

Hvilken betydning har avstand til barnehage for boligprisene i Kristiansand?

Helene Sofie Baalerud Tranum

Veileder

Anne Wenche Emblem

*Masteroppgaven er gjennomført som ledd i utdanningen ved
Universitetet i Agder og er godkjent som del av denne utdanningen.
Denne godkjenningen innebærer ikke at universitetet inntår for de
metoder som er anvendt og de konklusjoner som er trukket.*

Universitetet i Agder, 2014

Fakultet for Handelshøyskolen ved UiA

Institutt for Handelshøyskolen ved UiA

Forord

Denne masteroppgaven markerer slutten på en femåring siviløkonomutdanning med fordypning i økonomisk styring og prosjektledelse, ved Universitetet i Agder. Oppgaven utgjør 30 studiepoeng og har som mål å lære studenten å anvende teori og metode for å gjennomføre et selvstendig vitenskapelig arbeid.

Arbeidet med masteroppgaven har vært en lærerik og utfordrende prosess. Datainnsamlingen var tidkrevende, men til gjengjeld fikk jeg et godt og nøyaktig datamateriale som det var spennende å analysere.

Jeg vil rette en stor takk til veilederen min, Anne Wenche Emblem, for gode og raske tilbakemeldinger, og Theis Theisen for sitt engasjement til oppgaven og hjelpsomme tips underveis. I tillegg vil jeg takke familien min for støtte og motivasjon dette semesteret.

Helene Sofie Baalerud Tranum

Kristiansand 28.05.2014

Sammendrag

Hensikten med denne oppgaven er å undersøke hvilken betydning avstand til barnehage har for boligprisene. En bolig er et produkt med ulike egenskaper som forbrukerne etterspør. Det er sammensetningen av ulike egenskaper, og hvordan forbrukerne verdsetter hver og en av disse, som bestemmer prisen på en bolig, gitt et konstant tilbud. I min oppgave er jeg først og fremst ute etter å finne den implisitte prisen på attributtet «avstand til barnehage», men vil også teste flere andre variabler som antas å ha betydning for boligprisen.

Datamaterialet mitt består av totalt 5609 boligsalg registrert i perioden 14.02.2011 til 14.02.2014 i Kristiansand kommune. Den foretrukne semilogaritmiske modellen forklarer 73,9% av variasjonen i boligprisen, og jeg finner en signifikant negativ sammenheng mellom boligpris og avstand til barnehage som gir støtte til hypotesen om at boligpris synker med avstand fra barnehage. Jeg tester i tillegg en modell hvor avstand til barnehage er kategorisert inn i dummyvariabler, og finner at prisen på en bolig som ligger innen 50 meter i luftavstand fra en barnehage vil være lavere enn prisen på en bolig som ligger mellom 51-100 meter og 101-500 meter i luftavstand fra en barnehage. Dette er i tråd med teori om at nærhet til en fasilitet både kan ha en positiv og negativ eksternalitet da det vil være attraktivt å bo i nærheten av den på grunn av tilgjengeligheten, men ikke for nære på grunn av støy og andre ulemper det kan medføre. Begge modellene viser at både fellesgjeld, boareal, alder på bolig, boligtype, eierform og avstand til sentrum også har en signifikant betydning for boligprisen.

På bakgrunn av analysen konkluderer derfor jeg med at boligprisene øker med nærhet til barnehage, men at den synker innenfor 50 meter fra en barnehage.

Innholdsfortegnelse

1.	Innledning.....	1
2.	Bakgrunn	3
2.1.	Kristiansand	3
2.2.	Boligmasse og befolkning.....	4
2.3.	Barnehager	7
3.	Teori	8
3.1.	Kjennetegn ved boligmarkedet	8
3.2.	Alonso-Muth-Mills- modellen	9
3.3.	Den hedonistiske prisfunksjonen	13
3.3.1.	Likevekt på etterspørselssiden av markedet	14
3.3.2.	Likevekt på tilbudssiden av markedet	17
3.3.3.	Markedslikevekt	20
3.4.	Eksternaliteter	21
3.5.	Hypoteser	23
4.	Økonometrisk metode	25
4.1.	Valg av funksjonsform.....	25
4.1.1.	Enkel og multippel lineær regresjonsmodell	25
4.1.2.	Logaritmiske funksjonsformer.....	26
4.2.	Forutsetninger for OLS	28
4.3.	Modellens forklaringskraft.....	29
4.4.	Hypotesetesting	31
5.	Innhenting, bearbeiding og beskrivelse av datamaterialet	32
5.1.	Innsamling av datamaterialet	32
5.2.	Datarensing	35
5.3.	Beskrivelse av datamaterialet.....	36
6.	Analyse.....	46
6.1.	Korrelasjon.....	46
6.2.	Valg av funksjonsform.....	48
6.2.1.	Multippel lineær regresjon.....	48
6.2.2.	Multippel lineær regresjon med avstandsdummyer.....	57
6.2.3.	Dobbeltlogaritmisk regresjon	59

