



Språkkompetanse hos 4-åringer som har gått i barnehage

Foreløpige resultater fra forskningsprosjektet *Barns sosiale utvikling*

Atferdssenteret-Unirand

Rapport til Utdanningsdirektoratet

15. November 2013

Revidert versjon 3. Januar 2014

Henrik Daae Zachrisson, Eric Dearing, Imac M. Zambrana, & Ane Nærde

Atferdssenteret, Pb 7053 Majorstuen, 0306 Oslo.

E-post: hdzachrisson@atferdssenteret.no

Bakgrunn for rapporten

De fleste barn i Norge går i barnehagen. I 2012 hadde hele 90 prosent av alle landets 1- til 5-åringer barnehageplass. Internasjonal forskning, og enkelte resultater fra Norge, tyder på at det å gå i barnehage kan bidra positivt til barns språkutvikling. Det er imidlertid behov for mer forskning som spesifikt fokuserer på betydningen av barnehage for barns språkutvikling i Norge. Forskning om barn i norske barnehager kan gi viktig kunnskap som vil fungere som premissleverandør for videre utvikling av det nasjonale barnehagetilbudet og for samfunnsmessige prioriteringer.

Forskere ved Atferdssenteret har på oppdrag fra Kunnskapsdepartementet tidligere skrevet tre rapporter om bruk av barnehage, strukturelle kvalitetsindikatorer og barns sosiale og atferdsmessige utvikling. Denne rapporten, som skrevet på bestilling fra Utdanningsdirektoratet, er den første som undersøker betydningen av barnehage for barns språkutvikling.

Problemstillingene er utarbeidet av Utdanningsdirektoratet i samarbeid med prosjektgruppen ved Atferdssenteret. Prosjektgruppen består av Henrik Daae Zachrisson, seniorforsker ved Atferdssenteret, og hovedansvarlig for rapporten, Eric Dearing, Associate Professor ved Boston College og seniorforsker ved Atferdssenteret, Imac M. Zambrana, Post Doc., og Ane Nærde, forsker, begge ved Atferdssenteret. Rapporten er kvalitetssikret av Martina Narayanan, Post Doc. ved Atferdssenteret. Rapporten ble levert til Utdanningsdirektoratet 15. November 2013, og har blitt revidert etter tilbakemeldinger fra Direktoratet. Revisjonene besto av presiseringer i innledningen, metoddelen, og diskusjonen, noen små språklige endringer gjennom rapporten, samt en referanse i innledningen til en relevant studie som ble publisert etter at rapporten ble levert.

Analysene bygger på data fra det longitudinelle forskningsprosjektet *Barns sosiale utvikling*, ved Atferdssenteret, også kjent under det engelske akronymet BONDS (Behavior Outlook Norwegian Developmental Study). I prosjektet, som startet i 2006, har vi fulgt 1157 uselekterte barn og deres familier fra 5 ulike kommuner i Norge fra barna var 6 måneder gamle. Det er samlet inn omfattende multimetode informasjon om ulike forhold i familien og ved barnet, og om de barnehagene barna har gått i. De eldste barna i prosjektet er nå 7 år. *Barns sosiale utvikling* har fått støtte fra Norges Forskningsråd til å følge barna ut 2. klasse.

Innledning

Ett av de mer robuste funn fra internasjonal forskning om betydningen av barnehage for barns utvikling er at kvalitetsbarnehager fremmer kognitiv utvikling generelt, og språkutvikling spesielt (Zachrisson, Lekhal, & Schjølberg, 2010). Eksempelvis har en rekke artikler fra den omfattende National Institute of Child Health and Human Development (NICHD) The Study of Early Child Care and Youth Development (SECCYD) at barn som går i barnehager av god kvalitet bla. har bedre språkkompetanse enn barn som ikke går i barnehage eller som går i barnehager av dårlig kvalitet, både gjennom de tidlige barneårene (f.eks., NICHD Early Child Care Network, 2006), og opp i skolealderen (Belsky et al., 2007). I denne studien defineres barnehagekvalitet gjennom et globalt prosessmål som legger vekt på stimulerende og omsorgsfulle relasjoner mellom barn og voksne. Tilsvarende sammenhenger hva gjelder bruk av barnehage har også blitt funnet i Norge, hvor en artikkel har vist at barn som gikk i barnehage hadde lavere sannsynlighet for å ha forsinket språkutvikling ved 3 års alder (Lekhal, Zachrisson, Wang, Schjølberg, & von Soest, 2010).

I en sosiopolitisk kontekst som den Norske har de alle fleste barn noe erfaring med barnehage før de starter på skolen. Snarere enn å vektlegge betydningen av hvorvidt barn går i barnehage eller ikke, ser vi det som mest relevant å undersøke hvorvidt mengden av barnehageerfaring (dvs. hvor mange år barnet har gått i barnehage) er av betydning for språkutvikling. Noe internasjonal forskning har undersøkt betydningen av dette for barns kognitive utvikling, men så vidt vi kjenner til har kun én artikkel, nylig publisert, spesifikt omhandlet betydningen for språkutvikling. Artikkelen er basert på NICHD SECCYD, og viser at barn som gikk i barnehage av god kvalitet *både* i småbarnsalder (opp til 3 år) og i førskolealder (3-5 år) hadde bedre språklige, kognitive, og pre-akademiske ferdigheter enn barn som gikk i kvalitetsbarnehage i kun en av disse periodene (Li, Farkas, Duncan, Burchinal, Vandell, 2013). Videre fant Côté og kollegaer (Cote, Doyle, Petitclerc, & Timmins, 2013) i det Britiske The Millennium Cohort Study at barn som startet i barnehage eller andre barnepassordninger før de fylte 9 måneder hadde bedre generell kognitiv utvikling ved 3 års alder, men at denne sammenhengen senere forsvant. Loeb og kollegaer (Loeb, Bridges, Bassok, Fuller, & Rumberger, 2007) fant i Amerikanske data at de barna som startet i barnehage (pre-K) mellom 2 og 3 års alder hadde bedre skoleferdigheter i begynnelsen av Kindergarten (dvs. ca 5 års alder) enn barn som startet tidligere eller senere. Ettersom vi i Norge har ganske stor spredning i alder ved barnehagestart, men med en hovedvekt rundt ett år, er det derfor relevant å undersøke om barn som tilbringer flere eller færre år i barnehage har ulik språkkompetanse, og om det kan være «optimale aldre» for barnehagestart med hensyn til å fremme barns språkutvikling.

Det finnes videre en omfattende internasjonal forskning, både fra observasjonelle studier (f.eks., Dearing, McCartney, & Taylor, 2009) og meta-analyser av eksperimentelle studier (Barnett, 2011; Camilli, Vargas, Ryan, & Barnett, 2010), som viser at barn fra familier i sosial risiko har relativt sett større språklig utbytte av å gå i barnehager (inkludert pre-K) enn barn fra familier uten sosial risiko. Dette er en viktig

premiss i Norsk politikktutvikling (Ministry of Education, 2007), men har—etter hva vi kjenner til—ikke blitt undersøkt i Norge.

Som nevnt over er det et konsistent funn i internasjonal forskning at barnehagekvalitet er viktig for barns språkutvikling. I Norge er noen aspekter ved strukturell kvalitet i barnehagen regulert gjennom lov (bla. rammeplan), eller gjennom forskrifter (pedagogtethet). Det ser også ut som om strukturell kvalitet er relativt homogent på tvers av barnehager (Winsvold & Gulbrandsen, 2009). Likefullt er det andre aspekter ved strukturell kvalitet, så som alderssammensetning i barnegruppen og gruppestørrelse, som ikke er regulert.. Vi kjenner ikke til forskning som har undersøkt betydningen av disse faktorene spesifikt for språkutvikling. Det er imidlertid i språklitteraturen etablert kunnskap at eksponering for mye språklig input med av høy kvalitet i tidlig alder er relatert til bedre språkferdigheter (Hart & Risley, 1995; Oshima-Takane & Robbins, 2003). Sensitiv stimulering i foreldre-barn interaksjoner har typisk blitt knyttet positivt til barns tidlige språkutvikling (Fidalgo & Pereira, 2005; Huttenlocher, Vasilyeva, Waterfall, Vevea, & Hedges, 2007). Samtidig viser litteraturen at det å ha eldre søsken og dermed være del av større søskenflokker fordelt på de voksne der hjemme kan medføre mer samtaleerfaringer som kan gi sosiale og konversasjonelle fordeler (Oshima-Takane & Robbins, 2003; Pine, 1995; se også: Zambrana, Ystrom, & Pons, 2012). Nettopp på grunn av slike antakelser er det viktig å undersøke hvilken rolle barnehagen kan ha for barns språkutvikling, tatt i betraktning de unike sosiale erfaringene barn får der.

Videre viser forskning, bla. fra Norge, at gutter og jenter kan ha ulik språkutvikling gjennom småbarns- og førskolealderen (f.eks., Zambrana et al., 2012; Zambrana, Ystrøm, Schjølberg, & Pons, 2013a). I gjennomsnitt har jenter bedre språkkompetanse enn gutter i småbarnsalderen, mens kjønnsforskjellene minker etter dette (Fenson et al., 2007). Men selv om kjønnsforskjellene minsker over tid, kan miljøfaktorer (som f.eks. språkstimulering i en barnehagekontekst) ha ulik betydning for språkutviklingen til jenter og gutter. Gutteres økte risiko for språkforsinkelser kan dermed være knyttet til en økt sensitivitet for miljøpåvirkning sammenlignet med jenter (Zambrana et al., 2012), hvilket kan være relatert til biologiske eller genetiske forskjeller (Spinath, Price, Dale, & Plomin, 2004). For å kunne forstå mer om hvorvidt det sosiale barnehagemiljøet kan ha ulik betydning for språkkompetanse innad i jente- og guttegrupper selv om det ikke er forskjeller på tvers av kjønn, vil vi gjøre analyser separat for gutter og jenter.

I denne rapporten vil vi undersøke sammenhenger mellom barnehagebruk og språkkompetanse i et normalutvalg av Norske 4-åringer. Dernest vil vi undersøke om det er ulik sammenheng mellom barnehagebruk og språkkompetanse for barn fra familier i sosial risiko, sammenlignet med barn fra familier uten sosial risiko. Til slutt vil vi også undersøke sammenhengen mellom enkelte strukturelle kvaliteter ved barnehagen (gruppestørrelse og alderssammensetning i barnegruppen) og språkkompetanse.

Metode

Utvalg

Vi har brukt data fra *Barns sosiale utvikling*, en pågående longitudinell studie av 1157 barn (558 jenter og 559 gutter) fra 5 kommuner sør-øst i Norge (Porsgrunn, Skien, Bamble, Tinn og Drammen). Barna utgjør 79% av 1465 barn som ble invitert til å delta, og 60% av de 1931 barna hvis familier ble informert om prosjektet.

Foreldrene ble informert om studien på 5-månederskontrollen på den lokale helsestasjonen og forespurt av en helsesøster om de kunne tenke seg å bli kontaktet av en intervjuer. Inklusjonskriteriene var at barnet skulle være i passende alder, og at minst en av foreldrene skulle kunne delta på intervjuer uten tolk. Familiene som ønsket å delta oppga kontaktinformasjon, og ble senere kontaktet av en lokal intervjuer som gjorde avtale om et intervju når barnet ble 6 måneder. Begge foreldrene ble inviterte til å delta på dette første intervjuet. Fra 6-månedersintervjuet har vi data fra 1135 mødre og 679 fedre, hvorav 659 familier hvor vi har data fra begge foreldre. Familiene har videre blitt intervjuet når barnet fylte 1, 2, 3, og 4 år, samt gjennom flere korte telefonintervjuer med noen måneders mellomrom. Ved 1, 2, og 3 år ble det også gjennomført strukturerte samspills-observasjoner med forelder og barn. Ved 4 år gjennomførte barna testing av språkkompetanse, motorisk utvikling og selv-regulering. Testen av barnas språkkompetanse danner utgangspunktet for denne rapporten.

