er forklaringsvariabelen en indikatorvariabel for om læreren har antall studiepoeng i matematikk over medianen for den respektive lærergruppen, mens indikatorvariabelen i kolonne (3) viser det samme for summen av studiepoengene for smågruppelærerne og storgruppelærerne. Medianen for de tre gruppene var henholdsvis 30, 60 og 90 studiepoeng. I skoler med storgruppelærer som har flere studiepoeng i matematikk enn medianen, har elevene testskår som er 0,0763 standardavvik høyere enn andre skoler. Elever i skoler hvor smågruppelæreren har flere studiepoeng enn medianen, presterer 0,0563 standardavvik bedre enn andre skoler. I skoler hvor summen av begge lærernes studiepoeng er over medianen, presterer elevene 0,0826 standardavvik bedre enn elever andre skoler.

Forklaringsvariablene i kolonne (4) og (5) i Tabell 9 er indikatorvariabler for om læreren oppfyller kravet på 30 studiepoeng for å undervise i matematikk i barneskolen. Blant smågruppelærere er det en så liten andel som ikke oppfyller kravet at det ikke er mulig å måle noen effekter, mens 40 prosent av elevene i storgruppene har en lærer som ikke oppfyller kravet. Årsaken til dette er at kravet bare gjelder for lærere utdannet etter 1. januar 2014. Elever i skoler hvor storgruppelæreren oppfyller kravet, presterer 0,0426 standardavvik bedre enn andre skoler. Kolonne (6) og (7) viser effekt av dummyvariabler for kombinasjoner av antall studiepoeng for storgruppelærer og smågruppelærer. I skoler hvor både smågruppelæreren og storgruppelæreren har antall studiepoeng i matematikk over medianen for den respektive gruppen, presterer elevene 0,0298 standardavvik bedre enn andre skoler, mens forskjellen er langt større når kun smågruppelæreren har antall studiepoeng høyere enn medianen. Her er forskjellen 0,0938 standardavvik.

Det kan være flere årsaker til at de skolene med relativt høyt antall studiepoeng i matematikk hos smågruppelæreren i forhold til storgruppelæreren i stor grad kan påvirke elevens testresultater. En forklaring er at kontrasten mellom ordinær undervisning og smågruppeundervisning ble størst for disse elevene hvis dette i utgangspunktet er skoler

Tabell 9: Korrelasjoner mellom antall studiepoeng i matematikk blant lærere og skolens effektivitet

	Faste skoleeffekter						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Høyt antall studiepoeng storgruppe	0.0763*** (0.00758)						
Høyt antall studiepoeng smågruppe		0.0563*** (0.00980)					
Høyt antall studiepoeng totalt			0.0826*** (0.00936)				
Mattekrav storgruppelærer				0.0426*** (0.00831)			
Mattekrav smågruppelærer					-0.0201 (0.0142)		
Studiepoeng høy/høy						0.0298** (0.0128)	
Studiepoeng lav/høy							0.0938*** (0.0145)
Konstant	-0.0290*** (0.00492)	0.0308*** (0.00495)	0.0270*** (0.00598)	-0.0272*** (0.00700)	0.0632*** (0.0134)	0.0561*** (0.00507)	0.0500*** (0.00492)
Observasjoner	4 760	3 751	3 020	4 760	3 751	3 020	3 020
R^2	0.021	0.009	0.025	0.005	0.001	0.002	0.014

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser separate MKM-regresjoner for effekt av læreres studiepoeng i matematikk på et mål for en skoles effektivitet. Standardfeil i parentes.

med lavt presterende elever hvor lærerne har et lavt antall studiepoeng. Selv om testskår på pretesten er kontrollert for, kan det være andre egenskaper ved skolene som er korrelert med lærernes studiepoeng og skolens evne til å forbedre elevenes testresultater, for eksempel skoleledelsens evne til å rekruttere gode lærere. Hvis skolens evne til å rekruttere gode smågruppelærere er positivt korrelert med smågruppelærers studiepoeng og skolens evne til å påvirke elevers testresultat, vil estimert effekt i kolonne (7) være noe overestimert.

