

Kapitalisering av barnehagetjenester i eiendomsverdiene.

Øyvind Aas

Institutt for samfunnsøkonomi

Norges teknisk-naturvitenskaplige universitet

Desember 2016

Forord

Denne oppgaven markerer slutten på min 2-årige mastergrad i samfunnsøkonomi ved NTNU. Jeg ønsker å takke Lars-Erik Borge for god veiledning gjennom hele prosessen. Takk også til min samboer, familie og venner for god støtte.

Trondheim, 1. desember 2016.

Øyvind Aas

Innhold

1	Innledning	1
2	Bakgrunnsinformasjon	3
	2.1 Det norske boligmarkedet	3
	2.2 Norske barnehager	5
3	Teori	7
	3.1 Kapitalisering og Tibout-modellen	7
	3.2 Formelt rammeverk	8
	3.3 Tidligere studier	10
4	Data	14
	4.1 Boligvariabler	14
	4.2 Barnehagevariabler	14
	4.3 Kommunevariabler	19
5	Metodologi	21
	5.1 Empirisk spesifisering	21
	5.2 Metodologiske utfordringer	23
6	Resultater	28
	6.1 Minste Kvadraters Metode	28
	6.2 Robusthetsregresjoner	30
	6.3 Instrumentvariabelmetoden	34
7	Oppsummering og konklusjon	40

Tabeller

1	Prisvekst for boligtype og landsdel.	4
2	Boligvariable, deskriptiv statistikk.	14
3	Barnehagevariable, deskriptiv statistikk.	15
4	Barnehagevariablenes utvikling gjennom perioden.	16
5	Barnehagevariablenes gjennomsnitt for ulike landsdeler.	17
6	Korrelasjon mellom barnehagevariablene.	18
7	Barnehagevariablenes standardavvik, utvikling gjennom perioden.	19
8	Resultater fra “pooled” OLS, 2003-2006.	29
9	Effekt av boligvariabler.	30
10	År-for-år-regresjoner.	31
11	“Pooled” OLS-resultater for ulike innbyggertall.	32
12	“Pooled” OLS-resultater for ulike hustyper.	33
13	Regresjon med eiendomsskatt.	34
14	Regresjoner med færre barnehagevariabler.	35
15	Instrumentvariable, deskriptiv statistikk.	37
16	Resultater fra 2SLS.	38

Figurer

1	Nominell boligprisutvikling (2003-2006) i millioner kroner.	3
---	---	---

1 Innledning

På starten av 2000-tallet opplevde Norge en av de største satsingene på barnehage noensinne. Antall barn med barnehageplass økte markant, og det ble et uttalt politisk mål at alle som ønsket det skulle få gå i barnehage i hjemkommunen. Dette skulle oppnås blant annet gjennom økt statlig finansiering, pristak for foreldrebetaling og likebehandling av private og kommunale barnehager. Bakgrunnen for satsingen var at barnehagesektoren var et område med store forskjeller fra kommune til kommune, noe som førte til at tilgang på plass var svært avhengig av bosted. Det sto i kontrast til det øvrige offentlige tjenestetilbudet, hvor utformingen var mer universell. I henhold til Tiebout (1956) vil kommuner med ulikt tjenestetilbud tiltrekke seg husholdninger med forskjellige preferanser, noe som gjør barnehagesektoren egnet til å måle effekten av forskjeller i tilbudet. Men er etterspørselen etter barnehageplass så stor at norske foreldre er villige til å betale ekstra for å bo i en kommune med god barnehagedekning? Og hvis ja, avhenger denne effekten av hvor gamle barna er? Det er noen av spørsmålene denne oppgaven skal forsøke å besvare. Oppgavens problemstilling er: *Blir norske barnehagetjenester kapitalisert i eiendomsverdiene?* Med kapitalisering menes her at områder med et godt barnehagetilbud har høyere eiendomsverdier enn andre, alt annet likt. I tillegg til å se på tilgang på barnehageplass undersøker jeg også om kvalitetsfaktorer, slik som antall barn per ansatt og de ansattes utdanning, påvirker etterspørselen.

Oppgaven bygger videre på Borge og Rattsø (2014) sitt arbeid, hvor forfatterne fant at eiendomsskatt påvirker boligverdiene negativt, samtidig som god barnehagedekning har motsatt effekt. Jeg forsøker å utvide denne kunnskapen ved å dele opp målet på dekningsgrad mellom de yngste og eldste barna, og ved å inkludere faktorer som skal måle barnehagens kvalitet. For å besvare problemstillingen har jeg benyttet meg av data for alle boligtransaksjoner i Norge i perioden 2003 til 2006 og data for en rekke indikatorer for barnehagetilbudet til norske kommuner i samme periode. Jeg benytter Minste Kvadraters Metode, også kjent som “Ordinary Least Squares” (OLS). For å sikre at estimatene blir så presise som mulig kontrollerer jeg også for en rekke andre faktorer i kommunen som kan tenkes å påvirke boligprisene. Resultatene tyder på at dekningsgraden for både de eldste og de yngste barna, samt antall barn per ansatt, er faktorer som verdsettes av norske husholdninger, og som dermed kapitaliseres inn eiendomsverdiene. Jeg forsøker å gjøre disse resultatene så robuste som mulig ved å drøfte de metodologiske utfordringene analysen står overfor, og imøtegå disse ved hjelp av ulike økonomiske spesifikasjoner.

Resten av oppgaven er strukturert som følger: I del 2 utredes litt kortfattet bakgrunnsinformasjon for det norske boligmarkedet og barnehagene i den aktuelle perioden. I Del 3 utledes det et formelt økonomisk rammeverk som ligger til grunn

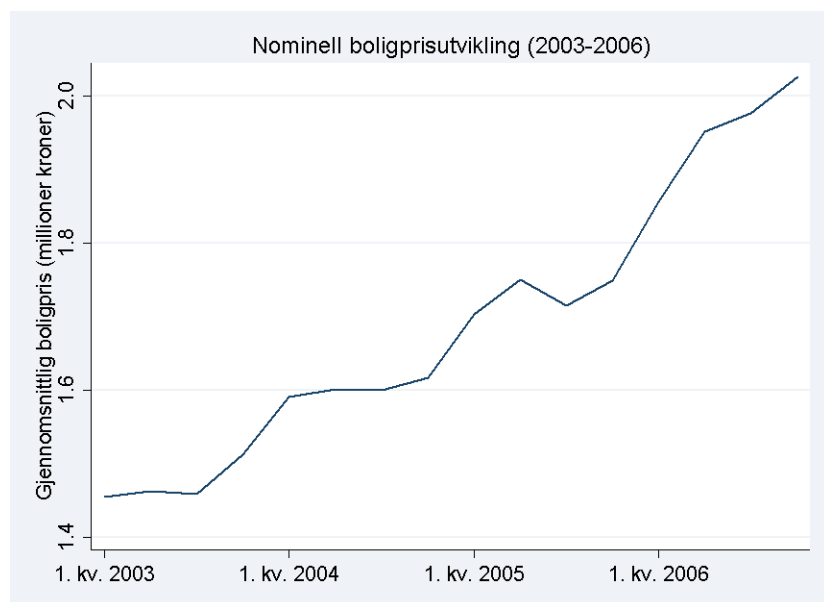
for teorien om kapitalisering av offentlige tjenester. I tillegg blir tidligere forskning gjennomgått. I del 4 redegjøres det for datamaterialet oppgaven benytter. Her legges det spesielt vekt på variablene som beskriver barnehagetilbudet, siden disse er de mest sentrale forklaringsvariablene. I del 5 formuleres det en empirisk spesifisering som legger grunnlaget for den kvantitative analysen. Her blir også en del metodologiske utfordringer oppgaven står overfor drøftet. Resultatene gjennomgås og diskuteres i del 6, før oppgaven rundes av med en oppsummering og konklusjon i del 7. Den siste delen inneholder også forslag til videre forskning på området.

2 Bakgrunnsinformasjon

2.1 Det norske boligmarkedet

Før man kan benytte data fra boligmarkedet til å si noe fornuftig om ulike etterspørselsfaktorer er det viktig å ha en forståelse for de generelle utviklingstrekkene i markedet. For Norge sin del har de siste årene vært preget av høy, og tilnærmet kontinuerlig, prisvekst. Med unntak av årene 2007 til 2009 har det vært en nesten sammenhengende vekst siden 1993 og fram til i dag (Statistisk sentralbyrå 2016a). Total prisvekst for perioden oppgaven ser på, 2003 til 2006, var på omtrent 34,2 %. Tilsvarende vokste konsumprisindeksen i samme periode med 4,3 %, noe som gir en realprisøkning på rundt 29,9 % (Statistisk sentralbyrå 2016b). Dette reflekterer at hele perioden var preget av økonomisk vekst generelt. Etter et par år med lavkonjunktur var bunnen nådd i starten av 2003, hvorpå norsk økonomi gikk inn i en høykonjunktur som varte ut 2007 (Statistisk sentralbyrå 2008a). Den kvartalsvise endringen i nominelle priser i perioden kan sees i figur 1. Dette er en figur som er beregnet ut i fra datmaterialet som benyttes i oppgaven, noe som innebærer at den ikke er justert for sesongmessige variasjoner. På fire år gikk altså gjennomsnittlig boligpris fra å være i underkant av 1,5 millioner kroner, til å være over 2 millioner kroner.

Figur 1: Nominell boligprisutvikling (2003-2006) i millioner kroner.



Kilde: Statistisk sentralbyrå, egne beregninger.

For å undersøke om veksten har vært forskjellig for ulike boligtyper og landsdeler presenterer jeg i tabell 1 nominelle veksttall som er oppdelt for disse kategoriene. Dette er også data basert på datasettet som benyttes gjennom hele oppgaven, og jeg har her

Tabell 1: Prisvekst for boligtype og landsdel.

Type hus	2003	2004	2005	2006
Enebolig	100.0	109.6	118.5	132.8
Småhus	100.0	107.8	119.2	134.5
Blokkleilighet	100.0	109.3	118.1	134.8
Del av landet				
Oslo og Akershus	100.0	109.2	118.3	134.4
Hedmark og Oppland	100.0	109.2	118.3	133.5
Østf/Busk/Vestf/Telem	100.0	109.3	118.4	133.9
Agder og Rogaland	100.0	109.2	118.4	134.0
Vestlandet	100.0	109.0	118.2	134.1
Trøndelag	100.0	109.0	118.3	134.1
Nord-Norge	100.0	108.5	117.7	133.6
Alle hus	100.0	109.2	118.4	134.2

Merk: Tallene er indeksert slik at 2003=100. *Kilde:* Statistisk sentralbyrå, egne beregninger.

indeksert tallene slik at 2003=100, noe som gjør det lettere å sammenligne. Total vekst i perioden kan da sees i kolonnen for 2006. Ser at eneboliger er den boligtypen som har hatt den flateste prisutviklingen, med en nominell vekst på rundt 32,8 %. Tilsvarende har blokkleiligheter hatt høyest vekst med 34,8 %. Geografisk er Oslo og Akershus den landsdelen som har hatt høyest prisvekst, mens vi finner den laveste veksten i Hedmark og Oppland. Prisutviklingen har likevel vært temmelig jevn over hele landet.

Det er en høy andel nordmenn som eier egen bolig. I 2015 bodde 82 % av nordmenn over 16 år i en bolig eid av husholdningen selv. Denne eierandelen har vært stabil i mange år, og var den samme i 1997 (Statistisk sentralbyrå 2015). For gruppen som er mest relevant for min analyse, par med barn i alderen 0-6 år, er eierandelen enda høyere. I perioden 2001 til 2007 har den ligget på mellom 88 og 90 % (Statistisk sentralbyrå 2008b). At så mange småbarnsfamilier eier i stedet for å leie er gunstig for den påfølgende analysen, siden det øker sannsynligheten for at kvalitet i barnehager skal slå ut i boligprisene. Dette er likevel ikke en avgjørende faktor, siden økt etterspørsel fra familier som leier uansett vil ha en indirekte effekt på boligprisen gjennom leiemarkedet. En annen faktor som kan bidra til at småbarnsforeldre har stor påvirkning på boligmarkedet er at denne gruppen er mer mobile enn foreldre med større barn. Blant personer som flyttet innenlands i 2007 var det nesten to og en halv ganger så mange i aldersgruppen 25 til 29 år som i gruppen 35 til 39 år (Statistisk sentralbyrå 2007a). Gjennomsnittlig alder for førstegangsfødende samme år var 28 år (Statistisk sentralbyrå 2007b). Til sist er det også bra for analysen at det var en økende eiendomsomsetning i perioden. Dette fordi høy omsetning gjør at etterspørselsendringer vil gi raskere utslag i markedet. Her gikk det fra 161 775 tinglyst omsatte eiendommer i 2003, til 179 820 i 2006 (Statistisk sentralbyrå 2008c).

2.2 Norske barnehager

Barnehagesektoren i Norge innehar stor variasjon i tilbud fra kommune til kommune, noe som gjør den egnet til å studere effekter av forskjeller i standard (Borge og Rattsø 2014). Fra å være et tiltak beregnet for barn som ikke får tilstrekkelig omsorg hjemme, er barnehagen i dag et tilbud som skal gis til alle som ønsker det. Ved innføringen av barnehageloven i 1975 ble barnehagene kommunenes ansvar, noe som bidrar til at de fleste barn går i barnehage i bostedskommunen (Barne- og familiedepartementet 2005). Denne sammenhengen har blitt ytterligere styrket gjennom at rett til barnehageplass ble lovfestet i 2009 (Utdanningsdirektoratet 2015). Andelen barn med barnehageplass steg markant i perioden oppgaven ser på. Dette må sees i sammenheng med den kraftige politiske satsingen som pågikk gjennom barnehagereformen. Reformen hadde sitt utspring i en avtale inngått i juni 2002 av opposisjonspartiene Arbeiderpartiet, Sosialistisk Venstreparti, Senterpartiet og Fremskrittspartiet. Sammen hadde disse partiene flertall i Stortinget. Hovedmålene i avtalen var full behovsdekning innen 2005 samt redusert foreldrebetaling (St. meld nr. 24 (2002-2003)). Dette ble vedtatt av et enstemmig Storting ett år senere i det såkalte barnehageforliket. For å oppnå målene ble blant annet statstilskuddet til barnehager økt, og det ble vedtatt at private og offentlige barnehager skulle likebehandles av kommunen (Innst. S. nr. 250 (2002-2003)).

Satsingen ga resultater, i 2003 var dekningsgraden for barn mellom 1 og 5 år på 69,1 %, mens den i 2006 var oppe på 80,4 %. De store forskjellene fra kommune til kommune ble noe mindre, men de var ikke i nærheten av å bli fullstendig utjevnet. På fylkesnivå var det i 2006 en forskjell på 10 prosentpoeng mellom fylket med lavest og høyest dekningsgrad. Disse var henholdsvis Sør-Trøndelag med nesten 87 %, og Oslo med i underkant av 77 % (Statistisk sentralbyrå 2007c). Den øvrige variasjonen i barnehagetilbudet blir gjennomgått grundigere i datakapittelet.

Når det gjelder målet om redusert foreldrebetaling var planen at dette skulle oppnås ved hjelp av en trinnvis innføring av maksimalpris. Denne skulle være på 2500 kr i 2004, og 1500 kr i 2005, hvor begge beløpene skulle beregnes ut i fra prisnivået i 2002. En forutsetning for å gjennomføre trinn 2 skulle være at målet om full barnehagedekning var oppnådd. Det ble også åpnet for at barnehager som ga et tilbud ut over det ordinære, samt barnehager som ville være truet av nedleggelse hvis prisen ble for lav, ville kunne overgå maksprisen. Dette måtte i så fall godkjennes av foreldreutvalget (Innst. S. nr. 250 (2002-2003)). Trinn 1 ble raskt gjennomført nesten som planlagt, med en innføring av maksimalsats på 2750 kr fra og med 1. mai 2004. Den nye satsen ble overholdt med få unntak, noe som førte til en kraftig reduksjon i foreldrebetalingen. Spesielt for private barnehager var reduksjonen merkbar, med en nedgang på 17 % fra august 2003 til august

2004 (Statistisk sentralbyrå 2004a). Målet for trinn 2 ble likvel ikke nådd av hverken den daværende eller etterfølgende regjeringer. Maksprisen for 2016 er 2655 kr, noe som ved hjelp av Statistisk sentralbyrå sin inflasjonskalkulator kan omregnes til 1997 kr i 2002-kroner (Utdanningsdirektoratet 2016, Statistisk sentralbyrå 2016b). Det vil si at målet for trinn 2 ikke er oppnådd den dag i dag. I 2006, som er det siste året som omfattes i min periode, ble maksprisen redusert til 2250 kr (Statistisk sentralbyrå 2006a). Etter dette har den kun blitt justert med konsumprisindeksen. Selv om prisene ikke har blitt redusert så mye som målsetningen i reformen, er det likevel ingen tvil om at det har vært en kraftig reduksjon i foreldrebetalingen. Dette kan illustreres av at de nominelle prisene i dag er lavere enn de var etter den første nedjusteringen for tolv år siden.

