

Sigurd Pedersen Bakken

# Peer- og rangeringseffekter i norsk barneskole

Lineære- og heterogene effekter fra 1+1  
prosjektet

Masteroppgave i Samfunnsøkonomi

Veileder: Colin Green

Mai 2023



Sigurd Pedersen Bakken

# **Peer- og rangeringseffekter i norsk barneskole**

Lineære- og heterogene effekter fra 1+1 prosjektet

Masteroppgave i Samfunnsøkonomi  
Veileder: Colin Green  
Mai 2023

Norges teknisk-naturvitenskapelige universitet  
Fakultet for økonomi  
Institutt for samfunnsøkonomi



Kunnskap for en bedre verden



## **Forord**

Jeg ønsker å takke min veileder Colin Green for veldig god veiledning og gode faglige diskusjoner underveis i arbeidet. Videre ønsker jeg å takke Ole Henning Nyhus ved Senter for økonomisk forskning, for å ha introdusert meg til 1+1 prosjektet, og for teknisk støtte. Datamateriell brukt i denne oppgaven er fra 1+1 prosjektet. Prosjektet gjennomføres av NIFU, Institutt for samfunnsforskning og Senter for økonomisk forskning, avdeling ved NTNU Samfunnsforskning (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å.; NTNU Samfunnsforskning, u.å.). Prosjekt har finansiell støtte fra Norges forskningsråd, med prosjektnummer 256217 (Forskningsrådet, u.å.).

Trondheim, 28. mai 2023.

Sigurd Pedersen Bakken



## Sammendrag

I forbindelse med 1+1 prosjektet, har elever i norsk barneskole gjennomført smågruppeundervisning i matematikk (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å; NTNU Samfunnsforskning, u.å.). Med utgangspunkt i dette tiltaket, analyserer denne oppgaven peer- og rangeringseffekter i samspill med hverandre. Inspirert av resultatene til Bertoni og Nisticò (2023), beviser jeg hvor essensielt det er å inkludere rangeringseffekter i estimeringen av peer-effektens påvirkning på elevprestasjoner, i et tilfelle der begge variablene baseres på akademiske resultater. Ved å estimere disse effektene lineært, rapporteres en sterkere peer-effekt når rangeringseffekten inkluderes i estimeringen, sammenlignet med en situasjon der rangeringseffekten ekskluderes fra analysen. Peer-effekten som estimeres lineært, har en positiv og signifikant effekt på elevprestasjoner. Den tilknyttede lineære rangeringseffekten som estimeres er noe svakere.

Videre estimeres heterogene effekter av disse utdanningsvariablene, og resultatene viser både hvordan peer- og rangeringseffekter har påvirkning på elevers prestasjoner i grupper med forskjellig karakteristika, og hvordan disse effektene varierer mellom elever av ulikt ferdighetsnivå. Ferdighetsnivåene blir kategorisert ut i fra akademiske resultater. I grupper med lite innad spredning i akademiske resultater, estimeres ingen peer- eller rangeringseffekter som er både positive og signifikant. Positive peer- og rangeringseffekter estimeres for to elevgrupper i grupper med tilstrekkelig innad spredning i akademiske resultater, der elevgruppene defineres etter ferdighetsnivå. Med tanke på å skape gode utdanningsforhold, er dette viktige resultater som er med på å indikere hvordan smågruppeundervisning kan optimaliseres.

## Abstract

Related to the 1+1 project, Norwegian students in elementary school have participated in small-group instruction (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å.; NTNU Samfunnsforskning, u.å.). This paper reports peer- and rank-effects from this intervention. Inspired by the results from Bertoni og Nisticò (2023), I show that including rank-effects in the estimation of how student performance gets influenced by peer-effects is essential, when both effects are determined by academic results. Linear results from this study shows that the estimated peer-effect is of greater magnitude when the rank-effect is included in the estimation, compared to when the rank-effect is excluded from the estimation. By looking at how these effects influence student performance, the results show a positive and significant peer-effect from a linear estimation. The following linear rank-effects is smaller.

Going further, I estimate heterogenous effects from these variables, and by looking at different group-characteristics and different student academic performance level, I show how peer-and rank-effects varies between these different groups and different students. For groups containing students with small differences in academic performance, I find no estimates of peer-and rank-effects that are both positive and significant. When the difference in student performance gets larger inside the groups, I find positive peer-and rank-effects for two student performance levels. This results are important for optimalization of small-group instruction, which could improve conditions related to education.



# Innholdsfortegnelse

<b>1</b>	<b>Introduksjon</b> .....	<b>1</b>
<b>2</b>	<b>Estimeringsutfordringer og tidligere litteratur</b> .....	<b>5</b>
2.1	Estimeringsutfordringer tilknyttet peer-effekter .....	5
2.1.1	Inndeling av peer-sammensetninger.....	5
2.1.2	Identifisering.....	8
2.2	Rangeringseffekter .....	8
2.3	Gruppekarakteristika .....	9
<b>3</b>	<b>Data, institusjonell setting og deskriptiv statistikk</b> .....	<b>12</b>
3.1	1+1 prosjektet.....	12
3.2	Deskriptiv statistikk og forklaring av variabler.....	12
3.3	Nødvendige justeringer av datasett .....	19
<b>4</b>	<b>Empirisk metode</b> .....	<b>21</b>
4.1	Identifisering av peer-effekter .....	21
4.2	Inkludering av rangeringseffekter .....	23
<b>5</b>	<b>Resultater</b> .....	<b>25</b>
5.1	Lineære peer- og rangeringseffekter .....	25
5.2	Heterogene peer- og rangeringseffekter .....	32
<b>6</b>	<b>Diskusjon</b> .....	<b>40</b>
<b>7</b>	<b>Konklusjon</b> .....	<b>46</b>
<b>8</b>	<b>Referanseliste</b> .....	<b>48</b>
<b>9</b>	<b>Appendix</b> .....	<b>54</b>

# 1 Introduksjon

Mennesker påvirkes av hverandre, i hverdagslivet, sosialt og i jobb- og utdanningsammenheng. Utdanning står svært sentralt for god samfunnsutvikling, og FN's bærekraftsmål nr. 4 handler om, i et bredt perspektiv, å skape gode utdanningsmuligheter (FN-sambandet, 2022). Forståelse av hvordan peer-effekter påvirker elevprestasjoner og utdanningsforhold er derfor et høyst relevant tema. Elever på samme skole møter hverandre ofte, både i undervisning, friminutt og fritid. Spørsmålene er derfor flere om hvorfor forskning ikke kan påvise sterkere påvirkningskraft av peer-effekter på elevprestasjoner. Videre viser nyere forskning at rangering påvirker elever på lengere sikt, som senere elevprestasjoner (Denning et al., 2021; Murphy & Weinhardt, 2020), utdanningsvalg (Denning et al., 2021; Elsner et al., 2021) og lønn (Denning et al., 2021). I denne analysen bestemmes elevenes rangering ut i fra elevenes akademiske resultater, som er vanligst i litteraturen (Delaney & Devereux, 2022, s. 3). Effektene fra elevens rangering omtales i den internasjonale litteraturen som «rank-effects», og i denne analysen omtaler jeg det som «rangeringseffekter». Bertoni og Nisticò (2023) viser at ved å estimere peer-kvalitet uten å kartlegge studentens rangering i studentens peer, vil effekten av peer-kvalitet gi en svakere sammenheng på studentprestasjoner. Dette er grunnet to motstridene effekter, da både peer-kvalitet og rangering antas å ha positiv påvirkning på elevprestasjoner, men de er negativt avhengig av hverandre. Med utgangspunkt i dette, presenterer denne oppgaven hvordan peer- og rangeringseffekter påvirker akademiske prestasjoner til elever i norsk barneskole.

På hvilken måte påvirker elevens rangering elevens prestasjoner i skolesammenheng? Marsh (1987) og Marsh og Parker (1984) argumenterer at med utgangspunkt i hvordan elever oppfatter sine egne ferdigheter i en utdanningsammenheng, vil det å interagere med elever av lavere ferdighetsnivå skape en positiv effekt og motsatt. Basert på dette, foreligger det en rimelig forventning av at høyere rangering gir positive effekter på elevprestasjoner gjennom økt selvbilde/selvkonsept, og en positiv sammenheng mellom selvtillit og elevprestasjoner kalkuleres i Murphy og Weinhardt (2020, Tabell 7). Flere faktorer kan medføre at rangering påvirker elevprestasjoner, og det kan være naturlig å tenke seg at elever som presterer bedre enn sine gruppe-medlemmer kan motta positive effekter av at man lærer bort egen kunnskap til andre

(Granger, u.å.; Paul, 2011; Shatz, u.å.). I første del av denne analysen undersøker jeg omfanget av skjevheten i estimeringen av peer- og rangeringseffekter hvis man utelater den ene effekten ved estimeringen av den andre.

Videre har flere estimeringsmetoder blitt benyttet for å estimere peer-effekter. For å undersøke hvordan peer-effekter differerer med og uten inkludering av rangeringseffekter i en estimering, følger jeg fremgangsmåten til Bertoni og Nisticò (2023), ved å bruke en lineær gjennomsnittsmoell. Ved å inkludere rangeringseffekter i estimeringen av peer-effekter, presenterer jeg signifikante resultater av peer-kvalitet på barneskolenivå, samtidig som rangeringseffekter i seg selv estimeres. Skjevheten foreligger også ved estimering av rangeringseffekter, uten inkludering av peer-effekter i estimeringen. Ved å ha estimert peer- og rangeringseffekter lineært, gir dette et videre utgangspunkt for å undersøke i hvilke gruppesammensettinger med ulike elevkarakteristika disse effektene eventuelt er sterke. I andre del av analysen undersøker jeg derfor heterogene effekter, der fokusområdet er på peer- og rangeringseffekter for ulike gruppesammensettinger, og for elever av ulikt ferdighetsnivå. Disse resultatene gir utgangspunktet for to viktige prinsipper. Ved å se på hvilke gruppesammensettinger som gir størst læringsutbytte, gir resultatene mine et utgangspunkt for å si noe om elever, i en smågruppeundervisningssituasjon, bør interagere med andre elever på samme ferdighetsnivå, eller med andre elever på ulikt ferdighetsnivå.

For det andre, er dette effekter som estimeres ikke-lineært. For å iverksette tiltak som er pareto-effektiv, må man i dette tilfelle, kunne sette sammen grupper på en måte som bedrer resultater til alle elever, eller ved å bedre resultatene til noen elever, uten at dette gir negativ effekt på resultatene til resterende elever (Hoxby & Weingarh 2005; Sacerdote, 2014; Vigdor & Ludwig 2010, referert i Paloyo (2020, s. 300-301)).

Ved å ta utgangspunkt i analysene til Bertoni og Nisticò (2023), presiserer jeg en todelt problemstilling:

-Ved å estimere peer- og rangeringseffekter lineært, undersøker jeg graden av skjevhet ved å utelatte rangeringseffekter i estimeringen av peer-effekter. Utelatelse av peer-effekter i estimering av rangeringseffekter undersøkes også.

-Ved å estimere peer- og rangeringseffekter heterogent, undersøker jeg hvordan disse effektene varierer mellom ulike grupper som skiller seg fra hverandre basert på differanse i ferdighetsnivå innad i gruppene, og hvordan effekten varierer mellom ulike ferdighetsnivå blant elevene. Deretter undersøker jeg hvordan peer- og rangeringseffekter varierer mellom elever av ulikt ferdighetsnivå mellom grupper med ulik spredning i ferdighetsnivå.

Estimering av peer-effekter i tidligere forskning har gitt noe blandede resultater i form av om det eksisterer noen effekt, og eventuelt graden av effekten, som også kan variere ut ifra hvor i utdanningsløpet elevene er (Paloyo, 2020, s. 292). Ved estimering av peers i et klasserom eller enda større sammenhenger, kan det oppstå estimeringsutfordringer basert på at man ikke klarer å skille mellom peers som faktisk påvirker elevens prestasjoner og ikke (Hong & Lee, 2017). Med utgangspunkt i smågruppeundervisning, der elever blir tatt ut av vanlig matematikkundervisning i klassen, og mottar undervisning i mindre grupper, (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å.; NTNU Samfunnsforskning, u.å.), minimeres dette problemet tilknyttet inkludering av peer-effekter som ikke er relevant.

I forbindelse med 1+1 prosjektet, gjennomførte elevene en prøve i matematikk i hver enda av skoleåret (Bonesrønning et al., 2022, s. 4). Dette gir mitt utgangspunkt for identifisering av peer- og rangeringseffekter på elevprestasjoner. Resultatene mine viser at inkludering av rangeringseffekter i estimering av peer-effekter i denne institusjonelle settingen, gir et sterkere estimat av peer-effekter. Skjevheten som avdekkes er tilnærmet presis. Ved å bytte på rollene, altså at peer-effektene utelattes i estimering av rangeringseffekter, oppstår det en skjevhet i rangeringseffekten som følger samme prinsipp. Avdekkingen av denne skjevheten tar som sagt utgangspunkt i forskningen til Bertoni og Nisticò (2023), og disse resultatene er med på å underbygge deres forskning ytterligere.

Ved å estimere heterogene-effekter av ulike typer gruppesammensettinger, viser resultatene mine at peer- og rangeringseffekter er sterkest i heterogene gruppesammensettinger, altså grupper bestående av elever med ulikt ferdighetsnivå i matematikk. En slik sammensetting gir peer- og rangeringseffekter som er sterkest og positiv for elever med lavt- og middels-ferdighetsnivå, samtidig som elever med høyere ferdighetsnivå ikke får noen sterkere peer-effekt eller

rangeringseffekt av å gjennomføre smågruppeundervisning i homogene grupper. For å optimalisere elevprestasjoner, gir dette resultater som går i en retning: fra et peer- og rangeringseffekt synspunkt der peer-effekter defineres som gruppens ferdighetsnivå, bør mer heterogene gruppesammensettinger prioriteres. I heterogene gruppesammensettinger har disse effektene positive påvirkningskraft på elevprestasjoner hos flere elevgrupper, samtidig som disse effektene ikke har noen betydelig negative innvirkninger på elevprestasjoner hos noen av elevgruppene. Når jeg omtaler begrepet «elevprestasjoner» tilknyttet resultatene i denne oppgaven, er dette relatert til hvordan elever presterer på avsluttende prøve i matematikk. Resultatene er ikke direkte tilknyttet hvordan elevene presterer i andre emner på skolen.

## 2 Estimeringsutfordringer og tidligere litteratur

Dette kapittelet tar for seg noen resultater av peer-effekter fra tidligere forskning, og første delkapittel belyser flere estimeringsmetoder benyttet for å håndtere de økonometriske utfordringene tilknyttet sammenhengen denne effekten har på utdanningsvariabler. Videre beskrives nyere litteratur som argumenterer for hvorfor inkludering av rangeringseffekter er sentralt for kausaliteten til estimatene av peer-effekter. Til slutt presenteres typiske argumentasjoner tilknyttet ulik sammensetning av elevgrupper i utdanningssammenheng.

### 2.1 Estimeringsutfordringer tilknyttet peer-effekter

Jeg velger å hovedsakelig ta utgangspunkt i estimeringsutfordringer og tidligere litteratur belyst i Paloyo (2020).