Det er ingen forskjell i gjennomsnittlig testskår på pretesten mellom skoler som hadde lærere med et høyt antall studiepoeng i matematikk og lærere med et lavt antall studiepoeng i matematikk. Likevel er det en forskjell i gjennomsnittlig testskår mellom de skolene som rekrutterte en lærer med høyt antall studiepoeng til smågruppeundervisning og skolene som rekrutterte en lærer med lavt antall studiepoeng. Tilsynelatende klarte høyt presterende skoler å rekruttere smågruppelærere med høyere antall studiepoeng i matematikk. Dette kan være med på å forklare den forholdsvis høye estimerte effekten i kolonne (7). Effekten av antall studiepoeng hos smågruppelæreren kommer tydeligst frem i høyt presterende skoler som rekrutterer en smågruppelærer med høyt antall studiepoeng, hvor storgruppelæreren har et lavt antall studiepoeng. Dette fordi den positive effekten av et høyt antall studiepoeng hos læreren ikke er hentet ut i disse høyt presterende skolene før det ansettes en smågruppelærer med høy kompetanse i matematikk.

Tabell 10 viser separate regresjoner for korrelasjoner mellom lærernes erfaring og skolens effektivitet. I kolonne (1)-(3) er indikatorvariablene konstruert på samme måte som i Tabell 9, hvor variabelen tar verdien én hvis læreren har erfaring høyere enn medianen i sin respektive lærerkategori. Medianerfaring for de tre ulike kategoriene er henholdsvis 11,5, 10 og 26. Sammenlignet med antall studiepoeng spiller lærererfaring en mindre rolle i skolens evne til å påvirke elevenes testresultat, og forklarer en mindre del av variasjonen. Likevel finner jeg flere betydelige og signifikante effekter.

I skoler med storgruppelærer som har flere års erfaring som matematikklærer enn medianen, har elevene testskår som er 0,0477 standardavvik høyere enn andre skoler. Elever i skoler hvor smågruppelæreren har flere studiepoeng enn medianen, presterer 0,0739 standardavvik dårligere enn andre skoler. Elever i skoler hvor summen av begge lærernes erfaring er over medianen, presterer 0,106 standardavvik dårligere enn andre skoler.

Effekten av erfaring i læreryrket antas å være avtakende jo flere års erfaring læreren har. I litteraturen nevnes tre år som en grense hvor mye av effekten som er uttømt (se for eksempel Nye et al. (2004)). I kolonne (4) tar forklaringsvariabelen verdien én hvis storgruppelæreren har mer enn tre års erfaring. Jeg finner at nesten halvparten av effekten av å være en storgruppelærer med antall års erfaring høyere enn medianen også gjelder for lærere som har over tre års erfaring. Effektene av erfaring for smågruppelærere er gjennomgående signifikant negativ på skolens evne til å påvirke elevenes testresultater,

Tabell 10: Korrelasjoner mellom lærers erfaring og skolens effektivitet

	Faste skoleeffekter					
	(1)	(2)	(3)	(4)	(5)	(6)
Lang erfaring storgruppe	0.0477*** (0.00730)					
Lang erfaring smågruppe		-0.0739*** (0.00851)				
Lang erfaring totalt			-0.106*** (0.00866)			
Erfaring > 3 år storgruppe				0.0211** (0.0103)		
Erfaring > 3 år smågruppe					-0.0429*** (0.0103)	
Erfaring > 3 år gjennomsnitt						0.241*** (0.0301)
Konstant	-0.0199*** (0.00513)	0.0803*** (0.00587)	0.111*** (0.00596)	-0.0144 (0.00952)	0.0786*** (0.00914)	-0.175*** (0.0297)
Observasjoner	4 825	3 751	3 085	4 825	3 751	3 085
R^2	0.009	0.020	0.046	0.001	0.005	0.020

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser separate MKM-regresjoner for effekt av en lærers erfaring på et mål for skolens effektivitet. Standardfeil i parentes.

Tabell 11: Partielle effekter av egenskaper ved lærere på skolenes effektivitet

	(1)	(2)	(3)
	FE	Studiepoeng smågruppe	Studiepoeng smågruppe
Høyt antall studiepoeng storgruppe	0.0491*** (0.00898)		
Høyt antall studiepoeng smågruppe	0.0726*** (0.0102)		
Lang erfaring storgruppe	-0.0149 (0.00912)		
Lang erfaring smågruppe	-0.0873*** (0.00906)		
Studiepoeng storgruppe		0.181*** (0.0273)	
Erfaring smågruppe			-1.503*** (0.0697)
Konstant	0.0739*** (0.00985)	52.79*** (1.315)	74.28*** (1.021)
Observasjoner	2 956	3 301	4 100
R^2	0.068	0.013	0.102

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser effekter av ulike indikatorvariabler på skolens effektivitet. I tillegg vises sammenheng mellom antall studiepoeng for storgruppelærer og smågruppelærer. Tabellen viser også sammenheng mellom antall års erfaring til smågruppelærer og antall studiepoeng for smågruppelærer. Alle modeller er estimert med MKM. Standardfeil i parentes.

hvilket indikerer at lærererfaring ikke alltid er et godt mål på lærerkvalitet.