I motsetning til i skolesektoren tar det offentlige i bruk private barnehager for å nå sine dekningsmål. Tallene ovenfor dekker altså både offentlige og private barnehager. Andelen barn i private barnehager økte i perioden fra 42 % i 2003 til 46 % i 2006, det vil si at disse sto for flest av de nye barnehageplassene (Statistisk sentralbyrå 2004b, 2007c). I den påfølgende analysen blir kommunale og private barnehager behandlet på lik linje, med unntak for tallene for antall barn per ansatt, hvor det kun var tilfredsstillende data for kommunale barnehager.

3 Teori

3.1 Kapitalisering og Tiebout-modellen

I offentlig økonomi er kapitalisering et begrep som oftest benyttes i forbindelse med eiendomsskatt. Det beskriver da en prosess hvor den forventede nåverdien av alle framtidige skattebetalinger blir hensyntatt i eiendomsprisene umiddelbart ved innføringen av skatten. Den fulle kostnaden av skatten blir dermed belastet den nåværende innehaveren av eiendommen (Rosen og Gayer 2010). Dette kalles full kapitalisering, hvis deler av skattebyrden blir lagt på framtidige eiere har vi delvis kapitalisering (Borge og Rattsø 2014). Kapitalisering av offentlige tjenester, som denne oppgaven omhandler, fungerer på samme måte. Tanken er at økt kvalitet på barnehagetilbudet kommer nåværende boligeiere til gode i form av høyere boligpriser. Siden det er vanskelig å fastslå nåverdi av en kvalitetsendring vil oppgaven ikke se på grad av kapitalisering, men kun på om det forekommer.

Den teoretiske begrunnelsen for at barnehagetjenestenes standard skal slå ut i boligprisene er basert på Tiebout (1956). I hans modell vil ulikt skatte- og servicenivå mellom lokalområder føre til at innflyttere etablerer seg i det området som speiler husholdningens preferanser best. Hvis det er tilstrekkelig mange lokale enheter, full informasjon, og ingen begrensninger på husholdningens mobilitet, så vil man nå ha en slags markedsløsning for produksjon og konsum av offentlige goder. Tiebout argumenterer også for at dette fører til et effektivt tilbud av offentlig tjenester uten innblanding fra høyere styringsnivå. Overført til praksis innebærer Tiebout-modellen at dersom norske husholdninger verdsetter barnehagedekning når de velger bolig, så vil det føre til at kommuner med god barnehagedekning, alt annet likt, har høyere boligpriser enn kommuner med dårlig dekning.

En kritikk mot modellen er at husholdningene ikke er så mobile som den forutsetter. For eksempel er det lett å tenke seg at forhold som jobbsituasjon og familie begrenser utvalget av potensielle boområder så mye at skatte- og servicenivå ikke lenger spiller inn. På den andre siden har urbaniseringen av samfunnet ført til at det ofte er et stort antall boligområder tilgjengelig i områdene rundt de sentrale arbeidsplassene, noe som gjør modellen mer realistisk (Oates 1969). Den norske strukturen med mange og små kommuner gjør dette argumentet spesielt relevant i denne oppgaven. Det er ikke vanskelig å tenke seg at en nyinnflyttet arbeidstaker i Trondheim kan vurdere å bo i nabokommunene Malvik eller Klæbu. Som vist tidligere er i tillegg unge voksne mer mobile enn eldre, noe som betyr at aldersgruppen som er mest relevant for oppgaven også er den som ligger nærmest mobilitetsforutsetningen i Tiebout-modellen.

En annen antakelse som kan kritiseres for å være urealistisk er forutsetningen om full informasjon. Denne oppgaven baserer seg på tall som er offentlig tilgjengelige, og som dermed er fullt mulig å innhente for eksempel av en barnefamilie på flyttefot. Det er likevel mulig at informasjonsinnhenting fra slike kilder krever en grad av interesse og spesialkunnskap som gjør at få familier faktisk legger inn den nødvendige innsatsen. Det er også sannsynlig at disse tallene ikke var like tilgjengelige for brukerne under perioden jeg ser på. Ved starten av perioden var bruken av internett i Norge i oppstartsfasen, men barnefamilier var overrepresentert blant husholdningene som hadde tilgang til dette relativt nye informasjonsverktøyet. I 2003 hadde 74 % av barnefamilier tilgang til internett, mot 55 % for alle husholdninger totalt (Statistisk sentralbyrå 2003). I 2006 hadde disse andelene økt til henholdsvis 86 og 69 % (Statistisk sentralbyrå 2006b). Med andre ord sto barnefamiliene i perioden overfor fallende kostnader ved informasjonsinnhenting, siden internett økte tilgjengeligheten til relevant informasjon. Nøyaktig hva som var tilgjengelig på internett på den tiden er likevel vanskelig å avdekke i dag. Barnehagedata fra Statistisk sentralbyrå har blitt publisert elektronisk siden 1999, men deres nettsider var kanskje ikke like lette å navigere som de er i dag. Brorparten av informasjonsinnhenting vil uansett foregå på et uformelt nivå, hvor familier som har barn i en barnehage forteller det videre hvis de er fornøyde eller misfornøyde. Selv om ingen offentlige data er tilgjengelige vil man altså likevel kunne få en kapitaliseringseffekt gjennom jungeltelegrafene. Til sist er det viktig å huske på at en effekt på boligprisene ikke bare trenger å skyldes innflyttere fra andre kommuner, det kan like gjerne være fordi husholdninger blir værende. For eksempel kan det være en familie som har behov for et større hus, og som kjøper dette i samme kommune fordi de er fornøyde med barnehagetilbudet.

3.2 Formelt rammeverk

Som teoretisk rammeverk i denne oppgaven benytter jeg en standard modell for kapitalisering av offentlige tjenester, definert i Ross og Yinger (1999) og Nguyen-Hoang og Yinger (2011). Modellen bygger opprinnelig på Ellickson (1971). Jeg ser på en mengde husholdninger som kan ha ulik inntekt og ulike preferanser, men som alle kan plasseres i en distinkt gruppe med andre som har samme kombinasjon av inntekt og preferanser. Det forutsettes at husholdningene er mobile, noe som innebærer at alle husholdninger i samme inntekts- og preferansegruppe må oppnå samme nyttenivå. Implisitt betyr dette at vi ikke har noen flyttekostnader. Modellen tar for seg et bebyggt område med mange kommuner. Disse har klare grenser og varierende kvalitet på de offentlige tjenestene. Husholdningen kan bare få tilgang til de offentlige tjenestene hvis den bor i kommunen, og alle innbyggere i kommunen mottar samme nivå av offentlig tjenester. Jeg antar at alle husholdningene eier sin egen bolig, men modellen kunne like gjerne omhandlet leiemarkedet.

Husholdningens nytte avhenger av boligkvalitet, H , kvalitet på offentlige tjenester, G , og et privat gode, Z . Dette gir følgende nyttefunksjon:

$$U(Z, H, G) = U^0(Y) \quad (1)$$

$U^0(Y)$ viser nyttenivået til en husholdning med inntekt Y . H inneholder her alle aspekter ved boligen, inkludert kvalitet på omgivelsene. Z er et sammensatt, privat gode som illustrerer alt annet en husholdning bruker penger på. Antar at alle disse godene gir positiv nytte, men at denne effekten er avtakende. Det vil si $U_Z > 0$, $U_{ZZ} < 0$, $U_H > 0$, $U_{HH} < 0$, $U_G > 0$ og $U_{GG} < 0$, hvor fotskriften betegner første- og andreordens partiellderiverte. Husholdningens budsjettbetingelse innebærer at inntekten er lik konsum, noe som gir:

$$Y = Z + PH + tV \quad (2)$$

P er her prisen på en enhet boligkvalitet. Dette kan framstå noe abstrakt, men viser altså hvor mye en gitt mengde av de samlede egenskapene ved en bolig, H , verdsettes i markedet. t er den effektive skatteraten på eiendom, det vil si nominell skatterate multiplisert med anslått boligverdi, og dividert på faktisk salgsverdi. Grunnen til at anslått og faktisk verdi må adskilles her er at eiendom beskattes ut ifra anslått verdi, samtidig som reell skattebelastning avhenger av hva eiendommen faktisk blir solgt for. Hvis boligprisene for eksempel øker mye uten at de anslåtte boligverdiene blir justert tilsvarende, så vil det innebære en reduksjon i den effektive skatteraten på eiendom. V viser her markedsverdien av et hus. Siden vi kalkulerer med framtidig markedsverdi og beskatning må disse faktorene neddiskonteres med en diskonteringsrente, r . Det gir da $V = \frac{PH}{r}$ og $t^* = \frac{t}{r}$, hvor t^* nå beskriver neddiskontert effektiv skatterate. For å oppsummere viser (2) at husholdningen bruker sin inntekt, Y , på en mengde private goder, Z , boligjenester, som totalt koster PH , og eiendomsskatt, tV . Ved første øyekast kan det her se ut som at utgiftene til eiendomsskatt avhenger av markedsverdi, ikke anslått verdi. At dette ikke er tilfelle kan raskt vises ved å benytte at definisjonen til den effektive skatteraten er $t = \frac{TA}{V}$, hvor T er nominell skatterate og A er anslått boligverdi. Ser da at faktiske utgifter til eiendomsskatt blir $\frac{TA}{V} * V = TA$. Kan skrive om (2) til:

$$Y = Z + PH + t \frac{PH}{r} = Z + PH + t^* PH = Z + PH(1 + t^*) \quad (3)$$

Problemet husholdningen står overfor er hvor mye de skal være villige til å betale for H , gitt kvaliteten på de lokale offentlige tjenestene, G , og den effektive skatteraten, t . Dette kan formuleres ved å løse (3) for P , og maksimere med hensyn på H og Z , hvor (1) er en bibetingelse. Det vil si at vi ser på hva som er maksprisen en husholdning er villig til å gi for en bolig i en kommune med et gitt nivå på offentlige tjenester, når vi holder nytten

konstant. Maksimeringsproblemet blir da som følger:

$$\begin{aligned} \max_{H,Z} P &= \frac{Y - Z}{H(1 + t^*)} \\ \text{gitt } U(Z, H, G) &= U^0(Y) \end{aligned} \quad (4)$$

Funksjonen som maksimeres i (4) kalles ofte for en “bid rent”-funksjon, og den viser hvordan de ulike faktorene påvirker prisen husholdningen er villig til å betale. Ser av telleren at husholdningen naturlig nok har mulighet til å betale mer jo høyere inntekt de har, og mindre hvis de bruker mye penger på andre private goder, det vil si hvis Z har en høy verdi. At H inngår positivt i nevneren kan se noe ulogisk ut, men her er det viktig å huske på at vi ser på pris per enhet boligjenester. Jo færre enheter husholdningen kjøper, jo høyere pris kan de betale per enhet. (4) gir følgende Lagrange-funksjon:

$$\mathcal{L} = \frac{Y - Z}{H(1 + t^*)} - \lambda(U(Z, H, G) - U^0(Y)) \quad (5)$$

Resultatet jeg er interessert i er hvordan prisen endres som følge av endret offentlig tjenestetilbud, det vil si et uttrykk for P_G . Siden G og t her blir behandlet som parametre vet jeg gjennom “envelope”-teoremet at $P_G = \mathcal{L}_G$ i optimum. Fra å derivere (5) med hensyn på G får jeg:

$$\mathcal{L}_G = -\lambda U_G \quad (6)$$

Løser for λ gjennom førsteordensbetingelsen med hensyn på Z :

$$\mathcal{L}_Z = -\frac{1}{H(1 + t^*)} - \lambda U_Z = 0 \quad (7)$$

$$\lambda = -\frac{1}{U_Z H(1 + t^*)} \quad (8)$$

(8) inn i (6) gir da

$$P_G = \frac{U_G/U_Z}{H(1 + t^*)} \quad (9)$$

Ser at effekten kvaliteten på offentlige tjenester har på boligprisene avhenger av den marginale substitusjonsraten mellom G og Z . Denne kan tolkes som marginalnyttens i kroner av økt kvalitet på offentlige tjenester, siden Z betegner alt annet en husholdning bruker penger på. Så lenge offentlige tjenester blir verdsatt av husholdningene ($U_G > 0$), så vil de da bli kapitalisert inn i boligprisen.

3.3 Tidligere studier

Oates (1969) var trolig den første som testet Tiebout-modellen empirisk. Han benyttet tversnittsdata fra New Jersey, USA, til å teste effekten eiendomsskatt og offentlig utgifter

til skole hadde på eiendomsprisene. Resultatene ga støtte for Tibouts hypotese om at husholdninger velger bosted ut i fra sine preferanser om skattenivå og kvalitet på offentlige tjenester. Et viktig poeng fra artikkelen er at kapitalisering av offentlige tjenester kan veie opp for den negative kapitaliseringseffekten av eiendomsskatt. Det vil si at hvis inntektene fra skatten går til å bedre offentlig tjenestetilbud, så kan innføring av eiendomsskatt faktisk øke eiendomsverdiene.

Oates sin artikkel inspirerte en rekke etterfølgende empiriske arbeid. De tidlige studiene på eiendomsskatt oppsummeres av Yinger m.fl (1988). Alle studiene finner at eiendomsskatt har en negativ påvirkning på eiendomsverdiene. Den anslåtte graden av kapitalisering varierer for de fleste studiene mellom 50 og 100 %. Disse anslagene avhenger naturligvis av hvilken diskonteringsrente som legges til grunn. Forfatterne argumenterer for at de fleste studiene overestimerer kapitaliseringsgraden, fordi det benyttes en diskonteringsrente på mellom 5 og 6 %, noe de mener er for høyt. Ved å nedjustere diskonteringsrenta til 3 % finner de at de fleste studiene gir en grad av kapitalisering på mellom 15 og 60 %. For det høyeste anslaget vil dette innebære at en innføring av eiendomsskatt fører til at nåværende boligeiere må betale 60 % av den totale skattebelastningen. Siden fokuset i oppgaven er på kapitalisering av offentlige tjenester, ikke eiendomsskatt, så går jeg ikke nærmere inn på dette i analysen.

Studier som omhandler kapitalisering av offentlige tjenester har i all hovedsak fokusert på effekten av skoletilbud. En av grunnene til dette er at skolesektoren inneholder et egnet kvalitetsmål i form av testresultater. Som påpekt av Bradford, Malt og Oates (1969) og Charney (1993) er offentlig pengebruk et upresist mål på kvalitet, fordi områder med vanskeligere forhold, for eksempel mye kriminalitet, også må bruke mer penger. Tilsvarende kan en kommune ha lav pengebruk på helse fordi innbyggerne i kommunen er forholdsvis friske. Ved å benytte helseutgifter som uavhengig variabel kunne man dermed risikert å konkludere med at økt satsing på helse gir mer sykdom. Dette er en viktig grunn til at mange studier på kapitalisering av skolekvalitet, for eksempel McDougall (1976), Rosen og Fullerton (1977) og Haurin og Bresington (1996), benytter elevenes testresultater som uavhengig variabel i sin regresjon. I stedet for å benytte en variabel som måler en innsatsfaktor i produksjon av skolekvalitet kan man nå benytte en variabel som man håper kan måle kvaliteten direkte. Alle disse studiene finner en positiv effekt av skolekvalitet på boligprisene. Ross og Yinger (1999) oppsummerer en rekke andre studier på området, felles for alle er at de også finner effekt av testresultater i skolen på boligprisene.