#### 2.1.1 Inndeling av peer-sammensetninger

Hvilke peers man interagerer med er ikke alltid tilfeldig. Kultur og kommuneeffekter kan variere fra område til område, noe som kan påvirke elevens hverdag, og også skille peers fra hverandre fra ulike områder. Eventuelt kan foreldre sende barna sine på privatskoler med spesialisert fokus, eksempelvis idrett og språk. Videre vil også trolig lærere kunne påvirke hvilke elever som interagerer med hverandre, i form av at gruppesammensetning og elevens plassering i klasserommet velges for å skape best mulig læring for alle, med tanke på hvem som samarbeider godt med hvem. Disse eksemplene samsvarer med forklaringen til McPherson et al. (2001), som forklarer at i flere tilfeller foreligger det fellestrekk mellom aktører som tilknyttes til hverandre, i f.eks. jobb- og skolesammenheng. Det å kunne skape en situasjon med total tilfeldig fordeling som gir utgangspunkt for estimering av peer-effekter kan være svært utfordrende. Basert på forklaringen til Manski (1993, s. 532-533) tilknyttet påvirkninger av gruppefaktorer på medlemmene i gruppa, vil en bestemmelse av gruppesammensetning som ikke baseres på tilfeldighet forårsake faktorer som kan påvirke utfallsvariabelen (i denne analysen, elevprestasjoner), slik at peer-effekter ikke nødvendigvis estimeres presist (Paloyo, 2020, s. 293).

En noe brukte strategi i litteraturen, for å kunne skape en eksogen peer-variabel til tross for at inndeling i peers ikke er tilfeldig, er estimeringer som baseres på bruken av fikserte effekter (Paloyo, 2020, s. 298). I en situasjon der man benytter paneldata med fikserte effekter som estimeringsmetode, ser man på hvordan estimerer ved gitte tidsperioder varierer fra gjennomsnittet, da effekter som ikke varierer over tid eller er tilnærmet lik over tid, har ingen avvik fra gjennomsnittet, og vil dermed ikke bli en del av regresjonslikningen (Wooldridge, 2020, s. 463). Dermed foreligger det ekstra fordeler ved bruken av fikserte effekter i estimering av peer-effekter, da dette eliminerer eventuell korrelasjon mellom uobserverbare tidsuavhengige effekter og elevprestasjoner. Som nevnt, kan disse variasjonene være lokalisert på kommunenivå, skolenivå og klassenivå. Kommuner kan eksempelvis variere i klima og kvalitet på skolevei, mens skoler kan variere knyttet til bestemmelser av hvordan generell undervisning skal gjennomføres og kvalitet på digitale utstyr med ulikt nivå av effektiv anvendelse. Videre kan klasser variere fra hverandre i hvor villig elevene i klassen er til å lære. Dette kan også variere helt ned på gruppenivå, slik at inkludering av gruppefikserte effekter kunne vært relevant i en slik analyse.

Bruken av en eksogen instrumentvariabel som estimeringsstrategi for å kunne presentere kausale resultater av peer-effekter i tilfeller ved seleksjon som ikke er tilfeldig, er også noe brukt i litteraturen (Paloyo, 2020, s. 297-298). Med utgangspunkt i forklaringen til Wooldridge (2020, s. 505-511), kan bruken av en eksogen instrumentvariabel som ikke er korrelert med peer-variabelen, som samtidig har relevans i form av den har en påvirkning på elevprestasjoner, medføre at man kan estimere peer-effekter eksogent. Bramoullé et al. (2009) uttrykker peer-effektens påvirkning på individer eksogent ved å estimere peer-effektene ut i fra eksogene faktorer som er korrelert med peer-effektene. Faktorene er eksogene hvis de ikke har direkte innflytelse på utfallsvariabelen for individet av interesse. I denne analysen velger jeg å ikke bruke instrumentvariabel, selv om dette er en relevant og kanskje en mulig prosess å gjennomføre med dataen tilgjengelig. I tilfelle med en dynamisk paneldatamodell der man studere langsiktige peer-effekter (over flere år), kunne potensielt metoden utviklet av Anderson og Hsiao (1981), som viser hvordan forrige tidsobservasjon av en endogen forklaringsvariabel eksogent kan uttrykke denne forklaringsvariabel, vært en relevant estimeringsstrategi. Med

utgangspunkt i dette kunne peer-kvalitet uttrykkes eksogent gjennom å bruke peer-kvalitet som et instrument for seg selv.

Paloyo (2020, s. 295-297) presenterer også eksempler der peer-effekter estimeres med bruk av regresjonsdiskontinuitetsdesign. Jeg baserer følgende forklaring av denne estimeringsstrategien på forklaringen til Lee og Lemieux (2010). Estimeringsstrategien tar utgangspunkt i at individer enten mottar eller ikke mottar en form for behandling ut ifra en gitt grense som defineres eksogent. Typisk i utdanningssituasjon kan dette være relatert til akademiske resultater, i form av en oppnåelse av et prøveresultat over en viss grense, slik at man mottar man en behandling (f.eks. en pris eller at man kan velge hvilken skole man vil gå på). Relevant er det å ta utgangspunkt i de individene som ligger nært grensen på begge sider, som i utgangspunktet har omtrentlige samme egenskaper i tilknytning til hva grensen definerer (Lee & Lemieux, 2010, s. 283). En utfordring tilknyttet denne estimeringen er hvorvidt utfallet tilknyttet om et individ mottar behandlingen er endogen eller eksogent. Individer kan ha ønske eller behov for å motta behandlingen, som kan endre individenes atferd, som igjen kan påvirke resultatet til individene. Dette medfører at den sammenhengen man i utgangspunktet ønsker å estimere, blir upresis, grunnet at grensen for behandling påvirker den avhengige variabelen av interesse, som medfører en skjev estimering av uavhengige variable av interesse. Gitt at man kan argumentere for at en gitt grense ikke påvirker elevprestasjoner, er denne estimeringsmetoden relevant for estimering av peer-effekter. Vardardottir (2013) bruker en slik estimeringsmetode for å identifisere peer-effekter i klasser der sannsynligheten tilknyttet hvilken klasse elever blir plassert i brukes som den overnevnte grensen, og med dette estimeres positive peer-effekter av relativt stor tallverdi.



### 2.1.2 Identifisering

Basert på fundamentene presentert av (Manski, 1993), er det å identifisere hvem man faktisk påvirkes av å ikke, utfordrende. Hong og Lee (2017) tar utgangspunkt i en institusjonell setting der sør-koreanske studenter innehar faste sitteplasser, slik at utfordringen nevnt ovenfor i identifisering av peer-effekter reduseres, ved at man i større grad kan argumentere for at den definerte peer-effekten faktisk har betydning på utfallsvariabelen. Estimerer man klasseeffekter på elevprestasjoner i sin helhet, medberegnes alle elevene i elevens klasse i estimeringen. Totalt sett, rapporterer Hong og Lee (2017) sterkere peer-effekter på elevprestasjoner når peer-kvalitet baserer seg på karakteristikk tilknyttet sidekameraten til eleven av interesse, sammenliknet med når peer-kvalitet inkluderer aktører som ikke er like relevant for individets prestasjoner. Dette impliserer at en sentral utfordring ved estimering av peer-effekter er å kunne identifisere hvem som faktisk påvirker hvem. I en større gruppe der peer-effekten baserer seg på gruppen som helhet, kan dette medføre at peer-effekter estimeres skjevt.

Disse fundamentene styrker min institusjonelle setting. Fra Tabell 1 og Tabell 2 rapporteres gjennomsnittlig gruppestørrelse på mellom 4 og 5 elever. Dette styrker argumentasjonen til at mine resultater kan tolkes presis med tanke på at i en smågruppesituasjon, vil elevene være tilknyttet hverandre i denne type undervisning. Når det er sagt, er det naturlig å tenke at grupper på 2 elever og grupper på 8 elever vil variere noe i dynamikk. Denne dynamikken kan gjenspeiles i en mye større påvirkningskraft fra sine gruppemedlemmer hvis gruppene består av et lavt antall elever, relativt til større grupper. Dette vil også igjen variere med personlighetstrekkene til gruppemedlemmene.

### 2.2 Rangeringseffekter

Som tidligere nevnt, presenterer Bertoni og Nisticò (2023) signifikante resultater som sier at det å utelatte rangeringseffekter fra estimering av peer-kvalitet medfører negativ skjevhet av peer-effekter på studentprestasjoner. Resultatene viser hvor essensielt inkludering av rangeringseffekter er i analyser av peer-kvalitets påvirkning på elevprestasjoner i matematikk og andre utdanningsvariabler, da dette påvirker både styrken og signifikansen til estimatene. Anta at vi studerer elev «i» i gruppe «j». En generell og fornuftig tanke er at elever blir positivt påvirket

av gruppe medlemmer med høy kompetanse, dermed forventes det en positiv sammenheng mellom økt gjennomsnittsferdigheter i gruppa og elevprestasjoner. Hvis gjennomsnittsferdighetene i gruppe «j» er høy, er det rimelig å anta at dette kan medføre positive effekter for individ «i's» elevprestasjoner. I kapittel 1 presenteres litteratur som argumenterer for en positiv sammenheng mellom rangeringseffekter og elevprestasjoner. Men i de fleste tilfeller, følger det intuitivt en negativt sammenheng mellom økt peer-kvalitet og økt rangering, særlig hvis begge er basert på akademiske resultater, og på bakgrunn av dette er det essensielt å kontrollere denne korrelasjonen, som vist i Bertoni og Nisticò (2023). I nevnte analyse bevises dette ved to forskjellige utdanningsnivå: data tilknyttet kenyanske barneskole-elever, som gav utgangspunkt for Duflo et al. (2011) sin undersøkelse knyttet hvordan elevprestasjoner differerer blant annet i ut ifra hvordan klasser settes sammen, og data fra nederlandske studenter, der Booij et al. (2017) undersøker peer-effekter ved tilfeldig gruppering av studenter ved universitetet i Amsterdam.

### 2.3 Gruppekarakteristika

Bakgrunnen for elevers plassering i skoler, klasser og grupper, kan variere rundt om i verden og fra utdanningssystem til utdanningssystem, og utdanningssystemer tilknyttet tidlig stadiet av utdanningsnivå i flere andre land, plasseres elever i skoler, og eller klasser/grupper basert på ferdigheter (Paloyo, 2020, s. 291-292). Dette vil da trolig skape en undervisningssituasjon der elevene har omtrentlig likt ferdighetsnivå, men kan variere i andre karakteristikk, som f.eks. hjemsted, kultur og språk.

I Norge eksisterer en opplæringslov, som gir utgangspunkt for at elever i grunnskolen skal få utdanningsplass tilknyttet sitt bosted (Opplæringslova-oppl, 1998, §8-1). Avstanden fra elevens bosted til selve skolen vil variere basert på kommuneeffekter, f.eks. demografi og økonomi.

Videre er det bestemt at i hovedsak skal gruppering av elever innad i skolene ikke baseres etter ferdigheter eller enkelte sosiale faktorer (Opplæringslova-oppl, 1998, § 8-2). Fra et forskningsperspektiv, gir dette utgangspunkt i et bredt spekter av ferdighetsnivåer på elevene, samtidig som elevkarakteristika kan preges av lokale faktorer som kultur, språk og foreldre- og familiekarakteristika, som igjen kan påvirke ferdighetsnivået og prestasjonene til elevene.

Et relevant spørsmål er derfor hvordan elever bør grupperes i en undervisningssituasjon for å skape best mulig læringsmiljø for alle. Innad i begrepet «grupperes», menes her fordeling av elever til skoler klasser og grupper innad i skolene. Både litteraturen og utdanningsaktører presenterer forskjellige resultater og argumenter for hvilken grupperingsmetode som gir best resultater. Jeg fokuserer på om gruppering av elever i en undervisningssituasjon bør være basert på at elever med like ferdigheter bør grupperes sammen eller ikke.

Det praktiseres flere måter for å gruppere elever etter ferdigheter, og i det amerikanske utdanningssystemet, hvis en slik gruppering skal finne sted, vil typisk gruppering etter ferdigheter på grunnskolenivå foregå innad i klasserommet, mens høyere utdanningsnivå praktiserer oftere en slik gruppering mellom klasser (Loveless, 2013, s. 13-14). I et tilfelle der slike grupperinger praktiseres, kan selve gruppene være tilfeldig fordelt i den forstand at hvilke elever av samme ferdighetsnivå som havner i samme gruppe ikke nødvendigvis er bestemt, men elevene kategoriseres inn i ferdighetsnivåer, og inndeles i grupper deretter. Dette skiller seg fra denne undersøkelsen, da gruppering i utgangspunktet ikke er basert på ferdighetsnivå, og hvorvidt randomisering av grupper deretter er praktisert eller ikke er uklart og kan variere. Mer om dette i kapittel 4. Duflo et al. (2011) finner at elever ved kenyanske barneskoler som inndeles grupper etter ferdigheter, presterer høyere enn elever ved kenyanske barneskoler som ikke grupperes etter ferdigheter. Totalt sett betyr dette at elever har fått positiv utbytte av et undervisningsmiljø bestående av medelever som kategoriseres til samme ferdighetsnivå. Booij et al. (2017) viser at det å gruppere studenter på universitet i Amsterdam etter ferdigheter har en generell positiv effekt på prestasjoner for studenter som ikke tilhører den øverste del av totalt tre kategorier definert etter ferdigheter.

Til tross for disse resultatene, har flere stater og skoler i USA ønsket seg bort fra gruppering etter ferdigheter, på bakgrunn av at en slik gruppering kan skape stort skille mellom elever, både i læringsutbytte og i sosiale faktorer (Rui, 2009). Rui (2009), presenterer en vurdering av resultater og argumenter tilknyttet om ferdigheter bør være en faktor for gruppesammensetning, og konkluderer at resultater fra forskning viser til positive effekter på prestasjoner til studenter som ikke tilhører gruppen med høyest ferdighetsnivå hvis elever ikke grupperes basert på ferdighetsnivå. Dette gir grunnlag for argumentasjon i motsatt retning av resultatene i avsnittet

over. En naturlig argumentasjon imot mer tilfeldig fordeling av grupper vil være på grunnlag av at høytpresterende elever svekkes av dette i form av at lavere peer-kvalitet, eller at lærere ikke har mulighet for å tilrettelegge undervisningen sammenlignet med grupper satt sammen etter ferdighetsnivå. Videre viser vurderingen til Rui (2009) at dette ikke nødvendigvis stemmer fra empiriske resultater. Her kommer viktigheten av rangeringseffekter inn i bilde. Som vist i Bertoni og Nisticò (2023, Tabell 5), eksisterer det rangeringseffekter for lavtpresterende elever i homogene grupper, mens høytpresterende elever opplever motsatt utvikling. For høytpresterende elever gruppert sammen, vil de positive effektene fra høyere peer-kvalitet svekkes grunnet svakere rangering. Estimering av rangeringseffekter er dermed svært sentralt, da mer heterogene gruppesammensetninger kan gi elever av høyere ferdighetsnivå positive rangeringseffekter som bidrar til at den totale effekten på elevprestasjoner ikke svekkes ved en slik grupperingsfilosofi. Samtidig er det ikke nødvendigvis slik at det er den tilpassede undervisning som fremkaller bedre elevprestasjoner i homogene gruppesammensetninger. Flere lavtpresterende elever vil oppleve positive rangeringseffekter, som like gjerne kan bidra til økte elevprestasjoner. Ved å ta alle disse synsvinklene i betraktning, er spørsmålet tilknyttet hvordan elever bør fordeles i peer-grupper svært komplisert.

### 3 Data, institusjonell setting og deskriptiv statistikk

#### 3.1 1+1 prosjektet

1+1 prosjektet undersøker hvordan elevprestasjoner påvirkes av smågruppeundervisning i matematikk, der halvparten av skolene i prosjektet får tildelt ekstra ressurser til å gjennomføre tiltaket (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å.; NTNU Samfunnsforskning, u.å.). Tildeling av ressurser til hvilke skoler er tilfeldig. Datamaterialet i denne oppgaven består av observasjoner av elever i tidlig stadige av norsk barneskole, på to forskjellige skoletrinn, som gjennomførte totalt to prøver i matematikk, en prøve i starten av skoleåret før selve smågruppeundervisningen begynte, og en prøve i senere faser av skoleåret (Bonesrønning et al., 2022, s. 4). I denne oppgaven kaller jeg de henholdsvis for innledende- og avsluttende prøve. Skolene deles inn i ulike delgrupper, og innad i disse er typisk fire eller seks skoler, både skoler som har mottatt smågruppeundervisning og ikke. Disse er like i karakteristikk basert på tidligere akademiske resultater, og derfor kontrollerer jeg for fikserte effekter tilknyttet gruppering av skoler, for å i større grad kontrollere for effekter som varierer mellom skolene. Etter at skolene fikk beskjed om de var behandlings- eller kontrollskole, trakk en skole seg fra prosjektet, alle skoler tilhørende vedrørende delgruppe utelukkes fra analysen.