Resultatene i kolonne (6) må tolkes med varsomhet da det kun var 82 elever som hadde storgruppelærer og smågruppelærer med under tre års erfaring i gjennomsnitt. Det er fordi i nesten alle tilfeller hvor én lærer hadde lite erfaring hadde den andre læreren mange års erfaring og trakk opp gjennomsnittet.

Tabell 11 viser effekten ulike egenskaper hos lærere har på skolens evne til å påvirke elevens testresultater, kontrollert for de andre variablene. Jeg finner også her signifikante positive effekter av antall studiepoeng hos både smågruppelærer og storgruppelærer. Effekten er imidlertid her størst for smågruppelærerne. Kontrasten fra Tabell 9 skyldes trolig en høy korrelasjon mellom studiepoengene til smågruppelærere og storgruppelærere. Denne korrelasjonen undersøkes i kolonne (2). For hvert ekstra studiepoeng for en storgruppelærer vil smågruppelæreren ha 0,181 ekstra studiepoeng. Som nevnt tidligere

rekrutterte skolene med høyest presterende elever smågruppelærere med høy kompetanse. Disse skolene hadde også lærere med høyere kompetanse i utgangspunktet. Derfor inkluderer den estimerte effekten av antall studiepoeng for storgruppelæreren i Tabell 9 også effekten av studiepoeng hos smågruppelæreren, og tallet blir noe inflatert. Dette fører til en lavere estimert effekt når det kontrolleres for studiepoengene til smågruppelæreren.

I tillegg til positiv korrelasjon mellom studiepoengene til smågruppelærer og storgruppelærer, er det også høy negativ korrelasjon mellom antall studiepoeng og antall års erfaring, særlig for smågruppelærere. Denne korrelasjonen vises i kolonne (3) i Tabell 11. For hvert ekstra års erfaring som matematikklærer for en smågruppelærer vil læreren ha 1,503 færre studiepoeng i matematikk. Dette demper estimert effekt av å ha en smågruppelærer med antall studiepoeng over medianen i Tabell 9. Den estimerte partielle effekten blir derfor langt høyere når det kontrolleres for erfaring. Jeg finner at elever i skoler hvor storgruppelæreren har flere studiepoeng enn medianen, presterer 0,0491 standardavvik bedre enn elever i andre skoler, mens forskjellen er 0,0726 standardavvik for smågruppelærere.

Generelt virker antall studiepoeng som et godt mål på skolens evne til å påvirke elevenes testresultater, særlig for smågruppelæreren. Dette understreker viktigheten av diskusjonen i kapittel 4.1. Kvaliteten på læreren som ansettes til det formål å lede smågruppeundervisning i matematikk er avgjørende for effekten av smågruppeundervisningen. Her nevnes antall studiepoeng som en indikator for lærerkvalitet, og andre egenskaper ved læreren vil påvirke utfallet i ytterligere.

9 Konklusjon

Denne oppgaven tar utgangspunkt i en randomisert kontrollert studie av smågruppeundervisning utført i andre og tredje klasse i ti store norske kommuner. Datamaterialet inkluderer informasjon fra lærere om gjennomføring av prosjektet og elevenes testresultater i tillegg til karakteristikker ved lærere. Formålet med oppgaven er å undersøke om effekten av smågruppeundervisning er ulik for lavt presterende og høyt presterende elever. Videre diskuteres resultatene i lys av ulike korrelasjoner innad i skolene som deltok i smågruppeundervisning.

Den første delen av analysen tar for seg den kausale effekten av smågruppeundervisning på elevresultater. Jeg finner en signifikant positiv effekt av smågruppeundervisning, og randomiseringen sørger for at resultatene er robuste. Elever ved tiltaksskolene, altså de elevene som deltok i smågruppeundervisning, forbedret i gjennomsnitt matematikkresultatene sine med 0,184 standardavvik mer enn elever ved sammenligningsskolene.