Nyere studier på skole oppsummeres av Nguyen-Hoang og Yinger (2011). Her viser de til empiriske studier fra åtte forskjellige land som alle viser at skolekvalitet blir

kapitalisert i eiendomsverdiene. Resultatene er sterkest for studiene som benytter testresultater som mål på kvalitet. Forfatterne framhever at et av hovedproblemene når OLS benyttes er å skille effekten som skyldes skolekvalitet fra effekter som skyldes andre kvaliteter boligområdet innehar. Dette forsøkes løst på mange forskjellige måter i litteraturen. Den mest intuitive metoden er å inkludere mange variable som skal kontrollere for andre aspekter ved området, slik som arbeidsledighet, klima, avstand til byer, naturområder og så videre. Men, det kan likevel være uobserverte effekter som man ikke får kontrollert for. Flere studier, for eksempel Downes og Zabel (2002), Gibbons og Machin (2003) og Rosenthal (2003), benytter derfor instrumentvariabelmetoden for å løse problemet. Andre benytter faste effekter for å filtrere ut den såkalte nabolagseffekten. Blant disse er Black (1999) en spesielt innovativ studie, hvor faste effekter kombineres med diskontinuitetsanalyse. For å unngå problemer med utelatte variable sammenligner forfatteren her boligprisene på hver sin side av grensene til skoledistriktene i Massachusetts. Kun skoledistrikter som entydig avgjør hvilken skole barn må gå på ut i fra husets lokasjon er inkludert i analysen. Siden hun kun benytter hus fra samme by filtreres alle faste effekter ut, slik som offentlig pengebruk og skattenivå. Ved å sammenligne hus som er i samme nabolag, men på hver sin side av grensa, filtrerer hun også ut de uobserverte nabolagseffektene. Resultatet blir dermed at kapitaliseringseffekten av skolekvalitet blir isolert på en presis måte. Andre studier benytter en mer tradisjonell “faste effekter”-metode. For eksempel benytter Clapp, Nanda og Ross (2008) variasjon innad i folketellingsområder i Connecticut, og Dee (2000) benytter variasjon innad i amerikanske stater som følge av ny lovgivning. Alle disse finner en kapitaliseringseffekt av skolekvalitet. Jevnt over er likevel effekten som anslås lavere i studiene som benytter faste effekter enn de som kun benytter tversnittsdata (Nguyen-Hoang og Yinger 2011). Til slutt er det et viktig poeng fra Carlsen m.fl (2009) at ingen av studiene som benytter testresultater har bidratt til å utvikle kvalitetsmål som kan benyttes for andre offentlige tjenester. Litteraturen på skole har altså ikke gitt noen bredere forståelse av hvordan lokale myndigheter kan endre tjenestesammensetningen for å tiltrekke seg nye husholdninger.

Av norske studier på området er det atskillig færre. Fiva og Kirkebøen (2011) benytter offentliggjøring av tidligere upubliserte skolerresultater til å undersøke effekten på boligprisene i Oslo. De finner en tydelig positiv effekt på kort sikt, men at prisene faller tilbake til det gamle nivået etter mellom to og tre måneder. Artikkelen bidrar likevel til å vise at skolekvalitet verdsettes av husholdningene. Borge og Rattsø (2014) benytter data fra norske kommuner til å påvise en negativ kapitaliseringseffekt av eiendomsskatt og en positiv effekt av barnehagedekning. De påpeker at boligprisene, eiendomsskatten og tjenestetilbudet kan være simultant bestemt, siden lokale myndigheter kan endre de to sistnevnte som følge av befolkningens flyttemønster. For å rette opp for dette benytter de instrumentvariabelmetoden. Her instrumenteres eiendomsskatten og barnehagedekningen

med en Hefindahl-indeks for partifragmentering i kommunestyret, andel sosialistiske representanter i kommunestyret, en dummyvariabel for bystatus og andel kvinner i kommunestyret. "Two Stage Least Squares" (2SLS) gir et mye sterkere resultat enn OLS, noe som tyder på at simultaniteten gir en skjevhet i forventet retning. Dette gjelder både variabelen for eiendomsskatt og variabelen for dekningsgrad.

Carlsen (2005) benytter data fra norske spørreundersøkelser til å fastslå hvilke lokale faktorer som verdesettes av husholdninger som planlegger å flytte. Han finner at offentlige tjenester, spesielt de som er rettet mot barnefamilier, samt kultur, er viktige faktorer ved valg av bosted. I en påfølgende studie finner Carlsen m.fl (2009) at kultur, helse og offentlig transport er de tjenestene som verdsettes høyest. De finner overraskende nok ingen signifikant effekt av barnehage eller skole på flyttebeslutningene. En svakhet med begge disse studiene er at man ikke vet sikkert om respondentene faktisk flyttet i etterkant av undersøkelsen.

Som denne litteraturgjennomgangen har vist er det kun Borge og Rattsø (2014) som spesifikt ser på barnehager i sin analyse. Den store mengden artikler som omhandler skole er likevel svært relevant for min analyse, på grunn av de store likhetene mellom sektorene. Både skole og barnehage er tilbud rettet mot barnefamilier, og de er tjenester som tilbys nært brukernes bosted. Ved estimering av kapitaliseringseffekten ved hjelp av OLS står jeg overfor de samme utfordringene som studiene over, og disse vil bli grundig drøftet i kapittel 5.

4 Data

4.1 Boligvariabler

Som mål på boligverdier har jeg benyttet et datasett fra Statistisk sentralbyrå for alle hustransaksjoner i Norge i perioden 2003-2006. Jeg hadde også tilgang til transaksjoner for årene 2001 og 2002, men disse ble utelatt fordi de manglet informasjon om husets alder. Datasettet inneholder informasjon om pris, antall kvadratmeter, husets byggeår og type hus. Type hus betyr at alle boligene er fordelt i en av tre kategorier, enebolig, småhus eller blokkleilighet. Enebolig omfatter her ren enebolig, generasjonsbolig og enebolig med hybel-/sokkelleilighet. Småhus er definert som tomannsboliger, rekkehus og andre småhus, og blokkleilighet er en leilighet i boligblokk med to etasjer eller flere (Statistisk sentralbyrå 2016a). Deskriptiv statistikk for disse variablene kan sees i tabell 2. Ser at over halvparten av de omsatte boligene i perioden er blokkleiligheter.

Tabell 2: Boligvariable, deskriptiv statistikk.

Variabel	Antall	Gj.snitt	St.avvik	Min	Median	Maks
Areal	236 709	100.71	52.06	15	88	500
Alder	206 091	36.55	24.44	0	34	106
Enebolig	75 817					
Småhus	31 484					
Blokkleilighet	135 634					

Merk: Areal måles i kvadratmeter, alder i antall år. *Kilde:* Statistisk sentralbyrå, egne beregninger.

4.2 Barnehagevariabler

Data for barnehager er alle hentet fra Statistisk sentralbyrå sin KOSTRA-database. Dette er en database som gir statistikk om ressursinnsats, prioriteringer og måloppnåelse i norske kommuner, bydeler og fylkeskommuner, hvor KOSTRA står for Kommune-Stat-Rapportering (Statistisk sentralbyrå 2016c). Jeg har benyttet data for barnehagedekning, styrers utdanning, de ansattes utdanning og barn per ansatt for alle norske kommuner i perioden 2003-2006. Som nevnt over er pengebruk et upresist mål på kvalitet i offentlige tjenester, noe som er grunnen til at dette ikke er inkludert i analysen.

Tallene jeg benytter for barnehagedekning er oppdelt mellom aldersgruppene 1 til 2 år og 3 til 5 år. Alle viser andelen barn med barnehageplass i forhold til populasjonen av barn i aldersgruppen. Siden populasjonen her omfatter alle barn bosatt i kommunen, ikke bare de som etterspør barnehageplass, så sier ikke disse tallene noe om hvor stor andel av etterspørselen som ble dekt opp av tilbudet. Det er likevel mest sannsynlig at kommuner

med lav dekningsgrad også har en høyere andel barn på venteliste enn kommuner hvor dekningsgraden er høy. Tallene for dekningsgrad er kun et mål på kvantitativ produksjon, og de sier ingenting om barnehagekvaliteten i kommunen.

De øvrige variablene som benyttes er inkludert for å måle kvalitet. Jeg har her ikke tilgang til data som måler kvaliteten direkte, slik som testresultater i litteraturen som omhandler skolekvalitet. Jeg må derfor forsøke å måle kvaliteten indirekte gjennom kvantitative tall for innsatsfaktorer som forventes å være kvalitetshevende. Et av disse er andel styrere og pedagogiske ledere med godkjent barnehagelærerutdanning. Den godkjente utdanningen er det 3-årige barnehagelærerstudiet som tilbys på norske høyskoler, og som tidligere het førskolelærerutdanning. Fra og med 1995 ble det lovpålagt at man skal ha denne utdanningen for å bli ansatt som styrer eller pedagogisk leder i en barnehage (Denk 2009). Jeg har brukt tilsvarende tall for de ansatte i barnehagen. Også her er målet som benyttes andel ansatte med godkjent barnehagelærerutdanning, og definisjonen for godkjent utdanning er den samme. Det stilles likevel ikke de samme kravene til de ansattes utdanning som til lederne, og jeg forventer derfor at andelen her er mye lavere. Til sist benytter jeg data for antall barn, korrigert for alder, per årsverk til basisvirksomhet i kommunale barnehager. Her er barn mellom 0 og 2 år vektet med 2, barn på 3 år er vektet med 1,5, og de mellom 4 og 5 år er vektet med 1. Grunnen til at tallene er vektet på denne måten er at yngre barn er mer ressurskrevende enn eldre. Følgelig vil det gå mer ut over de ansattes tid til å ta seg av de andre barna hvis det starter et nytt barn på 1 år i barnehagen, enn hvis dette barnet er 4 år.

Tabell 3: Barnehagevariable, deskriptiv statistikk.

Variabel	Antall	Gj.snitt	St.avvik	Min	Median	Maks
Dekningsgrad 1-2 år	1732	0.53	0.16	0	0.53	1.50
Dekningsgrad 3-5 år	1732	0.91	0.08	0.51	0.87	1.58
Ansattes utdanning	1727	0.31	0.08	0	0.32	0.82
Styrers utdanning	1729	0.87	0.15	0	0.91	1
Barn per ansatt	1724	7.15	1.12	3.60	7.10	17.20

Kilde: Statistisk sentralbyrå, egne beregninger.

Deskriptiv statistikk for de ulike barnehagevariablene presenteres i tabell 3. Ser at dekningsgraden er mye høyere for de eldste enn for de yngste barna, med landsgjennomsnitt på henholdsvis 91 og 53 %. Dekningsgraden for de yngste varierer også mye mer fra kommune til kommune, noe som går fram av at standardavviket er dobbelt så høyt som for barn mellom 3 og 5 år. At maksverdien til begge dekningsgradene er over 100 % virker ved første øyekast merkelig, men dette er kun tilfelle for noen små kommuner. I disse kommunene vil det faktisk at noen få barn fra nabokommunen krysser

kommunegrensa for å gå i barnehage være nok til å vippe dekningsgraden over 100 %. Et alternativ kunne vært å utelate observasjonene fra disse kommunene, men antallet omsatte hus i de aktuelle kommunene er så få relativt til resten av utvalget at det ikke ville påvirket resultatene uansett. Siden de høye dekningsgradene i disse kommunene ikke skyldes noen feil i datamaterialet mener jeg derfor det er ryddigst å beholde alle observasjonene.

Som forventet ser vi også av tabell 3 at andelen med godkjent barnehagelærerutdanning er mye høyere for styrere og pedagogiske ledere enn den er for de øvrige ansatte. Det er tydelig at barnehagesektoren har en temmelig høy andel ufaglært arbeidskraft, i gjennomsnitt har kun 31 % av de ansatte godkjent barnehagelærerutdanning. Siden det er lovpålagt at styrerne har godkjent utdanning er det noe overraskende at ikke andelen her er enda høyere enn gjennomsnittet på 87 %. Dette kan tyde på at kravet ikke følges like nøye over alt. Det er likevel mulig at en del av de som ikke har godkjent utdanning er personer som enten har en annen utdanning som blir betraktet som tilfredsstillende, eller som har drevet barnehage helt siden før kravet ble innført.

Tabell 4: Barnehagevariablenes utvikling gjennom perioden.

Variabel	2003	2004	2005	2006
Dekningsgrad 1-2 år	0.453	0.496	0.548	0.635
Dekningsgrad 3-5 år	0.879	0.904	0.924	0.940
Ansattes utdanning	0.311	0.315	0.320	0.310
Styrers utdanning	0.890	0.874	0.874	0.841
Barn per ansatt	7.36	7.23	7.12	6.88

Merk: Verdier er gjennomsnitt per kommune. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Barnehagevariablenes utvikling gjennom perioden går fram av tabell 4. Dette er gjennomsnittstall på kommunenivå, det vil si at for eksempel tallene for dekningsgrad ikke viser den totale dekningsgraden for landet, men norske kommuners gjennomsnittlige dekningsgrad. Ser at dekningsgraden har økt for begge aldersgruppene, men at det er dekningsgraden for de yngste som har økt mest i perioden. Økningen per år har for denne aldersgruppen vært akselererende, med en vekst på 4,4 prosentpoeng mellom de to første årene i perioden, og hele 8,7 prosentpoeng mellom de to siste. For de eldstes dekningsgrad har trenden vært omvendt, noe som kan tyde på at man i perioden har nådd tilnærmet full barnehagedekning for denne gruppen. Her vokste dekningsgraden med 2,5 prosentpoeng det første året, og 1,6 prosentpoeng det siste. At gjennomsnittet for norske kommuner er så høyt som 94 % for den eldste aldersgruppen i 2006 viser ikke bare at barnehagetilbudet har økt, men også at etterspørselen etter barnehageplass er stor i denne gruppen.

Tabell 4 viser også at antall barn per ansatt har hatt en jevn reduksjon gjennom

perioden. Dette er kanskje noe overraskende, og det viser at den økte dekningsgraden ikke er oppnådd ved å tøyne grensene for antall barn i barnehagen. Snarere tyder det på at den skyldes økt utbygging og flere ansatte. At andelen styrere med godkjent barnehagelærerutdanning har falt gjennom perioden kan tyde på at utbyggingen førte til at dette kravet ikke ble håndhevet like sterkt, siden hovedmålet var å øke dekningsgraden. Som vist over sto også private barnehager for den største veksten, og det er sannsynlig at kravet til godkjent utdanning ikke følges like nøye her som i kommunale barnehager. Andelen ansatte med godkjent utdanning har ligget ganske stabilt mellom 31 og 32 %. At den ikke har falt i takt med den kraftige satsingen tyder på at både nye barnehager og barnehager som har utvidet kapasiteten har hatt tilgang på kvalifisert arbeidskraft.

I analysen benytter jeg ikke bare tidsvariasjonen i barnehagevariablene, men også tversnittsvariasjonen mellom kommunene. Det kan derfor være interessant å se nærmere på hvordan variasjonen er mellom ulike deler av landet. Dette presenteres i tabell 5. Også her er tallene som oppgis et gjennomsnitt for kommunene i landsdelen, det vil si at for eksempel tallet 0.534 for de yngstes dekningsgrad i Oslo og Akershus ikke betyr at den totale dekningsgraden i denne landsdelen er 53,4 %. Som forventet er det temmelig stor variasjon mellom landsdelene for alle variablene. For de yngstes dekningsgrad er det hele 12,3 prosentpoengs forskjell mellom landsdelen med lavest kommunalt gjennomsnitt, som er Agder og Rogaland med 46,2 %, og landsdelen som har høyest, som er Nord-Norge med 58,5 %. For de eldste er det en forskjell på 6,7 prosentpoeng. Her har Hedmark og Oppland høyest gjennomsnittlig dekningsgrad med 94,1 %, og Oslo og Akershus har den laveste med 87,4 %.

Tabell 5: Barnehagevariablenes gjennomsnitt for ulike landsdeler.

Del av landet	Dekn. 1-2	Dekn. 3-5	Ansatte	Styrere	Barn pr.
Oslo og Akershus	0.534	0.874	0.289	0.809	6.54
Hedmark og Oppland	0.523	0.941	0.311	0.877	7.59
Østf/Busk/Vestf/Telem	0.512	0.906	0.327	0.910	7.31
Agder og Rogaland	0.462	0.883	0.328	0.908	7.43
Vestlandet	0.529	0.931	0.317	0.865	7.44
Trøndelag	0.565	0.940	0.326	0.892	7.61
Nord-Norge	0.585	0.893	0.293	0.819	6.20
Hele landet	0.533	0.912	0.314	0.870	7.15

Merk: Tallene viser gjennomsnitt for barnehagene i landsdelen. *Kilde:* Statistisk sentralbyrå, egne beregninger.