#### 3.2 Deskriptiv statistikk og forklaring av variabler

Peer-variabelen baseres på resultatet på de innledende prøvene, og for elev «i» defineres denne som gjennomsnittlig resultat på denne prøven i gruppe «j», ekskludert resultatet til elev «i». Denne variabelen representerer omtalte «peer-effekter» og «peer-kvalitet» i denne analysen. Rangeringsvariabelen, som indikerer rangeringen til elevene i gruppa, baseres også på resultatet tilknyttet den innledende prøven. Gruppestørrelse varierer på tvers av grupper totalt sett, og på tvers av gruppene for elever som deltar i forskjellige gruppesammensetninger, slik at en normalisering av denne variabelen er nødvendig, og derfor tar jeg utgangspunkt i fremgangsmåten til Murphy og Weinhardt (2020), også brukt av Bertoni og Nisticò (2023),

Delaney og Devereux (2022) og Denning et al. (2021), der rangeringen til individ  $i$ , i gruppe  $j$  ved skole  $s$  gis ved:

$$\text{Rangering}_{ijs} = \frac{r_{ijs}-1}{N_{js}-1} \quad (\text{A})$$

Her studeres gruppe « $j$ » ved skole « $s$ », der  $r_{ijs}$  er et heltall, og indikerer hvilken rangering elev « $i$ » har i gruppe « $j$ » ved skole « $s$ », og desto høyere tallet er, desto bedre rangering har eleven i gruppa.  $N$  viser hvor mange elever som tilhører gruppe « $j$ » ved skole « $s$ ». Likning (A) gir da en rangering til elev  $i$  mellom 0 og 1, og desto høyere verdi, desto høyere rangering. I flere tilfeller, innehar elever i samme gruppe likt resultat på den innledende prøven. I disse tilfellene blir begge elevene tildelt høyest tilhørende rangering. Dette skiller seg fra hvordan Bertoni og Nisticò (2023), Denning et al. (2021) og Murphy og Weinhardt (2020) håndterer dette, men de viser at for ulike håndteringer av elever som deler rangering, produserers ganske like resultater. Mer om dette i kapittel 5 og kapittel 6. Antall gruppe-medlemmer varierer fra 2 til 12, men med bruk av den mest hyppige verdien av gruppestørrelse for hver elev, varierer dette fra 2 til 8 gruppe-medlemmer, der 4-6 elever per gruppe er det vanligste.

Elever tilhører ikke nødvendigvis samme gruppe gjennom smågruppeundervisningen. Dette betyr at både variablene for både peer og rangering kan varierer for elevene mellom hver gang smågruppeundervisning gjennomføres på den respektive skole. Variabelen som indikerer peer-kvalitet/peer-effekter, kalkuleres ut i fra gjennomsnittlig resultat på den innledende prøven for de gruppene eleven har deltatt i, ekskludert elevens eget resultat. For å kalkulere elevens rangering, tar jeg den mest hyppige rangeringen til eleven gjennom smågruppeundervisningen, og det mest hyppige antall gruppe-medlemmer i gruppene eleven har deltatt i. Et alternativ hadde vært å brukt gjennomsnittsverdier av  $r_{ijs}$  og  $N_{js}$ , men dette skaper flere tilfeller av  $\text{Rank}_{ijs} > 1$ .

I gjennomføringen av analyser knyttet til heterogene peer- og rangeringseffekter, kategoriseres elever og grupper inn i kvartiler, basert på elevenes resultater på den innledende prøven. For å sikre nok spredning og for å få tydelig nok skille mellom elevene, vil de 25 % med lavest resultat kategoriseres som «lavtpresterende» og de med 25% høyest resultat kategoriseres som «høytpresterende». De resterende to kvartilene slås sammen til et intervall, slik at de resterende

50 % av elevene kategoriseres som «middelspresterende». Resultat fra alternativ inndeling for å sikre robusthet, presenteres i Tabell A3. Inndeling av grupper etter spredning i resultat på den innledende prøven følger samme prinsipp. De gruppene med 25% lavest spredning av totalt alle grupper kategoriseres som homogene grupper, og de gruppene med 25 % høyest spredning av totalt alle grupper kategoriseres som heterogene grupper. Resterende 50 % av gruppene kategoriseres som standard grupper. Resultat for alternativ inndeling for å sikre robusthet, presenteres i Tabell A3.

Variabel for peer-kvalitet/peer-effekter defineres som «Peer», og i likhet med resultat på avsluttende prøve og resultat på innledende prøve, standardiseres variabelen i alle regresjoner, resultater og figurer. For å styrke resultatene, inkluderer jeg flere kontrollvariabler som tradisjonelt sett påvirker studentprestasjoner. Følgende kontrollvariabler er valgt med utgangspunkt i de kontrollvariablene som ble brukt i Bonesrønning et al. (2022, online appendix D), bortsett fra at denne oppgaven ikke inkluderer en variabel for skolestørrelse, og inneholder enn litt annen kategorisering av utdanningsnivå til elevens foreldre:

Elev-født-utland: Dummyvariabelen indikerer om eleven er født i utlandet eller i Norge, og tar verdien 1 for utenlandskfødte elever, og 0 for norskfødte elever.

Foreldre-født-utland: Dummyvariabelen er knyttet til tilfeller der elevens fødeland er Norge, samtidig som begge foreldre ikke har Norge som fødeland. Variablene tar verdien 1 hvis dette er tilfelle, og 0 hvis ikke.

Kvinne: Dummyvariabel som indikerer elevens kjønn, og tar verdien 1 for jenter og 0 for gutter.

Utd-foreldre\_n: Jeg inkluderer fire dummyvariabler som sier noe om foreldrenes fullførte utdanning. I tilfeller der elevens foreldre har forskjellig nivå av fullført utdanning, vil variabelen ta utgangspunkt i den forelderen med det høyeste fullførte nivået.  $n \in [1,4]$ , der 4 tilsier fire år eller mer på universitet/høgskole, 3 tilsier opptil fire år på universitet/høgskole, 2 tilsier videregående-/fagskole og 1 tilsvarer grunnskoleutdanning, ingen utdanning eller ikke rapportert utdanning. Utdanningsnivået til foreldre ved elevens avsluttende prøve blir benyttet. Eventuelle

endringer i foreldres utdanningsnivå mellom gjennomføring av prøvene som påvirker elevprestasjoner, kontrolleres ikke for. Samtidig, siden dette er elever i 2. og 3. trinn, kan det argumenteres for at en eventuell endring i utdanningsnivå til foreldrene forventes å ha liten betydning, med mindre denne endringen skjer på de laveste utdanningsnivåene, med tanke på i hvilken grad dette har betydning på foreldrenes evne til å hjelpe elevene med skolearbeidet. I tillegg kan det argumenteres for at sannsynligheten er liten for at dette har en sterk korrelasjon med peer- og rangeringseffekter. Hvis dette skulle ha betydning, kan det være av en økt underliggende rangeringseffekt som ikke fanges opp i analysen, da elevers rangering er bestemt ut ifra resultater på den innledende prøve.

Tabell 1 viser deskriptiv statistikk for hele utvalget. Ved å sammenlikne gjennomsnittlig standardisert poengsum på prøvene, bemerker vi oss at gjennomsnittresultatet har økt. Dette betyr at det eksisterer en positiv behandlingseffekt, som dokumentert i Bonesrønning et al. (2022). Dette bekreftes også av Figur 1 som viser fordelingen av prøveresultatene. Figur 1 gir et oversiktsbilde over hvilke elever, med utgangspunkt i ferdighetsnivå, som har fått mest positiv effekt av smågruppeundervisning, og vi ser at den positive utviklingen i prøveresultater hovedsakelig skyldes at lavtpresterende elever øker sine poengsummer. Hvis man antar at deler av behandlingseffekten skyldes positive peer- og rangeringseffekter, vil Figur 1 indikere at disse effektene er størst for lavtpresterende elever. Hvis en opptelling av hvor mange elever som tilhører de ulike ferdighetsnivåene rett etter at prøvene hadde funnet sted, viser Figur 1 at andelen lavtpresterende elever er redusert, og andel middelspresterende elever har økt. For de elevene som presterer høyere enn gjennomsnittet på begge prøvene, er tendensene noe ulikt, da andelen elever som presterer helt i toppen har økt, mens andelen av elever med et standardisert resultat på omtrent 1 er redusert.

Videre, viser Tabell 1 at gjennomsnittlig gruppestørrelse er omtrent 5 elever, og gjennomsnittlig differanse mellom den med høyest og lavest poengsum i hver gruppe er på 10.56. Merk at denne verdien er ikke standardisert. Gjennomsnittlig rangering er 0.524, og kontrollvariablene viser at omtrent halvparten av elevene er jenter, og i overkant av 14% av elevene har en nær form for utenlandsk bakgrunn.



**Tabell 1: Deskriptiv statistikk.**

---

Observasjoner: 5,306

---

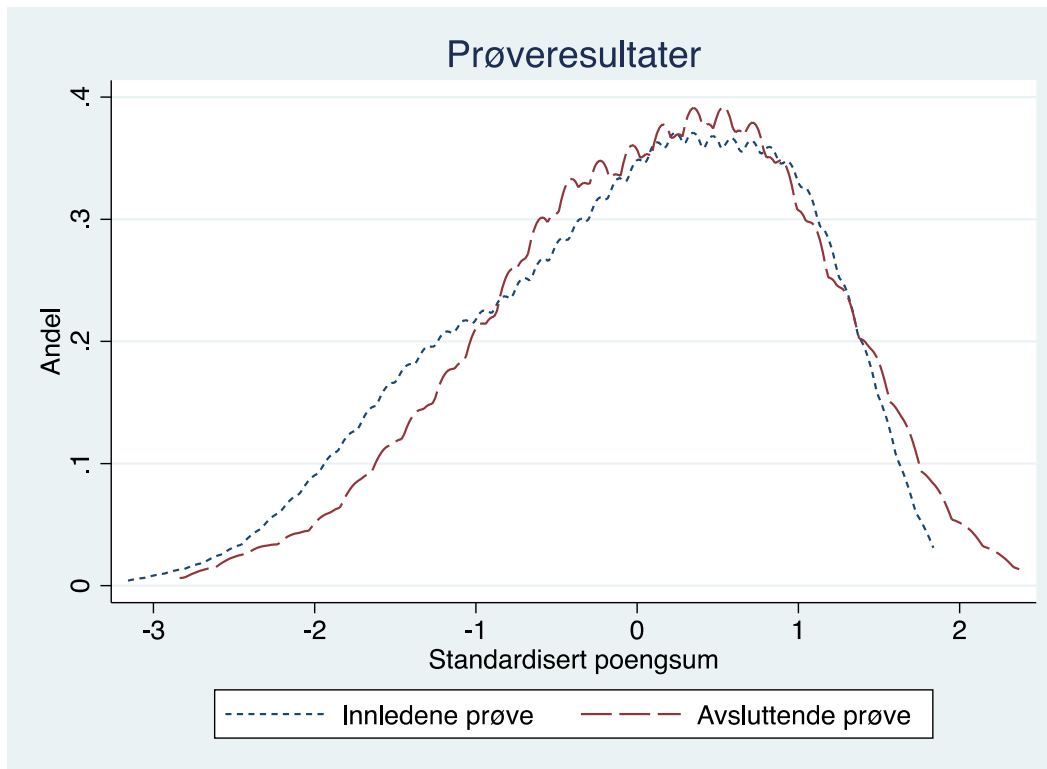
	(1)	(2)	(3)	(4)
VARIABLER	Gj	St.av	Min	Maks
Avsluttende prøve	0.094	0.976	-2.838	2.371
Innledende prøve	-0.069	1.010	-3.158	1.838
Peer	0.008	1.003	-3.214	2.148
Rangering	0.524	0.361	0.000	1.000
Elev-født-utland	0.055	0.229	0.000	1.000
Foreldre-født-utland	0.088	0.283	0.000	1.000
Kvinne	0.496	0.500	0.000	1.000
Utd-foreldre_1	0.057	0.233	0.000	1.000
Utd-foreldre_2	0.200	0.400	0.000	1.000
Utd-foreldre_3	0.387	0.487	0.000	1.000
Utd-foreldre_4	0.356	0.479	0.000	1.000
Gruppestørrelse	4.814	1.073	2.000	8.000
Differanse i gruppe	10.557	5.491	0.000	38.000

---

Tabellen viser deskriptiv statistikk for utvalget i hovedanalysen.

Variablene Avsluttende prøve, Innledende prøve og Peer er standardisert.

**Figur 1: Kumulativ fordeling av prøveresultater.**



Figuren viser fordelingen til resultater på innledende- og avsluttende prøve for hele utvalget, med standardiserte variabler.

Figur 2 viser fordeling av resultater på den innledende prøven innad i de ulike gruppekategoriene av interesse. Kombinert med Tabell 2, som viser deskriptiv statistikk for variabler innad i hver gruppekategori, ser vi at prøveresultater i homogene grupper fordeles mer mot høyre i figuren relativt til prøveresultater innad i de to andre gruppene. Dette trekker i retning av at det eksisterer flere homogene grupper som består av høytpresterende elever, enn homogene grupper bestående av lavtpresterende elever. Dette bekreftes også av peer-verdien i homogene grupper fra Tabell 2, som er betydelig høyere enn de to andre gruppekategoriene og tilsvarende verdi rapportert i Tabell 1. Når det gjelder de to andre gruppene, er fordelingen av elever ganske jevn basert på resultat på innledende prøve, slik at verken standard eller heterogene grupper karakteriseres av en særlig større andel av elever med et typisk prestasjonsnivå enn noe annet. Gjennomsnittlig differanse i resultat på innledende prøve innad i gruppene rapporteres også i Tabell 2 og ved at denne variabelen har et intervall  $[0, 38]$ , sikrer en kategorisering av grupper i tre kategorier, at

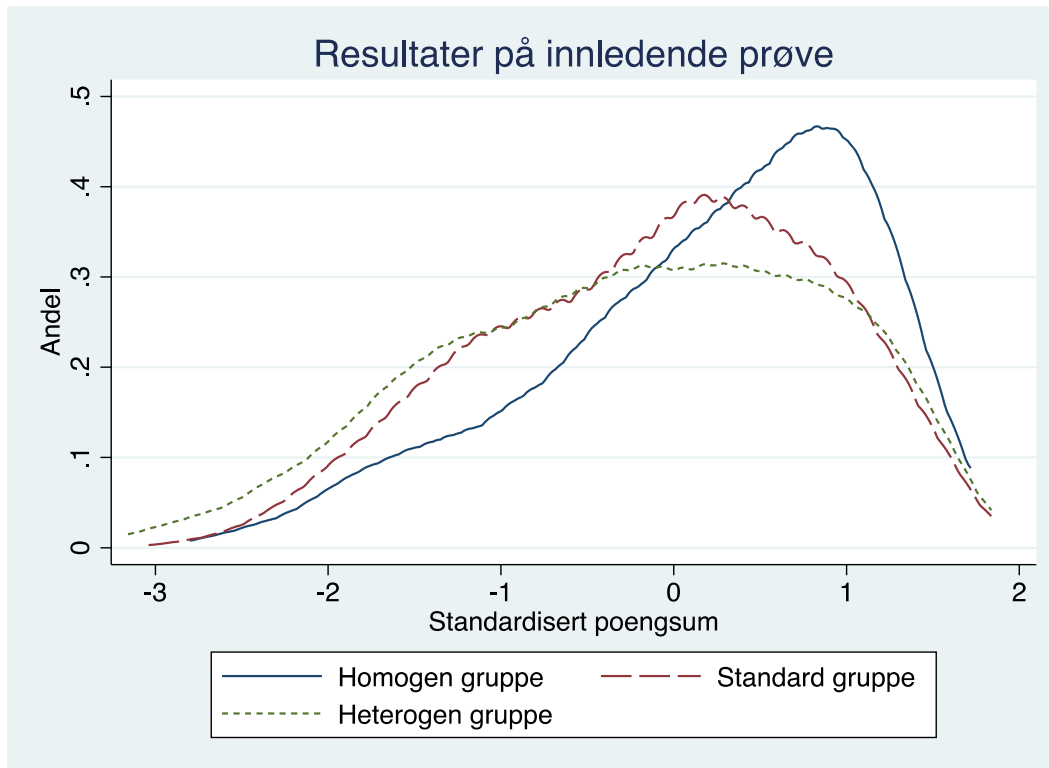
forskjellen mellom homogene og heterogene grupper blir stor. Dette medfører at effekter fra de ulike kategoriene som kan påvirke resultat på avsluttende prøver i sjeldnere grad overlappes, og heller fanges opp i den gruppenkategorien jeg kaller standard grupper.