Informasjon om barnehagestart har blitt innhentet fra foreldrene. Alle barnehagene som barna går i har også ved gjentatte anledninger bidratt med informasjon om barnet og om ulike forhold ved barnehagene. For samtlige barn som går i barnehage har pedagogisk leder blitt invitert til å besvare spørreskjemaer om barnets atferd og trivsel i barnehagen når dette var 2, 3, og 4 år. For en fullstendig oversikt over datainnsamling og design av studien *Barns sosiale utvikling*, se Nærde, Janson & Ogden (innsendt).

Mål

British Picture Vocabulary Scale. I forbindelse med foreldreintervjuene ved 4 år testet vi barnas språkferdigheter ved hjelp av en norsk oversettelse av det veletablerte språkmålet British Picture Vocabulary Scale, versjon II (BPVS-II; Dunn, Dunn, Whetton, & Burleu, 1997; se også; Lyster, Horn, & Rygvold, 2010). BPVS er basert på den amerikanske Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 1997), og er utformet for å teste variasjon i språkferdigheter hos barn mellom 3 og 16 år. Denne testen er derfor spesielt velegnet til bruk i longitudinelle studier hvor man ønsker å følge barns utvikling over tid. Den norske oversettelsen består av 12 sett (av opprinnelig 14) med oppgaver med stigende vanskelighetsgrad, hvorav hvert sett har 12 deloppgaver. I hver deloppgave presenteres barna for fire bilder av gangen, og de blir spurt om hvilket bilde som forestiller det ordet som intervjueren angir. Det er viktig å understreke at BPVS-II primært er et mål på impressivt vokabular eller begrepsforståelse. For det første må ikke barnet produsere ordene selv, men kan ta i bruk nonverbale gester som å peke eller indikere

korrekt bilde ved å svare for eksempel ”det” eller ”der”, slik at det ikke er et mål på språkproduksjon. For det andre, selv om vi for enkelthetskyld bruker begrepet «språkkompetanse» som betegnelse for BPVS-II skåre i rapporten, representerer begreper eller vokabular altså strengt tatt kun *et* aspekt ved språkkompetanse.

I administrasjonen av testen estimeres en totalskåre på barnets impressive vokabular ved at antall feil trekkes fra det totale antall mulige riktige svar (testen avsluttes på det settet hvor barnet får ≥ 8 feil). Testingen begynner på det settet som har en vanskelighetsgrad tilsvarende barnas alder. For vårt utvalg startet testen på sett 2 når tester antok at sett 1 ville være for lett, og i tråd med manualen ble sett 1 også gjennomført dersom barnet ble testet fra sett 2 og her fikk ≥ 2 feil. Testen tar ca. 30 minutter å fullføre (avhengig av hvor mange sett barnet fullfører). Validitet og reliabiliteten til den britiske versjonen av BPVS-II er svært robust (Dunn og kollegaer, 1997). Da Lyster og kollegaer (2010) testet ut den norske oversettelsen på et utvalg av 884 barn og unge i aldersspennet 3-16 år, fant de at reliabiliteten til denne versjonen (den indre konsistensen til deloppgavene for hvert sett) var tilfredsstillende ($\alpha = .98$). Videre sammenlignet de norske forskerne barnas BPVS-II skårer med de på TROG 2, som er et annet impressivt språkmål. Det var høye korrelasjoner mellom de to målene, noe som gir et bilde av validiteten, og det var en særlig sterk sammenheng på de laveste alderstrinnene (Lyster og kollegaer, 2010).

Behandling av manglende besvarelser. Det var totalt 942 barn på 4 år som ble testet med BPVS-II. Data fra i alt 14 barn ble slettet grunnet feiladministrasjon eller feilpunching, og vi har dermed BPVS-II data fra totalt 928 4-åringer. I 93 tilfeller der testingen ble startet på sett 2 og barna fikk ≥ 2 feil, ble allikevel ikke sett 1 administrert. For disse barna valgte vi derfor å estimere missing på sett 1 ved å benytte informasjonen fra sett 2 før vi slo sammen de ulike settene til en total skåre. Videre fikk to barn skåret et ekstra sett til tross for ≥ 8 feil, slik at de fikk en noe forhøyet skåre. Ni barn ble derimot ikke testet på neste sett, selv om de hadde < 8 feil, men da vi ikke fikk imputert disse dataene får barna en noe lavere skåre enn de skulle hatt. Disse to sistnevnte aspektene har imidlertid liten betydning for den totale variasjonen og dermed ingen nevneverdig innvirkning på det endelige resultatet.

Tid tilbrakt i barnehage. Vi definerer tid tilbrakt i barnehage som hvor lenge barn totalt sett har gått i barnehage frem til fylte fire år. I de personlige intervjuene med foreldrene ved 1, 2, 3 og 4 år har de rapportert om type barnepass, inkludert barnehage. I tillegg har foreldre oppgitt nøyaktig dato for barnehagestart. Ved hjelp av denne informasjonen har vi definert barn som tilhørende fem gjensidig utelukkende grupper; (1) helt uten barnehageerfaring ved 4 år; (2) mindre enn 1 års barnehageerfaring (dvs start etter fylte 3 år, men før fylte 4 år); (3) mindre enn 2 års barnehageerfaring (dvs start etter fylte 2 år, men før fylte 3 år); (4) mindre enn 3 års barnehageerfaring (dvs start etter fylte 1 år, men før fylte 2 år); og (5) mer enn 3 års barnehageerfaring (dvs start før fylte 1 år).

Familievariabler målt ved 6 mnd. Hovedhensikten med å inkludere bakgrunnsvariabler er å justere analysene for observerte tredjevariabler som kan påvirke sammenhengen mellom barnehagebruk/

barnehagevariabler og barnas språkkompetanse. Vi benytter derfor familievariabler som ble innhentet før barna hadde startet i barnehage, dvs. ved personlig intervju ved 6 måneders alder. Vi har valgt ut variabler som med rimelig grunn kan antas å være relatert til både barnehagebruk og barnets språkkompetanse. Dette inkluderer mors alder (i år), foreldrenes høyeste utdanning (i år) gjennomført av mor eller far, mors yrkesdeltagelse før svangerskapspermisjon og fars yrkesdeltagelse ved 6 mnd. intervju (begge kodet 0=ikke i arbeid, 1=i arbeid), kun en forelder i husholdningen (kodet 1= en forelder, 0=to foreldre), ikke-vestlig innvandrerbakgrunn (Tyrkia, Asia, Afrika, Sør- og Mellom-Amerika; kodet 1=ja, 0=nei), vestlig innvandrerbakgrunn (Europa eksklusive Tyrkia, Nord-Amerika, Oceania; kodet 1=ja, 0=nei), om mor mottar noen form for trygdeytelser i stedet for å arbeide/studere. Videre ble mors symptomer på angst og depresjon målt med en 13-spørsmåls versjon av Hopkins Symptom Check List (SCL-13). Spørsmålene er et utvalg fra 25-spørsmålsversjonen, som er en mye brukt skala for å måle psykiske vansker i ikke-kliniske utvalg (Strand, Dalgard, Tambs, & Rognerud, 2003; Tambs & Moum, 1993). Svarformatet er en Likert-skala fra ”ikke plaget i det hele tatt” (1) til ”veldig plaget” (4), Gjennomsnittet av de 13 spørsmålene reflekterer egenrapportert nivå av angst og depresjon. Familiens boligsituasjon måler vi i med en indeks som summerer om de er misfornøyde med boligsituasjonen, om de bor i ett- eller toromsleilighet, og om de leier bolig. Indeksen varierer fra 0 (god boligstandard) til 3 (møter alle kriteriene for lav boligstandard). For enkelte analyser konstruerte vi også en kumulativ risikoindeks, som summerer tilstedeværelsen av følgende risiki: Ikke-vestlig innvandrer, fødselsvekt < 2.5 kg, arbeidsløs far ved 6 mnd alder, mors alder < 21 år, foreldre uten fullført videregående skole, boligproblemer (skåre 3 eller 4), enslig forelder ved 6 mnd. alder. En skåre på 1 betyr at barnet har 1 av disse risikoene, en skåre på 2 at barnet har 2, osv.

Barnevariabler. Vi inkluderer følgende informasjon om barnet: Kjønn (kodet jente= 0, gutt= 1), eldre søsken i familien ved tidspunktet for barnets fødsel (nei=0, ja=1), og fødselsvekt i kg. Barnets generelle utviklingsnivå ved 6 mnd ble målt med Ages and Stages Questionnaire (ASQ; Squires, Bricker, & Potter, 1999) i norsk oversettelse (Janson & (with Smith, 2003). Mor besvarte spørsmål om barnets utvikling på en 3-punkts skala som ble skåret 0 (ikke ennå), 5 (noen ganger), eller 10 (ja). To av de 30 leddene i den originale skalaen ble ekskludert. Vi beregnet den totale ASQ-skåren som et gjennomsnitt av de 28 leddene, multipliserte med 30 og rundet opp til nærmeste 5-poengs nivå for å oppnå samme verdier som i den originale skalaen (Janson & (with Smith, 2003).

Barnehagevariabler. Spørreskjemaet om barnet som pedagogisk leder fylte ut ved 2, 3, og 4 år inneholdt også enkelte spørsmål om kvaliteter ved barnehagen og barnegruppen. Vi inkluderte gruppestørrelse ved alle tre tidspunkter og alderssammensetning i barnegruppen (småbarnsgruppe, [eldste barn i gruppen er 3 år], blandet aldersgruppe, samt rene storebarnsgrupper [yngste barn var 3 år]), ved 2 og 3 år.

Bakgrunnsinformasjon. Alle justerte analyser inkluderer fødselskohort (om barnet er født i 2006, 2007, eller 2008), kodet som dummy-variabler (0=nei, 1=ja). Videre inkluderte vi fødselskvartal, dvs. om barnet var født i hhv. perioden desember-februar, mars-mai, juni-august, eller september-november (kodet som dummy-variabler, 0=nei, 1=ja). Dette fordi barnas alder ved barnehagestart varierer som funksjon av

deres alder ved barnehagens hovedopptak 1. september (Dearing, Zachrisson, & Nærde, under arbeid). Endelig inkluderte vi også kommunetilhørighet (Bamble, Porsgrunn, Skien, Tinn, eller Drammen).

Analyser

Manglende verdier og imputering. For de aktuelle analysene var BPVS-II språkdata tilgjengelig for 928 barn. Når vi sammenlignet barna som henholdsvis gjennomførte og ikke gjennomførte denne testen, viste deg seg å være større frafall blant familier med yngre mødre ($t[1148]=-3.2$), $p<.01$, lavere utdanning ($t[945]=-3.3$), $p<.001$, vanskeligere boligsituasjon ($t[1146]=4.2$), $p<.001$, mor uten jobb $\chi^2(1)=5.6$, $p<.05$, enslige forsørgere $\chi^2(1)=12.5$, $p<.001$, ikke-vestlig innvandrere $\chi^2(1)=5.5$, $p<.05$, mor på trygd $\chi^2(1)=4.8$, $p<.05$, og gutter $\chi^2(1)=10.0$, $p<.01$. Vi benyttet multipl imputering (MICE: multiple imputation by chained equations i STATA 12; StataCorp LP, 2013) for å erstatte manglende verdier slik at analysene kunne gjøres på det fulle datasettet. Dette er den anbefalte teknikken for å erstatte manglende verdier, og gir mer korrekte estimater og standardfeil enn analyser kun basert på komplette data. Ved multipl imputering genereres et antall datasett (her: 20), hvor manglende informasjon erstattes med lineære kombinasjoner av alle data inkludert i imputeringsmodellen (vi inkluderte variablene som er beskrevet i Tabell 1 nedenfor). Analysene gjøres separat i alle datasettene og kombineres etter algoritmer som gir mest mulig korrekte estimater.