For de øvrige variablene er det interessant å merke seg at landsdelene med færrest barn per ansatt, det vil si Oslo og Akershus og Nord-Norge, også er de landsdelene som har lavest andel styrere med godkjent utdanning. For å undersøke om dette skyldes tilfeldigheter, eller om det er en faktisk korrelasjon mellom disse variablene, presenterer

jeg i tabell 6 toveis korrelasjon mellom alle barnehagevariablene. Tallene som presenteres er korrelasjonskoeffisienten variablene i mellom. Hvis variablene korrelerer samme vei, for eksempel hvis høyere dekningsgrad for de yngste er korrelert med høyere dekningsgrad også for de eldste, så vil korrelasjonskoeffisienten mellom dem være positiv. Tilsvarende vil den være negativ hvis høyere dekningsgrad for de yngste er korrelert med lavere dekningsgrad for de eldste. En korrelasjonskoeffisient på -1 eller 1 innebærer at variablene er perfekt korrelert, og en verdi på 0 viser at de er fullstendig ukorrelert (Wooldridge 2013:733).

Tabell 6: Korrelasjon mellom barnehagevariablene.

	Dekn. 1-2	Dekn. 3-5	Ansatte	Styrere
Dekn. 3-5	0.59			
Ansatte	-0.05	-0.004		
Styrere	-0.12	-0.05	0.70	
Barn pr.	-0.21	0.06	0.13	0.11

Merk: Verdiene i tabellen er korrelasjonskoeffisienten mellom variablene. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Tabell 6 viser at de fleste av barnehagevariablene er relativt svakt korrelert med hverandre. Unntakene er at de to variablene for dekningsgrad og utdanning er ganske sterkt korrelert, med en koeffisient på henholdsvis 0,59 og 0,70. At dekningsgradene er korrelert er ikke overraskende, siden satsinger på barnehage sjeldent prioriterer mellom aldersgrupper. Dette bekreftes av tabell 6, som viser at kommuner med god barnehagedekning for de minste oftere har god dekning også for de eldste. Tilsvarende har kommuner med en høy andel styrere med godkjent utdanning som regel også mange ansatte med godkjent utdanning. Dette kan kanskje skyldes en kombinasjon av at disse kommunene har god tilgang på utdannet arbeidskraft og at styrere med godkjent utdanning i større grad ansetter personer med samme utdanning som dem selv. Teorien om at variabelen for barn per ansatt og andel styrere med godkjent utdanning er korrelert, som var basert på landsdelstallene, ser ikke ut til å stemme. Korrelasjonskoeffisienten mellom disse variablene er bare 0,11. Dette viser hvor skummelt det er å konkludere uten å se nærmere på de underliggende tallene.

Til sist kan det være interessant å se nærmere på om satsingen på barnehage gjennom barnehagereformen har gjort forskjellene i tilbudet fra kommune til kommune mindre. I tabell 7 presenterer jeg derfor standardavviket til barnehagevariablene for de fire årene oppgaven ser på. Det er her tydelig at forskjellene i dekningsgrad har blitt mindre gjennom perioden for begge aldersklassene. Barnehagesektoren har altså tatt steg i retning av et mer universelt tilbud for hele landet. Men, forskjellene var fremdeles store også i 2006. Tallene i denne kolonnen kan tolkes som at to tilfeldig valgte kommuner gjennomsnittlig vil

Tabell 7: Barnehagevariablenes standardavvik, utvikling gjennom perioden.

Variabel	2003	2004	2005	2006
Dekningsgrad 1-2 år	0.158	0.151	0.150	0.131
Dekningsgrad 3-5 år	0.095	0.082	0.076	0.063
Ansattes utdanning	0.071	0.073	0.083	0.080
Styrers utdanning	0.138	0.148	0.154	0.158
Barn per ansatt	1.18	1.10	1.15	0.97

Merk: Verdier er standardavvik mellom norske kommuner. *Kilde:* Statistisk sentralbyrå, egne beregninger.

ha en forskjell på 13,1 prosentpoeng i dekningsgraden for de yngste, og 6,3 prosentpoeng for de eldste. Når det gjelder de øvrige variablene er det kun andel styrere med godkjent utdanning som har en tydelig trend i variasjonsendringen, her har variasjonen mellom kommunene blitt noe høyere. Dette viser at det ikke har vært en jevn nedgang i andel med utdanning som har gitt reduksjonen vi så i tabell 4, men at det sannsynligvis er enkeltkommuner med høy utbygging som har ansatt ledere uten utdanning og følgelig trekker ned gjennomsnittet. For de to siste variablene kan det godt være at tabell 7 bare viser naturlig variasjon. Det er likevel verdt å merke seg at også variabelen for barn per ansatt har sitt laveste standardavvik i 2006, noe som kan tyde på barnehagetilbudet har blitt mer universelt også på dette punktet.

4.3 Kommunevariabler

For at regresjonen skal måle en reell kausaleffekt er det viktig å kontrollere for øvrige faktorer i kommunen som kan påvirke boligprisen. Jeg benytter tall for populasjon, arbeidsledighet, boligtetthet, sentralitet, landsdel og klima. Med unntak av sistnevnte er alle tallene hentet fra Statistisk sentralbyrå. Tallet for populasjon viser totalt antall innbyggere i kommunen 1. januar det aktuelle året. Arbeidsledigheten er målt som andelen registrerte arbeidsledige i aldersgruppen 15-74 år som prosent av arbeidsstyrken. Målet på boligtetthet er andel av populasjonen som bodde i spredtbebygde strøk i 1990. For sentralitet benytter jeg en oppdeling definert av Statistisk sentralbyrå i 1994, hvor kommunene blir inndelt i syv kategorier basert på reiseavstanden til det nærmeste regionale senteret. Som regionalt senter menes her et tettsted som oppfyller funksjoner av høy orden, for eksempel postkontor og bank. Data for landsdel innebærer at kommunene blir fordelt i kategoriene “Oslo og Akershus”, “Hedmark og Oppland”, “Østfold/Buskerud/Vestfold/Telemark”, “Agder og Rogaland”, “Vestlandet”, “Trøndelag” og “Nord-Norge”. Tallene for klima er gjennomsnittlig vintertemperatur i perioden 1971 til 2000. De er innhentet av Fredrik Carlsen ved NTNU og basert på data fra 128 forskjellige værstasjoner.

Datasettet for boligtransaksjoner inneholdt totalt 254 431 observasjoner. Etter å ha fjernet

observasjoner fra kommuner med manglende data for noen av kontrollvariablene, hus med byggeår før 1900, samt observasjoner tilknyttet utdaterte kommunenumre, sto jeg igjen med 202 322 observasjoner som kunne benyttes i regresjonen. Med andre ord har jeg hatt tilgang til en omfattende mengde data som dekker hele landet over fire år.

5 Metodologi

5.1 Empirisk spesifikasjon

I analysen estimerer jeg følgende modell ved hjelp av OLS:

$$\log P_{ikqt} = \beta_0 + \beta_1 D12_{kt} + \beta_2 D35_{kt} + \beta_3 U_{kt} + \beta_4 L_{kt} + \beta_5 B_{kt} + \beta_6 X_{ik} + \beta_7 Q_{kt} + \alpha_q + \eta_t + \epsilon_{ikqt} \quad (10)$$

$\log P_{ikqt}$ er her den naturlige logaritmen til prisen bolig i i kommune k ble solgt for i kvartal q , år t . En slik modell, hvor prisen på et gode implisitt defineres av en rekke andre egenskaper og forhold, kalles for en hedonisk prismodell (Rosen 1974). I tilfeller slik som her, hvor variabelen kun kan ha en positiv pengeverdi, så er det vanlig å benytte logaritmen til variabelen i regresjonen. En fordel med denne transformasjonen er at utfallsrommet til variabelen blir snevret inn, noe som gjør OLS-estimatene mindre sensitive for ekstreme verdier (Wooldridge 2013:185). En slik semilogaritmisk modell er den som oftest blir benyttet i litteraturen som omhandler skolekvalitet (Clapp, Nanda og Ross 2008). (10) blir estimert ved hjelp av “pooled” OLS, det vil si at jeg benytter tversnittsdata fra flere år i en samlet regresjon (Wooldridge 2013:9).

$D12_{kt}$ måler dekningsgrad for barn mellom 1 og 2 år, $D35_{kt}$ måler det samme for barn mellom 3 og 5 år, U_{kt} viser andel ansatte med godkjent barnehagelærerutdanning, L_{kt} viser andelen styrere og pedagogiske ledere med godkjent barnehagelærerutdanning og B_{kt} måler antall barn per ansatt i kommunale barnehager. Dekningsgraden varierer stort fra kommune til kommune, noe som gjør det nærliggende å anta at det er en viktig faktor i bosetningsbeslutningen til barnefamilier. Borge og Rattsø (2014) har allerede vist at den totale dekningsgraden blir kapitalisert inn i boligprisene, derfor er det ekstra interessant å se på om denne effekten er avhengig av aldersgruppe. De ansattes utdanning, styrerens utdanning og antall barn per ansatt er inkludert i modellen fordi de antas å påvirke barnehagekvaliteten. Alle regnes blant de viktigste kvalitetsfremmende innsatsfaktorene i barnehagen (Barne og- familiedepartementet 2005). Her er det viktig å påpeke at formålet med analysen ikke er å finne ut om disse variablene faktisk påvirker kvaliteten, kun om norske husholdninger mener de gjør det. Min hypotese er at alle disse variablene kapitaliseres inn i eiendomsverdiene, det vil si at jeg forventer å finne $\beta_1 > 0$, $\beta_2 > 0$, $\beta_3 > 0$, $\beta_4 > 0$ og $\beta_5 < 0$. Sistnevnte forventes negativ fordi en økning i antall barn per ansatt vil gi lavere kvalitet, og dermed en reduksjon i etterspørselen.

X_{ik} er en vektor som kontrollerer for egenskaper ved boligen som forventes å påvirke prisen. For å unngå at utelatte variable gir forventningsskjevne estimat er det vanlig å inkludere et høyt antall variable i denne vektoren (Atkinson og Crocker 1987). For eksempel benytter

Borge og Rattsø (2014) alder, areal, antall bad, antall toalett, hvorvidt boligen inkluderer en garasje, og type bolig, som forklaringsvariabler i sin analyse. Dessverre har jeg ikke hatt tilgang til et like rikt datasett for årene 2003 til 2006 som de hadde for 1997 til 1999. X_{ik} er derfor begrenset til å inneholde areal i kvadratmeter, husets alder, og dummyvariabler som viser om boligen er en enebolig, småhus eller blokkleilighet. Eventuelle konsekvenser dette har for analysen drøftes i neste delkapittel. Da det er nærliggende å anta at areal har en positiv, men avtakende, effekt på boligprisen, har jeg inkludert et kvadratledd av areal. Jeg har også inkludert et kvadratledd av alder, da den negative effekten her forventes å avta etterhvert.

Q_{kt} er en vektor som inkluderer forhold i kommunen som forventes å påvirke boligprisene i seg selv. Som påpekt av Nguyen-Hoang og Yinger (2011) er en av hovedutfordringene ved analyser av skolekvalitet problemet med å skille kausaleffekten av skole fra effekter som skyldes andre faktorer ved nabolaget. Tilsvarende er det viktig at man regresjon skiller effekter som skyldes barnehagetilbudet fra andre effekter som skyldes forholdene i kommunen. Et viktig steg i retning av å løse denne utfordringen er å inkludere mange relevante variable i Q_{kt} . For å kontrollere for forhold tilknyttet kommunens arbeidsmarked er arbeidsledigheten inkludert som forklaringsvariabel. Det forventes at denne variabelen påvirker boligprisene negativt. I og med at datasettet dekker et stort, heterogent område er det også viktig å kontrollere for grunnleggende forskjeller mellom ulike områder. Dette er hensyntatt ved å inkludere syv dummyvariabler for landsdel, og like mange dummyvariabler for sentralitet. I tillegg inneholder Q_{kt} variabler for populasjon og bosetningstetthet. Til sist kontrollerer jeg for effekten av klima ved å inkludere en variabel for vintertemperatur. For å kontrollere for felles tidsvarierende effekter, slik som nasjonale makroøkonomiske trender, inkluderer jeg også dummyvariabler for kvartal, α_q , og år, η_t . ϵ_{ikqt} er et stokastisk restledd som omfatter alle uobserverte effekter.

For at OLS-estimering av (10) skal gi forventningsrette og lineære estimatorer med lavest mulig varians må følgende forutsetninger holde:

- i) Modellen i populasjonen kan uttrykkes på lineær form, slik som i (10).
- ii) Utvalget av observasjoner som benyttes er tilfeldig valgt fra populasjonen.
- iii) Ingen av de uavhengige variablene er konstante, og ingen av dem er perfekt kolineære. $Var(D12_{kt}, D35_{kt}, U_{kt}, L_{kt}, B_{kt}, X_{ik}, Q_{kt}, \alpha_q, \eta_t) > 0$ og $Cov(D12_{kt}, D35_{kt}, U_{kt}, L_{kt}, B_{kt}, X_{ik}, Q_{kt}, \alpha_q, \eta_t) \neq 1$.
- iv) Restleddet ϵ_{ikqt} har forventningsverdi lik null uansett verdi på de uavhengige variablene. $E(\epsilon_{ikqt} | D12_{kt}, D35_{kt}, U_{kt}, L_{kt}, B_{kt}, X_{ik}, Q_{kt}, \alpha_q, \eta_t) = 0$.

v) Restleddet ϵ_{ikqt} er homoskedastisk, det vil si at variansen er den samme uansett verdi på forklaringsvariablene. $Var(\epsilon_{ikqt}|D12_{kt}, D35_{kt}, U_{kt}, L_{kt}, B_{kt}, X_{ik}, Q_{kt}, \alpha_q, \eta_t) = \sigma^2$.

Disse kalles Gauss-Markov-antakelsene, hvor (i)-(iv) er tilstrekkelig for å gi forventningsrette og konsistente estimatorene (Wooldridge 2013:98). I og med at datasettet består av alle boligtransaksjoner for de aktuelle årene er det nærliggende å tro at forutsetning (ii) holder. Et ankepunkt kan eventuelt være at boliger som omsettes kan være gjennomgående annerledes enn de som ikke omsettes, og at man dermed ikke har et tilfeldig utvalg. En løsning på dette kunne vært å inkludere anslått verdi for boliger som ikke er omsatt, men disse ville sannsynligvis vært mindre egnet til å plukke opp etterspørselsendringer. Siden jeg kontrollerer for både faktorer ved eiendommen, beliggenhet i landet og omsetningstidspunkt, så mener jeg at dette uansett ikke vil gi problemer i regresjonen. Det er heller ingen grunn til å tro at (iii) ikke skal holde, og hvis det ikke er tilfelle vil det raskt gå fram av regresjonen. Men, det er overveiende sannsynlig at (iv) og (v) ikke er oppfylt. Hvilke konsekvenser dette kan få, og hvordan disse utfordringene kan løses, vil bli drøftet i det påfølgende underkapittelet.

5.2 Metodologiske utfordringer

Jeg kontrollerer for færre egenskaper ved eiendommen enn det en del av de lignende forskningsprosjektene fra litteraturen over gjør. Sammenlignet med Borge og Rattsø (2014) kontrollerer jeg ikke for antall bad, antall WC og tilstedeværelse av garasje. Det er ingen tvil om at dette er variable som påvirker boligprisen, og at deres utelatelse dermed fører til at Gauss-Markov-forutsetning (iv) ikke holder. Dette er det velkjente “utelatt variabel”-problemet. Når man skal vurdere hvor store konsekvenser dette får for regresjonen er det en rekke aspekter som må hensyntas. Butler (1982) sammenligner to boligprisregresjoner, hvor den ene kontrollerer for en stor mengde faktorer ved boligen og den andre er svært restriktiv. Han finner at den praktiske konsekvensen av en restriktiv spesifisering er liten, i motsetning til hva som tidligere har vært antatt. Et viktig poeng fra hans analyse er at “utelatt variabel”-problemet aldri kan utelukkes fullstendig kun ved å inkludere en mengde boligvariabler. Dette skyldes både at mengden potensielt relevante variabler er uhåndterbart høy, og at mange av boligvariablene er svært korrelert med hverandre. Hvis man mot formodning kunne inkludere alle relevante variable ville det altså uansett ikke vært ønskelig, fordi det sannsynligvis hadde gitt problemer med multikollinearitet. Når man sammenligner ulike modellspesifikasjoner ser man med andre ord på relativ grad av forventningsskjevhet, ikke på om en skjevhet eksisterer eller ikke. Også Atkinson og Crocker (1987) sammenligner de forskjellige boligvariablene som inkluderes i litteraturen. De finner at mange forskere undervurderer multikollinearitetsproblemene som kan oppstå, og at man som en følge av dette i mange tilfeller kan håndplukke de forklaringsvariablene man vil for å oppnå ønskede fortegn

på estimatorene. De påpeker likevel at disse problemene er større for variable som beskriver faktorer ved nabolaget enn for de som beskriver den fysiske eiendommen. Som en følge av dette anbefaler de uansett å inkludere en stor mengde faktorer ved huset som forklaringsvariable. Problemet er selvsagt at dette ofte er lettere sagt enn gjort, siden tilgangen på relevant datamateriale kan være begrenset.