**Tabell 2: Deskriptiv statistikk for ulike gruppekarakteristika.**

Grupper	A		B		C	
Observasjoner	1,327		2,611		1,368	
VARIABLER	(2) Gj	(3) St.av	(5) Gj	(6) St.av	(8) Gj	(9) St.av
Avsluttende prøve	0.192	0.946	0.088	0.991	0.010	0.968
Innledende prøve	0.183	0.946	-0.109	0.975	-0.235	1.090
Peer	0.320	1.089	-0.038	1.018	-0.207	0.793
Rangering	0.544	0.361	0.519	0.358	0.515	0.365
Elev-født-utland	0.054	0.227	0.059	0.236	0.050	0.217
Foreldre-født-utland	0.084	0.278	0.089	0.285	0.089	0.285
Kvinne	0.476	0.500	0.501	0.500	0.505	0.500
Utd-foreldre_1	0.060	0.237	0.054	0.227	0.061	0.240
Utd-foreldre_2	0.199	0.399	0.207	0.405	0.187	0.390
Utd-foreldre_3	0.408	0.492	0.370	0.483	0.398	0.490
Utd-foreldre_4	0.333	0.471	0.369	0.483	0.353	0.478
Gruppestørrelse	4.538	1.231	4.872	1.023	4.971	0.947
Differanse i gruppe	4.338	1.534	9.866	2.125	17.907	3.687

Tabellen viser deskriptiv statistikk i de ulike gruppene. Kolonne (1), (3) og (5) viser gjennomsnittsverdier for variabler i homogene, standard og heterogene grupper. Kolonne (2), (4) og (6) viser standardavvikene for de respektive gruppene. Gruppe A=homogene grupper, Gruppe B=standard grupper Gruppe C=heterogene grupper.

**Figur 2: Kumulativ fordeling av resultater på innledende prøve for ulike gruppekarakteristika.**



Figuren viser hvordan resultater på innledende prøve fordeles innad i de tre gruppene av interesse. Prøveresultat er standardisert.

### 3.3 Nødvendige justeringer av datasett

Grunnet manglende nødvendig informasjon hos flere elever, har nødvendige justeringer av data blitt gjennomført. Videre følger forklaring på hvorfor ikke alle elever som har deltatt på smågruppeundervisning ikke blir medberegnet i hovedanalysene.

Grunnet manglende nødvendig informasjon tilknyttet sentrale variabler fra deltakelsen i smågrupper hos en andel elever, må omtrent en tredjedel av observasjonene som i utgangspunktet skal ha mottatt smågruppeundervisning fjernes fra analysene.

Av elever som har registrert gjennomført smågruppeundervisning, har en liten andel ikke registrert innledende prøve og/eller avsluttende prøve. Videre har en betydelig andel av elevene ikke registrert gjennomsnittlig grupperesultat med/uten respektive elev, med påfølgende ingen

registret rangering, mens noen få elever har ikke fått registrert annen nødvendig informasjon til å medberignes i analysene. En videre nødvendig justering i analysen er tilknyttet rangeringsvariabelen. Ved å ta utgangspunkt i hyppigheten i rangering og antall gruppemedlemmer til hver elev, vil det oppstå tilfeller av elever som ikke har hyppig verdi for en eller begge variabler. Av totale elever med øvrig nødvendig informasjon tilknyttet deltakelse i smågruppeundervisning, utgjør dette mindre enn 5%. Det gjennomføres robusthetstester med bruken av maksimum- og minimumsverdier av disse variablene for å undersøke om dette gir utslag på resultatene. Med bruken av de mest hyppige observasjonene i likning (A), oppstår det noen svært få tilfeller der rangeringsvariabelen får større verdi enn «1». For å opprettholde kausalitet til resultater, medberignes ikke disse i analysen.

## 4 Empirisk metode

### 4.1 Identifisering av peer-effekter

Jeg tar utgangspunkt i en generell lineær gjennomsnittmodell (Bertoni & Nisticò, 2023, likning (1)), som estimerer sammenhengen mellom elevprestasjoner og peer-effekter som bygger på Manski (1993) sin argumentasjon knyttet til hvordan ulike effekter kan medføre at aktører som tilhører samme gruppe har en form for felles atferd. Med utgangspunkt den innledende prøven, selve gjennomføring av smågruppeundervisningen, og avsluttende prøve, presenteres følgende sammenheng:

$$Y_{ijs}^E = \alpha_0 + \alpha_1 \bar{Y}_{js}^F + \alpha_2 Y_{is}^F + \phi X'_i + e_{ijs} \quad (1)$$

Der  $Y_{ijs}^E$  er resultat på avsluttende prøve for individ «i», i gruppe «j», ved skole «s».  $\bar{Y}_{js}^F$  er gjennomsnittlig resultat på den innledende prøven for gruppe «j», ekskludert individ «i», ved skole «s».  $Y_{is}^F$  er gjennomsnittlig resultat på innledende prøven for gruppe j ekskludert individ i's resultat, og  $X'_i$  er en vektor av kontrollvariabler for individ.

Videre tar jeg utgangspunkt i teorien til Manski (1993, s. 532-533), for å vurdere om likning (1) gir et godt rammeverk for estimering av peer-effekter.

Ved å ta utgangspunkt i likning (1), kontrollerer jeg for at elever i en smågruppe kan påvirke hverandre ulikt. I løpet av skoleåret er det flere elever som interagerer med forskjellige gruppemedlemmer, og i hvilken grad de påvirkes av de ulike gruppemedlemmene er ikke observerbart. Hvor mye man påvirkes av de man er på gruppe med, varierer fra gruppe til gruppe, men med utgangspunkt i forklaringen til (Manski, 1993, s. 532-533), argumenterer jeg her at min peer-variabel reflekterer hvordan elevens resultater på avsluttende prøve påvirkes av ferdighetsnivået i gruppa eller andre karakteristika i gruppa, da peer-variabelen baseres på prøveresultater før gjennomføringen av selve smågruppeundervisningen. Peer-variabelen i denne undersøkelsen er derfor i utgangspunktet ikke noe mål på hvor mye man påvirkes av ulike elever med ulik karakteristika, men heller et mål på hvordan elevprestasjoner påvirkes av det å interagere med elever av høyt eller lavt ferdighetsnivå. Dette betyr at en positiv verdi på  $\alpha_1$ , viser en positiv lineær peer-effekt på elevprestasjoner der peer-effekten representerer det

gjennomsnittlige ferdighetsnivået i smågruppa, og etterhvert som denne øker, forventes også elevprestasjonene å øke.

I løpet av et skoleår, kan faktorer som har direkte påvirkning på elevprestasjoner endres, og eksogene sjokk kan inntreffe. Basert på det Manski (1993, s. 532-533) skriver, kan dette være individspesifikke sjokk, eller sjokk knyttet til smågruppa, klasse, skole osv. Eksempler på potensielle relevante eksogene sjokk kan være at lærerstaben endres, klassen/skolen får nye lokaler eller at undervisningen digitaliseres. Effektene fanges opp i restleddet  $e_{igst}$ , og det foreligger derfor en realistisk antakelse at

$$E(e_{ijs}, Y_{ijs}^E) \neq 0 \quad (a)$$

For identifisering av en uavhengig variabel (i dette tilfelle, peer-effekter) er dette ikke noe problem så lenge følgende betingelse holder:

$$E(e_{ijs}, \alpha_1) = 0 \quad (b)$$

(Wooldridge, 2020, s. 82).

Med utgangspunkt i eksogene sjokk som inngår i  $e_{ijs}$ , og sett bort i fra alle andre uobserverbare effekter, holder denne betingelse ved bruk av  $\bar{Y}_{js}^F$  som peer-variabel, da eksogene sjokk som inntreffer i løpet av smågruppeundervisningen ikke er korrelert med peer-effektvariabelen, som er eksogent gitt av resultater på innledende prøve.

Manski (1993, s. 533) forklarer også at det kan foreligge underliggende effekter basert på fellesfaktorer mellom elever med relasjoner til hverandre. I denne settingen er smågruppen en felles relasjon for elevene. Dette punktet drar oss tilbake til en komplisert utfordring i estimering av peer-effekter som ble mye diskutert i kapittel 2, tilfeller der peers ikke er bestemt tilfeldig. Kort sagt kontrolleres dette ikke for i min analyse, og bekymringer om at elevene ikke er tilfeldig fordelt inn i grupper er absolutt reelle. Likevel argumenterer jeg for at gjennomføringen av analysen medfører diskuterbare resultater, og valg av estimeringsmetode medfører flere styrkende argumentasjoner til kausaliteten til estimatene. Som poengtert i avsnittet over, med bruk av resultater som inntraff før smågruppeundervisningen startet, kontrolleres det for eksogene sjokk. Videre ved å bruke de første prøveresultatene etter et skoleår med smågruppeundervisning, unngås eventuelle gruppeendringer gjort av lærere eller skoleledelse

etter første skoleår med smågruppeundervisning. Med utgangspunkt i en realistisk antakelse at lærere ønsker å maksimere resultatene til elevene, kan de etter en periode organisere grupper basert på observasjoner over hvilke gruppesammensettinger som fungerer eller ikke fungerer, for å sikre en god læringsprosess for flere elever. Disse endringene kan medføre overestimering av peer-effekter, hvis endringene medfører positiv effekt på læring. Ved å studere peer-effektene over et lengere tidsperspektiv, ville dette vært en større bekymring. En mulig utvidelse av denne analysen, er å kontrollere for om lærere har fulgt anbefalinger av gruppesammensetning fra ledelsen av 1+1 prosjektet. Hvis anbefalingen har innvirkning på gruppesammensetningen, kunne det gitt en pekepinn på om inndeling av grupper er tilfeldig eller ikke.

## 4.2 Inkludering av rangeringseffekter

Som nevnt i kapittel 1 og kapittel 2, presenterer Bertoni og Nisticò (2023) signifikante resultater på en negativ korrelasjon mellom peer- og rangeringseffekter. Dette delkapittelet er basert på begrunnelse og empiriske teori presentert av Bertoni og Nisticò (2023), og jeg følger deres fremgangsmåte som gir likning (2) og likning (3) (henholdsvis likning (2) og likning (3) i Bertoni & Nisticò, 2023), og med utgangspunkt i likning (1) presenteres følgende sammenheng som er basismodellen i denne oppgaven.

$$Y_{ijs}^E = \delta_0 + \delta_1 \bar{Y}_{js}^F + \delta_2 Y_{is}^F + \gamma_1 \text{Rangering}_{ijs} + \varphi X'_i + \epsilon_{ijs} \quad (2)$$

Med utgangspunkt i likning (3) nedenfor, er inkluderingen av variabelen for rangering essensiell for å presentere den sanne peer-effekten. Rangering identifiseres med likning (3) som også gir resultatene i Tabell 3.

$$\text{Rangering}_{ijs} = \beta_0 + \beta_1 \bar{Y}_{js}^F + \beta_2 Y_{is}^F + \beta_3 X'_i + v_{ij} \quad (3)$$

Det forventes at  $E(\text{Rank}_{ij} | \alpha_1) < 0$ . Det at rangering er negativt avhengig av peer-kvalitet, gjør at dette er en underliggende effekt som ikke presiseres i likning (1). Denne effekten inngår i restleddet, og basert på hvordan effekten uttrykkes av Bertoni og Nisticò (2023, s. 3), uttrykker jeg effekten slik:



$$e_{ijs} = \gamma_1 \text{Rangering}_{ijs} + v_{ij} \quad (\text{i})$$

Det eksisterer altså en kovarians mellom variabelen for peer-kvalitet og restleddet i likning (1) uttrykt ved:

$$\text{Kov}(\alpha_1, e_{ijs}) < 0 \quad (\text{c})$$

som forklart i Wooldridge (2020, s. 310-312)

Jeg inkluderer derfor  $\gamma_1 \text{Rangering}_{ijs}$  i estimeringen for presise estimater av peer-effekter.

Hvor sterk peer- og rangeringseffekt som avdekkes, avhenger av den institusjonelle settingen. Det vi allerede her kan si noe om, er forventet avvik mellom peer-effekter estimert med og uten en variabel for rangering. Ved å nok en gang følge teorien presentert i Bertoni og Nisticò (2023) og teorien beskrevet i Pei et al. (2019) og Wooldridge (2020, s. 310-312), vil skjevheten som følge av utelatte rangeringseffekter uttrykkes ved (\*), og peer-effekter med inkludering av rangeringseffekter uttrykkes ved (\*\*) (hvis vi antar at restleddet kun består av utelatte rangeringsvariabler):

$$\text{Kov}(\alpha_1, e_{ijs}) = \alpha_1 - \delta_1 = \beta_1 * \gamma_1 \quad (*)$$

$$\delta_1 = -(\beta_1 * \gamma_1) + \alpha_1 \quad (**)$$

For å identifisere heterogene effekter, benyttes også likning (2), der separate regresjoner som skiller seg fra hverandre på gruppe- og/eller individkarakteristika estimeres.

Nivå 4 av foreldres utdanningsnivå blir utelatt i estimeringen. Årsaken er at de fire kategoriene av foreldres utdanning er funksjoner av hverandre, og fra Tabell 1 kan vi se at summen av alle gjennomsnittsverdiene for disse dummyvariablene blir 1, slik at for å etablere et sammenligningsgrunnlag kan ikke alle kategoriene av foreldres utdanningsnivå inkluderes. (Wooldridge, 2020, s. 228-230).

## 5 Resultater

### 5.1 Lineære peer- og rangeringseffekter

Jeg presenterer resultater fra lineære estimeringer av peer- og rangeringseffekter i Tabell 3 og Tabell 4. Resultatene i tabellene er svært tilknyttet hverandre, og det gir derfor mening og forklare disse i en og samme kontekst. Tabell 3 viser at det å definere peer-kvalitet basert på ferdighetsnivå, har en negativ og signifikant effekt på rangering. Anta elev  $i$ , som mottar undervisning i gruppe  $j$ . Hvis gruppe  $j$  har høyt gjennomsnittlig ferdighetsnivå, består gruppen, av flere elever med høyt ferdighetsnivå. Dette vil ha negativ innvirkning på individets rangering, sammenliknet med en gruppe med lavt gjennomsnittlig ferdighetsnivå. Fra likning (3) og Tabell 3 ser vi at  $\beta_1 = -0.367$  standardavvik. Ved å kun studere peer- og rangeringseffekter opp mot hverandre, møter man derfor to effekter som går i motsatt retning (Bertoni & Nisticò, 2023, s. 1-2). I dette tilfelle knyttes positive peer-effekter til høyere gjennomsnittsresultater i gruppa ekskludert individ  $i$ . Dette forventes å gi positiv effekt på elevprestasjoner for individ  $i$ . Samtidig antas også at høyere rangering medfører positive effekter på elevprestasjoner. Høyere rangering medfører lavere peer-kvalitet og dermed reduserte og eventuelt negative peer-effekter, samtidig som høyere peer-kvalitet medfører lavere rangering, som igjen medfører reduserte og eventuelt negative rangeringseffekter.