Hovedanalyser. Analyser er gjort med STATA versjon 12. Som nevnt har vi i prosjektet *Barns sosiale utvikling* samlet data fra fem ulike kommuner i Norge. Det er viktig å være oppmerksom på at barn fra samme kommune potensielt vil ha mer til felles enn det de vil ha med barn fra andre deltakende kommuner (som en funksjon av særegne karakteristika ved hver kommune). Barn fra samme kommune går i barnehager som er underlagt liknende administrasjon og som følgelig kan ha mer til felles enn det de har med barnehager i andre kommuner. Slik avhengighet i data kan skape feilaktige estimater og standardfeil (og derfor signifikansnivåer), og må av den grunn undersøkes spesifikt. Innledende analyser viste da også forskjeller mellom kommunene i assosiasjoner mellom tid tilbrakt i barnehagen og barns språkkompetanse. Vi tok høyde for dette i de videre analysene ved å gjennomføre flernivåanalyser med kommune som nivå to, og med alle andre variabler som nivå en. Vi la derfor inn tilfeldig varierende startnivå (random intercept) mellom kommunene i analysene, som medfører at både estimater og standardfeil tar høyde for at det er forskjeller mellom kommunene. Vi hadde for få kommuner til å kunne estimere andre kommune-nivåeffekter. Estimaterne i modellene kan tolkes som alminnelige regresjonskoeffisienter, hvor ulikheter mellom kommunene er tatt høyde for.

Bruk av barnehage er ikke tilfeldig, men derimot en funksjon av hver families unike karakteristika, ønsker, og prioriteringer (Zachrisson, Janson, & Nærde, 2013). Dette kompliserer studier av barnehagens betydning for barns utvikling. For å kunne trekke slutninger om effekter av barnehage er det nødvendig å isolere betydningen av barnehage per se fra den av andre faktorer som både påvirker barnehagebruk og utfallet (for oss språkkompetanse), i størst mulig grad. I observasjonelle studier gjøres

dette ofte gjennom kovariatjusteringer i multiple regresjonsmodeller. Denne teknikken har imidlertid vist seg å være utilstrekkelig (McCartney, Bub, & Burchinal, 2006). En komplementær teknikk, som gir noe mer konservative estimater, er bruk av matching med såkalte propensity skårer (Rosenbaum & Rubin, 1989). Dette er en tre-trinns prosedyre som kan redusere seleksjonsskjevhetene i observasjonelle data med inntil 90% (Leon & Hedeker, 2007). I første trinn blir «behandlingen» (dvs. barnehagebruk) regrediert på et sett relevante kovariater, vanligvis ved logistisk regresjon. Fra denne analysen lagres en estimert sannsynlighet for å «motta behandlingen» (her: benytte barnehage). Deretter matches personer med samme sannsynlighet for å «motta behandlingen», men som i realiteten hhv. mottar- eller ikke mottar behandlingen. I tredje trinn estimeres effekten av «behandlingen» innad i de matchede gruppene. En videreutvikling av propensity skår matching er å bruke propensity skårer i studier av ordinale dose-respons-sammenhenger ved hjelp av vekting (Imbens, 2000). Vektene lages ved at «dosen av behandling» (her: år tilbrakt i barnehage) regredieres på et sett kovariater ved hjelp av ordinal logistisk regresjon. Deretter lages vektene ved å invertere den predikerte sannsynligheten for å tilhøre den «doseringen» der hver enkelt person befinner seg. Dette gir dermed en høy vekt til barn som (gitt bakgrunnsvariablene) har liten sannsynlighet for å tilbringe en bestemt mengde tid i barnehage (f.eks. mer enn 3 år), og en tilsvarende lav vekt til barn som har høy sannsynlighet for å tilbringe den gitte tiden i barnehagen. Analysene vektes følgelig for propensity skåren, slik at de med høy sannsynlighet for å være i barnehage i et bestemt omfang «teller mindre», mens de med lav sannsynlighet «teller mer». Se (Dearing et al., 2009) for eksempler på slik bruk av vekting med propensity skåre.

Resultater

Preliminære og deskriptive analyser

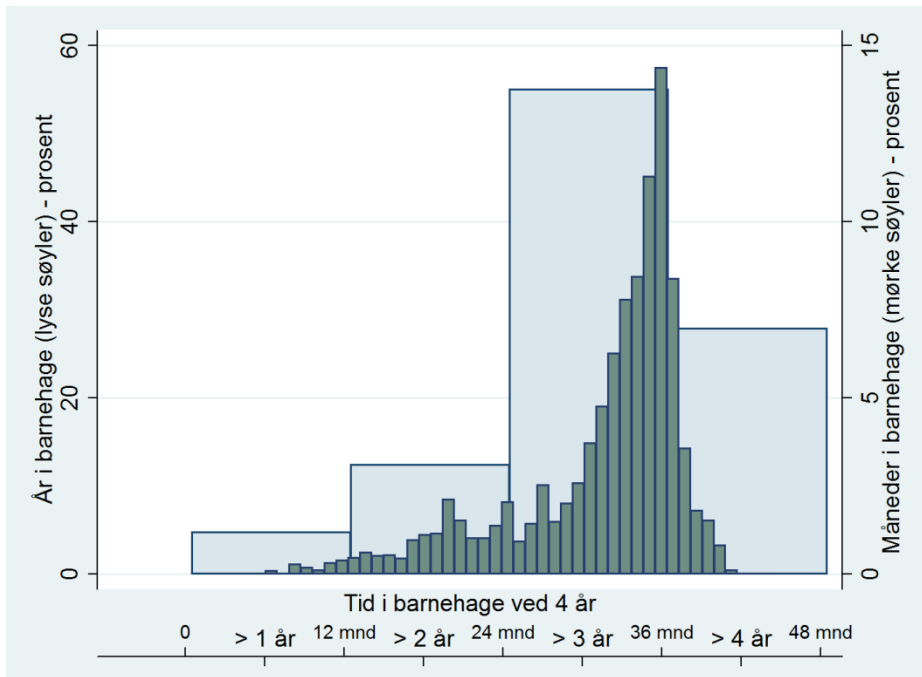
Tabell 1 viser deskriptiv statistikk (prosent/gjennomsnitt, standardavvik, min og maks-verdier) for alle variablene som inngår i analysene, samt prosentvis manglende verdier. Tabellen illustrerer at barn ved 4-års alder i gjennomsnitt hadde gått 31.37 måneder i barnehage, og at 27.4% av utvalget ved dette tidspunktet hadde gått mer enn 3 år i barnehage (dvs. startet før ett års alder). Videre hadde 54.1% gått mellom 2 og 3 år i barnehage (dvs. startet mellom 1 og 2 års alder), 12.2% mellom 1 og 2 år (dvs. startet mellom 2 og 3 års alder), mens 4.7% av 4-åringene hadde gått mindre enn 1 år i barnehage (dvs. startet etter fylte 3 år). I de videre analysene benyttet vi i hovedsak de kategorielle variablene (dvs. om barnet hadde tilbrakt > 3 år, >2 år, >1 år, eller < 1 år i barnehage). Dermed unngikk vi å forutsette en bestemt form på sammenhengen mellom barnas tid i barnehage og språkkompetanse (lineær eller kurvlineær), og muliggjorde bruk av propensity skåre vekting (som beskrevet over), for å redusere påvirkningen av seleksjonsvariabler på estimatene.

Tabell 1. viser at kun 1.6% av barna (18 barn) ikke hadde tilbrakt noe tid i barnehage før språktestingen ved 4 år. Ettersom denne gruppen er for liten til meningsfulle sammenligninger med barn som gikk i barnehage valgte vi å ekskludere disse 18 barna i videre analyser (dvs at det endelige utvalget i analysene besto av 1139 barn). Videre så vi at svært få barn (<.7 %) begynte i barnehage for så å slutte igjen. Følgelig konsentrerte vi analysene knyttet til barnehagebruk til betydningen av *samlet tid* tilbrakt i barnehagen (målt i måneder/år) gjennom barnets første fire leveår. Dette er i tråd med deler av den internasjonale forskningen på feltet. Figur 1 viser distribusjonen av tid tilbrakt i barnehage i henholdsvis måneder (mørke, smale søyler), og i «påbegynte» år (lyse, brede søyler). Figuren illustrerer at hovedvekten av barna har gått over 3 år i barnehagen ved 4 års alder (dvs. at de startet i barnehagen rett etter 1 års alder), mens det til sammenligning er relativt få som har gått mindre enn 2 år i barnehagen når de er 4 år (dvs. som starter etter at de er 2 år).

Tabell 1. Deskriptiv statistikk for alle variabler (N=1,157), inkludert prosent manglende besvarelser.

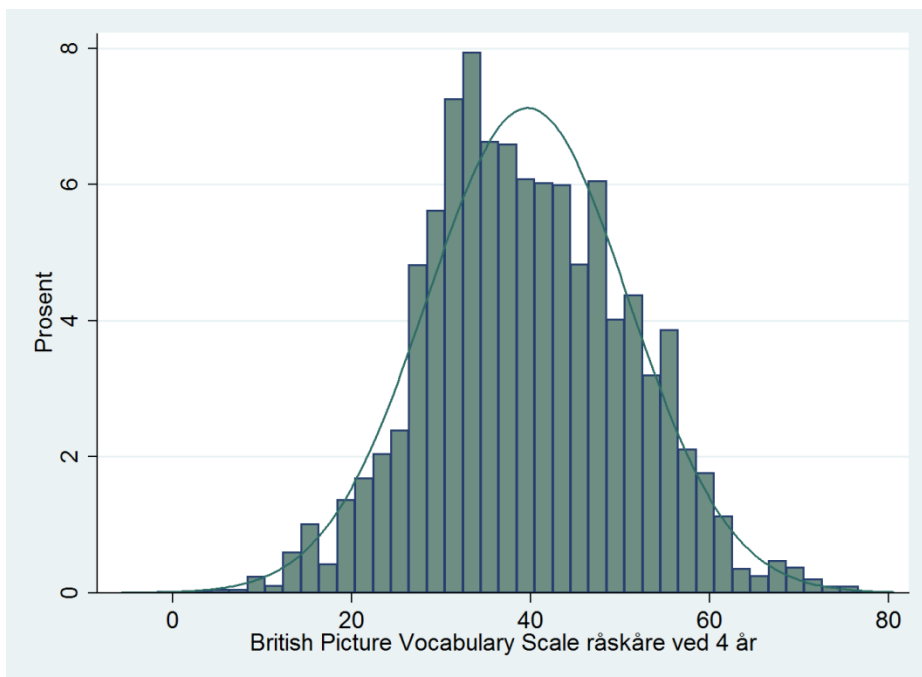
Variabel	Manglende	Gj.snitt/ %	Std. Avv.	Min	Maks
Primære variabler					
BPVS-II ved 4 år	19.8%	39.66	11.25	10	75
Avvik alder	0.1%	0.19	.63	-2.50	2.96
Bhg < 1 år	7,8%	4.7%			
Bhg < 2 år	7.5%	12.2%			
Bhg < 3 år	5.8%	54.1%			
Bhg 3 år +	4.4%	27.4%			
Måneder i Bhg	9.9%	31.37	6.49	6.41	41.00
Aldri i Bhg	0.3%	1.6%			
Familievariabler målt ved 6 mnd.					
Mors alder (år)	0.6%	30.79	4.87	19	42
Foreldres høyeste utdanning	18.2%	15.01	2.41	9	18
Mor i jobb (før permisjon)	2.1%	87.6%			
Far i jobb	41.5%	93.1%			
Enslig forelder	0.8%	5%			
Ikke-vestlig innvandrerbakgrunn	1.5%	6.5%			
Vestlig innvandrerbakgrunn	1.5%	6.8%			
Boligsituasjon	0.8%	0.24	.55	0	3
Mor trygdet	2.2%	3.5%	.18		
Mors angst & depresjon (SCL)	2.5%	1.34	.37	1	3.77
Barnevariabler					
Gutt	0%	51.8%			
Eldre søsken	0.8%	58.5%			
Generell utvikling (ASQ)	1.1%	231.11	38.33	50	336.6662
Fødselsvekt (kg)	4.2%	3.50	0.59	.924428	5.4
Barnehagevariabler					
Gruppestørrelse 2 år	35.6%	14.07	4.04	3	30
Gruppestørrelse 3 år	28.7%	17.28	4.72	3	30
Gruppestørrelse 4 år	39.5%	20.67	4.94	6	43
Småbarnsgruppe 2 år	35.6%	83%			
Småbarnsgruppe 3 år	29.04%	38%			
Bakgrunnsinformasjon					
Fødselskohort 1	0%	37.4%			
Fødselskohort 2	0%	45.7%			
Fødselskohort 3	0%	19,65%			
Født kvartal 1	0%	28.3%			
Født kvartal 2	0%	23.0%			
Født kvartal 3	0%	18.4%			
Født kvartal 4	0%	30.3%			
Bamble	0%	12.4%			
Porsgrunn	0%	27.0%			
Skien	0%	12.9%			
Tinn	0%	8.6%			
Drammen	0%	39,1%			

Note: BPVS-II=British Picture Vocabulary Scale, Bhg= Barnehage.