6.2.4. Dobbellogaritmisk regresjon med avstandsdummyer	63
6.2.5. Semilogaritmisk regresjon	64
6.2.6. Semilogaritmisk regresjon med avstandsdummyer	67
6.2.7. Valg av modell.....	68
6.3. Testing av hypotesene	69
7. Nærmere drøfting av hovedproblemstillingene.....	73
7.1. Undersøkelse av verdiendring.....	73
7.2. Kritiske vurderinger av analysen	76
8. Konklusjon	77
9. Litteraturhenvisninger	79

Figurliste

Figur 1: Bydeler i Kristiansand	3
Figur 2: Sirkulær monosentrisk by (Robertsen, 2013).....	9
Figur 3: Husleiegradienten (Dipasquale & Wheaton, 1996).....	10
Figur 4: Virkninger av befolkningsvekst (Dipasquale & Wheaton, 1996)	11
Figur 5: Husholdningenes budfunksjoner (Osland, 2001)	16
Figur 6: Produsentenes offerfunksjon (Osland, 2001)	19
Figur 7: Markedsliekevekt (Osland, 2001)	20
Figur 8: Positiv og negativ effekt på boligpriser (Li & Brown, 1980).....	21
Figur 9: Nettoeffekten av eksternaliteter (Li & Brown, 1980)	22
Figur 10: Oversikt over solgte boliger mellom 2011-2014 etter postnummer.....	38
Figur 11: Gjennomsnittspris solgte boliger, 2011-2014.....	39
Figur 12: Gjennomsnittlig fellesgjeld solgte boliger, 2011-2014	40
Figur 13: Gjennomsnittlig alder solgte boliger, 2011-2014.....	41
Figur 14: Gjennomsnittlig boareal solgte boliger, 2011-2014	42
Figur 15: Gjennomsnittsavstand til sentrum	43
Figur 16: Gjennomsnittsavstand til barnehage, 2011-2014	44
Figur 17: Residualplott, multippel lineær regresjon.....	53
Figur 18: Normalfordeling residualer, multippel lineær regresjon	56
Figur 19: Normalskråplott, multippel lineær regresjon.....	56
Figur 20: Prikkdiagrammatrise, lineær modell	57
Figur 21: Residualplott, dobbeltlogaritmisk regresjon.....	60
Figur 22: Normalskråplott, dobbeltlogaritmisk regresjon.....	62
Figur 23: Prikkdiagrammatrise, dobbeltlogaritmisk modell	62
Figur 24: Residualplott, semilogaritmisk regresjon	65
Figur 25: Normalskråplott, semilogaritmisk regresjon	66
Figur 26: Prikkdiagrammatrise, semilogaritmisk modell.....	67
Figur 27: Avstand til barnehage	74
Figur 28: Avstand til barnehage- kategorisering.....	76