**Tabell 3: Lineær sammenheng mellom rangering og peer.**

Avhengig variabel: rangering i smågruppe.

VARIABLER	(1)
Peer	-0.367*** (0.012)
Innledende prøve	0.433*** (0.013)
Kontrollvariabler	Ja
FE (skolegruppering)	Ja
Konstant	0.559*** (0.006)
Observasjoner	5,306
R <sup>2</sup>	0.628

OLS-regresjon som estimerer en lineær sammenheng mellom rangering og peer. Avhengig variabel er elevenes rangering i smågruppe. Standardavvik klustres på skolenivå, rapportert i parenteser.  
\*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

Tabell 4 avdekker peer- og rangeringseffekter i en lineær gjennomsnittsmo­dell. Det første man kan bemerke er signifikante og positive peer-effekter på elevprestasjoner, både med og uten rangeringseffekter i estimeringen. I tillegg rapporteres en noe svak positiv signifikant effekt av høyere rangering på elevprestasjoner. Kolonne (4) viser at hvis et individ opplever økt peer-kvalitet på 1.000 standardavvik, øker elevens resultat på avsluttende prøve med 0.153 standardavvik, med kontroll for rangeringseffekter og kontrollvariabler. En sammenlikning av kolonne (1) og (2), gjør oss oppmerksomme på at inkluderingen av kontrollvariabler er relevant, da en liten del av peer-effekter på avsluttende prøve drives av effekter fanget opp av en eller flere kontrollvariabler. Tabell A2 i Appendix rapporterer effektene fra kontrollvariablene. Kolonne (3) og (4) i Tabell 4 bekrefter en negativ korrelasjon mellom peer- og rangeringseffekter, som medfører underestimering av peer-effekter i kolonne (1) og (2).

**Tabell 4: Lineære estimeringer av peer- og rangeringseffekter.**

Avhengig variabel er resultat på avsluttende prøve.

VARIABLER	(1)	(2)	(3)	(4)
Peer	0.131*** (0.021)	0.121*** (0.021)	0.162*** (0.030)	0.153*** (0.030)
Innledende prøve	0.505*** (0.017)	0.489*** (0.017)	0.468*** (0.028)	0.452*** (0.028)
Rangering			0.085* (0.048)	0.085* (0.047)
Kontrollvariabler	Nei	Ja	Nei	Ja
FE (skolegruppering)	Ja	Ja	Ja	Ja
Konstant	0.127*** (0.024)	0.241*** (0.026)	0.080** (0.035)	0.193*** (0.036)
Observasjoner	5,306	5,306	5,306	5,306
R <sup>2</sup>	0.415	0.425	0.415	0.426

Fire OLS regresjoner som rapporterer lineære peer-effekter med og uten kontrollvariabler. I kolonne (1) og (2) ekskluderes rangeringseffekter fra estimeringene, i kolonne (3) og (4) inkluderes rangeringseffekter i estimeringene. Avhengig variabel er elevens resultat på avsluttende prøve. Standardavvik klustres på skolenivå, rapportert i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01

Når det kommer til rangeringseffekter, har det en positiv effekt på 0.85 standardavvik på resultat på avsluttende prøve. Ved å sammenligne kolonne (1) med kolonne (3), og kolonne (2) med kolonne (4), estimeres sterkere peer-effekter når rangeringseffekter inkluderes i estimeringen. Dette er i samsvar med undersøkelsene gjort av Bertoni og Nisticò (2023), og videre kan en

vurdering av kausaliteten til estimatene i tabellen gjøres. Med utgangspunkt i Tabell 3, og Tabell 4, kolonne (2) og kolonne (4) kan man si noe om fremgangsmåten er korrekt fra et økonometrisk perspektiv. Som vist i Bertoni og Nisticò (2023) og argumentert i kapittel 4, skal avviket fra peer-effekten med og uten inkludering av rangeringseffekter tilsvare den negative effekten peer har på rangering multiplisert med estimatet på peer-effekten uten rangeringseffekter inkludert. Produktet i denne analysen blir tilnærmet lik forventet verdi noe som betyr at estimeringen er ganske nære ved å kontrollere for de negative effektene rangering har på peer-effekter. I dette tilfelle er differansen i estimatet av peer-effekter mellom kolonne (2) og kolonne (4) lik 0.032 standardavvik. Produktet mellom rangeringseffekten i kolonne (4) og absoluttverdien av peer-effekten fra Tabell 3 er omtrentlig 0.031 standardavvik. En medvirkende årsak til at sammenhengen ikke er helt presist, er antall desimaler i estimatene. Ved inkludering av tilstrekkelig mengde desimaler, kunne sammenhengen estimeres mer presist. Uavhengig av dette, kan peer- og rangeringsvariablene kritiseres noe i form av hvordan disse er beregnet. Den institusjonelle settingen gjør at variabelen for rangering tilsvare i flere tilfeller en tilnærmet verdi av den sanne rangeringen elevene har hver gang de er i en smågruppeundervisningssituasjon. Deltakelse i smågruppeundervisning der man blir utsatt for ulike gruppesammensettinger, medfører at rangeringen til en elev kan differere. Estimeringen baserer seg i utgangspunktet på kun en rangering, som med konstante gruppesammensettinger ville vært sann i hele behandlingsperioden med utgangspunkt i innledende prøver. Tabell A3 rapporterer en estimering der det benyttes maksimum- og minimumsverdier av den mest hyppige observasjonen, som i prinsippet betyr at de elevene som ikke ble inkludert i hovedanalysen grunnet ingen hyppig observasjon av rangering, inkluderes i estimeringen. Vi ser at estimatet for peer-kvalitet er uforandret ved bruk av maksimumsverdier av hyppigheten til rangeringen og antall gruppemedlemmer, mens estimatet forandres noe ved bruk av minimumsverdier av hyppigheten til rangeringsvariabelen og antall gruppemedlemmer. Rangeringsvariabelen forandres i begge tilfeller. En tilsvarende analyse med disse verdiene kunne gitt en pekepinn på hvor stor betydning framgangsmåten for å bestemme individers rangering i fravær av konstante gruppesammensettinger. Videre kunne en robusthetsanalyse med gjennomsnittlig verdi av rangeringsvariabelen blitt benyttet, men dette medfører flere observasjoner utenfor intervallet  $Rangering_{ig} \in [0,1]$ , da gjennomsnittlig verdi for gruppestørrelse kan i enkelte tilfeller overskride den numeriske heltallsverdien av rangeringen eleven i den respektive gruppen har. En alternativ

estimeringsmetode ville derfor vært med utgangspunkt i paneldata, da man observerer rangering og peer-kvalitet ved hver registrerte deltakelse i smågruppeundervisning. Dette forutsetter at man hadde observasjoner på elevenes resultater etter hver økt med smågruppeundervisning. En positiv ringvirkning av paneldata er at man kan kontrollere for fikserte effekter som ville gitt mulighet for å kontrollere for skolefikserte-effekter, eventuelt kommunefikserte-effekter.

Tabell 4 bekrefter at det å beregne peer-effekter uten rangeringseffekter der peer-kvalitet måles i prøveresultater, medfører skjeve estimater. Dette betyr at en estimering der man bytter avhengig- og uavhengig variabel også er relevant. Tabell A1 beregner rangeringseffekter på avsluttende prøve der peer-effekten ekskluderes fra likning (2). Kolonne (1) viser en sterk positiv sammenheng mellom individets rangering og elevprestasjoner. I Kolonne (2) inkluderes resultat på innledende prøve i estimeringen, og effekten blir negativt. Inkludering av innledende prøver vil naturlig redusere rangeringseffekten, da jeg kontrollerer for elevenes prestasjoner før iverksetting av smågruppeundervisning, og da det eksisterer flere grupper med en del spredning i resultat på innledende prøver, er det større sannsynlighet for at høytpresterende elever har en høy rangering i gruppa, enn lavtpresterende. Inkluderes resultater på innledende prøve i estimeringen, blir effekten av rangering negativ, og høyere rangering gir negativ effekt på elevprestasjoner. Man skulle forvente at denne effekten var positiv, knyttet til forventninger om positive effekter av høyere rangering på elevprestasjoner. Det eksisterer en skjevhet i denne estimeringen, som er knyttet til utelatelse av peer-effekter i estimeringen, på samme måte som utelatelse av rangeringseffekter fra estimeringen i Tabell 4 skaper skjevhet. Produktet av rangeringseffekten i kolonne (3) fra Tabell A1 og peer-effekten i kolonne (4) fra Tabell 4 tilsvarer 0.221 standardavvik, og differansen i rangeringseffekt fra kolonne (4) i Tabell 4 og kolonne (2) i Tabell A1 tilsvarer 0.222 standardavvik. En helt presist fremstilling er det dermed ikke.

Et viktig spørsmål er knyttet til virkningsområde til peer- og rangeringseffektene.

Gruppestørrelsene varierer fra 2 til 8 elever, og hvor sterk disse effektene er kan variere ut fra gruppestørrelsen. Hong og Lee (2017) viser resultater som påpeker viktigheten av en institusjonell setting som kun inkluderer peer-effekter som faktisk påvirker elevprestasjoner, og ikke inkludere peer-effekter som i realiteten ikke påvirker elevprestasjoner. Min analysen består hovedsakelig av elever som deltar i grupper med hyppig størrelse på mellom 4 og 6 elever, men i

noen tilfeller er hyppig gruppestørrelse på 8 elever. Basert på resultatene til Hong og Lee (2017), kan grupper bestående av mange elever skape kritikkverdige estimater av peer-effekter. Hvorvidt grupper med få elever skaper en sterk nok positivt effekt av å være høyere rangert enn sine medelever, kan også diskuteres med en synsvinkel på at en eventuell selvtilit man får av å være sterkere faglig enn sine medelever øker desto flere medelever med lavere rangering i gruppa. Tabell 1 bekrefter en gjennomsnittlig gruppestørrelse på mellom 4 og 5 elever. Likevel tar jeg alle disse synspunktene på gruppestørrelse i betraktning, og Tabell A4 i Appendix viser at ujevn gruppestørrelse ikke har vesentlig påvirkning på selve peer- og rangeringseffektene.

Et viktig poeng å utdype, er at denne analysen baserer seg ikke på grupper som er tilfeldig sammensatt, en problematikk mye omhandlet i litteraturen (Paloyo, 2020). Hvis lærere, eventuelt skoleledelse eller elever, er med på å bestemme gruppekonstruksjonen, kan dette medføre upresise estimater, basert på teorien om interaksjon mellom aktører i Manski (1993, s. 532-533) og McPherson et al. (2001). I denne analysen karakteriseres peer-effektene av ferdighetsnivå. I et tilfelle der elever selv får velge, eventuelt ønske hvem de vil være på gruppe med, er det naturlig for en elev å ønske seg noen man får økt nytte av å være på gruppe med. På barneskolenivå, kan det diskuteres om elever oppfatter bedre prestasjoner på skolen som økt nytte, men likevel en mulig eksisterende kilde til skjeve estimater i Tabell 4. Det samme kan sies om det er lærere som setter sammen grupper basert på å maksimere prestasjonene til flest mulig elever i klassen. Elever som lærerne vet samarbeider godt i utgangspunktet, eller elever som læreren vet får andre positive effekter av å interagere med hverandre, kan medføre at noe av estimatet av peer- og rangeringseffektene drives av andre effekter enn at høyere gjennomsnittresultater i gruppa gir positive effekter på avsluttende prøve. Viktigheten av denne kilden til mulige skjeve estimater gjelder for alle regresjonsresultater i denne analysen.



## 5.2 Heterogene peer- og rangeringseffekter

De videre resultatene undersøker heterogene peer- og rangeringseffekter på elevprestasjoner. Som tidligere nevnt vil signifikante heterogene effekter kunne gi et godt utgangspunkt for diskusjon av pareto-effektive tiltak knyttet til gruppesammensetning i en undervisningssituasjon, noe lineære resultater ikke avdekker (Paloyo, 2020, s. 300-301). Disse resultatene rapporterer effekter innad i ulike spesifikasjoner av grupper og elever, men resultatene fra en kategori sier ikke noe om effekten relativt til en annen kategori. Med andre ord, rapportere effekter gjelder innad i en kategori av gruppekarakteristika og/eller elevers ferdigheter, og kan ikke sammenlignes med utgangspunkt i relativ effekt mellom definerte kategorier i denne analysen.

Tabell 5 undersøker graden av peer- og rangeringseffekter til elever i ulike grupper, der gruppene defineres ut i fra differansen mellom den eleven som presterte best og den eleven som presterte svakest på innledende prøve. Hver kolonne rapporterer totale peer- og rangeringseffekter i den respektive gruppetypen. Peer-effektene er positive og signifikante i både standard og heterogene grupper. Tabellen rapporterer størst peer-effekt i heterogene grupper, der en økning på 1.000 standardavvik i peer-kvalitet øker elevprestasjoner med 0.236 standardavvik. Videre avdekkes en veldig sterk positiv rangeringseffekt for elever i heterogene grupper. Det at effekten er så sterk, gir intuitivt mening da forskjellene i prestasjonsnivå er større i heterogene grupper. Hvis det foreligger tilstrekkelig differanse i spredningsnivå av ferdigheter, viser tabellen at de mekanismene som bidrar til at rangering har en positiv effekt, blir gjeldene, eventuelt sterkere. Resultatet viser at en enhets økning i standardisert rangering medføre en økning i resultat på avsluttende prøve med 0.402 standardavvik for elever i heterogene grupper. Kolonne (2) rapporterer en rangeringseffekt med omtrentlig lik styrke som rapportert for hele utvalget, men denne er ikke signifikant. Antall observasjoner er redusert, og standardavviket har økt. Kolonne (1) viser at det eksisterer ingen signifikante effekter av peer eller rangering i homogene grupper, og estimatet er også negativt. Alle gruppene sett under ett, impliserer tabellen at det å øke peer-kvalitet i homogene grupper ikke er et effektivt virkemiddel for å øke elevprestasjoner, da det medfører negative peer- og rangeringseffekter på avsluttende prøve. Tabellen viser at tilstrekkelig spredning er nødvendig for å iverksette positive og signifikante peer-effekter, og hvis spredningen i ferdighetsnivå blir stort, medfører dette svært sterke rangeringseffekter.