Figur 1. Distribusjon av tid tilbrakt i barnehage ved 4 år

Tabell 1 viser for øvrig at familiene som deltar i *Barns sosiale utvikling* i gjennomsnitt er relativt velutdannede, og at det er en forholdsvis lav andel familier med eneforsørger, der mor er på trygd, har innvandrerbakgrunn, og andre sosioøkonomiske risikoer. Figur 2 viser distribusjonen av språkmålet, British Picture Vocabulary Scale, som er tilnærmet normalfordelt i vårt utvalg. Distribusjonen er lik for gutter og jenter, og det er ikke signifikante gjennomsnittsforskjeller i de to gruppene, $t(926) = .098, p = .92$.



Figur 2. Distribusjon av språkmålet, British Picture Vocabulary Scale-II (BPVS-II) ved 4 år

Ujusterte analyser

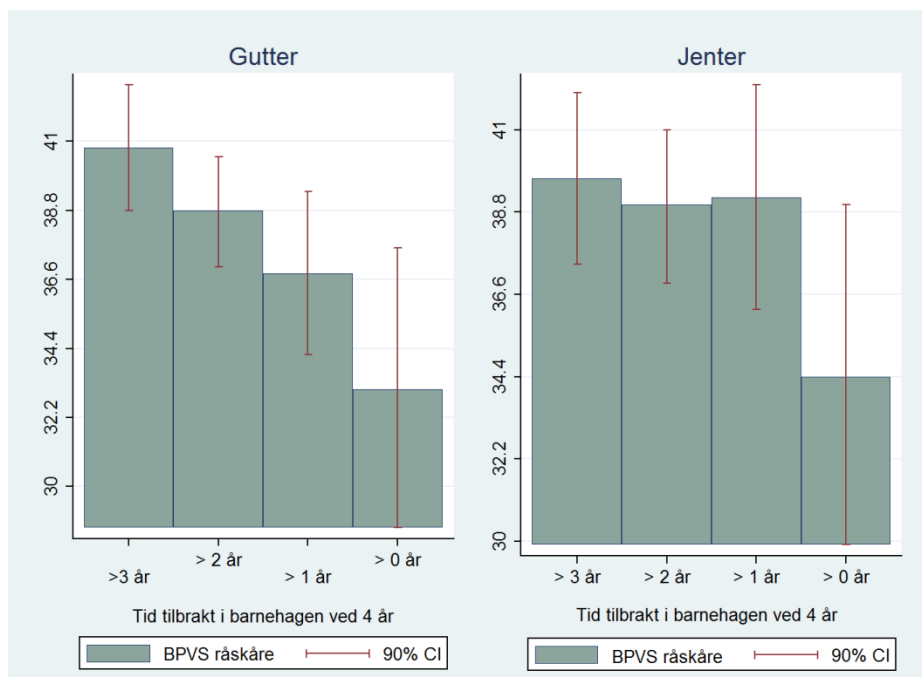
Som første skritt i analysene undersøkte vi den bivarierte (ujusterte) assosiasjonen mellom barnas tid tilbrakt i barnehage og deres språkkompetanse. Dette gir en beskrivelse av de faktiske sammenhengene i data, og er grunnlaget for videre undersøkelser ved kovariatjustering og propensity score vektning. Som nevnt i innledningen så vi det som meningsfullt å gjøre separate analyser for gutter og jenter. Selv om det ikke er nivåforskjeller mellom kjønnene kan imidlertid sammenhengene med prediktorene være ulike for jenter og gutter. Analysene er justert for barnets faktiske alder ved språktesting, ettersom denne varierte fra ca. 2.5 måneder før fylte 4 år til ca. 3 måneder etter. Gruppen med barn som hadde gått lengst i barnehage (> 3 år), ble satt som referansegruppe i analysene. Koeffisientene for de andre gruppene representerer dermed forskjeller i språkkompetanse (målt i prosent av standardavvik), fra denne referansegruppen. Når det gjelder guttene, viser Tabell 2 at, sammenlignet med de som har gått lengst i barnehage (dvs. startet før ett års alder), har gutter med mindre enn ett år i barnehage 68% av et standardavvik lavere skåre på BPVS-II ($p < .01$) når de er 4 år. Tilsvarende har guttene som har gått i barnehage mellom ett og to år, 35% av et standardavvik lavere skåre ($p < .05$), mens de som har gått mellom to og tre år i barnehage, har 17% av et standardavvik lavere skåre ($p < .1$). Ut fra de ujusterte analysene for guttene ser det dermed ut til at vi har en dose-respons-sammenheng mellom tid tilbrakt i barnehage og språkskåre ved 4 år, selv om de to gruppene med lengst tid i barnehage kun er grensesignifikant forskjellige. Blant jentene er det bare de som har gått minst i barnehage (< 1 år), som adskiller seg nevneverdig fra referansegruppen med lengst tid (> 3 år), idet de har 47% av et standardavvik lavere skåre på BPVS-II ($p < .1$) ved 4 års alder. Vi testet også interaksjonseffekten av kjønn og tid tilbrakt i barnehagen, men denne var ikke signifikant. Dette betyr at gutter og jenter ikke er signifikant forskjellige med hensyn til sammenhengen mellom tid tilbrakt i barnehage frem til fylte 4 år og språkkompetanse som målt ved BPVS-II. Resultatene viser imidlertid at gutters språkkompetanse ved 4 års alder varierer signifikant som en funksjon av hvor lenge de har gått i barnehage.

Tabell 2. Ujusterte^a sammenhenger mellom antall år i barnehage og BPVS for gutter og jenter fra multilevel-modeller som estimerer tilfeldige startpunkter (intercepts) for hver kommune i BONDS. BPVS er standardisert, koeffisientene er forskjellig i standardavvik fra referansegruppen (>3 år i barnehage). Resultatene er basert på 20 imputerte datasett, N=1139

År i Bhg	Gutter		Jenter	
	Coef. (SE)	Sign.	Coef. (SE)	Sign.
> 0 år	-.68 (.08)	0.008	-.47 (.07)	0.055
> 1 år	-.35 (.16)	0.015	.05 (.16)	0.762
> 2 år	-.17 (.11)	0.085	-.07 (.11)	0.525
> 3 år	referanse		referanse	

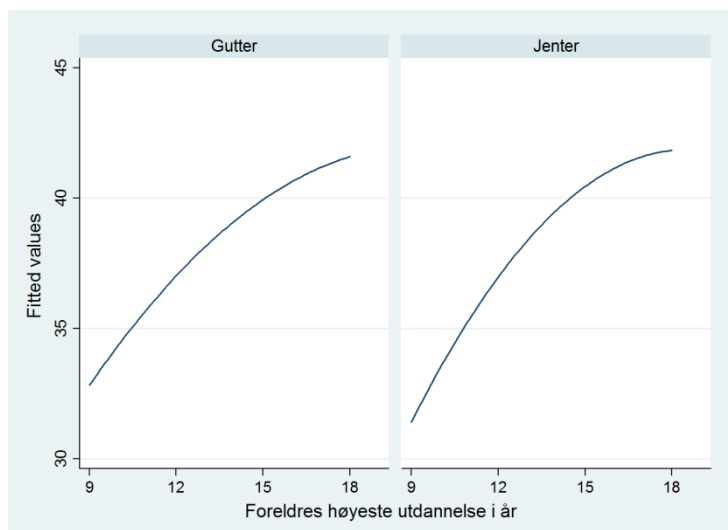
^aJustert for barnets alder ved testing.

Figur 3a viser forskjeller i BPVS-II råskåre som funksjon av antall år i barnehage for gutter og jenter. Den fulle Y-aksen fra 30-41 representerer ca. ett standardavvik. De guttene som har gått minst i barnehage (< 1 år) frem til 4-års alderen, skårer i gjennomsnitt ca. 33 poeng. Den tilsvarende skåren for guttene som har gått mest i barnehage (> 3 år), er til sammenligning på knappe 41 poeng.

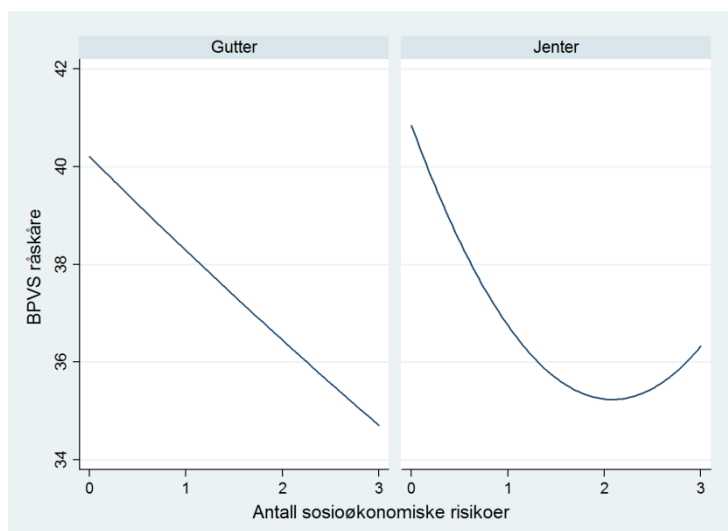


Figur 3. Ujusterte sammenhenger mellom tid tilbrakt i barnehage og BPVS for gutter og jenter. CI= konfidensintervall. Det er 20% av et standardavvik (2.2 poeng) mellom de horisontale hjelpelinjene.

Figurene 4 og 5 illustrerer hvordan barnas testresultater fra BPVS-II relaterer seg til to sentrale sosiale prediktorer for språkutvikling, nemlig foreldres utdanning og en kumulativ indeks for sosioøkonomisk risiko. Figur 4 viser at det er en tilnærmet lineær sammenheng mellom foreldrenes nivå av utdanning og barnets BPVS råskåre; jo lenger utdanning foreldrene har, desto høyere er barnas gjennomsnittlige språkskåre. I Figur 5 illustreres likeledes en negativ sammenheng mellom antall sosiale risiko (se metodedelene for spesifisering av hvordan variabelen er konstruert), og BPVS-skåre ved 4 år. For gutter er sammenhengen helt lineær, mens den flater ut, eller reverseres noe, for jentene. Dette kan være reelle effekter, men det kan også forklares ved at relativt få barn har såpass mange som 3 risikoer, og at gruppen derfor er mer sårbar for enkelte avvik i data.



Figur 4. Ubetinget estimert sammenheng mellom foreldres høyeste fulførte utdanning (i år) og BPVS-II skåre.



Figur 5. Ubetinget estimert sammenheng mellom kumulativ sosioøkonomisk risiko-indeks og BPVS-II skåre. Risikoindeksen summerer tilstedeværelsen sosioøkonomiske risikoer.