Jeg finner en signifikant ulik effekt av smågruppeundervisning avhengig av prestasjonsnivået til en elev. Det er særlig middels presterende elever som skiller seg ut i prosjektet ved at de har den klart høyeste effekten av smågruppeundervisning. Disse elevene forbedrer testresultatene med 0,233 standardavvik relativt til gjennomsnittseleven i en sammenligningsskole. I tillegg antydes det at lavt presterende elever har noe høyere utbytte av smågruppeundervisning enn høyt presterende elever.

Til tross for endogenitetsproblemer som i stor grad skyldes læreres subjektive vurderinger, konkluderer diskusjonen i kapittel 8.1 med at resultatene i kapittel 7 om heterogene programeffekter er robuste. Ytterligere analyser i kapittel 8.3 avdekker en høy variasjon i skolers evne til å påvirke testresultater. Forskjellen i testresultater mellom en skole ett standardavvik under og over gjennomsnittskolen estimeres til å være 2,8 ganger større enn den gjennomsnittlige forskjellen mellom en tiltaksskole og en sammenligningsskole. Dette indikerer at det spiller langt større rolle hvilken tiltaksskole en elev er plassert i enn om eleven deltar i smågruppeundervisning eller ikke.

Det er likevel viktig å huske på at smågruppeundervisning er et presist tiltak med en klar positiv effekt som er forholdsvis enkel å implementere i mange kommuner. Generell variasjon i skolens evne til å påvirke testresultater er mer sammensatt, og det kreves ytterligere analyser for å avdekke årsaker til variasjonen. I denne oppgaven forklares noe av variasjonen med læreres studiepoeng i matematikk og erfaring som matematikklærer. Jeg finner gjennomgående signifikant positiv korrelasjon mellom lærernes studiepoeng og skolens evne til å påvirke elevers testresultat.

Faglitteraturen på dette området er liten, og analysene i denne oppgaven er ikke tidligere blitt utført med dette datamaterialet. Derfor kommer denne oppgaven med flere viktige bidrag til debatten. Fremtidige forsøk vil særlig ha nytte av å randomisere an-

tall uker smågruppeundervisning for å avdekke om mengden smågruppeundervisning har positiv signifikant effekt på elevresultater, og eventuelt om effekten er avtakende. Dette vil blant annet være aktuelt for kostnad-nytte-analyser for å avgjøre hvor mange uker smågruppeundervisning som er samfunnsøkonomisk lønnsomt før effekten er uttømt.

Kan resultatene i denne analysen generaliseres til flere norske kommuner? Et problem er mangel på kvalifiserte lærere. Dette prosjektet ble gjennomført i store kommuner med den begrunnelse at det skulle være god tilgang på kvalifiserte lærere med matematikkbakgrunn. Det er ikke sikkert at tilbudet av kvalifiserte lærere er like stort i alle kommuner.

Referanser

- Andersen, S., Beuchert, L., Nielsen, H. & Kjærgaard Thomsen, M. (2018, 01). The effect of teacher's aides in the classroom: Evidence from a randomized trial. *SSRN Electronic Journal*.
- Angrist, J.D. & Lavy, V. (1999). Using maimonides' rule to estimate the effect of class size on scholastic achievement. *The Quarterly Journal of Economics*, 114(2), 533–575.
- Angrist, J.D., Pathak, P.A. & Walters, C.R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, 5(4), 1–27.
- Chetty, R., Friedman, J.N. & Rockoff, J.E. (2014a, September). Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, 104(9), 2593-2632. Hentet fra <http://www.aeaweb.org/articles?id=10.1257/aer.104.9.2593> doi: 10.1257/aer.104.9.2593
- Chetty, R., Friedman, J.N. & Rockoff, J.E. (2014b, September). Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood. *American Economic Review*, 104(9), 2633-79. Hentet fra <http://www.aeaweb.org/articles?id=10.1257/aer.104.9.2633> doi: 10.1257/aer.104.9.2633
- Chris, R. (2017). Measurement of peer effects. *Australian Economic Review*, 50(1), 121-129. Hentet fra <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-8462.12213> doi: 10.1111/1467-8462.12213
- Cook, P.J., Dodge, K., Farkas, G., Fryer Jr, R.G., Guryan, J., Ludwig, J., . . . Steinberg, L. (2015). Not too late: Improving academic outcomes for disadvantaged youth. *Institute for Policy Research Northwestern University Working Paper WP-15-01*.
- Dobbie, W. & Fryer Jr, R.G. (2013). Getting beneath the veil of effective schools: Evidence from new york city. *American Economic Journal: Applied Economics*, 5(4), 28–60.
- Finn, J.D., Gerber, S.B., Achilles, C.M. & Boyd-Zaharias, J. (2001). The enduring effects of small classes. *Teachers College Record*, 103(2), 145–183.
- Finn, J.D., Pannozzo, G.M. & Achilles, C.M. (2003). The “why’s” of class size: Student behavior in small classes. *Review of Educational Research*, 73(3), 321–368.
- Fryer Jr, R.G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics*, 129(3), 1355–1407.