### Tabell 5: Heterogene gruppeeffekter.

Avhengig variabel er resultat på avsluttende prøve

VARIABLER	(1) Homogen	(2) Standard	(3) Heterogen
Peer	-0.046 (0.126)	0.174*** (0.044)	0.236*** (0.050)
Innledende prøve	0.680*** (0.146)	0.469*** (0.051)	0.344*** (0.046)
Rangering	-0.126 (0.102)	0.091 (0.075)	0.402*** (0.104)
Kontrollvariabler	Ja	Ja	Ja
FE (skolegruppering)	Ja	Ja	Ja
Konstant	0.257*** (0.082)	0.209*** (0.051)	0.053 (0.068)
Observasjoner	1,327	2,611	1,368
R <sup>2</sup>	0.448	0.453	0.409

Tre OLS regresjoner, der kolonne (1), (2) og (3) rapporterer estimater for henholdsvis homogene, standard og heterogene grupper. Standardavvik klustres på skolenivå, rapportert i parenteser. Avhengig variabel er elevens resultat på avsluttende prøve.  
\*p<0.1, \*\*p<0.05, \*\*\*p<0.01

Tabell 6 presenterer peer- og rangeringseffekter der elever kategoriseres etter ferdigheter i matematikk. Elevene kategoriseres i tre ulike grupper, og hver kolonne rapporterer totale peer-

og rangeringseffekter for elever tilhørende det respektive ferdighetsnivået. Kolonne (1) og (2) viser positive og signifikante peer-effekter for lavt- og middelspresterende elever, samtidig som kolonne (3) rapporterer ingen signifikante peer-effekter for høytpresterende elever. Tabellen viser at elever som ikke tilhører de høyeste prestasjonsnivået i denne undersøkelsens kategorisering, får en positiv effekt av økt peer-kvalitet i sin smågruppe. Effekten for lavt- og middelspresterende elever er ganske lik, og 1.000 standardavvik økning i peer-kvalitet medfører en økning på 0.142 og 0.147 standardavvik på avsluttende prøve for henholdsvis lavtpresterende elever og middelspresterende elever. Interessant nok er rangeringseffekten ikke signifikant for disse elevene. Den er riktig nok positiv, men svakere enn peer-effekten, som driver i retning av at peer-effekter har sterkere påvirkningskraft for elevens resultater enn det rangeringseffekten har. Dette samsvarer med tidligere resultater i denne analysen. Samspillet mellom ferdighetsnivå og gruppekarakteristika undersøkes nærmere i Tabell 7.

**Tabell 6: Heterogene ferdighetseffekter.**

Avhengig variabel: resultat avsluttende prøve.

VARIABLER	(1) Lavt- presterende	(2) Middels- presterende	(3) Høyt- presterende
Peer	0.142** (0.060)	0.147*** (0.045)	0.056 (0.061)
Innledende prøve	0.374*** (0.074)	0.480*** (0.054)	0.686*** (0.101)
Rangering	0.114 (0.115)	0.063 (0.067)	-0.043 (0.089)
Kontrollvariabler	Ja	Ja	Ja
FE (skolegruppering)	Ja	Ja	Ja
Konstant	0.071 (0.107)	0.154*** (0.044)	0.185* (0.096)
Observasjoner	1,413	2,730	1,163
R <sup>2</sup>	0.188	0.182	0.137

Tre OLS-regresjoner, der kolonne (1), (2) og (3) rapporterer estimater for henholdsvis lavtpresterende, middelspresterende og høytpresterende elever. Avhengig variabel er resultat på avsluttende prøve. Standardavvik klustres på skolenivå, rapportert i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

Hvordan gruppene og elevene kategoriseres, har ikke noe fasitsvar i denne institusjonelle settingen. En alternativ definering av variablene kan være å skille grupper og elever på gjennomsnittet. Dette betyr to kategorier for hver variabel, henholdsvis homogene og heterogene

grupper, og høyt og lavtpresterende elever. Tabell A5 undersøker noe av betydningen av valg av kategorisering. Kolonne (1) og (2) rapporterer at peer-effekter fortsatt er sterkest i heterogene grupper, og rangeringseffekten er fortsatt signifikant og positiv, men noe redusert. Med utgangspunkt i innledende prøve, er gjennomsnittlig avstand fra den eleven som presterte best til den eleven som presterte dårligst innad i gruppene i kategorien heterogene grupper, nå blitt mindre, noe som forklarer hvorfor rangeringseffekten er redusert. Fra kolonne (1), avdekkes et svakt signifikant positivt estimat av peer-effekter i homogene grupper, som kan forklares med at forskjellen i prestasjonsnivå i de homogene gruppene er nå større med denne kategoriseringen, og inkluderer derfor noe av de positive effektene fra standard grupper avdekket i kolonne (2) i Tabell 5. Kolonne (3) og (4) viser et noe annet resultat enn Tabell 6. Høytpresterende elever har nå en signifikant peer-effekt, og estimatet er høyere enn kolonne (3) fra Tabell 6. Estimaten forteller at ved estimering av peer- og rangeringseffekter er det viktig å skille de som prester godt fra de som presterer svært godt. Elever som ligger over gjennomsnittlig ferdighetsnivå, men ikke presterer blant de 25% høyeste, har positive utslag på avsluttende prøve av å motta smågruppeundervisning i grupper med høyere ferdighetsnivå. Fra Tabell 6 vet vi at de 25% med høyest ferdighetsnivå ikke har signifikant effekt av dette.

I Tabell 7 undersøkes peer- og rangeringseffekter til elever med ulik ferdighetsnivå i ulike typer grupper. I forbindelse med en diskusjon knyttet til hvilke gruppesammensettinger som gir best utgangspunkt for gode elevprestasjoner, og hvordan ulike elever blir påvirket i ulike typer gruppesammensettinger, gir tabellen utgangspunkt for argumentasjoner tilknyttet dette. Homogene gruppesammensettinger etter ferdighetsnivå, argumenteres som positivt fordi undervisningen kan tilpasses til det respektive ferdighetsnivået, samtidig som et negativt argument er at særlig elever med lavt ferdighetsnivå kan oppleve negative effekter av en slik gruppering (Hallinan, 1994; Rui, 2009). Kolonne (1) viser ingen signifikante peer-effekter for noen av ferdighetsnivåene for elever i homogene grupper, og vi bemerker en sterk negativ effekt hos elevene med lavt ferdighetsnivå. Dette er ikke overraskende, da lavtpresterende elever i homogene grupper tilsier et lavt gjennomsnittlig ferdighetsnivå i slike grupper, slik at hvis en ny lavtpresterende elev ankommer gruppa og gjennomsnittlig ferdighetsnivå i gruppa øker, vil dette medføre negative peer-effekter på elevprestasjonene til de andre lavtpresterende elevene i gruppa. Noe overraskende er det at estimatet for rangeringseffekter for lavtpresterende elever i

homogene grupper er såpass negativt, da en homogen gruppesammensetning gir utgangspunkt for at lavtpresterende elever kan oppnå høyere rangering. Merk at effektene ikke er signifikant. Her kontrolleres det ikke for om lærerne tilpasser undervisning tilknyttet ferdighetsnivået i gruppen, slik at en eventuell positiv effekt av dette kunne vært fanget opp i peer-effektene. Det som uansett kan konkluderes, er at ingen av elevgruppene har noen signifikante positive peer- eller rangeringseffekter av å interagere med elever av tilsvarende ferdighetsnivå. Merk at for høytpresterende elever i homogene grupper, har den negative rangeringseffekten større verdi enn den positive peer-effekten. Hadde begge estimatene vært signifikant, kunne dette dratt i retning av en konklusjon som tilsier at høytpresterende elever vil oppleve større negative effekter av at gjennomsnittlig ferdighetsnivå øker i gruppen, hvis gruppen er homogen, da den negative rangeringseffekten dominerer den positive peer-effekten. Videre til kolonne (2) rapporteres positive peer-effekter i standardgrupper for alle ferdighetsnivå, og signifikante for middelspresterende elever, og både peer- og rangeringseffektene har positiv utvikling for alle ferdighetsnivå. Beveger vi oss videre til kolonne (3), estimeres signifikante og sterkere positive peer-effekter for alle tre prestasjonsgruppene i heterogene grupper. Gitt at rangeringen til en elev øker i en heterogen gruppe, viser resultatene at alle ferdighetsnivåer opplever positive effekter av dette med forskjellig signifikans. Peer-effektene er sterkest for middelspresterende elever, og rangeringseffektene er omtrent like for lavt- og middelspresterende elever, med stor verdi. Et resultat som er ekstra bemerkelsesverdig, er positive peer-effekter for høytpresterende elever. Disse peer-effektene, drives ikke av rangeringseffekter, da disse rangeringseffektene er kontrollert for. Nå er det slik at heterogene grupper baseres i dette tilfelle på differansen i ferdighetsnivå fra den mest høytpresterende til den mest lavtpresterende eleven i gruppa, som betyr i prinsippet at en høytpresterende elev kan ha en lavtpresterende elev og tre høytpresterende elever i sin gruppe, og samtidig være rangert som den nest svakeste eleven i gruppa. En mer robust estimering ville derfor vært å kontrollere for antall lavt/middels/høytpresterende elever i hver gruppe. Sett bort i fra dette, vises et tydelig resultatet ut i fra et helhetlig perspektiv, at høytpresterende elever oppnår sterkere peer-effekter ved å oppleve høyere gjennomsnittlig ferdighetsnivå på medelever i smågruppeundervisning med stor spredning fra den svakeste til den sterkeste eleven i gruppa.

**Tabell 7: Kombinasjon av gruppekarakteristika og ferdigheter.**

Avhengig variabel: resultat avsluttende prøve.

VARIABLER	(1) Homogen	(2) Standard	(3) Heterogen
Peer-lavtpresterende	-0.396 (0.331)	0.134 (0.099)	0.223** (0.094)
Rangering-lavtpresterende	-0.347 (0.290)	0.087 (0.175)	0.414* (0.238)
Peer-middelspresterende	-0.022 (0.166)	0.175*** (0.066)	0.263*** (0.072)
Rangering-middelspresterende	-0.117 (0.138)	0.059 (0.098)	0.403** (0.166)
Peer-høytpresterende	0.051 (0.160)	0.065 (0.128)	0.163* (0.088)
Rangering-høytpresterende	-0.118 (0.123)	-0.077 (0.177)	0.221 (0.349)
Innledende prøve	Ja	Ja	Ja
Kontrollvariabler	Ja	Ja	Ja
FE (skolegruppering)	Ja	Ja	Ja

Totalt ni OLS-regresjoner. Kolonne (1) rapporterer estimater for elever i homogene grupper, kolonne (2) rapporterer estimater for elever i standard grupper, kolonne (3) rapporterer estimater for elever i heterogene

grupper. Avhengig variabel er elevens resultat på avsluttende prøve. Standardavvik klustres på skolenivå, rapportert i parenteser. \* $p < 0.1$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ .

Antall lavtpresterende i følgende grupper med tilhørende  $R^2$ -verdi (avrundet til 3 desimaler) til regresjoner:  
homogene (Obs = 231 ,  $R^2 = 0.287$ ), standard (Obs = 725 ,  $R^2 = 0.209$ ) , heterogene (Obs = 457 ,  $R^2 = 0.302$ )

Antall middelspresterende i følgende grupper med tilhørende  $R^2$ -verdi (avrundet til 3 desimaler) til regresjoner:  
homogene (Obs= 699 ,  $R^2 = 0.245$ ) , standard (Obs = 1388 ,  $R^2 = 0.226$ ) , heterogene (Obs = 643 ,  $R^2 = 0.196$ )

Antall høytpresterende i følgende gruppe: homogene (Obs = 397 ,  $R^2 = 0.232$ ) , standard (Obs = 498 ,  $R^2 = 0.237$ )  
heterogene (Obs = 268 ,  $R^2 = 0.221$ )

Det at høytpresterende elever i heterogene grupper opplever positive peer- og rangeringseffekter, i motsetning til høytpresterende elever i homogene grupper, er et svært interessant resultat. Hvis man antar at ved høyere gjennomsnittlig ferdighetsnivå i gruppen, kommer underliggende positive effekter som ikke er redegjort i denne analysen, er dette resultatet noe overraskende. For lavtpresterende elever, avdekkes sterke og negative effekter av homogene grupper, og selv om estimatet ikke er signifikant, gir det en indikasjon at homogene gruppesammensetninger skaper negative effekter på elevprestasjonene til de lavtpresterende elevene. De typiske argumentene for at homogene gruppesammensetninger medfører positive effekter på elevprestasjoner, som beskrevet i kapittel 2.3 og i avsnittet ovenfor, får ikke noen støtte basert på resultatene i denne analysen. Rangeringseffektene for lavtpresterende elever i homogene grupper er negativ, og den positive peer-effekten for høytpresterende elever er høyere i etterhvert som gruppene blir mer heterogene. Med tanke på pareto-effektive peer- og rangeringseffekter, gir Tabell 7 svært interessante resultater. Fra et «peer- og rangeringseffekt synspunkt», bør skoler tilpasse smågruppeundervisning slik at gruppene har tilstrekkelig spredning i ferdighetsnivå, da elever uavhengig av ferdighetsnivå kommer bedre ut prestasjonsmessig av dette. En mer heterogen gruppetilpasning har betydningsfulle positive effekter på lavt- og middelspresterende elever, og effektene får en positiv utvikling etterhvert som vi hopper i nivå av spredning i ferdighetsnivå innad i gruppene, uavhengig av ferdighetsnivå.



## 6 Diskusjon

Resultatene fra Tabell 3 og Tabell 4 rapporterer signifikante lineære peer- og rangeringseffekter. En mulig utfordring før gjennomføring av denne analysen, er spørsmålet knyttet til i hvilken grad barneskoleelever er bevisst på sin egen rangering sammenlignet med sine medelever. En hypotetisk utfordring, spesielt i forsøket på å avdekke skjevheten av peer- effekter uten rangeringseffekter i estimeringen, er nettopp dette. Grunnen til dette er ubevissthet tilknyttet egen rangering muligens medførte skjeve estimater av rangeringseffekter, og som igjen korreleres med estimeringen av peer-effektene. Vurderinger i norsk barneskole baseres ikke på karakter (Forskrift til opplæringslova, 2006, § 3-4). For elevene, kan dette skape et mer utydelig bilde av eget ferdighetsnivå sammenlignet med medelevers ferdighetsnivå. Det at rangeringsvariabelen i tillegg kun baseres på et prøveresultat, samtidig som ferdigheter i andre emner kan virke inn på oppfatning av egen rangering, gir grunnlag til argumentasjon i retning av at rangeringsvariabelen ikke nødvendigvis gjenspeiler et helhetlig bilde av hvordan elever oppfatter seg selv knyttet sine egne medelever. Et annet viktig poeng er at flere vurderingsgrunnlag hadde gitt mer nøyaktige estimater på den sanne rangeringsfordeling i gruppen. Som nevnt i kapittel 3, gir like poengsummer lik rangering i gruppen for elever, og flere vurderingsgrunnlag hadde gjort det mulig å skille disse elevene fra hverandre. Samtidig kunne dette medføre en overestimering av rangeringseffekten, da det er uvisst hvor stor forskjellen mellom elevene må være for rangeringen skal gi en signifikant effekt på prøveresultatene. Fra estimeringene av heterogene effekter, rapporteres sterke rangeringseffekter i heterogene gruppesammensetninger, og ingen i homogene grupper. Ingen signifikante effekter av rangering i homogene grupper tyder på at elever på omtrentlig samme ferdighetsnivå ikke opplever noen rangeringseffekt overfor hverandre. Dette taler til fordel for å gi elever med samme resultat/omtrentlig likt ferdighetsnivå samme rangering.

Som nevnt i kapittel 5, gir den institusjonelle setningen noen utfordringer knyttet til nøyaktigheten av verdiene av rangeringene til elevene. Grunnet ulike grupperinger og at elever utsettes for grupper med ulik størrelse, medfører dette en alternativ beregning av individets rangering. Hvis elevene testes i matematikk i tidsrommet mellom de avslutter en runde med smågruppeundervisning og før de går inn i en ny, samt at tilnærmet de samme prøvene gis til alle

elever, ville dette gitt grunnlaget for en mer presis estimering av rangeringseffekter, hvis skolene praktiserer at smågruppene ikke er konstante. I dette tilfelle ville det vært naturlig å estimere sammenhengene med bruk av fikserte effekter. Alternativt ville det å beholde de samme gruppene i tidsrommet mellom innledende- og avsluttende prøve løst denne utfordringen. Det samme gjelder også beregningen av peer-variabelen, som estimeres på bakgrunn av gjennomsnittlig resultat i gruppa, da ekskludert individ «i». Skjevheten i disse beregningene er trolig noe mindre, da dette estimeres på bakgrunn av en variabel, mens rangeringen til elevene estimeres på bakgrunn av to variabler, resultat på innledende prøve, og antall elever i gruppa. I en situasjon der gruppene ikke er tilfeldig satt sammen (som her), men man beholder de samme gruppene i hver tidsperiode, og en tidsperiode varer fra gjennomføringen av en prøve til gjennomføringen av den neste, kunne bruken av estimatoren presentert i Arellano og Bond (1991) være relevant. Dette er med utgangspunkt i teorien til Anderson og Hsiao (1981) der prøveresultater fra perioden før neste evaluering kunne fungert som instrument for endogen peer-sammensetning. Dette ville bidratt til eksogene peer- og rangeringseffekter. Med få tidsperioder og mange individ, kunne dette vært en interessant estimeringsstrategi for peer- og rangeringseffekter i 1+1 prosjektet.