Analysert justert for kovariater og vektet med propensity skåre

Tabell 3 viser resultater fra analysene hvor sammenhengen mellom antall år barna har gått i barnehage og deres BPVS skårer ved 4 år er justert for sentrale bakgrunnsvariabler (øverste del av tabellen), og i tillegg vektet for propensity skåre (nederste del av tabellen). Igjen er resultatene presentert separat for gutter og jenter. For lesbarhetens skyld er koeffisientene for kovariatene som inngår i modellen utelatt fra tabellen, og variablene er heller listet opp i en note. Resultatene fra de justerte analysene gjenspeiler i all hovedsak de ujusterte med en tilsvarende dose-respons-sammenheng mellom hvor mange år guttene har gått i barnehagen og deres BPVS-skåre ved 4 år. Sammenlignet med referansegruppen (> 3 år barnehage), har gruppen av gutter med minst barnehage-erfaring (< 1 år), 55% av et standardavvik lavere BPVS-skåre ($p < .05$), de med 1-2 år i barnehage har 29% av et standardavvik lavere skåre ($p < .1$), og

endelig har guttene som har gått 2-3 år i barnehage 20% av et standardavvik lavere skåre på språktesten ($p < .1$). Merk at forskjellen mellom referansegruppen og de to gruppene med mer enn henholdsvis ett og to år i barnehage, kun var grensesignifikant. Dette betyr at disse forskjellene er beheftet med større usikkerhet enn hva gjelder forskjellen mellom referansegruppen og den gruppen som har vært kortest tid i barnehagen. For jenter fant vi ingen sammenheng mellom hvor lenge de har gått i barnehage og deres BPVS-skåre ved 4 år, etter justering for kovariater. Vi fant heller ingen signifikante forskjeller mellom gutter og jenter i sammenhengen mellom tid tilbrakt i barnehage og språkkompetanse.

Tabell 3. Sammenhenger mellom antall år i barnehage og BPVS-II skårer for gutter og jenter. Sammenhengene er estimert med kovariater og med propensity skåre vektning og kovariater i multilevel-modeller som estimerer tilfeldige startpunkter (random intercepts) for hver kommune. BPVS-II er standardisert, koeffisientene er justert forskjell i standardavvik fra referansegruppen (> 3 år i barnehage). Resultatene er basert på 20 imputerte datasett, N=1139.

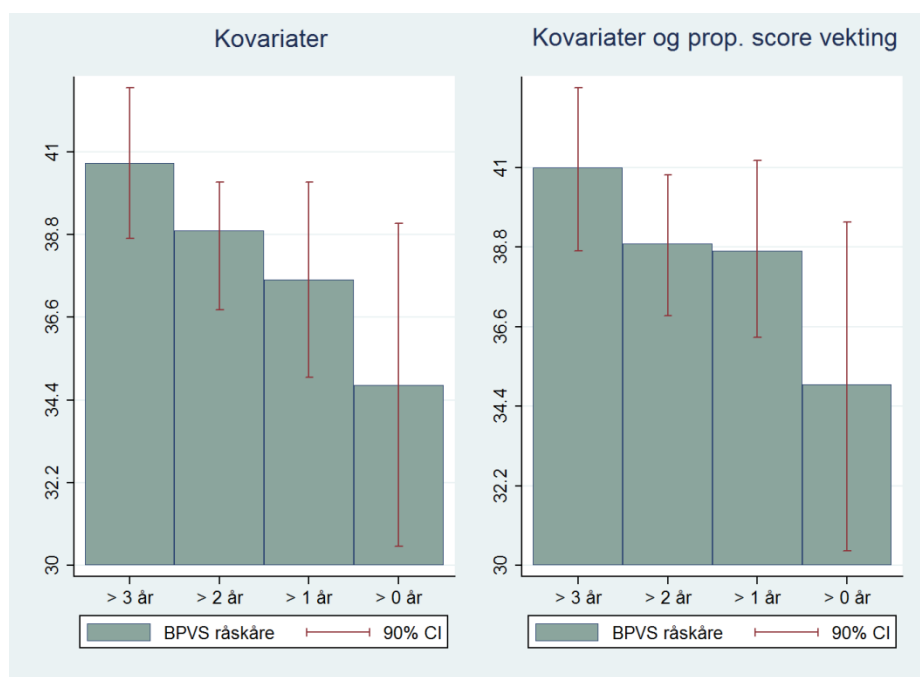
Justert	Kovariatjustering ^a			
	Gutter		Jenter	
	Coef. (SE)	Sign.	Coef. (SE)	Sign.
Antall år i bhg				
> 0 år	-.55 (.25)	0.029	-.22 (.24)	0.366
> 1 år	-.29 (.15)	0.061	.10 (.16)	0.540
> 2 år	-.20 (.11)	0.072	-.01 (.11)	0.913
> 3 år	Referanse		referanse	
	Propensity score vektning og kovariatjustering ^a			
	Gutter		Jenter	
	Coef. (SE)	Sign.	Coef. (SE)	Sign.
Antall år i bhg				
> 0 år	-.53 (.32)	0.096	.02 (.24)	0.922
> 1 år	-.20 (.16)	0.204	.19 (.18)	0.287
> 2 år	-.19 (.09)	0.035	-.01 (.13)	0.943
> 3 år	Referanse		referanse	

Note: ^aKovariater i modellene er: barnets alder ved testing, mors alder (år), foreldres høyeste utdanning, mor i jobb (for permisjon), far i jobb, enslig forelder, mor på trygd, ikke-vestlig innvandrerbakgrunn, vestlig innvandrerbakgrunn, boligsituasjon, mor trygdet, mors angst & depresjon (SCL), eldre søsken, generell utvikling (ASQ), fødselsvekt (kg), fødselskohort, fødselskvartal. Tabell 3. Sammenhenger mellom tid tilbrakt i barnehage og BPVS for gutter og jenter. Sammenhengene er estimert med kovariater og med propensity skåre vektning og kovariater i multilevelmodeller som estimerer tilfeldige startpunkter (random intercepts) for hver kommune i BONDS. BPVS er standardisert, koeffisientene er justert forskjell i standardavvik fra referansegruppen, som har tilbrakt mer enn 3 år i barnehage. Resultatene er basert på 20 imputerte datasett, N=1139.

Den nederste delen av Tabell 3 viser resultatene fra analysene med vektning for propensity skårer i tillegg til kovariatjustering. Innledende analyser av propensity skårene viste at vi har et kovariatsett med god prediksjon for barnas tid tilbrakt i barnehage. For gruppen med < 1 år i barnehage, klassifiserte vi korrekt (ved bruk av logistiske regresjonsmodeller), for 95.3% av barna. Tilsvarende klassifiserte vi korrekt for 87.6% av dem med 1-2 års barnehage-erfaring, 63.6%, for dem med 1-3 år, og endelig 73.9% for dem med > 3 år (de to sistnevnte gruppene er også de største). Vi kunne også korrekt klassifisere 98.3% av barna som ikke gikk i barnehage i det hele tatt, men merk at de ikke er inkludert i analysene. De barna som ut fra kovariatene hadde lavest sannsynlighet for å være i den gruppen de faktisk tilhørte, ble vektet opp i analysene. Det vil si at de ble «tatt mer hensyn til», sammenlignet med de barna som hadde stor

sannsynlighet for å være i den gruppen de faktisk tilhørte. Resultatene fra analysene med kovariater og samtidig vektning for propensity skåre må anses som de mest konservative estimatene mht. seleksjonseffekter.

Resultatene fra disse siste analysene var i hovedsak konsistente med dem fra de ujusterte og justerte analysene, i det de viste det samme mønsteret av dose-respons-sammenheng for guttene, men ikke for jentene. Imidlertid endret bruken av vekter presisjonen i estimatene (standardfeilene), og følgelig signifikansnivået. Sammenlignet med guttene som hadde vært mer enn tre år i barnehage, hadde de med mindre enn ett års erfaring 53% av et standardavvik lavere skåre på BPVS ($p < .1$), de med mellom ett og to år 20% av et standardavvik (ikke signifikant), mens guttene med mellom to og tre år hadde 19% av et standardavvik ($p < .05$) lavere skåre. For jenter var det ingen forskjell i skåre på BPVS-II som funksjon av hvor lenge de hadde gått i barnehagen i løpet av sine fire første leveår.

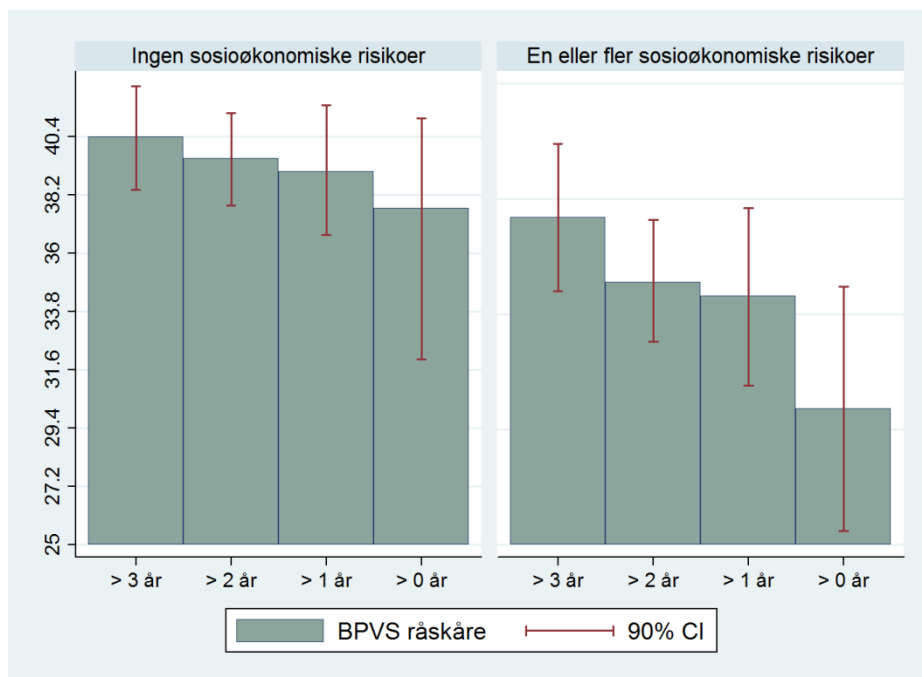


Figur 6. Sammenhengen mellom antall år i barnehage og BPVS-II skårer for gutter, justert for kovariater og med propensity skåre vektning og justering for kovariater. CI= konfidensintervall. Det er 20% av et standardavvik (2.2 poeng) mellom de horisontale hjelpelinjene.

Det venstre panelet i Figur 6 viser estimater for forskjeller i BPVS-II råskåre som funksjon av antall år i barnehage for gutter, etter justering for kovariater. Panelet til høyre viser tilsvarende estimater fra modellene justert for kovariater og vektet med propensity skårer. Den fulle Y-aksen fra 30-41 utgjør ca. et standardavvik. I begge modellene skåret guttene som har gått færrest år i barnehage (< 1 år) i gjennomsnitt ca. 35 poeng, sammenlignet med knappe 41 poeng for de som har gått flest år i barnehage (> 3 år). De to grafene er imidlertid ulike mht. forskjellene i språkskårer mellom guttene som har gått henholdsvis 1-2, og 2-3 år i barnehage. I analysene som også inkluderer vektning med propensity skåre var estimatene for disse gruppene tilsynelatende like, mens færre år i barnehagen var assosiert med lavere estimert språkskåre i analysene som kun var kovariat-justert.

Er antall år i barnehagen viktigst for språkkompetansen til barn i sosial risiko?