For å avdekke hvilken konsekvens ekskludering av de aktuelle boligvariablene får er det viktig å se på hvilke variable som blir forventningsskjevne som følge av utelatelsen. I (10) er jeg bare interessert i fortegn og størrelse på estimatene foran barnehagevariablene, alle de andre variablene er kun inkludert for å gjøre de førstnevnte mest mulig presise. Det har altså ingen direkte konsekvens for analysen hvorvidt estimatet foran for eksempel blokkleilighet-variabelen er riktig eller ikke. Så lenge barnehagevariablene hverken er korrelert med de inkluderte eller ekskluderte eiendomsvariablene direkte, så vil de utelatte variablene kun føre til forventningsskjevne estimatører for de inkluderte eiendomsvariablene, ikke barnehagevariablene (Wooldridge 2013:86). Det er likevel ingen grunn til å tro at barnehagevariablene og eiendomsvariablene skal være helt ukorrelerte. Jeg fant at utelatelse av samtlige boligvariable reduserte den estimerte effekten av dekningsgraden mellom 1 og 2 år med i underkant av en tredjedel. Men, alle variablene hadde fremdeles samme fortegn, og endringen var mindre for de andre barnehagevariablene. Selv om barnehage- og eiendomsvariablene ikke er ukorrelerte, så ser altså ikke korrelasjonen ut til å være så veldig sterk. Det er også grunn til å tro at “utelatt variabel”-problemet blir proposjonalt mindre for hver boligvariabel som inkluderes. Dette fordi sterk korrelasjon mellom de inkluderte og de ekskluderte boligvariablene vil bidra til at de ekskluderte variablene kontrolleres for gjennom en “proxy”-effekt (Butler 1982). For eksempel er det nærliggende å anta at brorparten av effekten antall bad og WC har på boligprisen allerede blir kontrollert for ved at areal er inkludert som forklaringsvariabel. Da har det ikke noe å si at koeffisienten foran areal blir overestimert, så lenge korrelasjonen med barnehagevariablene er svak. En sammenligning av mine og Borge og Rattsø (2014) sine estimerte effekter av boligkarakteristika gir støtte for denne hypotesen. Estimert effekt av areal er nesten dobbelt så stor i min regresjon, det samme er t-verdien. Dette tyder på at min regresjon indirekte kontrollerer for en del av effekten av de utelatte variablene gjennom areal. Selv om jeg ideelt sett skulle hatt tilgang på et grundigere datasett for bolig, så ser det altså ut som at min restriktive formulering har forholdsvis liten konsekvens for regresjonens resultater.

Ved oppstart av arbeidet tenkte jeg at utdanningsnivå og inntekt også var variable som burde inkluderes i Q_{kt} . Dette er i tråd med Bayer, Ferreira og McMillan (2007), som inkluderer disse variablene for å kontrollere for uobserverte effekter ved nabolaget. De finner at dette reduserer den estimerte kapitaliseringseffekten av skolekvalitet drastisk,

og konkluderer følgelig med at studier som ikke kontrollerer for et like rikt sett med variable innehar en positiv skjevhet. På samme måte opplevde jeg at inklusjon av disse variablene ga store endringer i den estimerte effekten av barnehagevariablene. Ved nærmere ettertanke var det likevel ikke like tydelig at utdanningsnivå og inntekt var variabler som burde være med i modellen. Dette fordi disse variablene i seg selv påvirker både etterspørselen etter boligkvalitet og etterspørselen etter barnehagetjenester. Selv om variabelen i utgangspunktet er inkludert for å kontrollere for kvaliteter ved kommunen, vil det som påpekt av Butler (1982) være umulig å skille mellom effekter som skyldes dette, og de som skyldes variabelens rolle som etterspørselsfaktor. At utdanningsnivå er en viktig faktor i etterspørselen etter barnehagetjenester i Norge går for eksempel fram av Borge, Borge Johannesen og Tovmo (2010). I tillegg vil inntektsnivå som forklaringsvariabel føre til et simultanitetsproblem, siden inntekt både påvirker boligprisene direkte ved at økt inntekt gir økte priser, samtidig som prisene i seg selv bidrar til å allokere husholdninger med ulik inntekt til ulike boligområder (Straszheim 1973). På bakgrunn av dette har jeg valgt å ikke inkludere inntekt- og utdanningsnivå som forklaringsvariable i OLS-regresjonen.

En svakhet med min empiriske spesifikkasjon er at skattesatsen på eiendom ikke inngår noen plass i (10). Hvis kommuner med eiendomsskatt jevnt over har høyere dekning og kvalitet på barnehagetjenestene, så vil dette føre til at kapitaliseringseffekten jeg finner underestimeres. Dette fordi jeg ved å utelukke eiendomsskatt i regresjonen ikke justerer for en faktor som trekker boligprisene nedover i disse kommunene. Grunnen til at eiendomsskatt ikke inngår i regresjonen er at det dessverre ikke finnes tilfredsstillende tall for perioden jeg ser på. Unntaket er året 2005, hvor jeg har tilstrekkelig med data. Jeg vil derfor foreta en separat regresjon kun for året 2005, med eiendomsskatt som en ekstra forklaringsvariabel, og drøfte i hvilken grad dette endrer resultatene.

Siden jeg benytter data på kommunenivå, så er det naturlig å tro at restleddet i (10), ε_{ikqt} , inneholder en felles gruppe-effekt. Det vil si at boliger innad i en kommune har noen felles uobserverte aspekter ved seg som påvirker boligprisen, og som ikke er kontrollert for:

$$\varepsilon_{ikqt} = \gamma_k + \theta_{ikqt} \quad (11)$$

hvor γ_k er kommune-effekten og θ_{ikqt} er det øvrige restleddet. Som vist i Wooldridge (2003) vil et restledd med form som i (11) kunne gi kraftig korrelasjon innad i gruppen, noe som gjør at standardavvikene til OLS-estimatorene blir kunstig små. Dette fører igjen til at t-verdiene blir høyere, slik at estimatene framstår mer signifikante enn de burde vært. Heldigvis er det enkelt å justere for dette i STATA ved å benytte standardavvik som er “cluster robust”, noe jeg har gjort i min analyse. Jeg har da benyttet kommune som

“cluster”-variabel. Dette gjør også standardavvikene robuste for heteroskedastisitet, slik at det ikke har noe å si for analysen om forutsetning (v) holder.

Selv om jeg kontrollerer for en stor mengde kommunespesifikke variable gjennom Q_{kt} , så er det en viss fare for det fremdeles er uobserverte effekter ved kommunen som ikke er kontrollert for, og som fører til en “utelatt variabel”-skjevhet i min regresjon. Jeg har derfor forsøkt å omgå dette problemer ved hjelp av “faste effekter”-metoden. Dette er en metode som kun kan brukes når man har paneldata, det vil si data for de samme enhetene over tid. Ved at man kun utnytter variasjonen innad i hver enhet vil man da kunne filtrere ut all variasjon som er tidskonstant. I og med at jeg hadde data for de samme kommunene over flere år var planen min derfor å benytte faste effekter til å isolere kapitaliseringseffekten og filtrere bort uobserverte kommuneffekter. En forutsetning for at dette skal fungere er likevel at variablene av interesse innehar tilstrekkelig variasjon over tid (Wooldridge 2013:467). Vanligvis vil man tro at fire år er for kort tid til at variabler som måler offentlig tjenestetilbud skal inneha nok variasjon, men mitt håp var at de store endringene som fulgte av barnehagereformen skulle rette opp for dette. Dessverre viste det seg å ikke være tilfelle, og mine “faste effekter”-estimat ga ingen signifikante effekter. Utvidelse av perioden ved å inkludere årene 2001 og 2002, og dermed fjerne husets alder som forklaringsvariabel, endret ikke på dette resultatet. Estimering av “tilfeldige effekter” i stedet ga heller ingen effekt. Jeg har derfor ikke valgt å inkludere disse resultatene under. Her kan man selvsagt påstå at mangelen på estimert effekt skyldes at kapitaliseringseffekten ikke eksisterer i hele tatt. I og med at alle barnehagevariablene også fikk motsatt fortegn av hva økonomisk resonnement skulle tilsi vil jeg likevel påstå at manglende tidsvariasjon er den klart mest sannsynlige årsaken.

Kanskje den største utfordringen med å analysere kapitaliseringseffekter ved hjelp av vanlig OLS er at man sannsynligvis står overfor et simultanitetsproblem. Dette ble tidlig påpekt av Oates (1969). Siden forklaringsvariablene av interesse kan påvirkes fra politisk hold, så er det vanskelig å utelukke en effekt hvor endringer i den avhengige variabelen gir endringer i de uavhengige. For eksempel kan det være at en kommune som opplever stor fraflytning bestemmer seg for å øke satsingen på barnehage for å motvirke dette. I et slikt tilfelle vil boligprisen og dekningsgraden være simultant bestemt, noe som vil gi forventningsskjevne estimatører (Wooldridge 2013:535). Instrumentvariabelmetoden kan benyttes for å omgå dette problemet. I så fall må man finne en variabel som er ukorrelert med restleddet i den opprinnelige modellen, samtidig som den er korrelert med forklaringsvariabelen som er simultant bestemt. Sistnevnte betingelse kan testes ved hjelp av vanlig OLS, mens den førstnevnte som regel kun kan begrunnes ved hjelp av økonomisk intuisjon (Wooldridge 2013:492). Jeg vil i det følgende forsøke å rette opp for potensiell simultanitet ved hjelp av instrumentvariabelmetoden. Grunnsteinen i min analyse er

likevel vanlig OLS, og jeg starter derfor resultatdelen av oppgaven med regresjoner som er utført med denne metoden.

6 Resultater

6.1 Minste Kvadraters Metode

I tabell 8 presenteres resultater hvor ulike spesifikasjoner av (10) er estimert ved hjelp av “pooled” OLS. Modell B er her den fulle spesifikasjonen fra (10), i A kontrollerer jeg kun for boligvariable og tidsdummyer. Viktigheten av å kontrollere for kommunespesifikke variable illustreres tydelig ved å se på endringene i estimert effekt av barnehagevariablene fra modell A til B. Ser at man overestimerer kapitaliseringseffekten av barnehage svært mye hvis man ikke kontrollerer for andre aspekter ved kommunen. For dekningsgradene er estimatene i modell B litt under halvparten så store som de i A, og for de øvrige barnehagevariablene er reduksjonen enda kraftigere. Alle variablene i modell B har fortegn som forventet, med unntak av andel styrere med godkjent barnehagelærerutdanning hvor den estimerte effekten er negativ. Dette estimatet er uansett ikke statistisk signifikant på 5 % signifikansnivå, noe som gjør at jeg tillegger det liten vekt. Det er likevel interessant å se på hva som skjer i overgangen fra spesifikasjon B til C, hvor den eneste endringen jeg gjør er å droppe variabelen for styrers utdanning fra modellen. Her går andel ansatte med godkjent barnehagelærerutdanning fra å ha en estimert positiv effekt som nesten er signifikant på 5 % signifikansnivå, til å ha en mye mindre effekt som er fullstendig insignifikant. Hva dette skyldes er vanskelig å si, men det viser hvertfall at jeg ikke kan benytte resultatene fra modell B til å si noe fornuftig om kapitaliseringseffekten av personalets utdanning. Vet også fra tabell 6 i datakapittelet at de to utdanningsvariablene er ganske sterkt korrelert, noe som gjør at man får et “utelatt variabel”-problem hvis bare en av dem inkluderes i modellen. Følgelig blir disse to variablene droppet i analysen videre. Estimert effekt av de gjenværende variablene vises i modell D, og disse endrer seg ikke nevneverdig mye i forhold til estimatene i B. Dette tyder på at utdanningsvariablenes utelatelse i liten grad påvirker estimatorenes forventningsrettethet. Det er også i tråd med resultatene fra tabell 6, hvor det går fram at utdanningsvariablene er temmelig ukorrelerte med de øvrige barnehagevariablene. At R^2 holder seg uendret i den nye spesifikasjonen viser også at utdanningsvariablene forklarer lite av variasjonen i boligpriser. Jeg tar derfor utgangspunkt i modell D videre i analysen.

Resultatene fra modell D tyder på at dekningsgraden og antall barn per ansatt i de kommunale barnehagene er faktorer som påvirker eiendomsetterspørselen i kommunen. Estimaten for dekningsgraden mellom 1 til 2 år og barn per ansatt er begge statistisk svært signifikante. Størrelsen på estimatene viser også at effekten på prisen er langt fra neglisjerbar. For eksempel vil en økning i dekningsgraden for de minste barna på 10 prosentpoeng her gi en estimert prisøkning på tilnærmet 4,1 %. Tilsvarende vil ett barn færre per ansatt gi en estimert prisøkning på omtrent 4,6 %. For de eldste barna estimeres

Tabell 8: Resultater fra “pooled” OLS, 2003-2006.

Variabel	A	B	C	D
Dekningsgrad 1-2 år	0.901 (4.61)	0.396 (3.79)	0.411 (3.97)	0.412 (3.98)
Dekningsgrad 3-5 år	-0.730 (-1.63)	0.347 (1.77)	0.341 (1.72)	0.338 (1.71)
Ansatte med utdanning	2.10 (4.35)	0.326 (1.94)	0.097 (0.72)	
Styrer med utdanning	-1.81 (-5.20)	-0.204 (-1.80)		
Barn per ansatt	-0.22 (-7.90)	-0.049 (-4.16)	-0.048 (-4.05)	-0.046 (-4.09)
Innbyggertall (log)		0.131 (15.44)	0.131 (15.44)	0.131 (15.53)
Andel som bor spredtbebygde		-0.301 (-5.12)	-0.301 (-5.10)	-0.310 (-5.50)
Arbeidsledighetsrate		-9.15 (-9.49)	-9.02 (-9.61)	-8.96 (-9.57)
Vintertemperatur		0.011 (2.03)	0.011 (2.08)	0.011 (2.10)
Boligvariabler	Ja	Ja	Ja	Ja
Sentralitet	Nei	Ja	Ja	Ja
Landsdel	Nei	Ja	Ja	Ja
Tidsdummyer	Ja	Ja	Ja	Ja
Antall kommuner	415	415	415	415
Antall observasjoner	202 322	202 322	202 330	202 337
R^2	0.552	0.708	0.708	0.708

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på “cluster”-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

det at en økning i dekningsgraden på 10 prosentpoeng gir rundt 3,4 % høyere boligverdier, men dette estimatet er kun signifikant med et signifikansnivå på 10 %. Resultatene for dekningsgrad er ikke langt unntak med hva Borge og Rattsø (2014) finner. De estimerer i sin “pooled” OLS-regresjon at en like stor økning i den samlede dekningsgraden mellom 1 og 5 år gir en økning i boligprisene på 3,3 %.

De øvrige variablene i regresjonen har også estimatorer med forventet fortegn. Boligprisene øker med innbyggertall og vintertemperatur, og de faller med arbeidsledigheten og andelen som bor i spredtbebygde strøk. Alle disse er signifikante på 1 % signifikansnivå, med unntak av vintertemperatur som er signifikant på 5 % nivå. De estimerte effektene av boligvariablene går fram av Tabell 9. Dette er resultater fra regresjon D i tabell 8. Som forventet har areal en estimert positiv effekt på prisen, og husets alder har en negativ effekt. Fortegnene foran kvadratleddene viser at begge disse effektene beregnes å være avtakende. Ser også at type hus i seg selv ikke estimeres å ha noen effekt på prisen når andre faktorer blir kontrollert for, da dummyvariablene som viser boligtype er insignifikante. Effekt av de øvrige dummyvariablene som inngår i modell D presenteres ikke her, men de fleste av disse har en signifikant effekt. Alle dummyene for år og kvartal er signifikante på 1 % signifikansnivå. Det samme gjelder variablene for landsdel, med unntak av Hedmark/Oppland og Nord-Norge. Disse er kun signifikante på 20 % nivå, noe som kanskje skyldes at dette er de to landsdelene med lavest eiendomsomsetning. Alle dummyvariablene for sentralitet med unntak av én er signifikante på 5 % nivå. R^2 for modellen er på 0.708, noe som betyr at de inkluderte variablene forklarer 70,8 % av den

totale variasjonen i boligprisene.

Tabell 9: Effekt av boligvariabler.