Tabelloversikt

Tabell 1: Beboede boliger etter bygningstype (Statistisk Sentralbyrå, 2011a).....	4
Tabell 2: Privathusholdninger etter eierstatus (Statistisk Sentralbyrå, 2011b)	4
Tabell 3: Privathusholdninger etter husholdningstype (Statistisk Sentralbyrå, 2011c)	4
Tabell 4: Nettoinnflytting Kristiansand (Statistisk Sentralbyrå, 2013b).....	5
Tabell 5: Gjennomsnittlig nettoinnflytting 1993-2011 (Kristiansand Kommune, 2014c).....	5
Tabell 6: Antall barn 0-15 år i de ulike bydelene i 2012 (Kristiansand Kommune, 2014b).....	6
Tabell 7: Antall barnehager og barnehageplasser 2014 (Kristiansand Kommune, 2014d)	7
Tabell 8: Regresjon boareal og bruttoareal	35
Tabell 9: Utelatte observasjoner.....	36
Tabell 10: Kontinuerlige variabler	36
Tabell 11: Kategoriske variabler	37
Tabell 12: Solgte boliger etter boligtype, 2011-2014	43
Tabell 13: Solgte boliger etter eierform, 2011-2014.....	44
Tabell 14: Solgte boliger etter avstandskategorier	45
Tabell 15: Korrelasjonsmatrise	47
Tabell 16: Lineær multippel regresjon	50
Tabell 17: Variance Inflation Factor (VIF)- test	51
Tabell 18: Breusch- Pagan- test, multippel lineær regresjon	54
Tabell 19: Robuste standardfeil, multippel lineær regresjon	55
Tabell 20: Multippel lineær regresjon med avstandsdummyer	58
Tabell 21: Dobbellogaritmisk regresjon.....	59
Tabell 22: Breusch- Pagan- test, dobbeltlogaritmisk regresjon	61
Tabell 23: Robuste standardfeil, dobbeltlogaritmisk regresjon	61
Tabell 24: Dobbellogaritmisk regresjon med avstandsdummyer.....	63
Tabell 25: Semilogaritmisk regresjon	64
Tabell 26: Robuste standardfeil, semilogaritmisk regresjon	66
Tabell 27: Semilogaritmisk regresjon med avstandsdummyer	68
Tabell 28: Undersøkelse av verdiendring.....	74
Tabell 29: Undersøkelse av verdiendring, kategorisering av avstand.....	75

Vedleggsliste

Vedlegg 1: Barnehager i drift per 01.01.2014.....	81
Vedlegg 2: Barnehager stengt de siste tre årene	84
Vedlegg 3: Solgte boliger i de ulike postnumrene og bydelene.....	85
Vedlegg 4: Gjennomsnittlig boligpris etter postnummer, 2011-2014.....	86
Vedlegg 5: Gjennomsnittlig fellesgjeld etter postnummer, 2011-2014	87
Vedlegg 6: Gjennomsnittlig alder på bolig etter postnummer, 2011-2014.....	88
Vedlegg 7: Gjennomsnittlig boareal etter postnummer, 2011-2014	89
Vedlegg 8: Avstand til sentrum fra de ulike postnumrene.....	90
Vedlegg 9: Boligtyper solgt mellom 2011-2014 etter postnummer.....	91
Vedlegg 10: Eierform på boliger solgt mellom 2011-2014 etter postnummer	92
Vedlegg 11: Gjennomsnittlig avstand til barnehage, 2011-2014 etter postnummer.....	93
Vedlegg 12: Avstandskategorier etter postnummer	94
Vedlegg 13: Fullstendig korrelasjonsmatrise	95
Vedlegg 14: Fullstendig semilogaritmisk regresjonsmodell.....	97
Vedlegg 15: Fullstendig semilogaritmisk regresjonsmodell med avstandsdummyer.....	98
Vedlegg 16: Robuste standardfeil, fullstendig semilogaritmisk regresjon	99
Vedlegg 17: Robuste standardfeil, fullstendig semilogaritmisk regresjon med avstandsdummyer.....	100
Vedlegg 18: Kommandoer i Stata	101

1. Innledning

Boligkjøp er muligens den største investeringen mange av oss gjør i løpet av livet, og det er flere faktorer som spiller inn på beslutningen om hvilken bolig man skal velge. Beliggenheten til boligen er kanskje en av de faktorene som betyr mest, siden mange bruker timevis i trafikken til og fra jobb hver uke. Har man i tillegg små barn som må fraktes til barnehagen medfører dette ekstra tid. Det er derfor nærliggende å tenke at det vil være positivt å bo nærme en barnehage, på samme måte som at det er positivt å bo nær arbeidsplassen. Under forutsetning om en monosentrisk by, der alle arbeidsplassene er plassert i sentrum, kan man forvente at boligprisene vil synke med avstand fra sentrum grunnet økte transportkostnader. Videre finnes det forskning på at det å bo for nære en støykilde kan påvirke boligprisen negativt igjen. Det er derfor interessant å se om det samme gjelder for avstand til barnehage.

Problemstillingen min er som følger:

Hvilken betydning har avstand til barnehage for boligprisene i Kristiansand?

Resultatene kan være av interesse for eiendomsutviklere, boligbyggere og offentlige planleggere ved beslutning av hvor nære en barnehage man skal bygge boliger, hva slags boliger som skal bygges nær en barnehage, eller lokalisering av barnehage. Samtidig kan også resultatene være interessante for boligkjøpere og eiendomsmeglere.