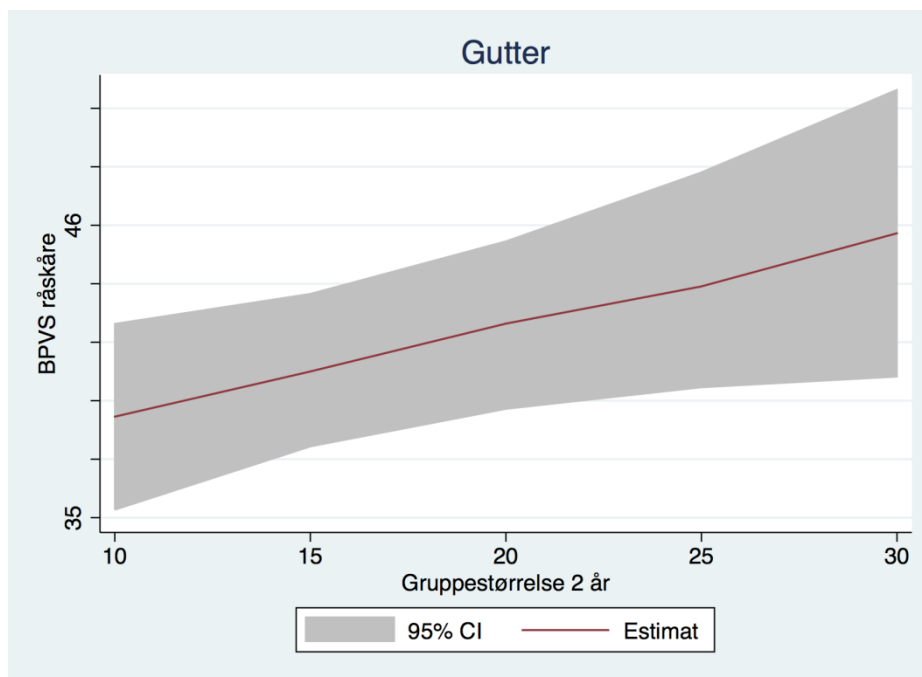
Vi brukte flere analytiske tilnærminger for å undersøke om antall år i barnehagen var viktigere for barn med sosial risiko enn for dem uten. Innledningsvis konsentrerte vi oss om foreldrenes utdanning, ved at vi testet en serie interaksjoner mellom antall år i barnehage og foreldres utdanningsnivå. Vi gjennomførte analyser som antok en lineær effekt av foreldrenes utdanning (ved å inkludere utdanning som en kontinuerlig variabel), og testet ulike ikke-lineære modeller (ved også å inkludere en kvadrert utdannings-variabel, samt en dummy-variabel for foreldre med mindre enn videregående utdanning). Ingen av disse analysene ga signifikante interaksjonseffekter med antall år i barnehagen, verken for hele gruppen eller for gutter og jenter separat. Deretter konstruerte vi en kumulativ indeks for sosial risiko, som beskrevet i metoddelen. En slik indeks kan sies å innbefatte en mer holistisk tilnærming til sosial risiko idet barn kategoriseres etter *antallet* risiki i hjemmet. Samtidig antar vi, ved å bruke en slik risikoindeks, at alle inkluderte risikofaktorer er av tilsvarende styrke, og at betydningen av ulike kombinasjoner av risiki har samme betydning. Vi fant ingen signifikante interaksjonseffekter mellom antall år i barnehage og risikoindeksen, verken for hele gruppen eller for gutter og jenter separat. Ettersom kun 8.3% av barna hadde mer enn en risiko, slo vi sammen alle barna med en eller flere risiki til en gruppe, som til sammen utgjorde 25.3% av utvalget. Vi fant heller ingen signifikante interaksjoner mellom denne dikotome risikoindeksen og antall år i barnehage, verken for hele gruppen eller for gutter og jenter separat. På et deskriptivt nivå så det imidlertid ut til at det var en sterkere sammenheng mellom antall år i barnehage og språkkompetanse for barna med en eller flere risikoer, sammenlignet med for de uten noe sosial risiko, se Figur 7. Det er imidlertid lav presisjon på estimatene (dvs. brede konfidensintervaller), noe som medfører at disse interaksjonseffektene ikke er signifikante. Sammenhengene i Figur 6 må derfor betraktes som trender i data som er beheftet med til dels stor usikkerhet.



Figur 7. Sammenhengen mellom antall år i barnehage og BPVS-II-skåre, justert for bakgrunnsvariabler, for barn uten sosioøkonomiske risiko og for barn med en eller flere risiki. Det er 20% av et standardavvik (2.2 poeng) mellom de horisontale hjelpelinjene.

Betydningen av kvaliteter ved barnegruppen for språkkompetanse

I det siste settet av analyser undersøkte vi om de strukturelle kvalitetene ved barnegruppen vi har tilgjengelige mål på i *Barns sosiale utvikling* (alderssammensetning i barnegruppen og gruppestørrelse), hadde sammenheng med barnas språkkompetanse ved 4 års alder. Vi fant ingen sammenheng mellom alderssammensetning, verken ved 2 eller 3 år, og språkkompetanse ved 4 år, heller ikke når vi undersøkte dette separat for gutter og jenter. Det var imidlertid en grensesignifikant positiv sammenheng mellom gruppestørrelse ved 2 år og språkkompetanse for gutter som begynte i barnehagen før fylte 2 år i analysene som justerte for kovariater, og denne sammenhengen var signifikant ved vekting for propensity skåre. En økning i gruppestørrelse på 10 barn var her assosiert med en økning i 30% av et standardavvik i språkkompetanse ($p < .05$). Sammenhengen (illustrert i Figur 8), ser ut til å være lineær i de justerte modellene (dvs. at ikke-lineære modeller ikke var signifikante). Videre fant vi ingen sammenheng mellom gruppestørrelse ved 3 og 4 år og barnas språkkompetanse, verken for hele gruppen eller for gutter og jenter separat. Vi testet også interaksjonseffekter mellom antall år i barnehagen og hhv. alderssammensetning og gruppestørrelse, dvs. at vi undersøkte om sammenhengen mellom kvalitetsvariablene og barnas språkkompetanse varierte som funksjon av *hvor lenge* barna hadde vært i barnehagen. Det var ingen slike interaksjonseffekter.

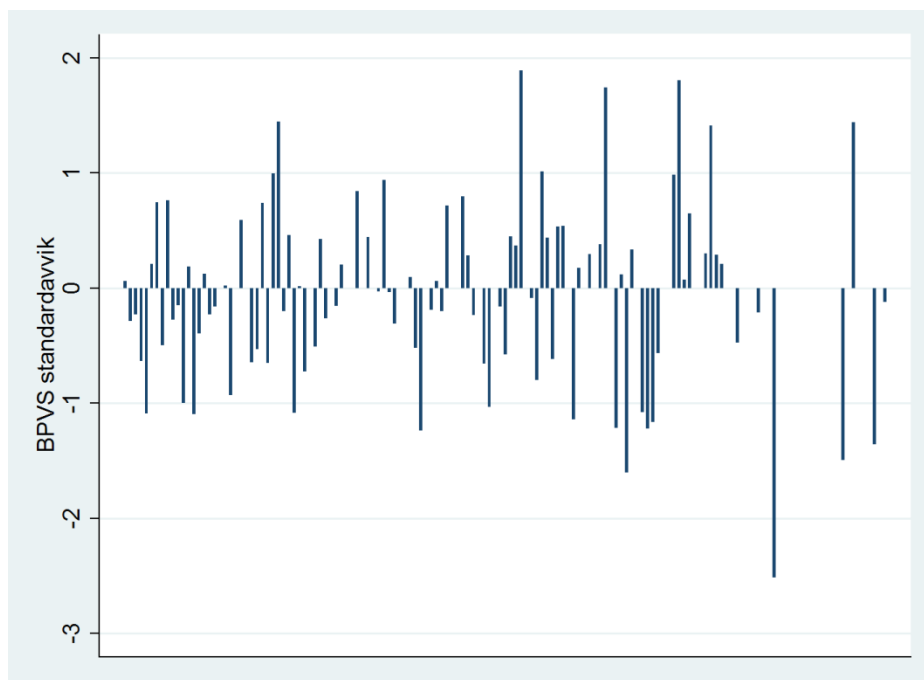


Figur 8. Sammenhengen mellom gruppestørrelse ved 2 år og BPVS-II-skåre justert for bakgrunnsvariabler og vektet med propensity skåre, for gutter som startet i barnehage før de fylte 2 år. Det er 20% av et standardavvik (2.2 poeng) mellom de horisontale hjelpelinjene.

Derneft undersøkte vi om dose-respons-sammenhengen er konsistent for barn som går i samme barnehage ved hjelp av en fixed-effect. Denne analysen tester om forskjellene i språkkompetanse mellom barn *i samme barnehage* er assosiert med forskjeller i hvor lang tid barna har gått i barnehagen. En slik analyse tar dermed som funksjon av designet høyde for alle kvalitetsforskjeller (observerte og ikke-observerte) mellom barnehager, ved at kun barna i samme barnehage sammenlignes med hverandre. For guttene fant vi at barn som gikk mer enn 3 år i barnehage skåret 35% av et standardavvik bedre på BPVS-II enn gutter som hadde gått mellom 2 og 3 år ($p=.012$). Det var ingen signifikant forskjell fra referansegruppen for gutter som hadde vært i barnehagen mellom 1 og 2 år, (men estimatet var 40% av et standardavvik). Det var ikke nok gutter som hadde mindre enn 1 år i barnehage til at denne forskjellen kunne estimeres. Resultatet var robust på tvers av kommuner.

Ettersom vi ikke hadde tilgang til data om andre kvaliteter i barnehagene enn barnegruppens alderssammensetning og størrelse, undersøkte vi avslutningsvis om det var systematiske forskjeller i språkkompetanse hos barna i ulike barnehager. Dette ble gjort ved å lagre residualene fra en regresjonsanalyse av sammenhengen mellom barnas BPVS skårer og det fulle settet av kovariater. Residualene representerer hvert individs avvik i språkskåre fra den lineære prediksjonen fra kovariatene. Barn med positive residualverdier har bedre språkkompetanse enn forventet ut fra kovariatene i modellen. Tilsvarende har barn med negative residualverdier dårligere språkkompetanse. Deretter plottet vi gjennomsnittlig residualverdi innad i barnehagene (til sammen 145 barnehager), se Figur 9. Hver stolpe i denne grafen representerer en spesifikk barnehage. I barnehager der stolpene peker oppover fra den horisontale midtlinjen har barna i gjennomsnitt *bedre* språkkompetanse enn utvalgets gjennomsnitt (målt i

standardavvik på y-aksen). Til sammenligning har barna i barnehager hvor stolpen peker nedover i gjennomsnitt *dårligere* språkkompetanse enn utvalgets gjennomsnitt. Grafen viser dermed at det generelt sett er stor variasjon i språkkompetanse mellom de 145 barnehagene, fra nærmere 2 standardavvik over gjennomsnittet, til mer enn 2 standardavvik under. Vi finner ingen systematikk i at store avvik, hverken positive eller negative, fra gjennomsnittet kun forekommer i barnehager med bare ett eller få barn.



Figur 9. Forskjeller i BPVS-II-skåre mellom barnehager, justert for bakgrunnsvariabler. Hver stolpe i diagrammet representerer det gjennomsnittlige avviket innad i den enkelte barnehage fra den gjennomsnittlige BVPS-skåren, etter justering for alle bakgrunnsvariabler i analysene presentert i Tabell 3, samt gruppestørrelse og alderssammensetning ved 2 år.

Diskusjon

Hovedfunnet i denne rapporten er at det var en dose-respons-sammenheng mellom antall år barn har vært i barnehagen og deres språkkompetanse ved 4 år, men kun for gutter. Jo lenger tid guttene hadde gått i barnehagen, desto bedre språkkompetanse (dvs. begrepsforståelse) hadde de. Disse resultatene er imidlertid beheftet med vesentlige usikkerheter i form av lav presisjon i flere av estimatene. Blant jentene var det lite som tydet på at tid i barnehagen hadde sammenheng med språkkompetanse. Videre fant vi en ikke-signifikant trend i retning av at dose-respons-sammenhengen mellom antall år i barnehage og språkkompetanse er sterkere hos barn i sosioøkonomisk risiko enn hos barn uten slik risiko. Vi undersøkte også betydningen av visse typer kvaliteter ved barnegruppene (der vi hadde tilgjengelige data), og fant en sammenheng mellom gruppestørrelse i barnehagen ved 2 års alder og gutters språkkompetanse ved 4 år; jo større barnegruppe, jo bedre språkkompetanse. Endelig tyder resultatene på at det er stor variasjon mellom de 145 barnehagene når det gjelder barnas gjennomsnittlige språkkompetanse. Vi vil diskutere disse funnene i det videre.