Variabel	Koeffisient
Areal	0.012 (13.21)
Areal kvadrert	-0.00002 (-10.64)
Alder	-0.012 (-32.05)
Alder kvadrert	0.0001 (11.41)
Småhus	-0.011 (-0.57)
Blokkleilighet	0.032 (1.26)

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på “cluster”-robuste standardavvik i parentes. Øvrige kontrollvariable er de samme som i modell D, tabell 8. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Resultatene fra OLS-regresjonen tyder altså på at dekningsgraden blir kapitalisert inn i eiendomsverdiene, men det er kun dekningsgraden for de yngste barna som slår ut signifikant. Det kan være flere grunner til dette. En mulig årsak kan være at foreldre til yngre barn i større grad må lete aktivt etter kommuner med barnehageplass, siden dekningsgraden for de yngste barna i utgangspunktet er mye lavere enn for de eldste. En annen kan være at foreldre til yngre barn oftere kjøper ny bolig, siden de er i en etableringsfase og barna ennå ikke har fått noen tilknytning til nåværende boligområde. Dette argumentet vil kanskje være noe begrenset av at det har mest forklaringskraft for førstegangsforeldre. Man skal uansett være forsiktig med å trekke for mye ut av resultatene i tabell 8 alene. I et forsøk på å oppnå enda mer robuste resultater vil jeg derfor estimere flere regresjoner basert på ulike utvalg. Alle vil ta utgangspunkt i modell D fra tabell 8.

6.2 Robusthetsregresjoner

For å avdekke om det er markante forskjeller på estimatene mellom år estimerer jeg separate OLS-regresjoner for hvert av årene i perioden. Forklaringsvariablene er de samme som i modell D, tabell 8, med unntak av at jeg selvsagt ikke kunne inkludere dummyvariable for år. Som følge av dette er det noe risiko for at eventuelle forskjeller mellom årene kan skyldes tidseffekter, slik som for eksempel makroøkonomiske trender. Disse vil likevel bli kontrollert for til en viss grad gjennom at kvartalsdummyer fremdeles er inkludert. For 2003 måtte to av disse ekskluderes som følge av multikollinearitet, noe som nok skyldes at antallet observasjoner er mye lavere for det året. For de øvrige årene er alle kvartalsdummyene svært statistisk signifikante. At 2003 har færre observasjoner skyldes ikke at antallet omsatte eiendommer var så mye lavere, men at en større andel av observasjonene fra dette året mangler informasjon om husets byggeår. Hvis det ikke er tilfeldig hvilke observasjoner som mangler denne informasjonen, så vil det være et brudd på forutsetning (ii) om tilfeldig utvalg, noe som igjen vil gi forventningsskjevne estimater

for det året. En sammenligning av estimatene for de øvrige forklaringsvariablene ga likevel ingen tydelige forskjeller mellom 2003 og de andre årene, og det er derfor ingen grunn til å tro at de utelatte observasjonene følger noe spesielt mønster.

Tabell 10: År-for-år-regresjoner.

Variabel	2003	2004	2005	2006
Dekningsgrad 1-2 år	0.162 (0.99)	0.297 (2.20)	0.633 (4.56)	0.475 (3.56)
Dekningsgrad 3-5 år	0.523 (2.53)	0.497 (2.43)	0.161 (0.58)	0.446 (1.34)
Barn per ansatt	-0.052 (-3.41)	-0.058 (-4.87)	-0.045 (-3.78)	-0.055 (-3.54)
Boligvariabler	Ja	Ja	Ja	Ja
Kommunevariabler	Ja	Ja	Ja	Ja
Kvartalsdummyer	Ja	Ja	Ja	Ja
Antall kommuner	335	382	381	389
Antall observasjoner	22 692	54 198	60 362	65 085
R^2	0.701	0.693	0.702	0.708

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på "cluster"-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Resultatene av år-for-år-regresjonene presenteres i tabell 10. Ser her at det er svært interessante forskjeller mellom dekningsgradene for de ulike årene. Den estimerte effekten av dekningsgraden for de eldste går fra å være stor og signifikant de to første årene, til å bli insignifikant de to siste. For de yngstes dekningsgrad er utviklingen omvendt, her har de to siste årene de største og mest signifikante effektene. Det er verdt å merke seg at det eneste året hvor begge dekningsgradene er signifikante er 2004, samme år som barnehagereformen trådte i kraft for fullt. Disse resultatene er i tråd med hypotesen over om at effekten av dekningsgraden for de yngste slår klarere ut fordi antallet barn som ikke har plass er høyere. Som vist i tabell 4 øker veksten i dekningsgraden mellom 1 og 2 år gjennom perioden, mens veksten for de mellom 3 og 5 år er avtakende. I 2005 har gjennomsnittlig dekningsgrad for de eldste blitt over 92 %, det vil si at de aller fleste som ønsker plass får det. Det er da logisk at dekningsgraden for denne aldersgruppa ikke lenger gir like tydelig utslag i boligprisene, siden nesten alle nå får plass uansett hvor de bor. Tilsvarende øker effekten av dekningsgraden for de yngste, siden dette er en gruppe hvor mange ennå ikke har fått plass, og hvor det er stor vekst i antall plasser. Når det gjelder variabelen for barn per ansatt er den estimerte effekten svært signifikant alle årene, noe som er betryggende.

En svakhet ved min hypotese for dekningsgradene er at den ikke fanger opp hvorfor dekningsgraden for de yngste ikke slår signifikant ut i 2003, til tross for at den da er på sitt laveste. En mulig forklaring på dette kan være at det først var gjennom den kraftige utbyggingen som fulgte av barnehagereformen at etterspørselen etter barnehageplass for

Tabell 11: “Pooled” OLS-resultater for ulike innbyggertall.

Variabel	Over 5000	Over 10 000	Over 20 000	Over 30 000
Dekningsgrad 1-2 år	0.448 (3.73)	0.455 (3.02)	0.389 (2.06)	-0.013 (-0.06)
Dekningsgrad 3-5 år	0.301 (1.33)	0.370 (1.41)	0.602 (1.69)	0.893 (2.03)
Barn per ansatt	-0.046 (-3.68)	-0.055 (-4.03)	-0.052 (-3.95)	-0.044 (-2.13)
Boligvariabler	Ja	Ja	Ja	Ja
Kommunevariabler	Ja	Ja	Ja	Ja
Tidsdummyer	Ja	Ja	Ja	Ja
Antall kommuner	193	103	46	25
Antall observasjoner	196 660	184 655	158 982	136 905
R^2	0.711	0.721	0.732	0.737

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på “cluster”-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

de minste skjøt fart. Foreldrene til de yngste barna hadde tidligere hovedsaklig vært mottakere av kontantstøtte, noe som kan ha ført til at barnhageplass ikke ble vektlagt like mye ved valg av boligområde (Statistisk sentralbyrå 2012). Som alltid skal man være forsiktig med å konkludere for bastant ut i fra en enkelt OLS-regresjon, men hypotesen om at disse tre barnehagevariablene blir kapitalisert inn i boligprisen blir hvertfall ikke svekket av resultatene i tabell 10.

For å undersøke om effekten er forskjellig i store og små kommuner har jeg også estimert separate regresjoner for ulike innbyggertall. Her har jeg estimert regresjoner for innbyggertall over 5000, 10 000, 20 000 og 30 000, med samme kontrollvariabler som i modell D. Det eneste avviket fra kontrollvariablene i denne modellen er at flere av sentralitetsdummyene naturlig nok blir irrelevante etterhvert som innbyggertallet økes. Resultatene presenteres i tabell 11, og igjen er det svært interessant variasjon i den estimerte effekten av dekningsgrad-variablene. Den estimerte effekten av dekningsgraden mellom 3 til 5 år øker både i størrelse og signifikans etterhvert som små kommuner utelukkes, og den er kun signifikant på 5 % signifikansnivå når innbyggertallet er over 30 000. For dette utvalget, som utgjør de 25 største kommunene i landet, så estimeres det at en økning i de eldstes dekningsgrad på 10 prosentpoeng gir nesten 9 % høyere boligpriser, alt annet likt. Borge og Rattsø (2014), som benytter den samlede dekningsgraden i sin analyse, finner også at effekten er høyest i de største kommunene. De foreslår at dette kan skyldes at urbane områder både har høyere lønninger og høyere arbeidsdeltakelse blant kvinner, noe som kan øke husholdningenes verdsettelse av barnehager. Denne hypotesen passer bra med mine resultater for de eldstes barnehagedekning, men den forklarer ikke hvorfor dekningsgraden for de yngste ikke gir effekt i de største kommunene. Når det gjelder effekten av barn per ansatt er den igjen signifikant over hele linja, uten å vise noe tydelig mønster for de ulike innbyggertallene.

Tabell 12: “Pooled” OLS-resultater for ulike hustyper.

Variabel	Enebolig	Småhus	Blokkleilighet
Dekningsgrad 1-2 år	0.428 (4.18)	0.332 (2.73)	0.393 (2.61)
Dekningsgrad 3-5 år	0.442 (2.34)	0.490 (2.05)	0.197 (0.69)
Barn per ansatt	-0.056 (-3.96)	-0.044 (-2.68)	-0.030 (-2.26)
Boligvariabler	Ja	Ja	Ja
Kommunevariabler	Ja	Ja	Ja
Tidsdummyer	Ja	Ja	Ja
Antall kommuner	414	274	285
Antall observasjoner	63 396	26 793	112 148
R^2	0.732	0.734	0.686

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på “cluster”-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Jeg undersøker også om effekten av barnehagevariablene er forskjellig for ulike hustyper. Dette er en interessant spesifisering, fordi det er nærliggende å tro at husholdninger med barn, i allefall hvis de har flere en ett, i større grad vil bo i eneboliger og småhus enn i blokkleiligheter. Resultatene fra disse regresjonene presenteres i tabell 12. Som tidligere benyttes samme kontrollvariable som i modell D. I tråd med forventningene er den estimerte effekten av dekningsgraden for de eldste barna ikke signifikant når utvalget begrenses til kun blokkleiligheter. For småhus og eneboliger estimeres det derimot en effekt som er signifikant på 5 % signifikansnivå. Dette er betryggende, og det styrker hypotesen om at dekningsgraden blir kapitalisert inn i boligprisene. Siden godt over halvparten av observasjonene i utvalget mitt er blokkleiligheter, kan det også bidra til å forklare hvorfor dekningsgraden for barna mellom 3 og 5 år ikke er signifikant i regresjonen i tabell 8. For de minste barna er det ikke en like tydelig forskjell mellom de forskjellige hustypene, men dette gir mening siden behovet for større bolig vil være større jo eldre husholdningens barn er. Det er også mer sannsynlig at foreldre med barn mellom 3 og 5 år har flere barn. Også hypotesen om at antall barn per ansatt påvirker boligprisen kommer styrket ut av denne spesifiseringen. Den estimerte effekten er størst og mest signifikant for eneboliger, og lavest og minst signifikant for blokkleiligheter. Alle de estimerte effektene for denne variabelen er likvel signifikante på 5 % signifikansnivå.

For å undersøke om resultatene påvirkes mye av at regresjonene ikke inneholder eiendomsskatt som forklaringsvariabel, har jeg estimert en separat regresjon for 2005 hvor eiendomsskatt er inkludert. Dette er det eneste året i min periode hvor jeg har tilfredstillende data for eiendomsskatt i norske kommuner. Jeg benytter her et datamateriale som jeg har fått oppgitt av Lars-Erik Borge ved NTNU. Tallene viser hvor mye en huseier med et gjennomsnittlig verdsatt hus betaler årlig i eiendomsskatt. Av de 68

742 husene som er omsatt i 2005 ligger 17 366 av dem i en kommune med eiendomsskatt. Årlige utgifter til eiendomsskatt for et hus med gjennomsnittlig verdi i disse kommunene varierer fra 1050 kr til 5600 kr. For å begrenset antallet nuller i estimatoren har jeg i regresjonen benyttet en variabel hvor den årlige utgiften er dividert på 1000. Det vil si at verdiene i kommuner med eiendomsskatt varierer mellom 1,05 og 5,6.

Regresjonen med eiendomsskatt presenteres i tabell 13, under overskriften 2005*. For sammenligning har jeg også inkludert resultatene fra regresjonen for 2005 fra tabell 10. Med unntak av eiendomsskatt kontrollerer de to regresjonene for nøyaktig de samme variablene. Ser at eiendomsskatten som forventet estimeres å ha en negativ effekt på boligprisene, men at denne effekten kun er signifikant på 10 % signifikansnivå. Den estimerte effekten er også forholdsvis liten, en økning i gjennomsnittlige utgifter til eiendomsskatt på 1000 kr estimeres å gi en reduksjon i eiendomsverdiene på 0,008 %. Den estimerte effekten av barnehagevariablene endrer seg lite når eiendomsskatt blir introdusert som forklaringsvariabel. Dette er noe overraskende, og tyder på at kommuner med eiendomsskatt ikke bruker denne inntekten systematisk til å bedre barnehage tilbudet. Man skal likevel være forsiktig med å legge for mye i dette, siden det like gjerne kan skyldes at estimatet for eiendomsskatt ikke er signifikant. Uansett tyder resultatene i tabell 13 på at utelatelse av eiendomsskatt som forklaringsvariabel i de øvrige regresjonene ikke øker estimatorenes forventningsskjevheter i nevneverdig grad.

Tabell 13: Regresjon med eiendomsskatt.

Variabel	2005	2005*
Dekningsgrad 1-2 år	0.633 (4.56)	0.643 (4.66)
Dekningsgrad 3-5 år	0.161 (0.58)	0.125 (0.45)
Barn per ansatt	-0.045 (-3.78)	-0.043 (-3.54)
Eiendomsskatt		-0.008 (-1.71)
Boligvariabler	Ja	Ja
Kommunevariabler	Ja	Ja
Kvartalsdummyer	Ja	Ja
Antall kommuner	381	381
Antall observasjoner	60 362	60 362
R^2	0.702	0.703

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på "cluster"-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

6.3 Instrumentvariabelmetoden

For å unngå den potensielle simultaniteten mellom den avhengige og de uavhengige variablene, kan det være lurt på benytte 2SLS i stedet for vanlig OLS. Dette innebærer at man må finne én eller flere egnede instrumentvariable for de uavhengige variablene

som er simultant bestemt. I min regresjon fra modell D er det grunn til å tro at alle de tre barnehagevariablene kan bli endret fra politisk hold for å påvirke tilflytningen til kommunen, det vil si at alle tre potensielt kan være endogene variable. Dette skaper en stor utfordring, siden instrumentering av flere variable ofte er mer krevende enn å instrumentere én. Dette skyldes både at man nå trenger flere instrumenter som alle må være korrelert med de endogene variablene, og at 2SLS i seg selv gir estimatorer med høyere varians enn vanlig OLS (Wooldridge 2013:508-509). Jo flere variable som instrumenteres, jo større er sjansen for at den økte variansen gir estimatorer som er for upresise til at de kan gi et presist estimat.

Som følge av de potensielle utfordringene knyttet til instrumentering av flere variable har jeg valgt å bare instrumentere én variabel. Dette gir igjen en utfordring i form av at de to barnehagevariablene som ikke instrumenteres nå må utelates fra modellen. For å undersøke hvor mye de tre variablene påvirkes av at de andre to droppes fra modellen har jeg i tabell 14 estimert tre ulike spesifikasjoner hvor hver av dem inngår alene. Resultatene fra modell D i tabell 8 er inkludert for sammenligning, og alle spesifikasjonene kontrollerer for de samme forklaringsvariablene som denne modellen. Ser at estimert effekt av de to dekningsgradvariablene øker markant både i størrelse og signifikans når de andre variablene droppes fra modellen. Dette samsvarer med resultatene fra tabell 6 i datakapittelet, hvor jeg fant at disse to variablene var korrelert med hverandre. Estimert foran barn per ansatt holder seg derimot fullstendig uendret, noe som også gir mening, siden denne variabelen er tilnærmet ukorrelert med variablene for dekningsgrad. Dette betyr igjen at jeg kan utelukke de øvrige barnehagevariablene, og kun instrumentere barn per ansatt alene, uten at det vil gi store konsekvenser for analysen. Jeg vil også påstå at det er ekstra interessant å øke robustheten til resultatene for denne variabelen, siden den ikke inngår i Borge og Rattsø (2014) sin analyse.

Tabell 14: Regresjoner med færre barnehagevariabler.