- Grønmo, L.S., Onstad, T., Nilsen, T., Hole, A., Aslaksen, H. & Borge, I.C. (2011). Framgang, men langt fram. *Norske elevers prestasjoner i matematikk og naturfag i TIMSS*.
- Hanushek, E.A. (2002, July). *The long run importance of school quality* (Working Paper nr. 9071). National Bureau of Economic Research. Hentet fra <http://www.nber.org/papers/w9071> doi: 10.3386/w9071
- Hanushek, E.A. & Woessmann, L. (2012). Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of economic growth*, 17(4), 267–321.
- Kjærnsli, M. & Jensen, F. (2016). *Stø kurs. norske elevers kompetanse i naturfag, matematikk og lesing i pisa 2015*. Universitetsforlaget.
- Kjærnsli, M. & Olsen, R.V. (2013). *Fortsatt en vei å gå. norske elevers kompetanse i matematikk, naturfag og lesing i pisa 2012*. Oslo: Universitetsforlaget.
- Konstantopoulos, S. (2008). Do small classes reduce the achievement gap between low and high achievers? evidence from project star. *The Elementary School Journal*, 108(4), 275–291.
- Krueger, A.B. (1999). Experimental estimates of education production functions. *The quarterly journal of economics*, 114(2), 497–532.
- Krueger, A.B. & Whitmore, D.M. (2001). The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star. *The Economic Journal*, 111(468), 1–28.
- Kunnskapsdepartementet. (2015). *Tett på realfag: Nasjonal strategi for realfag i barnehagen*.
- Kunnskapsdepartementet. (2017a). *Høringsnotat. forslag til endringer i opplæringsloven og friskoleloven (plikt til å tilby intensiv opplæring og plikt til flerfaglig samarbeid)*.
- Kunnskapsdepartementet. (2017b). *Lærelyst – tidlig innsats og kvalitet i skolen*. Oslo: Kunnskapsdepartementet. (Meld. St. 16 2016-2017)
- Nye, B., Hedges, L.V. & Konstantopoulos, S. (2002). Do low-achieving students benefit more from small classes? evidence from the tennessee class size experiment. *Educational Evaluation and Policy Analysis*, 24(3), 201–217.
- Nye, B., Konstantopoulos, S. & Hedges, L.V. (2004). How large are teacher effects? *Educational evaluation and policy analysis*, 26(3), 237–257.

- Roland G. Fryer, J. (2017, May). *Management and student achievement: Evidence from a randomized field experiment* (Working Paper nr. 23437). National Bureau of Economic Research.
- Steffensen, K., Ekren, R., Zachrisen, O.O. & Kirkebøen, L.J. (2017). Er det forskjeller i skolers og kommuners bidrag til elevenes læring i grunnskolen? en kvantitativ studie. *Statistisk sentralbyrå*.
- Todd, P.E. & Wolpin, K.I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113(485).
- Utdanningsdirektoratet. (2016). *Utdanningsspeilet 2016*.
- Webster, R., Blatchford, P. & Russell, A. (2013). Challenging and changing how schools use teaching assistants: findings from the effective deployment of teaching assistants project. *School leadership & management*, 33(1), 78–96.
- Weiss, M.J., Bloom, H.S. & Brock, T. (2014). A conceptual framework for studying the sources of variation in program effects. *Journal of Policy Analysis and Management*, 33(3), 778–808.