Oppgaven innledes med litt statistikk om boligmassen, befolkningen og barnehagene i Kristiansand. I det neste kapittelet presenteres kjennetegn ved boligmarkedet, samt Alonso-Muth– Mills modell om det urbane totemarkedet, og den hedonistiske prisfunksjonen som forklarer hvordan prisen dannes som et resultat av samspillet mellom tilbyderne og etterspørerne i markedet. Avslutningsvis formulerer jeg noen hypoteser basert på teorien jeg har gjennomgått.

Videre følger en kort innføring i økonometri som beskriver metoden min. Deretter vil jeg i det neste kapittelet forklare hvordan jeg samlet inn data, hvilke variabler jeg har med i analysen og til slutt presentere noe deskriptiv statistikk av datamaterialet mitt.

I analysedelen tester jeg ut en lineær, dobbeltlogaritmisk og semilogaritmisk funksjonsform. Jeg tester ut to modeller for hver funksjonsform. Den ene har avstand til barnehage som en

kontinuerlig variabel, mens den andre har avstand til barnehage i kategorier eller såkalte «avstandsdummyer». Deretter velger jeg den modellen som best kan forklare dataene og tester hypotesene før jeg gjør en tolkning av resultatene ved å lage stiliserte regneeksempler. Til slutt kommer jeg med en konklusjon basert på funnene mine.

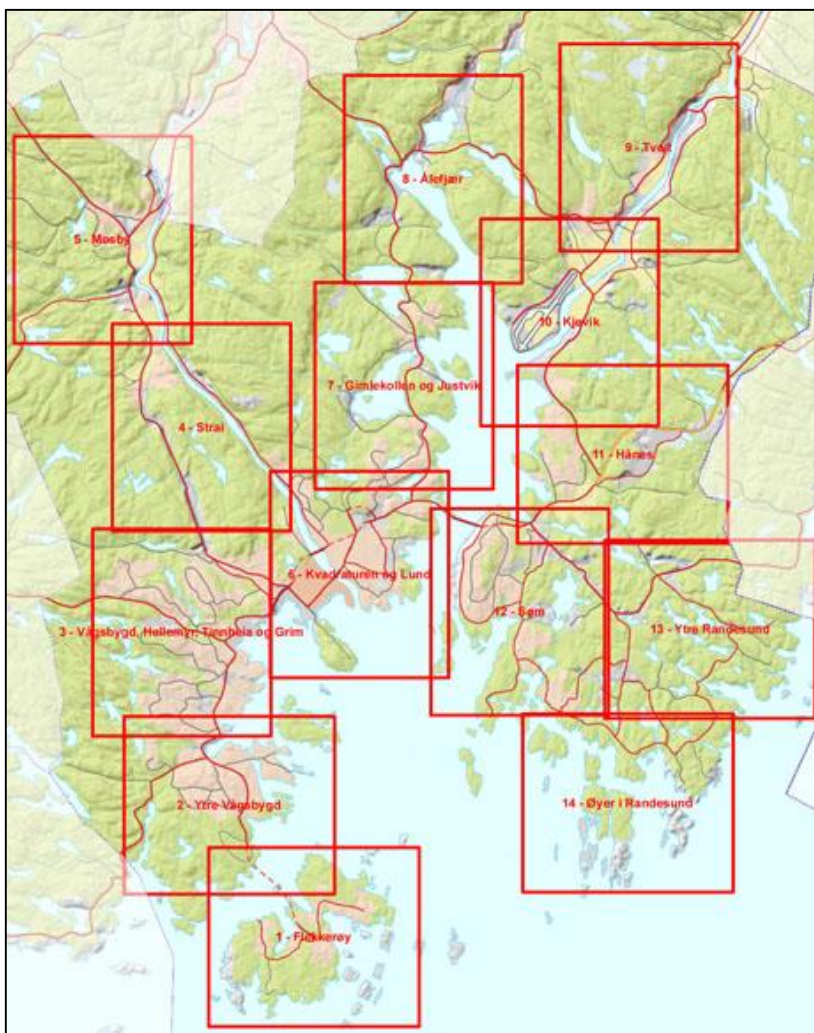
2. Bakgrunn

2.1 Kristiansand

Kristiansand ligger i Vest-Agder og er ifølge statistisk sentralbyrå Norges femte største by med sine 84.476 innbyggere per 1. januar 2013 (Statistisk Sentralbyrå, 2013a).