Videre kan senere analyser av heterogene peer- og rangeringseffekter i 1+1 prosjektet, og andre institusjonelle settinger, studere effektene relativt mellom gruppekategoriene og elevene ferdighetsnivå. Dette gir trolig et enda større sammenligningsgrunnlag for å vurdere «optimal» sammensetning av smågrupper i en undervisningssituasjon med utgangspunkt i peer- og rangeringseffekter. Denne analysen gir også et sammenligningsgrunnlag, men grunnet at kategoriene av grupper og ferdighetsnivå estimeres hver for seg, kan det argumenteres at sammenligningsgrunnlaget er noe svakere. En analyse av sammenligning av relative effekter med tydelige resultater, ville spesielt være nyttig i Tabell 7, da dette ville gitt en sterkere sammenligningsgrunnlag mellom elever og grupper. Likevel vil jeg argumentere for at denne analysen gir en tydelig indikasjon på hvor og for hvem peer- og rangeringseffekter har størst påvirkningskraft på elevprestasjoner. Med tanke på optimale gruppesammensetninger, gir resultatene utgangspunkt for å argumentere for heterogene gruppesammensetninger. Et viktig prinsipp å bemerke, er hvordan en optimalisering av disse resultatene skal foregå i praksis. Med tanke på en smågruppesituasjon, kan det være utfordrende å skape grupper med tilstrekkelig

spredning der høytpresterende elever opplever økt peer-kvalitet, da både elever og peer-kvalitet karakteriseres etter prøveresultater. Eksempelvis, hvis gruppen består av få elever, kan det være utfordrende å skape en situasjon der ferdighetsnivået i gruppen er høy nok til at en høytpresterende elev opplever positiv peer-effekter av dette, samtidig som spredningen i ferdighetsnivå er tilstrekkelig stor innad i gruppen. Videre eksisterer det en vurdering tilknyttet en ytterligere sammenligning av peer- og rangeringseffekter opp i alt dette. Fra Tabell 7 er verdien av de positive rangeringseffektene sterkere for alle elevgrupper, og signifikant for lavt- og middelspresterende. Generelt sett, trekker dette i retning av at rangeringseffektene har sterkere effekt på elevprestasjoner sammenlignet med peer-effekter i heterogene grupper. Basert på resultatene i Tabell 7, vil fremtidige politiske og analytiske vurderinger knyttet til hvordan man kan sikre heterogene gruppesammensetninger med positive peer- og rangeringseffekter for flest mulig elever være høyst relevant.

Elevers evner er et viktig begrep i estimeringen av elevprestasjoner, og studier, blant annet, Chen et al. (2012) og Shi og Qu (2021), viser en sammenheng mellom evner og sentrale utdanningsvariabler. Disse er vanskelig å kontrollere for på en generell basis, og i denne analysen inkluderes ikke noen direkte kontroll for elevers evner. Gitt at evner er relevant, ville en mer utdypende vurdering av hva innledende prøve faktisk kalkulerer knyttet til elevers ferdigheter og evner, skape større robusthet for analysen. For nærmere informasjon om prøvene gjennomført i 1+1 prosjektet, se Bonesrønning et al. (2022, online appendix D). Anta nå at innledende prøve ikke kalkulerer for relevante underliggende effekter som påvirker elevprestasjoner, så er spørsmålet hvorvidt disse effekten er korrelert med peer- og rangeringseffekter, og hvor sterk effekt det har på elevenes prøveresultater. Evnen til å absorbere den positive effekten av det å interagere med høytpresterende elever må være tilstede for å reflektere en positiv peer-effekt på resultatene på avsluttende prøve. Som tidligere diskutert er det naturlig å argumentere for at selvtilliten som oppstår av høyere rangering er en drivkraft for at høyere rangering medfører positive effekter (Bertoni & Nisticò, 2023; Delaney & Devereux, 2022; Denning et al., 2021; Murphy & Weinhardt, 2020). Men denne varierer også på individnivå, og er vanskelig å kontrollere for. Dermed foreligger det en usikkerhet på hvilke underliggende effekter som eventuelt medfører en rangeringseffekt på prøveresultater. En eventuell kontroll av elevers selvtillit kunne gi en mer omfattende bekreftelse knyttet til hvorfor

høyere rangering medfører positive effekter på elevprestasjoner i enkelte situasjoner. Nå viser resultatene fra heterogene grupper, at rangeringseffekter er mindre tilstedeværende hos høytpresterende elever enn for de to andre elevgruppene, som muligens kan forklares med at selvtilliten allerede er eksisterende for høytpresterende elever.

Shi og Qu (2021) presenterer hypoteser tilknyttet til hvordan ulike evner påvirker elevprestasjoner, og sammenhenger som er positive og signifikante presenteres. Eksempler på innholdet i noen av disse hypotesene er relatert til elevens hukommelse, i hvilken grad eleven klarer å behandle informasjon og elevens evne til logisk tenking. Dette beskriver hvor mange uobserverbare faktorer som kan tenkes at påvirker elevprestasjoner. Med tanke på peer-effekter, må elever ha evnen til å utnytte de positive effektene av smågruppeundervisning i grupper med høyt ferdighetsnivå. En måte å utnytte dette på, er evnen til å fokusere på undervisningen i smågruppa, og ikke bli forstyrret av noe annet eller bli ukonsentrert. Dette er nært knyttet til selvkontroll, som igjen er nært tilknyttet elevprestasjoner (Duckworth et al., 2019). Dette kontrolleres ikke for i min analyse, og kan være en kilde til skjevhet i peer-variabelen. Samme prinsipp gjelder også for rangeringsvariabelen. I avsnittet over ble de positive effektene av økt selvtillit diskutert, og tidligere i denne analysen ble positive effekter av å kunne lære bort til andre nevnt som en mulig effekt av høyere rangering. For å kunne ha den positive effekten av å ha muligheten til å lære bort til medelever, følger det intuitivt at man må ha en viss evne til å dele egen kunnskap til sine medelever. Dette varierer på individnivå, og det er naturlig å stille spørsmålet i hvilken grad elever i tidlig grunnskole har denne evnen. Videre vil gruppene variere ut ifra hvordan smågruppelæreren tilpasser undervisningen, og dermed vil man påvirkes av sin peer og sin rangering på ulikt vis.

Videre kontrolleres det ikke for kvaliteter tilknyttet lærere i denne analysen. Hvor mye elever får utbytte av en «type» lærer sammenlignet med en annen type lærer, kan variere fra elev til elev, og er vanskelig å observere. En direkte måte å måle lærerkvalitet, uten at dette nødvendigvis bestemmer om en lærer er god eller ikke, vil eksempelvis være gjennom lærerens utdanningsnivå og erfaring. Dette er observerbare faktorer, og både smågruppelærere og kontaktlærere vil ha direkte påvirkning på elevprestasjoner. Disse varierer typisk fra skole til skole, men kan også variere fra gruppe til gruppe. I utgangspunktet ansatte behandlingsskolene en ekstra lærer for å

gjennomføre smågruppeundervisningen (Institutt for samfunnsforskning, 2023; Nordisk institutt for studier av innovasjon, forskning og utdanning, u.å.; NTNU Samfunnsforskning, u.å.). Dette impliserer at man til en viss grad kan forvente at samme lærer underviser alle gruppene på samme skole, med mindre kontaktlærer og smågruppelærer veksler mellom dette, eller at smågruppelæreren slutter underveis og må erstattes. Det å kontrollere for hvilke karakteristika som skiller de ulike smågruppelærerne kan være relevant i denne analysen, da dette og elevprestasjoner kan være tilknyttet hverandre (Harris & Sass, 2011). Men enda viktigere her, er det å poengtere spørsmålet om hvordan lærerressursene fordeles. Francis et al. (2019) viser til noe begrensede resultater, at lærere med visse positive karakteristika blir til en viss grad prioritert til undervisning av miljøer bestående av karakteristika som forbindes med høyere ferdigheter. Situasjonen er nødvendigvis ikke den samme på grunnskolen, og kanskje ikke like betydningsfull, men sammensetting av grupper bør ikke kun gjøres med utgangspunkt i karaktertrekkene til elevene i gruppa, men man må også vurdere lærerne. I videre studier, der lærerressurser tildeles undervisningssituasjoner i mindre grupper, er det muligens en nødvendig faktor å kontrollere for lærerkvalitet, slik at estimatene av peer-effekter ikke blir skjeve, spesielt da rettet mot lavtpresterende elever i homogene grupper.

I lys av de svært interessante resultatene fra analysen av heterogene peer- og rangeringseffekter, bekrefter ikke denne analysen i hvilken grad rangeringseffekter påvirker kausaliteten til estimeringen av heterogene peer-effekter. En utvidelse av robustheten til analysen er å gjennomføre den samme analysen som Tabell 3 og Tabell 4 på de heterogene effektene, tilsvarende som Bertoni og Nisticò (2023) sin analyse som bekrefter skjevheten også i alternative modellspesifikasjoner og sammenhenger. Gitt at dette også stemmer for denne analysen, presenterer jeg svært interessante resultater som gir grunnlag for optimalisering og pareto-effektiv strukturering av gruppearbeid og smågruppeundervisning i norsk barneskole. Tabell 5 gir utgangspunktet for en argumentasjon av at heterogen sammensetting av grupper maksimerer, til den grad det er mulig, de positive peer- og rangeringseffektene. Fra Tabell 7 ser vi at de største driverne av dette resultatet er lavt- og middelspresterende elever. En relevant argumentasjon basert på dette er at en pareto-effektiv organisering av grupper, med fokus på å maksimere peer- og rangeringseffekter, er en organisering som sikrer stor nok spredning i ferdigheter mellom den eleven som presterer sterkest til den eleven som presterer svakest. Dette skaper positive og

signifikante peer-effekter for alle tre ferdighetsnivå. Det at peer-effekten i heterogene grupper er positiv og noe signifikant for høytpresterende elever er kanskje overraskende, men en like viktig observasjon er at den ikke er negativ og signifikant, eller at det eksistere positive og signifikante peer-effekter for høytpresterende elever i mer homogene gruppesammensettinger. Dette gjør at ingen kommer dårligere ut av en mer heterogen gruppesammensetning, fra et «peer- og rangeringseffekt perspektiv». Resultatene gir derfor utgangspunkt for en pareto-optimal organisering av grupper, som kan medføre ferdighetsnivå til elever i matematikk øker, spesielt hos lavt og middelspresterende elever.

Et annet svært interessant og viktig resultat som rapporteres i Tabell 7, er de sterke rangeringseffektene for lavt- og middelspresterende elever i heterogene grupper. Med utgangspunkt i argumentasjonen i denne analysen på hvorfor rangering medfører positive effekter, forklares disse estimatene med at lavt- og middelspresterende elever opplever en betydningsfull økning av selvtillit eller selvbilde, eller en positiv effekt av å kunne formidle egen kunnskap til elever med lavere ferdighetsnivå enn seg selv. Som nevnt i kapittel 5, betyr disse positive rangeringseffektene at man ikke nødvendigvis trenger mange høytpresterende elever på gruppa for å sikre positive effekter. En lavt og middelspresterende elev får positive effekter av høyere gjennomsnittlig ferdighetsnivå i sin gruppe, samtidig som vedkommende får positive effekter ved å ha høyere ferdighetsnivå enn noen av sine gruppemedlemmer. Rent hypotetisk sett, i en gruppe på 4 elever, kan en elev som ikke tilhører gruppen høytpresterende elever oppleve positive peer-effekter av å ha en høytpresterende elev, og samtidig oppleve positive rangeringseffekter ved å ha 2 elever som presterer lavere enn seg selv, i sin gruppe.

## 7 Konklusjon

I denne analysen estimeres både lineære- og heterogene peer- og rangeringseffekter i 1+1 prosjektet. For det første, viser resultatene at det å estimere peer-effekter uten rangeringseffekter lineært, medfører en underestimering av peer-effekter, når peer-kvalitet defineres som ferdighetsnivå til de man interagerer med. Undersøkelsen gjennomført i forbindelse med dette er basert på analysen til Bertoni og Nisticò (2023), og resultatene fra denne analysen er knyttet til deres analyse, i form av at de begge støtter opp om at det eksisterer en relevant sammenheng mellom peer- og rangeringseffekter. Dette bør kontrolleres for i videre analyser av slike effekter. I min analyse er skjevheten, som oppstår ved ekskludering av den ene effekten i estimering av den andre, tilnærmet presist estimert, med et avvik på 0.001 standardavvik. Disse resultatene estimeres lineært, og jeg finner positive og signifikante peer-effekter, og til en viss grad positive og signifikante rangeringseffekter.

Videre undersøkes heterogene peer- og rangeringseffekter, der lineære regresjoner gjennomføres basert på definisjoner av gruppekarakteristika og elevers ferdighetsnivå. Disse kan gi utgangspunkt for diskusjon for direkte undervisningstiltak for å optimalisere elevprestasjoner i grunnskolen. Resultatene trekker i retning av at tilstrekkelig spredning mellom den svakeste og sterkeste eleven i gruppa skaper positive peer- og rangeringseffekter. Uten å skille mellom prestasjonsnivået til elevene, er det som defineres som heterogene grupper i denne analysen, som skaper sterkest positive effekter. Videre estimeres sterkest peer-effekter for lavt- og middelspresterende elever. For å komme nærmere diskusjonen om hvordan grupper bør optimaliseres, undersøkes hvordan graden av peer- og rangeringseffekter i de ulike elevgruppene påvirker elever med ulikt ferdighetsnivå. Disse effektene er ikke signifikante for alle elever i homogene grupper. I såkalt standard-grupper, opplever middelspresterende elever positive og signifikante peer-effekter, og i heterogene grupper, estimeres positive peer-effekter for alle 3 kategoriene av ferdighetsnivå, med ulik grad av signifikans. Videre estimeres positive rangeringseffekter for elever som ikke presterer på det høyeste nivået. Fra et «peer-effekt perspektiv», skaper grupper med tilstrekkelig spredning i ferdighetsnivå innad i gruppene, positive og signifikante peer-effekter for elever som ikke presterer på det høyeste nivået. Samtidig som det ikke skaper noen signifikante negative effekter på elevprestasjonene til de elevene som i utgangspunktet presterer best.

Denne undersøkelsen bidrar særlig på to områder. For det første, underbygges viktige forutsetninger for å estimere signifikante og kausale peer- og rangeringseffekter. I likhet med resultatene til Bertoni og Nisticò (2023), viser min analyse til en skjevhet i estimerer ved utelatelse av relevante variabler. Dette er en svært essensiell faktor å ta i betraktning for videre undersøkelser av disse effektene i utdanningssammenheng der peer- og rangeringseffekter defineres ut i fra akademiske resultater. Peer- og rangeringseffekter kan kalkuleres med utgangspunkt i andre områder enn akademiske resultater, og i hvilken grad denne skjevheten er relevant på andre definisjonsområder, er et annet spørsmål.