Dose-respons-sammenheng og signifikansnivåer

En dose-respons-sammenheng betyr vanligvis at jo større dose en person får av en gitt «behandling», jo større blir effekten. Både de ujusterte og de kovariat-justerte analysene viser et slikt mønster for guttene: jo lenger tid de hadde gått i barnehagen, desto bedre synes språkkompetansen å være. Forskjellen mellom guttene som har gått kortest og lengst tid i barnehagen er betydelig. I de justerte modellene tilsvarer denne over 50% av et standardavvik, som for eksempel er større enn forskjellen i språkkompetanse mellom barn av foreldre med 9 og 18 års utdanning. Selv om effektstørrelsene er betydelige, må de imidlertid tolkes med varsomhet. Dette skyldes at estimatenes standardfeil er forholdsvis store, dvs. at estimatene har en relativt lav presisjon (det er stor variasjon innad i gruppene), hvilket igjen har betydning for signifikansnivåene til estimatene. I de kovariat-justerte analysene er språkkompetansen til guttene som har gått >1 år i barnehage signifikant forskjellig fra referansegruppen med >3 år i barnehagen, tilsvarende 55% av et standardavvik lavere, $p=.029$. I motsetning til dette er de to ”mellomgruppene”, dvs. >1 og >2 år i barnehage, kun «grensesignifikant» forskjellig fra referansegruppen, med $p=.061$ og $p=.072$. Analysene som også er vektet med propensity skårer viser imidlertid at det er signifikant forskjell ($p=.035$) mellom guttene i de to gruppene som har gått lengst i barnehagen (referansegruppen og >2 år), at >1 år ikke var signifikant, mens >0 år gruppen var grensesignifikant ($p=.096$). Forskjellen mellom de to analysene gjeldende hvilke estimer som er signifikant forskjellig (med og uten propensity skåre vektning), skyldes at vektningen i noen grad endrer standardfeilene på estimatene slik at de blir litt mer presise i de store gruppene og litt mindre presise i de små gruppene.

Tolkningen av signifikansverdiene, og hvilken vekt som kan tillegges «grensesignifikante» resultater (dvs. p -verdier mellom .1 og .05), er viktige for hvordan resultatene i rapporten skal forstås. I forskningslitteraturen er det vanlig at slike funn tolkes, men at de tillegges mindre verdi enn resultater med

lavere signifikansnivå (vanligvis er $p < .05$ den ”magiske grensen”). De «grensesignifikante» resultatene har marginalt høyere sannsynlighet for å være en funksjon av tilfeldig variasjon i populasjonen. Dette utgjør et tolkningsmessig dilemma, fordi vi har en større risiko enn det vi normalt foretrekker for å begå en Type I-feil, dvs. at vi sier at vi har et funn når vi faktisk ikke har det. Samtidig vil det å fullstendig avfeie alle resultater som er «grensesignifikante» være forbundet med en risiko for å begå en Type II-feil, dvs. å si at vi ikke har et funn når vi i realiteten har det. Balansen mellom Type I og Type II-feil er ofte en substansiell avveining; hva er konsekvensen av å akseptere noe som et funn versus det å ikke gjøre det? I vårt tilfelle handler det om hvorvidt disse resultatene skal sees som et argument for at gutter, i større grad enn jenter, profiterer språklig på å begynne tidlig i barnehage, eller ikke. Vi mener at en nøktern konklusjon tillegger resultatene fra analysene som er vektet for propensity skåre størst verdi, ettersom de er mest konservative mht. risiko for seleksjonseffekter. Disse analysene viser at det er forskjell i språkkompetanse mellom guttene i de to største gruppene i utvalget; de som har gått i barnehage i mer enn 3 år, og de med mellom to og tre års barnehage-erfaring, der guttene som har gått lengst tid i barnehagen i løpet av sine første fire leveår har best språk. Vi kan ikke med samme sikkerhet si at guttene som kun har gått enten mellom 1 og 2 år, eller mindre enn 1 år i barnehage, har dårligere språkkompetanse enn dem med lengst tid. Disse gruppene er forholdsvis små, og følgelig er standardfeilene i estimatene større, men at data tenderer til å vise at det er en dose-respons-sammenheng som diskutert over.

Hvorfor finner vi sammenheng hos gutter men ikke jenter?

Som vi nevnte i innledningen er det kjent i litteraturen at gutter og jenter kan ha ganske forskjellig språkutvikling. Det kan være mange årsaker til dette. En studie gjennomført av en av rapportens medforfatter (IMZ), viste blant annet at jentene hadde høyere nivå av språkforståelse enn guttene mellom 1.5 og 3 år. Det var imidlertid guttene som viste størst endring, eller vekst, i sine ferdigheter i løpet av denne perioden, slik at kjønnsforskjellene minsket over tid (Zambrana et al., 2012; se også; Spinath, et al., 2004). Og selv om gutter synes å ha en forhøyet risiko for et forsinket språkutviklingsforløp i forhold til jenter mellom 3 og 5 år (Spinath et al., 2004; Zambrana, Pons, Eadie og Ystrøm, et al., 2013b), ser det derimot ikke ut til at gutter har økt risiko for senere språkrelaterte vansker slik som dysleksi (Hawke, Olson, Willcut, Wadsworth, & DeFries, 2009). Resultatene fra våre analyser kan dermed være i overensstemmelse med muligheten for at gutters språkutvikling er mer sensitiv for miljøpåvirkning i denne alderen, slik at miljøet forklarer mer av variasjonen hos gutter enn jenter (Zambrana et al., 2012). Eksempelvis fant Zambrana et al. (2012), at lav utdanning hos mor var signifikant mer negativt for gutters enn for jenters språkutvikling, mellom 1.5 og 3 år.

En mulig forklaring på dose-responssammenhengen hos gutter er at det er dose i barnehage per se som bidrar til deres språkutvikling, det vil si at mer alltid er bedre. Denne tolkningen impliserer at guttene stimuleres språklig i barnehagen på en måte som ikke skjer hjemme, og er konsistent med forskningslitteraturen diskutert innledningsvis som viser at barn i kvalitetsbarnehager har bedre kognitiv- og språkutvikling, enn barn som ikke går i barnehage. At vi ikke finner dette på tvers av kjønn indikerer

allikevel at det eksisterer kjønns spesifikke prosesser som også må vurderes. Det kan blant annet tenkes at gutter har sensitive perioder for språkutvikling, og at de som starter i barnehagen før fylte ett år blir stimulert på et særlig gunstig utviklingsmessig tidspunkt, og at de som starter senere er mindre mottagelige for slik stimulering. Dette er en supplerende hypotese som indikerer at ikke kun dose, men også timing, er av betydning. (Da praktisk talt ingen barn i vårt utvalg slutter i barnehagen har vi ikke anledning til å undersøke dette, men det er et element som må tas med i vurderingen av resultatene). Hvis resultatene skal sees som et argument for at gutter profiterer språklig på å begynne tidlig i barnehagen, må det også være rimelig sterke argumenter for at det ikke er tredjevariabler (observerte eller ikke-observerte), som kan være forklaringer på en slik sammenheng. Vi har to elementer i analysene som er konsistente med en kausal sammenheng mellom antall år i barnehage og språkkompetanse. For det første er det en ganske tydelig dose-respons-sammenheng. En tredjevariabel som i så fall skal kunne forklare denne sammenhengene må derfor være relatert til både tid i barnehage og språkkompetanse i en tilsvarende dose-respons-sammenheng. Det er for eksempel en viss mulighet for at enkelte gutter, i større grad enn andre, er «klare» for de utfordringer og krav som barnehagehverdagen medfører. Dersom foreldre er bevisste på at deres gutt enda ikke er klar for barnehagen, kan det være visse kjennetegn ved utviklingen til de guttene som holdes tilbake fra barnehagen som vi ikke har fanget opp, og som dermed kan forklare de sammenhengene vi finner. En annen alternativ forklaring kan være at foreldre som sender barna tidlig i barnehagen søker å kompensere for dette i den tiden de har sammen med barna gjennom å legge vekt på å stimulere barna. Hvis dette ikke fanges opp av kovariater og vekting for propensity skårer, vil det kunne forklare de sammenhengene vi finner, gitt at denne kompenserende stimuleringen følger et dose-responsmønster. Slike tredjevariabler kan aldri fullstendig utelukkes i observasjonelle studier. Dette gjør at vi ikke, mer enn ad logisk vei, kan etablere argumenter for en viss grad av rimelighet i forhold til kausale tolkninger av resultater. Vi kan således ikke utelukke at det er årsaksfaktorer eller prosesser som kan forklare en ulik kjønnsseleksjon inn i barnehage og som også spesielt er knyttet til gutters utvikling, som vi ikke har tatt høyde for i våre analyser.

Det andre elementet som er konsistent med en kausal forklaring er at dose-respons-sammenhengene (estimatene, ikke standardfeilene), er relativt konstante på tvers av ulike analyser. Vi har et omfattende sett med kovariater, og et unikt datasett med detaljer informasjon om barna og deres familier fra før barnehagestart (for eksempel ASQ som mål på tidlig generell utviklingsrisiko). Vektingen med propensity skåre er et konservativt grep for ytterligere å redusere risikoen for at tredjevariabler forklarer sammenhengene. Når heller ikke et slikt grep i nevneverdig grad endrer estimatene, fremstår disse som robuste tatt i betraktning de kovariatene som er inkludert i analysene. Til tross for dette, kan vi ikke utelukke at måten vi for eksempel har målt foreldrenes utdanning på (nivåinndelt), ikke innbefatter de arvelige eller miljømessige aspektene ved utdanning som kan være relaterte både til barnehagebruk og til barns språkkompetanse. For eksempel kan retningen og kvaliteten på utdanningen, og ikke minst oppnådde resultater i utdanningen, være alternative og informative kovariater i en slik sammenheng.

Sosioøkonomisk risiko og språkutvikling

Et konsistent funn i den internasjonale forskningslitteraturen er at barnehager av god kvalitet har mer betydning for språkutviklingen hos barn i sosial risiko enn for andre barn (f.eks., Dearing et al., 2009). Vi fant imidlertid ingen slike effekter som var signifikante i våre analyser. Som vi illustrerer i Figur 7 er det trender i data som peker i retning av at vi faktisk kan ha en slik mekanisme, at barn i sosial risiko profiterer mer på å være mye i barnehagen enn andre, men at den er forbundet med stor usikkerhet i estimatene. En rimelig forklaring på dette er at utvalget i *Barns sosiale utvikling* ganske enkelt har relativt få barn i sosial risiko. Vi vet bla. at det gjennomsnittlige utdanningsnivået i vårt utvalg er høyere enn i befolkningen ellers. Selv om vår kumulative risikoindeks viser at ca. 25% av utvalget har mer enn én risiko, er indeksen et sammendrag av et bredt spekter av ulike risikoer som kan virke ulikt selv om de behandles likt, hvilket kan bidra til lav presisjon i estimatene. Snevrere definisjoner av sosial risiko gir mindre grupper, som gir samme problem mht. presisjon. Vi vil imidlertid i fremtidig forskning undersøke denne tematikken nærmere, bla. ved å søke om koblinger til registerdata som kan gi bedre mål på sosial risiko.