Variabel	D	E	F	G
Dekningsgrad 1-2 år	0.412 (3.98)	0.544 (7.27)		
Dekningsgrad 3-5 år	0.338 (1.71)		0.772 (5.35)	
Barn per ansatt	-0.046 (-4.09)			-0.046 (-3.66)
Boligvariabler	Ja	Ja	Ja	Ja
Kommunevariabler	Ja	Ja	Ja	Ja
Tidsdummyer	Ja	Ja	Ja	Ja
Antall kommuner	415	416	416	415
Antall observasjoner	202 337	202 426	202 426	202 337
R^2	0.708	0.706	0.704	0.701

Merk: Avhengig variabel er logaritmen til boligprisen. T-verdier basert på “cluster”-robuste standardavvik i parentes. Øvrige kontrollvariable er de samme som i modell D, tabell 8. *Kilde:* Statistisk sentralbyrå, egne beregninger.

Spesifikasjonen i modell G er som følger:

$$\log P_{ikqt} = \delta_0 + \delta_1 B_{kt} + \delta_2 X_{ik} + \delta_3 Q_{kt} + \alpha_q + \eta_t + \omega_{ikqt} \quad (12)$$

Det vil si den samme som (10) bortsett fra at alle barnehagevariablene med unntak av barn per ansatt, B_{kt} , er ekskludert. Hvis vi mistenker at $\log P_{ikqt}$ og B_{kt} er simultant bestemt, for eksempel fordi norske kommuner øker bemanningen i barnehager hvis de opplever fraflytning, så vil variabelen for barn per ansatt være korrelert med restleddet, det vil si $Cov(B_{kt}|\omega_{ikqt}) \neq 0$. I et slikt tilfelle vil estimatoren av interesse, δ_1 , være forventningsskjev (Wooldridge 2013:535-536). For å omgå dette ved hjelp av instrumentvariabelmetoden må jeg finne minst én variabel, V_k , som ikke inngår i (12), og som både er korrelert med barn per ansatt, og ukorrelert med restleddet (Wooldridge 2013:492). Variabelen må altså oppfylle både $Cov(V_k|B_{kt}) \neq 0$ og $Cov(V_k|\omega_{ikqt}) = 0$. Sistnevnte innebærer at den aktuelle variabelen ikke kan ha noen selvstendig påvirkning på boligprisen i kommunen.

Borge og Rattsø (2014) benytter i sin analyse instrumentvariabelmetoden til å instrumentere både eiendomsskatt og dekningsgrad mellom 1 og 5 år. De tre variablene som er signifikant korrelert med dekningsgraden i deres førstestegsregresjon er andel kvinner i kommunestyret, partifragmentering i kommunestyret, og andel representanter i kommunestyret fra sosialistiske parti. Siden det er temmelig vanskelig å finne andre instrumenter som ikke har en selvstendig påvirkning på boligprisen, så har jeg valgt å benytte de samme variablene i min analyse. Intuisjonen bak at andel kvinner i kommunestyret skal være korrelert med dekningsgraden er basert på Svaleryd (2009). Hun finner at en høyere kvinneandel i svenske kommunestyre gir økte utgifter til barnehage og utdanning. Siden en kommune som bruker mer penger på barnehage sannsynligvis også vil ha flere ansatte per barn er dette argumentet også relevant når barn per ansatt er variabelen som skal instrumenteres. Når det gjelder partifragmentering og andel representanter fra sosialistiske parti var de i Borge og Rattsø (2014) hovedsaklig antatt å være korrelert med eiendomsskatt. At disse variablene er inkludert som instrument har altså ikke en like tydelig teoretisk begrunnelse. Siden de likevel viste seg å være korrelert med dekningsgraden mellom 1 og 5 år i deres førstestegsregresjon gir det likevel mening å inkludere dem som instrumentvariabler også i min analyse.

Som vist over er det ikke nok at instrumentene jeg benytter er korrelert med variabelen for barn per ansatt, de kan heller ikke ha en selvstendig påvirkning på boligprisen. Jeg vil i likhet med Borge og Rattsø (2014) argumentere for at det er usannsynlig at kommunestyrets sammensetning vil ha en selvstendig påvirkning på boligprisen i

kommunen. Allikevel kan det tenkes at sammensetningen har en effekt som følge av at den er korrelert med andre variable som er utelatt fra modellen. Sannsynligheten for dette minsker likevel jo flere kommunespesifikke variabler som er inkludert i modellen. Jeg kontrollerer for en rekke slike variabler, noe som gjør sannsynligheten for at styresammensetningen er korrelert med relevante utelatte variable relativt liten. Men, man kan aldri være helt sikker på at et instrument er valid, og jeg må følgelig også ta hensyn til dette når resultatene tolkes.

Data for norske kommunestyres sammensetning er hentet fra Statistisk sentralbyrå, hvor jeg har benyttet tall som er innhentet etter kommunevalget i 2003. Kommunestyrene har altså hatt lik sammensetning gjennom hele perioden jeg ser på. For sosialistandelen benyttes en vid definisjon av begrepet “sosialist”, her inngår alle representantene fra partiene Arbeiderpartiet, Sosialistisk Venstreparti, Rød Valgallianse og Norges Kommunistiske Parti. For å måle partifragmentering har jeg konstruert en Herfindahl-indeks ut i fra følgende formel fra Borge (2005):

$$HERF = \sum_{p=1}^P SH_p^2 \quad (13)$$

hvor SH_p er andelen representanter fra parti p , og P er det totale antall partier som er representert i kommunestyret. Hver kommune får altså en indeksverdi mellom 0 og 1, hvor fragmenteringen i kommunestyret er høyere jo lavere verdien er. Hvis alle representantene kommer fra samme parti vil vi ha $HERF = 1$, hvis alle partiene har like mange representanter vil vi ha $HERF = \frac{1}{P}$.

Tabell 15: Instrumentvariable, deskriptiv statistikk.

Variabel	Gj.snitt	St.avvik	Min	Median	Max
Kvinneandel	0.35	0.08	0.17	0.35	0.62
Sosialistandel	0.38	0.15	0	0.38	1
Herfindahl-indeks	0.26	0.10	0.14	0.23	1

Kilde: Statistisk sentralbyrå, egne beregninger.

Deskriptiv statistikk for de tre instrumentvariablene jeg benytter presenteres i tabell 15. Ser at kvinneandelen i norske kommunestyres gjennomsnittlig er på 35 %, og at andelen som utgjør partiene fra venstresiden gjennomsnittlig ligger på 38 %. En gjennomsnittlig verdi på 0.26 på Herfindahl-indeksen kan tolkes som at det er 26 % sjans for at to tilfeldig utvalgte kommunestyreprerentanter tilhører samme parti. At maksverdien her er 1 kan kanskje virke rart, siden dette betyr at det finnes kommunestyres som kun består av ett parti. Grunnen til dette er at fellestater behandles som et eget parti i tallene fra Statistisk

Tabell 16: Resultater fra 2SLS.

Variabel	Første steg	Andre steg
Kvinneandel	-1.193 (-1.22)	
Sosialistandel	0.785 (1.35)	
Herfindahl-indeks	-0.677 (-0.68)	
Barn per ansatt		-0.530 (-1.43)
Boligvariabler	Ja	Ja
Kommunevariabler	Ja	Ja
Tidsdummyer	Ja	Ja
Antall kommuner	415	415
Antall observasjoner	202 337	202 337
R^2	0.417	0.497

Merk: Avhengig variabel i første steg er antall barn per ansatt. I andre steg er det logaritmen til boligprisen. I andre steg blir antall barn per ansatt instrumentert med variablene som inngår i det første steget. T-verdier basert på “cluster”-robuste standardavvik i parentes. *Kilde:* Statistisk sentralbyrå, egne beregninger.

sentralbyrå. Kommunene Kvitsøy, Modalen og Bjarkøy hadde kun felleslister i styret, og de fikk dermed en indeksverdi på 1. Siden dette kun gjelder tre kommuner er det ingen grunn til å tro at det påvirker resultatene i nevneverdig grad. Kommunen med det mest fragmenterte kommunestyret er Vestvågøy, med en indeksverdi på 0.14.

Kan nå benytte 2SLS til å instrumentere variabelen for barn per ansatt i (12). I førstesteget estimeres følgende ligning:

$$\hat{B}_{kt} = \pi_0 + \pi_1 F_k + \pi_2 S_k + \pi_3 HERF_k + \pi_4 X_{ik} + \pi_5 Q_{kt} + \alpha_q + \eta_t + \nu_{ikqt} \quad (14)$$

hvor F_k er kvinneandelen, S_k er sosialistandelen og $HERF_k$ er Herfindahl-indeksen for partifragmentering. I andresteget estimeres så

$$\log P_{ikqt} = \mu_0 + \mu_1 \hat{B}_{kt} + \mu_2 X_{ik} + \mu_3 Q_{kt} + \alpha_q + \eta_t + \nu_{ikqt} \quad (15)$$

Ser at ligning (15) er identisk med (12), bortsett fra at B_{kt} er byttet ut med den estimerte verdien \hat{B}_{kt} fra (14). Hvis antakelsen om at de tre instrumentvariablene er ukorrelert med restleddet holder, så vil vi nå ha $Cov(\hat{B}_{kt} | \nu_{ikqt}) = 0$, og estimering av (15) vil gi forventningsrette og konsistente estimatorene. Dette innebærer likevel at modellen må være identifisert i førstesteget, det vil si at man må kunne forkste følgende nullhypotese fra (14) ved hjelp av en F-test:

$$H_0 : \pi_1 = 0, \pi_2 = 0, \pi_3 = 0 \quad (16)$$

Hvis denne nullhypotesen ikke kan forkastes er ikke variabelen for barn per ansatt korrelert med noen av instrumentvariablene, og resultatene i andresteget kan følgelig ikke brukes til noe fornuftig. Merk at selv om man kan forkaste H_0 i (16) er det fremdeles ikke sikkert

at instrumentene er sterke nok, men det er et nyttig utgangspunkt før man eventuelt går videre til mer avanserte tester av instrumentene.

Resultatene av 2SLS-regresjonen presenteres i tabell 16. Ser umiddelbart at korrelasjonen mellom variabelen for barn per ansatt og de potensielle instrumentvariablene er forholdsvis svak. For å forkaste H_0 fra (16) med 95 % sannsynlighet må F-verdien for de tre instrumentvariablene i førstesteget være over 2,60. Utregning i STATA ga en “cluster”-robust F-verdi på 0,71, det vil si at nullhypotesen ikke kan forkastes. Med andre ord kan ikke resultatene fra andresteget benyttes til å anslå kapitaliseringseffekten av antall barn per ansatt i barnehagene. Ser også at den estimerte effekten nå er over ti ganger så stor, noe som virker urealistisk mye. Resultatet blir altså at instrumentvariabelmetoden i dette tilfellet ikke førte til noen økt robusthet av mine regresjoner.

7 Oppsummering og konklusjon

Denne oppgaven har forsøkt å avdekke om standarden i norske barnehager er en faktor som påvirker boligetterspørselen i kommunene. Problemstillingen var: *Blir norske barnehagetjenester kapitalisert i eiendomsverdiene?* For å besvare denne har jeg benyttet et omfattende datasett for norske boligtransaksjoner i perioden 2003 til 2006, og kombinert dette med data som måler forskjellige aspekter ved barnehagetilbudet i de ulike kommunene. I tillegg har jeg benyttet data for en rekke kommunespesifikke variable som skal kontrollere for forholdene i kommunen. Datasettet har blitt analysert ved hjelp av OLS, hvor målet har vært å avdekke om en eventuell kapitaliseringseffekt av dekningsgraden avhenger av aldersgruppe, og om kvalitetsindikatorer som personalets utdanning og antall barn per ansatt påvirker etterspørselen.

Analysen utnytter hovedsaklig tversnittsvariasjon mellom kommunene. Offentlige tjenester er ofte et område som endrer seg lite over tid, noe som gjør at dette er vanligst også i den øvrige litteraturen. Allikevel er det i tillegg en viss tidsserievariasjon i mitt datasett, noe jeg har forsøkt å utnytte. Grunnen til at tidsvariasjonen er høyere enn vanlig for en såpass kort tidsperiode er at oppgaven dekker en periode med kraftig politisk satsing på barnehagesektoren. Som følge av dette har den gjennomsnittlige dekningsgraden økt markant gjennom perioden. Forskjellene fra kommune til kommune har også blitt noe mindre, i tillegg til at gjennomsnittlig antall barn per ansatt i kommunale barnehager har blitt redusert. Sistnevnte antas å være et utviklingstrekk som har økt kvaliteten på barnehagetjenestene.

Resultatene av analysen tyder på at kommunens dekningsgrad for både de yngste og de eldste barna, samt antall barn per ansatt i de kommunale barnehagene, er faktorer som påvirker etterspørselen etter bolig i kommunen. Det vil si at dette er aspekter ved barnehagetjenestene som blir kapitalisert i eiendomsverdiene. Ved hjelp av “pooled” OLS estimerer jeg at en økning med 10 prosentpoeng i dekningsgraden for barn mellom 1 og 2 år gir en økning i kommunens boligpriser på 4,1 %. For barn mellom 3 og 5 år estimeres det at en tilsvarende økning gir 3,4 % høyere boligpriser. Den eneste kvalitetsfaktoren som ser ut til å påvirke etterspørselen er antall barn per ansatt. For de ansattes eller styrernes utdanning finner jeg ingen signifikant effekt. Det estimeres at en reduksjon i antall barn per ansatt med ett barn gir en økning i boligprisene på 4,6 %. Resultatene for dekningsgrad samsvarer bra med resultatene til Borge og Rattsø (2014), hvor forfatterne i sin “pooled” OLS-regresjon finner at en økning i den samlede dekningsgraden for begge gruppene på 10 prosentpoeng gir en estimert økning i boligprisene på 3,3 %.

Det er flere metodologiske utfordringer knyttet til å estimere kapitalisering kun ved

hjelp av vanlig OLS. Ved tolkning av resultatene fra min regresjon må man derfor være bevisst på at deler av effekten av barnehagevariablene kan skyldes at relevante variable som påvirker boligprisen, og som er korrelert med barnehagevariablene, er utelatt fra regresjonen. Det er også mulig at simultanitetsproblemer gir en skjevhet i estimatene. Mitt forsøk på å filtrere bort tidskonstante effekter ved hjelp av “faste effekter”-metoden var ikke suksessfullt, noe som sannsynligvis skyldes for liten tidsvariasjon i forklaringsvariablene som er av interesse. Jeg forsøker likevel å minimere “utelatt variabel”-problemet ved å inkludere forklaringsvariabler som beskriver både boligen og forholdene i kommunen. For å øke resultatenes robusthet estimerer jeg også en rekke regresjoner som er avgrenset for ulike underkategorier av utvalget.

Ved å dele opp regresjonene for hvert enkelt år fant jeg at kapitaliseringseffekten av dekningsgraden til de yngste var mer signifikant for de siste årene i perioden, og at effekten for de eldste fulgte en omvendt trend. Jeg foreslår at dette kan skyldes den generelle utviklingen til den gjennomsnittlige dekningsgraden i perioden, hvor den for de eldste barna passerte 90 % i 2004. I 2006 hadde den nådd 94 %, noe som betyr at de aller fleste barn mellom 3 og 5 år som ønsket barnehageplass dette året fikk det. Dette førte igjen til reduserte forskjeller mellom kommunene for denne aldersgruppen, noe som kan ha vært medvirkende til at den estimerte effekten av de eldstes dekningsgrad ble redusert. Ved å dele opp utvalget etter kommunens innbyggertall finner jeg også at den estimerte effekten av de eldstes dekningsgrad er høyere jo flere innbyggere som bor i kommunen. Dette er i tråd med Borge og Rattsø (2014) sine resultater. At jeg ikke finner noen tilsvarende effekt for de minste barna viser likevel at man skal være forsiktig med å tolke disse resultatene for bastant.

Det kanskje mest interessante resultatet blant robusthetsregresjonene er effektene som estimeres når utvalget deles opp ut i fra boligtype. Her finner jeg en signifikant effekt av de eldstes dekningsgrad for eneboliger og småhus, men ikke for blokkleiligheter. Effekten for de minste barna er signifikant for alle boligtyper. Dette kan tas som et argument for at resultatene i hovedregresjonen ikke skyldes utelatte variable, siden familier med større barn sjeldnere vil bo i blokkleilighet. Man kan selvsagt ikke helt utelukke at effekten skyldes en annen variabel som både er korrelert med boligpris, barnehagevariablene og boligtype, men det er vanskelig å tenke seg til hvilken variabel det eventuelt skulle være.