Appendiks A: Fordeling i kvintiler

Tabell A1: Fordeling av elevene i kvintiler

Pretestskår	Kvintiler					Total
	1	2	3	4	5	
0	32.0	0.0	0.0	0.0	0.0	32.0
1	21.0	0.0	0.0	0.0	0.0	21.0
2	18.0	0.0	0.0	0.0	0.0	18.0
3	35.0	0.0	0.0	0.0	0.0	35.0
4	28.0	0.0	0.0	0.0	0.0	28.0
5	56.0	0.0	0.0	0.0	0.0	56.0
6	75.0	0.0	0.0	0.0	0.0	75.0
7	99.0	0.0	0.0	0.0	0.0	99.0
8	105.0	0.0	0.0	0.0	0.0	105.0
9	142.0	0.0	0.0	0.0	0.0	142.0
10	128.0	0.0	0.0	0.0	0.0	128.0
11	181.0	0.0	0.0	0.0	0.0	181.0
12	252.0	0.0	0.0	0.0	0.0	252.0
13	252.0	0.0	0.0	0.0	0.0	252.0
14	303.0	0.0	0.0	0.0	0.0	303.0
15	338.0	0.0	0.0	0.0	0.0	338.0
16	361.0	0.0	0.0	0.0	0.0	361.0
17	399.0	10.0	0.0	0.0	0.0	409.0
18	194.0	231.0	0.0	0.0	0.0	425.0
19	192.0	230.0	0.0	0.0	0.0	422.0
20	218.0	255.0	0.0	0.0	0.0	473.0
21	0.0	484.0	0.0	0.0	0.0	484.0
22	0.0	548.0	0.0	0.0	0.0	548.0
23	0.0	549.0	0.0	0.0	0.0	549.0
24	0.0	551.0	8.0	0.0	0.0	559.0
25	0.0	302.0	282.0	0.0	0.0	584.0
26	0.0	0.0	646.0	0.0	0.0	646.0
27	0.0	0.0	713.0	0.0	0.0	713.0
28	0.0	0.0	681.0	0.0	0.0	681.0
29	0.0	0.0	652.0	10.0	0.0	662.0
30	0.0	0.0	359.0	373.0	0.0	732.0

31	0.0	0.0	0.0	706.0	0.0	706.0
32	0.0	0.0	0.0	694.0	0.0	694.0
33	0.0	0.0	0.0	693.0	17.0	710.0
34	0.0	0.0	0.0	405.0	307.0	712.0
35	0.0	0.0	0.0	363.0	315.0	678.0
36	0.0	0.0	0.0	0.0	629.0	629.0
37	0.0	0.0	0.0	0.0	545.0	545.0
38	0.0	0.0	0.0	0.0	451.0	451.0
39	0.0	0.0	0.0	0.0	356.0	356.0
40	0.0	0.0	0.0	0.0	244.0	244.0
41	0.0	0.0	0.0	0.0	99.0	99.0
42	0.0	0.0	0.0	0.0	13.0	13.0
Total	3 429.0	3 160.0	3 341.0	3 244.0	2 976.0	16 150.0

Tabellen viser testresultatene fra pretest fordelt i kvintiler. På grunn av ulik poengskala på hvert trinn er det overlapp mellom noen resultater.

Appendiks B: Utvidelse av heterogene programeffekter

Som en ytterligere robusthetssjekk av resultatene i kapittel 7 estimeres her partielle effekter av smågruppeundervisning på hvert av kvintilene. Modellspefikasjonen er gitt på følgende form:

$$\begin{aligned} A_{1, is} = & a + \lambda_1 Q2 + \lambda_2 Q3 + \lambda_3 Q4 + \lambda_4 Q5 + \gamma Tiltaksskole \\ & + d_1 Q2 Tiltaksskole + d_2 Q3 Tiltaksskole \\ & + d_3 Q4 Tiltaksskole + d_4 Q5 Tiltaksskole + \omega_{is}, \end{aligned} \tag{7}$$

hvor $Q2$ er indikatorvariabel for om eleven er i andre kvintil, $Q3$ er indikatorvariabel for om eleven er i tredje kvintil, $Q4$ er indikatorvariabel for om eleven er i fjerde kvintil og $Q5$ er indikatorvariabel for om eleven er i femte kvintil. Første kvintil blir da referansekategori. Resultatene rapporteres i Tabell B1.