Betydningen av kvaliteter ved barnegruppen

I *Barns sosiale utvikling* har vi enda ikke anledning til å studere andre kvaliteter ved barnehagen enn gruppetype (småbarnsgruppe vs. storebarnsgruppe) og gruppestørrelse. Dette skyldes utfordringer knyttet til å koble informasjon fra barnehagene til det enkelte barn, som er et pågående arbeid. Vi fant ingen sammenheng mellom gruppetype og barnas språkkompetanse. Dette er i seg selv interessant, ettersom gruppetype dels er en indikator på aldersspennet i barnegruppen, men også på voksentetthet (som er høyere i småbarnsgrupper). Tolkningen av dette null-funnet er derfor usikkert. Det kan bety at voksentetthet ikke er av betydning for barns språkutvikling, eller at det er relatert til alders-spennet i barnegruppen. Vi fant imidlertid at gruppestørrelse ved 2 år var assosiert med bedre språkkompetanse ved 4 år for gutter. Jo større barnegrupper guttene var i ved 2 år, jo bedre språkkompetanse hadde de ved 4 år. Det var ingen sammenheng med gruppestørrelser på senere tidspunkt, og heller ikke for jenter på noe tidspunkt. Funnet står i kontrast til forskning om språkutvikling i familien, hvor store søskengrupper er assosiert med senere strukturell språkutvikling (f.eks. tilegnelse av vokabular), fordi dette reduserer omfanget av språklige en-til-en interaksjoner mellom foreldre og barn (Oshima-Takane & Robbins, 2003; Pine, 1995; Zambrana et al., 2012). Disse effektene er derimot sterkere i barnas første leveår, og det er mangel på forskning som har studert betydningen av størrelsen på søskenflokken etter 2-3 år. Det gir derfor rom for å spekulere om mekanismene som fremmer språkkompetanse i barnehagen snarer er knyttet opp mot tilgjengeligheten av jevnaldrende, heller enn mot voksen-barn-interaksjoner (som presumptivt er mindre frekvent i større barnegrupper), og at slike sosiale erfaringer blir viktigere med alderen. Det kan naturligvis ikke utelukkes at sammenhengen skyldes seleksjonseffekter i familien eller barnehagen, dvs. at gutter som er sterkere språklig plasseres i større barnegrupper. Den positive sammenhengen mellom gruppestørrelse og språkkompetanse bør også sees i sammenheng med en tidligere rapport fra Atferdssenteret (Zachrisson, Backer-Grøndahl, Nærde, & Ogden, 2012), hvor

resultatene viste at mindre barnegrupper, primært ved 3 år, var assosiert med bedre sosial kompetanse hos 3-åringene. Et tema for videre forskning kan derfor være å se på den unike betydningen av barnehage for språk og sosial kompetanse gjennom barneårene, som blant annet vil bli gjort gjennom bevilgninger vi har fått fra Norges Forskningsråds program UTDANNING2020. På det nåværende tidspunkt forstår vi i liten grad mekanismen bak at gutter som går i større grupper ved 2 års alder skårer høyere på BPVS-II når de er 4 år. Ettersom dette funnet også står i kontrast til våre tidligere rapporterte funn vedrørende sosial kompetanse ved 3 års alder, er vi tilbakeholdne med å konkludere at større grupper for toåringene samlet sett er positivt for barns språkutvikling.

Når vi sammenlignet språkkompetansen hos barn innad i samme barnehage, fant vi at for gutter er dose-respons-sammenhengene robuste, i hvert fall når de som har gått mer enn tre år sammenlignes med dem som har gått mellom to og tre år. Dette betyr at kvaliteten i den enkelte barnehagen, dvs. den delen av barnehagekvalitet som påvirker alle barn på samme måte, ikke ser ut til å viske ut betydningen av år tilbragt i barnehage. Det kan tolkes som at vi sannsynligvis ikke har en systematisk trend i retning av at f.eks. de som starter i barnehagen før 1 år begynner i barnehager av spesielt god kvalitet, hvilket kunne vært en potensiell alternativ forklaring på hovedfunnene i rapporten.

Som illustrert i Figur 8 ser vi at det er stor variasjon i den gjennomsnittlige språkkompetansen mellom barnehagene, etter at alle bakgrunnsvariabler (inkludert gruppestørrelse og type) er justert for. Slik vi vurderer det kan det tenkes tre ulike scenarier som forklarer denne variasjonen. For det første kan *Barns sosiale utvikling* ha rekruttert barn som språklig sett er spesielt sterke, eller svake, og som deretter begynner i de samme barnehagene. Mens dette er et lite sannsynlig scenario, kan det allikevel ikke helt utelukkes. Det kan også tenkes at det er systematiske forskjeller i opptaksområdet til enkelte barnehager, dvs. at noen barnehager i større grad har barn fra miljøer hvor flere barn er språklig sterke (for eksempel områder hvor mange foreldre har høy utdanning eller mange ressurser), og at vi ikke får justert vekk effekten av dette ved å inkludere sosiodemografiske karakteristika for den enkelte familie. Til slutt kan det tenkes at variasjonen skyldes kvalitetsforskjeller mellom barnehagene, og at enkelte barnehager er spesielt gode til å stimulere barns språkutvikling, mens andre er langt dårligere. Dette er et tema vi kommer til å undersøke nærmere når vi får tilgang til mer data om barnehagene.

Konklusjon

Vi fant en dose-respons-sammenheng mellom tid tilbragt i barnehage og språkkompetanse (dvs. språkforståelse) ved 4 år for gutter. Dernest fant vi også en sammenheng mellom gruppestørrelse i barnehagen ved 2 år og gutters språkkompetanse ved 4 år, og betydelige forskjeller mellom barnehager i barns språkkompetanse. Samtidig er det klare heftelser ved disse konklusjonene, knyttet til presisjonen i estimatene og til potensialet for ikke-observerte tredjevariabler som kan forklare sammenhengen, som muligens ville ha justert våre effekter ned. Funnene kan både være knyttet til at gutter er mer sensitive for spesifikke miljøfaktorer i denne perioden, slik som barnehage, alternativt kan dette forklares gjennom kjønns spesifikke seleksjonsprosesser inn i barnehagen. Videre så vi tendenser i data som tyder på at det kan være en sterkere dose-respons-sammenheng for barn i sosial risiko, men vi har for lave nivåer av sosial risiko i vårt utvalg til å finne en slik effekt. Samlet sett vil det derfor være ønskelig at funnene repliseres i andre datasett og at alternative sammenhenger undersøkes videre før absolutte konklusjoner trekkes.

Referanser

- Barnett, W. S. (2011). Effectiveness of early educational intervention. *Science*, *333*, 975-978.
- Belsky, J., Vandell, D. L., Burchinal, M., Clarke-Stewart, K. A., McCartney, K., & Owen, M. T. (2007). Are there long-term effects of early child care? *Child Development*, *78*, 681-701.
- Camilli, G., Vargas, S., Ryan, S., & Barnett, W. S. (2010). Meta-analysis of the effects of early education interventions on cognitive and social development. *Teachers College Record*, *112*, 579-620.
- Cote, S. M., Doyle, O., Petitclerc, A., & Timmins, L. (2013). Child Care in Infancy and Cognitive Performance Until Middle Childhood in the Millennium Cohort Study. *Child Dev.*
- Dearing, E., McCartney, K., & Taylor, B. A. (2009). Does Higher Quality Early Child Care Promote Low-Income Children's Math and Reading Achievement in Middle Childhood? *Child Development*, *80*, 1329-1349.
- Dunn, L. M., Dunn, D. M., Whetton, C., & Burlew, J. (1997). *British Picture Vocabulary Scale - 2nd ed.* Windsor England: NFER-Nelson.
- Dunn, L. M. & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test-3rd edition (PPVT-III)*. Circle Pines, MN.: American Guidance Service.
- Fidalgo, Z. & Pereira, F. (2005). Socio-cultural differences and adjustment of mothers' speech to their children's cognitive and language comprehension abilities. *Learning and Instruction*, *15*, 1-21.
- Hart, B. & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children*. Baltimore, MD: Brookes Publishing Company, Inc.
- Hawke, J. L., Olson, R. K., Willcutt, E. G., Wadsworth, S. J., & DeFries, J. C. (2009). Gender ratios for reading difficulties. *Dyslexia*, *15*, 239-242.
- Huttenlocher, J., Vasilyeva, M., Waterfall, H. R., Vevea, J. L., & Hedges, L. V. (2007). The varieties of speech to young children. *Dev. Psychol.*, *43*, 1062-1083.
- Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, *87*, 706-710.
- Janson, H. & (with Smith, L. (2003). *Norwegian manual supplement for the Ages and Stages Questionnaires [Norsk manualsupplement til Ages and Stages Questionnaires]*. Oslo, Norway: Regional Center for Child and Adolescent Mental Health, Halth region East/South.
- Lekhal, R., Zachrisson, H. D., Wang, M. V., Schjølberg, S., & von Soest, T. (2010). Does universally accessible child care protect children from late talking? Results from a Norwegian population-based prospective study. *Early Child Development and Care, iFirst Article*, 1-13.
- Leon, A. C. & Hedeker, D. (2007). Quintile stratification based on a misspecified propensity score in longitudinal treatment effectiveness analyses of ordinal doses. *Computational Statistics & Data Analysis*, *51*, 6114-6122.
- Li, W., Farkas, G., Duncan, G.J., Burchinal, M.R., & Vandell, D.L. (2013). Timing of high-quality child care and cognitive, language, and preacademic development. *Developmental Psychology*, *49*, 1440-1451.
- Loeb, S., Bridges, M., Bassok, D., Fuller, B., & Rumberger, R. W. (2007). How much is too much? The influence of preschool centers on children's social and cognitive development. *Economics of Education Review*, *26*, 52-66.
- Lyster, S.-A. H., Horn, E., & Rygvold, A.-L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge: Resultater fra en utprøving av British Picture Vocabulary Scale II, Second Edition (BPVS II). *Spesialpedagogikk*, *74*, 35-43.

- McCartney, K., Bub, K. L., & Burchinal, M. (2006). Selection, reflection, detection. *Monographs of the Society for Research in Child Development, 71*, 105-126.
- Ministry of Education. (2007). St.mld. nr. 16 (2006-2007) ...og ingen sto igjen. Tidlig innsats for livslang læring). Ref Type: Bill/Resolution
- NICHD Early Child Care Network (2006). Child-care effect sizes for the NICHD Study of Early Child Care and Youth Development. *American Psychologist, 61*, 99-116.
- Oshima-Takane, Y. & Robbins, M. (2003). Linguistic Environment of Secondborn Children. *First Language, 23*, 21-40.
- Pine, J. M. (1995). Variation in Vocabulary Development as a Function of Birth Order. *Child Development, 66*, 272-281.
- Rosenbaum, P. R. & Rubin, D. B. (1989). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.
- Spinath, F. M., Price, T. S., Dale, P. S., & Plomin, R. (2004). The genetic and environmental origins of language disability and ability. *Child Dev., 75*, 445-454.
- Squires, J., Bricker, D. D., & Potter, L. (1999). *The ASQ user's guide - Second edition*. Baltimore, MD: Brookes Publishing Company.
- StataCorp LP (2013). STATA 12 [Computer software]. College Station, TX.
- Strand, B. H., Dalgard, O. S., Tambs, K., & Rognerud, M. (2003). Measuring the mental health status of the Norwegian population: a comparison of the instruments SCL-25, SCL-10, SCL-5 and MHI-5 (SF-36). *Nord.J Psychiatry, 57*, 113-118.
- Tambs, K. & Moum, T. (1993). How well can a few questionnaire items indicate anxiety and depression? *Acta Psychiatr.Scand., 87*, 364-367.
- Winsvold, A. & Guldbrandsen, L. (2009). *Kvalitet og kvantitet - kvalitet i en barnehagesektor i sterk vekst (Quality and quantity - Quality in a child-care sector in rapid growth)* Oslo: NOVA.
- Zachrisson, H. D., Backer-Grøndahl, A., Nærde, A., & Ogden, T. (2012). *Smått er godt: Sosial kompetanse og atferd hos 3-åring - sammenhenger med barnehagebruk og kjennetegn ved barnegruppen* Oslo: Atferdssenteret.
- Zachrisson, H. D., Janson, H., & Nærde, A. (2013). Predicting early center care utilization in the context of universal access. *Early Childhood Research Quarterly, 28*, 74-82.
- Zachrisson, H. D., Lekhal, R., & Schjølberg, S. (2010). Barnehage og psykisk helse hos sped- og småbarn. In V.Moe, M. Bergum-Hansen, & K. Slinning (Eds.), *Norsk håndbok i sped- og småbarns psykiske helse* (Oslo: Gyldendal.
- Zambrana, I. M., Ystrøm, E., & Pons, F. (2012). Impact of gender, maternal education, and birth order on the development of language comprehension: A longitudinal study from 18 to 36 months of age. *Journal of Developmental and Behavioral Pediatrics, 33*, 146-55.
- Zambrana, I. M., Ystrom, E., Schjolberg, S., & Pons, F. (2013a). Action imitation at 1(1/2) years is better than pointing gesture in predicting late development of language production at 3 years of age. *Child Dev., 84*, 560-573.
- Zambrana, I. M., Pons, F., Eadie, P., & Ystrom, E. (2013b). Trajectories of language delay from age 3 to 5: persistence, recovery and late onset. *Int. J. Lang Commun Disord*, Online first..