På grunn av datamangel er ikke eiendomsskatt inkludert som forklaringsvariabel i regresjonen. For å estimere hvor stor konsekvens dette har for analysen benytter jeg tall for eiendomsskatt i 2005 til å estimere en separat regresjon for dette året. Jeg finner ingen relevante endringer i de estimerte effektene, og foreslår følgelig at utelatelse av eiendomsskatt i regresjonen ikke medfører store avvik i estimatene. Variabelen for antall

barn per ansatt endrer seg forholdsvis lite i de ulike robusthetsspesifikasjonene, og den er i alle tilfeller signifikant på 5 % signifikansnivå. Også for denne variabelen er den estimerte effekten lavest når utvalget er begrenset til kun blokkleiligheter, noe som øker sannsynligheten for at det er en faktisk effekt av barnehagekvalitet som måles.

En bekymring som ikke minskes så mye ved hjelp av robusthetsregresjoner er at barnehagevariablene kan være simultant bestemt med boligprisen. Dette kan for eksempel oppstå ved at politikere øker satsingen på barnehage for å motvirke økt utflytting fra kommunen. Jeg forsøker å omgå dette problemet ved hjelp av instrumentvariabelmetoden. Siden antall barn per ansatt er den barnehagevariabelen som er minst korrelert med de andre forsøker jeg å instrumentere denne variabelen, og dermed utelate dekningsgradene fra modellen. Instrumentene som benyttes er kvinneandel i kommunestyret, sosialistandel i kommunestyret og partifragmentering i kommunestyret. Dessverre viser førstestegsregresjonen at korrelasjonen mellom de potensielle instrumentene og variabelen som skal instrumenteres er for svak til at resultatene kan benyttes.

Hverken “faste effekter”- eller instrumentvariabelmetoden bidro til å redusere de metodologiske utfordringene analysen sto overfor. Borge og Rattsø (2014) viser at kapitaliseringseffekten av både eiendomsskatt og dekningsgrad blir underestimert av vanlig OLS, og det er ingen grunn til å tro at dette ikke er tilfelle i min analyse. Hovedårsaken til dette er at OLS-regresjonen ikke korrigerer for potensiell simultanitet. Likevel viser resultatene av min analyse at barnehagetjenester med stor sannsynlighet blir kapitalisert inn i boligprisene. Robusthetsregresjonene gir som vist resultater som styrker troverdigheten til OLS-estimatene i den opprinnelige modellen. I tillegg er det mest sannsynlig at eventuelle skjevheter i analysen vil trekke i retning av at kapitaliseringseffekten underestimeres, ikke overestimeres. Dette gjelder både det mulige problemet med at eiendomsskatt ikke er inkludert som forklaringsvariabel og problemet med simultanitet. I førstnevnte tilfelle skyldes dette at en eventuell korrelasjon mellom barnehagevariablene og eiendomsskatt trolig vil være positiv, det vil si at kommuner med eiendomsskatt vil ha bedre barnehagestandard. For problemet med simultanitet tilsier økonomisk resonnement at politikere som opplever fallende boligpriser vil øke satsingen på offentlig sektor, ikke redusere den. Dette vil i så fall gi en underestimering av kapitaliseringseffekten. På bakgrunn av dette vil jeg påstå at det er mer sannsynlig at jeg i denne oppgaven estimerer en effekt som er mindre enn i virkeligheten, enn at jeg estimerer en effekt som ikke eksisterer i praksis.

Oppgaven har forsøkt å videreutvikle arbeidet på kapitalisering av barnehagetjenester i Norge ved å dele opp målet for dekningsgrad ut i fra aldersgruppe, og ved å inkludere

kvalitetsindikatorer i analysen. Den estimerte effekten av dekningsgraden for de yngste er både større og mer statistisk signifikant enn for de eldste. Jeg vil likevel ikke påstå at dette tyder på at norske husholdninger er mest opptatt av barnehagedekningen for de minste barna. Snarere viser robusthetsregresjonene at dette resultatet like gjerne kan skyldes at dekningsgraden for de eldste i løpet av perioden ble så høy generelt at den ikke lenger var like relevant ved valg av bokommune. I tillegg kan det faktum at over halvparten av datasettet består blokkleiligheter bidra til at den statistiske effekten av de eldstes dekningsgrad blir mindre, noe som støttes av regresjonen hvor utvalget er delt opp etter boligtype. Analysen gir med andre ord ikke noe tydelig grunnlag for å si at kapitaliseringseffekten av dekningsgrad avhenger av barnets alder. Jeg tror derfor heller ikke at framtidig forskning på området vil gå glipp av mye relevant informasjon ved å benytte et samlet mål på dekningsgrad.

At antall barn per ansatt i kommunale barnehager er en faktor som trolig påvirker kommunens boligpriser er et interessant resultat. En mulig årsak til at denne variabelen ga klarere utslag enn personalets utdanning kan være at antall barn per ansatt er en enklere kvalitetsfaktor for foreldre å observere. Det kan også være så enkelt som at dette er den faktoren som faktisk påvirker kvaliteten mest. Siden den totale dekningsgraden i Norge har økt kraftig er det sannsynlig at fokuset på kvalitet vil bli enda større i årene framover. I den sammenheng kan det være interessant for framtidig forskning å styrke forholdet mellom kvantitative kvalitetsindikatorer og faktisk kvalitetsoppfattelse, for eksempel ved hjelp av intervjuer med norske foreldre. Dette kan både gi kommunene en pekepinn på hvilke områder de burde satse på, og øke fruktbarheten til kapitaliseringsstudier som denne. Et annet område som kan være interessant å se nærmere på i framtiden er hvor stor påvirkning tilgangen på relevant informasjon har på kapitaliseringen av barnehagetjenester. Nettopp for å gjøre det lettere for barnefamilier å innhente informasjon lanserte Kunnskapsdepartementet i 2015 nettstedet barnehagefakta.no (Nybø og Edvardsen 2015). Her kan man ved hjelp av et enkelt søk finne relevant informasjon om enkeltbarnehager eller kommuner. Et mulig framtidig prosjekt kan derfor være å benytte “difference-in-difference”-analyse til å undersøke om den estimerte kapitaliseringseffekten i perioden etter lansering av nettstedet er høyere enn den var i perioden før.

Referanseliste

- Atkinson, Scott E. og Thomas D. Crocker (1987). "A Bayesian Approach to Assessing the Robustness of Hedonic Property Value Studies." *Journal of Applied Econometrics* 2(1), 27-45.
- Barne- og familiedepartementet (2005). *Klar, Ferdig, Gå! Tyngre satsing på de små!* Rapport fra Arbeidsgruppe om kvalitet i barnehagen. Oslo: Barne- og familiedepartementet.
- Bayer, Patrick, Fernando Ferreira og Robert MacMillan (2007). "A unified framework for measuring preferences for schools and neighborhoods." *Journal of Political Economy* 115(4), 588-638.
- Black, Sandra E. (1999). "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114(2), 577-599.
- Borge, Lars-Erik (2005). "Strong politicians, small deficits: evidence from Norwegian local governments." *European Journal of Political Economy* 21, 325-344.
- Borge, Lars-Erik, Anne Borge Johannesen og Per Tovmo (2010). *Barnehager i inntektssystemet for kommunene*. SØF-rapport nr. 02/10. Trondheim: Senter for økonomisk forskning AS.
- Borge, Lars-Erik og Jørn Rattsø (2014). "Capitalization of Property Taxes in Norway." *Public Finance Review* 42(5), 635-661.
- Bradford, D.F, R.A Malt og W.E Oates (1969). "The rising cost of local public services: some evidence and reflections", *National Tax Journal* 22(2), 185-202.
- Butler, Richard V. (1982). "The Specification of Hedonic Indexes for Urban Housing." *Land Economics* 58(1), 96-108.
- Carlsen, Fredrik (2005). "Migration Plans, Local Fiscal Variables, and Local Economic Conditions", *FinanzArchiv* 61(2), 154-177.
- Carlsen, Fredrik, Bjørg Langset, Jørn Rattsø og Lasse Stambøl (2009). "Using Survey Data to Study Capitalization of Local Public Services." *Regional Science and Urban Economics* 39(6), 668-695.
- Charney, Alberta H. (1993). "Migration and the public sector: a survey." *Regional Studies* 27(4), 313-326.

- Clapp, John M., Anupam Nanda og Stephen L. Ross (2008). "Which School Attributes Matter? The Influence of School District Performance and Demographic Composition on Property Values." *Journal of Urban Economics* 63(2), 451-466.
- Dee, Thomas S. (2000). "The Capitalization of Education Finance Reforms." *Journal of Law and Economics* 43(1), 185-214.
- Denk, Ingrid M. (2009). "Førskolelærer." *Store Norske Leksikon*, 14. feb. Hentet 18. nov. 2016 fra <https://snl.no/førskolelærer>.
- Downes, Thomas A. og Jeffrey E. Zabel (2002). "The impact of school characteristics on house prices: Chicago 1987–1991." *Journal of Urban Economics* 52(1), 1-25.
- Ellickson, Bryan (1971). "Jurisdictional fragmentation and residential choice." *American Economic Review* 61(2), 334-339.
- Fiva, Jon H. og Lars Kirkebøen (2011). "Information Shocks and the Dynamics of the Housing Market." *Scandinavian Journal of Economics* 113(3), 525-552.
- Gibbons, Steve og Stephen Machin (2003). "Valuing English primary schools." *Journal of Urban Economics* 53(2), 197-219.
- Haurin, Donald R. og David Brasington (1996). "School Quality and Real House Prices: Inter- and Intrametropolitan Effects." *Journal of Housing Economics* 5, 351-368.
- Innst. S. nr. 250 (2002-2003). *Innstilling til Stortinget fra familie-, kultur- og administrasjonskomiteen om barnehagetilbud til alle, -økonomi, mangfold og valgfrihet*. Oslo: Familie-, kultur- og administrasjonskomiteen.
- McDougall, Gerald S. (1976). "Local public goods and residential property values: some insights and extensions." *National Tax Journal* 29(4), 436-447.
- Nguyen-Hoang, Phuong og John Yinger (2011). "The capitalization of school quality into house values: a review." *Journal of Housing Economics* 20, 30-48.
- Nybø, Katrine og Ingvild Edvardsen (2015). "No kan du sjekke kvaliteten på barnehagen på nett." *NRK*, 11. aug. Hentet 29. sept. 2016. <https://www.nrk.no/norge/no-kan-du-sjekke-kvaliteten-pa-barnehagen-pa-nett-1.12493052>.
- Oates, William E. (1969). "The Effects of Property Taxes and Local Public Spending on Property Values: An Empirical Study of Tax Capitalization and the Tiebout Hypothesis." *Journal of Political Economy* 77(6), 957-971.
- Rosen, Harvey S. og David J. Fullerton (1977). "A Note on Local Tax Rates, Public Benefit Levels, and Property Values." *Journal of Political Economy* 85(2), 433-440.

- Rosen, Harvey S. og Ted Gayer (2010). *Public Finance*. 9. utgave. New York: McGraw-Hill.
- Rosen, Sherwin (1974). "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition." *Journal of Political Economy* 82(1), 34-55.
- Rosenthal, Leslie (2003). "The Value of Secondary School Quality." *Oxford Bulletin of Economics and Statistics* 65(3), 329-355.
- Ross, Stephen og John Yinger (1999). "Sorting and voting: a review of the litterature on urban public finance." I *Handbook of Urban and Regional Economics*, redigert av Paul Chesire og Edwin S. Mills, 2001-2060. 3. utgave. North Holland: Elsevier Science B.V.
- Statistisk sentralbyrå (2003). "Bruk av IKT i husholdningene, 2003, 2. kvartal." Hentet 18. nov 2016 fra <https://www.ssb.no/teknologi-og-innovasjon/statistikker/ikthus/aar/2003-11-06>.
- Statistisk sentralbyrå (2004a). "Eigenbetaling i barnehagar, nivåatal, 15. august 2004, januar." Hentet 17. nov 2016 fra <https://www.ssb.no/utdanning/statistikker/barnegenniv/aar/2004-11-01>.
- Statistisk sentralbyrå (2004b). "Barnehager, 2003, endelige tall." Hentet 17. okt 2016 fra <https://www.ssb.no/utdanning/statistikker/barnehager/aar-endelige/2004-06-15>.
- Statistisk sentralbyrå (2006a). "Eigenbetaling i barnehagar, nivåatal, 15. august 2006, januar." Hentet 17. nov 2016 fra <https://www.ssb.no/utdanning/statistikker/barnegenniv/aar/2006-11-01>.
- Statistisk sentralbyrå (2006b). "Bruk av IKT i husholdningene, 2006, 2. kvartal." Hentet 18. nov 2016 fra <https://www.ssb.no/teknologi-og-innovasjon/statistikker/ikthus/aar/2006-09-14>.
- Statistisk sentralbyrå (2007a). "Flyttinger, 2006." Hentet 17. nov 2016 fra <https://www.ssb.no/befolkning/statistikker/flytting/aar/2007-05-03>.
- Statistisk sentralbyrå (2007b). "Fødte, 2006." Hentet 17. nov 2016 fra <http://www.ssb.no/befolkning/statistikker/fodte/aar/2007-04-19>.
- Statistisk sentralbyrå (2007c). "Barnehager, 2006, endelige tall." Hentet 7. okt 2016 fra <https://www.ssb.no/utdanning/statistikker/barnehager/aar-endelige/2007-06-15>.
- Statistisk sentralbyrå (2008a). "Det svinger i norsk økonomi." Hentet 17. nov 2016 fra <https://www.ssb.no/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/det-svinger-i-norsk-okonomi>.

- Statistisk sentralbyrå (2008b). *Statistikkbanken: Boforhold, levekårsundersøkelsen*.
Tabell 05263: Husholdninger, etter eierform og husholdningstype. Hentet 5. okt 2016 fra <https://www.ssb.no/statistikkbanken/>.
- Statistisk sentralbyrå (2008c). “Eiendomsomsetning, 1. kvartal 2008, foreløpige tall.” Hentet 5. okt 2016 fra <http://ssb.no/bygg-bolig-og-eiendom/statistikker/eiendomsoms/kvartal/2008-05-22>.
- Statistisk sentralbyrå (2012). “Stadig færre mottakere av kontantstøtte.” Hentet 6. nov. 2016 fra <https://www.ssb.no/sosiale-forhold-og-kriminalitet/artikler-og-publikasjoner/stadig-faerre-mottakere-av-kontantstotte>.
- Statistisk sentralbyrå (2015). “Boforhold, levekårsundersøkelsen, 2015.” Hentet 5. okt 2016 fra <https://ssb.no/bygg-bolig-og-eiendom/statistikker/bo/hvert-3-aar/2015-11-25>.
- Statistisk sentralbyrå (2016a). “Boligprisindeksen, 2. kvartal 2016.” Hentet 5. okt 2016 fra <https://www.ssb.no/priser-og-prisindekser/statistikker/bpi/kvartal/2016-07-13>.
- Statistisk sentralbyrå (2016b). “Konsumprisindeksen, oktober 2016.” Hentet 17. nov 2016 fra <http://www.ssb.no/kpi/>.
- Statistisk sentralbyrå (2016c). *Kommune-Stat-Rapportering. KOSTRA*. Hentet 5. okt 2016 fra <https://www.ssb.no/offentlig-sektor/kostra>.
- Straszheim, Mahlon R. (1973). “Estimation of the Demand for Urban Housing Services from Household Interview Data.” *Review of Economics and Statistics* 55(1), 1-8.
- St.meld. nr. 24 (2002-2003). *Barnehagetilbud til alle - økonomi, mangfold og valgfrihet*. Oslo: Barne- og familiedepartementet.
- Svaleryd, Helena (2009). “Women’s representation and public spending.” *European Journal of Political Economy* 25: 186-198.
- Tiebout, Charles M. (1956). “A Pure Theory of Local Expenditures.” *Journal of Political Economy* 64(5): 416-424.
- Utdanningsdirektoratet (2015). “Rett til barnehageplass.” Hentet 17. okt 2016 fra <http://www.udir.no/regelverk-og-tilsyn/finn-regelverk-barnehage/rett-til-barnehageplass/rett-til-barnehageplass2/>.
- Utdanningsdirektoratet (2016). “Foreldrebetaling.” Hentet 17. nov 2016 fra <http://www.udir.no/regelverk-og-tilsyn/barnehage/foreldrebetaling/>.

- Wooldridge, Jeffrey M. (2003). "Cluster-Sample Methods in Applied Econometrics." *American Economic Review* 93(2), 133-138.
- Wooldridge, Jeffrey M. (2013). *Introductory Econometrics: A Modern Approach*. 5. utgave. Boston: Cengage Learning.
- Yinger, John, Howard S. Bloom, Axel Börsch-Supan og Helen F. Ladd (1988). *Property Taxes and Housing Values: The Theory and Estimation of Intra-jurisdictional Property Tax Capitalization*. San Diego: Academic Press.