Tabell B1: Effekt av smågruppeundervisning for kvintiler med interaksjonsledd

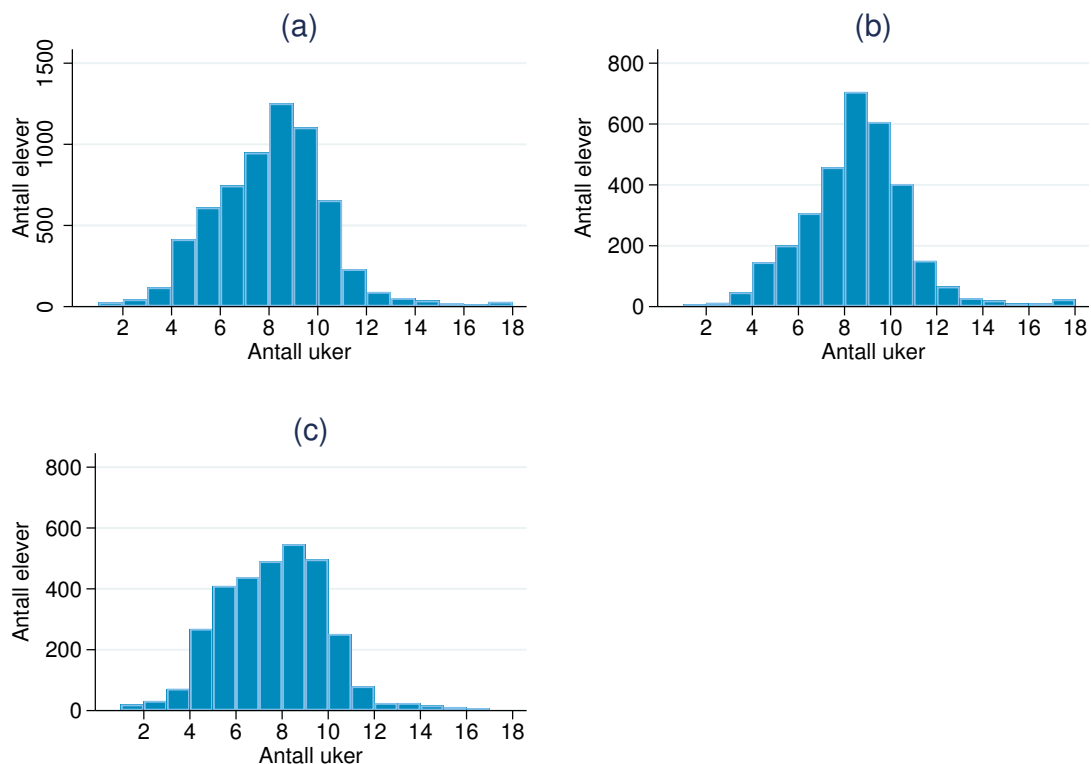
	(1) Posttest
2. kvintil	0.570*** (0.0346)
3. kvintil	0.965*** (0.0340)
4. kvintil	1.376*** (0.0333)
5. kvintil	1.871*** (0.0339)
Tiltaksskole	0.173*** (0.0315)
2. kvintil x Tiltaksskole	0.0238 (0.0446)
3. kvintil x Tiltaksskole	0.0589 (0.0437)
4. kvintil x Tiltaksskole	0.0246 (0.0437)
5. kvintil x Tiltaksskole	-0.0639 (0.0447)
Konstant	-1.046*** (0.0248)
Observasjoner	12 643
R^2	0.401

***= $p < 0,01$, **= $p < 0,05$, *= $p < 0,1$. Tabellen viser regresjon estimert med MKM. Standardfeil i parentes.

Resultatene i Tabell B1 viser at ingen av interaksjonsleddene mellom indikator for kvintil og indikator for tiltaskole er signifikante. Modellen er trolig for rigid til å avdekke noen effekter.

Appendiks C: Fordeling av antall uker smågruppeundervisning

Figur C1: Fordeling av antall uker smågruppeundervisning



Panel (a) viser fordeling for hele utvalget, panel (b) viser fordeling for elever som presterte under medianen, og panel (c) viser fordeling for elever som presterte over medianen

Figur C1 viser fordeling for hvor mange uker ulike grupper av elever deltok i smågruppeundervisning. Figuren viser en viss grad av variasjon i antall uker for alle grupper, og høyt presterende elever deltok gjennomgående mindre i smågruppeundervisning enn lavt presterende elever.