For det andre, gir denne analysen viktige resultater knyttet til å optimalisere elevprestasjoner, spesielt for elever som ikke presterer på det høyeste nivået. Bonesrønning et al. (2022) viser til en signifikant positiv effekt av smågruppeundervisning på elevprestasjoner. Jeg presenterer resultater som kan bidra til optimalisering av positive effekter av smågruppeundervisning, og fra et «peer- og rangeringseffekt perspektiv», kan elevprestasjoner bedres ved å tillate nok spredning i prestasjonsnivå mellom elevene innad i gruppene. Dette er på bakgrunn av at jeg estimerer flere positiv og signifikante effekter på lavt- og middelspresterende elever i heterogene grupper, samtidig som ingen negative peer- og rangeringseffekter for høytpresterende elever rapporteres av dette.



## 8 Referanseliste

Anderson, T. W. & Hsiao, C. (1981). Estimation of Dynamic Models with Error Components.

*Journal of the American Statistical Association*, 76(375), 598-606.

<https://doi.org/10.2307/2287517>

Arellano, M. & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo

Evidence and an Application to Employment Equations. *The Review of Economic*

*Studies*, 58(2), 277-297. <https://doi.org/10.2307/2297968>

Bertoni, M. & Nisticò, R. (2023). Ordinal rank and the structure of ability peer effects. *Journal*

*of Public Economics*, 217, Artikkel 104797.

<https://doi.org/10.1016/j.jpubeco.2022.104797>

Bonesrønning, H., Finseraas, H., Hardoy, I., Iversen, J. M. V., Nyhus, O. H., Opheim, V.,

Salvanes, K. V., Sandsør, A. M. J. & Schøne, P. (2022). Small-group instruction to

improve student performance in mathematics in early grades: Results from a randomized field experiment. *Journal of Public Economics*, 216, Artikkel 104765.

<https://doi.org/10.1016/j.jpubeco.2022.104765>

Booij, A. S., Leuven, E. & Oosterbeek, H. (2017). Ability Peer Effects in University: Evidence

from a Randomized Experiment. *The Review of Economic Studies*, 84(2), 547-578.

<https://doi.org/10.1093/restud/rdw045>

Bramoullé, Y., Djebbari, H. & Fortin, B. (2009). Identification of peer effects through social

networks. *Journal of Econometrics*, 150(1), 41-55.

<https://doi.org/10.1016/j.jeconom.2008.12.021>

- Chen, S.-K., Hwang, F.-M., Yeh, Y.-C. & Lin, S. S. J. (2012). Cognitive ability, academic achievement and academic self-concept: Extending the internal/external frame of reference model. *British Journal of Education Psychology*, 82(2), 308-326.  
<https://doi.org/10.1111/j.2044-8279.2011.02027.x>
- Delaney, J. M. & Devereux, P. J. (2022). Rank Effects in Education: What Do We Know So Far. *IZA DISCUSSION PAPER SERIES*, Artikkel 15128. <https://docs.iza.org/dp15128.pdf>
- Denning, J. T., Murphy, R. & Weinhardt, F. (2021). Class Rank and Long-Run Outcomes. *The Review of Economics and Statistics*, 1-45. [https://doi.org/10.1162/rest\\_a\\_01125](https://doi.org/10.1162/rest_a_01125)
- Duckworth, A. L., Taxer, J. L., Eskreis-Winkler, L., Galla, B. M. & Gross, J. J. (2019). Self-Control and Academic Achievement. *Annual Review of Psychology*, 70, 373-399.  
<https://doi.org/10.1146/annurev-psych-010418-103230>
- Duflo, E., Dupas, P. & Kremer, M. (2011). Peer Effects, Teacher Incentives, and the Impact of Tracking: Evidence from a Randomized Evaluation in Kenya. *American Economic Review*, 101(5), 1739-1774. <https://doi.org/10.1257/aer.101.5.1739>
- Elsner, B., Ispording, I. E. & Zölitz, U. (2021). Achievement Rank Affects Performance and Major Choices in College. *The Economic Journal*, 131(640), 3182-3206.  
<https://doi.org/10.1093/ej/ueab034>
- FN-sambandet. (2022, 15. desember). *God utdanning*. <https://www.fn.no/om-fn/fns-baerekraftsmaal/god-utdanning>
- Forskningsrådet. (u.å.). *Small group Instruction in Mathematics for Pupils Level 1-4: Effects of a randomized controlled trial intervention study*.  
<https://prosjektbanken.forskningsradet.no/project/FORISS/256217?Kilde=FORISS&distribution=Ar&chart=bar&calcType=funding&Sprak=no&sortBy=score&sortOrder=desc&resultCount=30&offset=0&Fritekst=256217>

Forskrift til opplæringslova. (2006). *Forskrift til opplæringslova* (FOR-2006-06-23-724).

Lovdata. <https://lovdata.no/LTI/forskrift/2006-06-23-724>

Francis, B., Hodgen, J., Craig, N., Taylor, B., Archer, L., Mazenod, A., Tereshchenko, A. & Connolly, P. (2019). Teacher 'quality' and attainment grouping: The role of within-school teacher deployment in social and educational inequality. *Teaching and Teacher Education*, 77, 183-192. <https://doi.org/10.1016/j.tate.2018.10.001>

Granger, B. (u.å.). *The Protégé Effect: How Doing for Others Makes Us Do More*. Supporti. <https://www.getsupporti.com/post/protege-effect-doing-for-others>

Hallinan, M. T. (1994). Tracking: From Theory to Practice. *Sociology of Education*, 67(2), 79-84. <https://doi.org/10.2307/2112697>

Harris, D. N. & Sass, T. R. (2011). Teacher training, teacher quality and student achievement. *Journal of Public Economics*, 95(7-8), 798-812. <https://doi.org/10.1016/j.jpubeco.2010.11.009>

Hong, S. C. & Lee, J. (2017). Who is sitting next to you? Peer effects inside the classroom. *Quantitative Economics*, 8(1), 239-275. <https://doi.org/10.3982/QE434>

Institutt for samfunnsforskning. (2023, 27.april ). *1+1 prosjektet*. Hentet 3. mai 2023 fra <https://www.samfunnsforskning.no/prosjekter/11-prosjektet.html>

Lee, D. S. & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature*, 48(2), 281-355. <https://doi.org/10.1257/jel.48.2.281>

- Loveless, T. (2013). *HOW WELL ARE AMERICAN STUDENTS LEARNING? With sections on the latest international tests, tracking and ability grouping, and advanced math in 8th grade* (2). The 2013 Brown Center Report on American Education. <https://www.brookings.edu/wp-content/uploads/2016/06/2013-brown-center-report-web-3.pdf>
- Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies* 60(3), 531-542. <https://doi.org/10.2307/2298123>
- Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of Educational Psychology*, 79(3), 280-295. <https://doi.org/10.1037/0022-0663.79.3.280>
- Marsh, H. W. & Parker, J. W. (1984). Determinants of student self-concept: Is it better to be a relatively large fish in a small pond even if you don't learn to swim as well? *Journal of Personality and Social Psychology*, 47(1), 213-231. <https://doi.org/10.1037/0022-3514.47.1.213>
- McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27, 415-444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Murphy, R. & Weinhardt, F. (2020). Top of the Class: The Importance of Ordinal Rank. *The Review of Economic Studies*, 87(6), 2777-2826. <https://doi.org/10.1093/restud/rdaa020>
- Nordisk institutt for studier av innovasjon, forskning og utdanning,. (u.å.). *I+I prosjektet*. <https://www.nifu.no/projects/11-prosjektet/>
- NTNU Samfunnsforskning. (u.å.). *Smågruppeundervisning i matematikk for elever på småskoletrinnet*. <https://samforsk.no/prosjekter/smagruppeundervisning-i-matematikk-for-elever-pa-smaskoletrinnet>

- Opplæringslova-oppl. (1998). *Lov om grunnskolen og den vidaregåande opplæringa* (LOV-1998-07-17-61). Lovdata. <https://lovdata.no/lov/1998-07-17-61>
- Paloyo, A. R. (2020). Peer effects in education: recent empirical evidence. I S. Bradley & C. Green (Red.), *THE ECONOMICS OF EDUCATION* (2. utg., s. 291-305). Academic Press. <https://doi.org/10.1016/B978-0-12-815391-8.00021-5>
- Paul, A. M. (2011, 30. november). The Protégé Effect. Why teaching someone else is the best way to learn. *TIME*. <https://ideas.time.com/2011/11/30/the-protege-effect/>
- Pei, Z., Pischke, J.-S. & Schwandt, H. (2019). Poorly Measured Confounders are More Useful on the Left than on the Right. *Journal of Business & Economic Statistics*, 37(2), 205-216. <https://doi.org/10.1080/07350015.2018.1462710>
- Rui, N. (2009). Four decades of research on the effects of detracking reform: Where do we stand?—A systematic review of the evidence. *Journal of Evidence-Based Medicine*, 2(3), 164-183. <https://doi.org/10.1111/j.1756-5391.2009.01032.x>
- Shatz, I. (u.å.). *The Protégé Effect: How You Can Learn by Teaching Others*. Effectiviology. <https://effectiviology.com/protege-effect-learn-by-teaching/>
- Shi, Y. & Qu, S. (2021). Cognitive Ability and Self-Control's Influence on High School Students' Comprehensive Academic Performance. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.783673>
- Vardardottir, A. (2013). Peer effects and academic achievement: a regression discontinuity approach. *Economics of Education Review*, 36, 108-121. <https://doi.org/10.1016/j.econedurev.2013.06.011>
- Wooldridge, J. M. (2020). *Introductory Econometrics: A Modern Approach* (7. utg.). Cengage Learning.



## 9 Appendix

**Tabell A1: Lineær estimering av rangeringseffekter**

VARIABLER	(1) Avsluttende- prøve	(2) Avsluttende- prøve	(3) Peer
Rangering	0.575*** (0.033)	-0.136*** (0.034)	-1.454*** (0.041)
Innledende prøve		0.600*** (0.021)	0.967*** (0.015)
Kontrollvariabler	Ja	Ja	Ja
FE (skolegruppering)	Ja	Ja	Ja
Konstant	0.060 (0.036)	0.324*** (0.031)	0.858*** (0.023)
Observasjoner	5,306	5,306	5,306
R <sup>2</sup>	0.143	0.421	0.809

Tre OLS-regresjoner. Kolonne (1) viser rangeringseffekter uten inkludering innledende prøve der avhengig variabel er resultat på avsluttende prøve. Kolonne (2), viser samme regresjon, bortsett fra at innledende prøve inkluderes. Kolonne (3) viser sammenheng mellom rangering og peer, der peer er avhengig variabel. Standardavvik klustres på skolenivå, rapportert i parenteser. \*p<0.1, \*\*p<0.05,\*\*\* p<0.01

**Tabell A2: Detaljerte estimater fra Tabell 4.**

VARIABLER	(1)	(2)	(3)	(4)
Peer	0.131*** (0.021)	0.121*** (0.021)	0.162*** (0.030)	0.153*** (0.030)
Innledende prøve	0.505*** (0.017)	0.489*** (0.017)	0.468*** (0.028)	0.452*** (0.028)
Elev-født-utland		-0.053 (0.049)		-0.051 (0.049)
Foreldre-født-utland		-0.133*** (0.042)		-0.133*** (0.042)
Kvinne		0.003 (0.022)		0.003 (0.022)
Utd-foreldre_1		-0.295*** (0.069)		-0.296*** (0.069)
Utd-foreldre_2		-0.227*** (0.036)		-0.227*** (0.036)
Utd-foreldre_3		-0.101*** (0.024)		-0.100*** (0.024)
Utd-foreldre_4		-		-
Rangering			0.085* (0.048)	0.085* (0.047)
FE (Skolegruppering)	Ja	Ja	Ja	Ja
Konstant	0.127*** (0.024)	0.241*** (0.026)	0.080** (0.035)	0.193*** (0.036)
Observasjoner	5,306	5,306	5,306	5,306
R <sup>2</sup>	0.415	0.425	0.415	0.426

Tabellen er en mer detaljert versjon av Tabell 4, med de samme regresjonene. Standardavvik klustres på skolenivå, rapportert i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.



**Tabell A3: Alternativ definisjon av rangering og gruppestørrelse.**

Avhengig variabel: resultat på avsluttende prøve

VARIABLER	(1)	(2)
Peer	0.153*** (0.030)	0.155*** (0.029)
Innledende prøve	0.452*** (0.028)	0.449*** (0.028)
Rangering-max	0.086* (0.046)	
Rangering-min		0.090* (0.046)
Kontrollvariabler	Ja	Ja
FE (skolegruppering)	Ja	Ja
Konstant	0.186*** (0.036)	0.185*** (0.034)
Observasjoner	5,519	5,519
R <sup>2</sup>	0.426	0.427

Tabellen viser to OLS-regresjoner, der elever som ikke hadde en hyppig observasjon av rangering og/eller gruppestørrelse inkluderes i estimeringene. Avhengig variabel er elevens resultat på avsluttende prøve. Kolonne (1) viser maksimums-verdier, og kolonne (2) viser minimumsverdier. Standardavvik klustres på skolenivå, rapporteres i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

**Tabell A4: Effekt av gruppeintervaller.**

Avhengig variabel: resultat på avsluttende prøve

VARIABLER	$N \in [2, 6]$		$N \in [4, 6]$		$N \in [4, 8]$	
	(1)	(2)	(3)	(4)	(5)	(6)
Peer	0.123*** (0.022)	0.154*** (0.030)	0.133*** (0.024)	0.165*** (0.033)	0.130*** (0.024)	0.163*** (0.033)
Innledende prøve	0.487*** (0.017)	0.451*** (0.028)	0.483*** (0.019)	0.445*** (0.032)	0.486*** (0.019)	0.447*** (0.032)
Rangering		0.084* (0.047)		0.087 (0.053)		0.089* (0.053)
Kontrollvariabler	Ja	Ja	Ja	Ja	Ja	Ja
Fe (skolegruppering)	Ja	Ja	Ja	Ja	Ja	Ja
Konstant	0.242*** (0.027)	0.195*** (0.036)	0.249*** (0.027)	0.200*** (0.040)	0.248*** (0.026)	0.199*** (0.040)
Observasjoner	5,262	5,262	4,688	4,688	4,732	4,732
R <sup>2</sup>	0.424	0.424	0.424	0.425	0.426	0.426

Tabellen viser totalt seks OLS-regresjoner som undersøker hvordan peer- og rangeringseffekter påvirkes av ulike intervaller på gruppestørrelse. Avhengig variabel er elevens resultat på avsluttende prøve. Standardavvik klustres på skolenivå, rapporteres i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

**Tabell A5: Alternativ definisjon av elev- og gruppekarakteristika.**

Avhengig variabel: resultat på avsluttende prøve.

VARIABLER	(1)	(2)	(3)	(4)
	Homogen	Heterogen	Lavt- presterende	Høyt- presterende
Peer	0.088* (0.052)	0.202*** (0.044)	0.150*** (0.050)	0.123*** (0.038)
Innledende prøve	0.538*** (0.059)	0.407*** (0.038)	0.414*** (0.053)	0.505*** (0.049)
Rangering	-0.008 (0.068)	0.210** (0.084)	0.071 (0.089)	0.052 (0.060)
Kontrollvariabler	Ja	Ja	Ja	Ja
FE (skolegruppering)	Ja	Ja	Ja	Ja
Konstant	0.214*** (0.052)	0.164*** (0.061)	0.109* (0.064)	0.223*** (0.050)
Observasjoner	2,900	2,406	2,320	2,986
R <sup>2</sup>	0.450	0.414	0.231	0.206

Tabellen viser 4 OLS-regresjoner for alternativ spesifisering av grupper og prestasjonsnivå til elevene, som skiller på gjennomsnittsverdier. Avhengig variabel er elevens resultat på avsluttende prøve. Standardavvik klustres på skolenivå, rapporteres i parenteser. \*p<0.1, \*\*p<0.05, \*\*\*p<0.01.