Byen som da het «Christianssand» ble grunnlagt av Christian IV i år 1641. Kristiansand er vendt mot Skagerrak og er en av Norges travleste internasjonale havner. Stadig flere bedrifter i byen leverer utstyr og tjenester til oljeindustrien i Nordsjøen (Kristiansand Kommune, 2014d).

Kristiansand er delt inn i bydeler som vist i figur 1 (Kristiansand Kommune, 2014a):



Figur 1: Bydeler i Kristiansand

2.2 Boligmasse og befolkning

I 2011 var det totalt 37 039 boliger i Kristiansand. Av tabell 1 ser man at boligtypen som dominerer er eneboliger med 40,4% av den totale boligmassen.

Tabell 1: Beboede boliger etter bygningstype (Statistisk Sentralbyrå, 2011a)

I alt	Prosent	Enebolig	Tomannsbolig	Rekkehus, kjedehus og andre småhus	Boligblokk	Bygning for bofelleskap og annet bygningstype
37 039	100%	40,4%	11,1%	18,6%	25,4%	4,5%

Tabell 2 gir en oversikt over eierstatus i Kristiansand hvor man ser at så mange som 60,1% av alle boligene er selveid.

Tabell 2: Privathusholdninger etter eierstatus (Statistisk Sentralbyrå, 2011b)

	Selveier alene/sameie	Borettslag/Aksjeleilighet	Leier
2011	60,10 %	15,19 %	24,10 %

Tabell 3 gir en oversikt over husholdningstypene i Kristiansand. Nesten 11% av befolkningen har barn i barnehagealder (0-5 år), og da er ikke husholdninger med aleneforeldre telt med. Aleneboende utgjør det største andelen av husholdningstypene.

Tabell 3: Privathusholdninger, etter husholdningstype (Statistisk Sentralbyrå, 2011c)

I alt	100%
Aleneboende	40,4%
Par uten hjemmeboende barn	20,9%
Par med små barn 0-5 år	10,8%
Par med større barn 6-17 år	11,3%
Mor/far med barn 0-17 år	5,7%
Par eller mor/far med voksne barn over 18 år	7,9%
Flerfamiliehusholdninger	3%

Tabell 4 viser nettoinnflyttingen (innflytting minus utflytting) til Kristiansand mellom 2008 og 2012. Man kan se at den positive nettoinnflyttingen har gått nedover helt til 2012 da den nesten doblet seg fra året før.

Tabell 4: Nettoinnflytting Kristiansand (Statistisk Sentralbyrå, 2013b)

År	2008	2009	2010	2011	2012
Kristiansand	760	705	692	445	884

Det er også interessant å få et bilde av hvem som flytter til Kristiansand, og hvor de velger å flytte. Tabell 5 gir en oversikt over gjennomsnittlig nettoinnflytting fra 1993-2011 inndelt i bydeler og alder.

Tabell 5: Gjennomsnittlig nettoinnflytting 1993-2011 (Kristiansand Kommune, 2014c).

	0-6år	7-15år	16-17år	18-19år	20-24år	25-29år	30-39år	40-49år	50-59år	60-69år	70-79år	80år+	Sum
Flekkerøy	14	5	1	0	-1	8	15	4	2	3	0	-2	48
Ytre Vågsbygd	16	3	2	-3	-8	-2	14	1	-4	-3	-5	4	15
Midtre Vågsbygd	6	4	1	1	5	1	12	-1	0	3	5	2	41
Slettheia	-5	-2	2	2	1	-2	-5	1	-1	-8	-10	-7	-33
Hellemyr	12	-4	0	-2	-14	5	6	-8	-10	-4	-3	-1	-24
Tinnheia	-3	0	1	-2	5	3	-4	1	2	0	-5	-5	-7
Grim	-23	3	3	9	27	7	-17	2	-1	-5	-6	-4	-6
Kvadraturen/ Eg	-27	6	15	34	70	-6	-39	23	25	21	30	9	162
Lund/ Sødal	-14	9	7	27	50	0	-16	17	13	3	-1	-11	84
Kongs./Gimlekollen	23	9	3	2	1	-6	27	8	3	-1	-1	7	74
Strai	7	4	-1	1	2	3	10	3	0	1	3	11	43
Mosby	3	-2	0	-2	-2	2	0	0	-1	-1	-3	-3	-10
Justvik	10	-2	0	-2	0	8	12	-1	-3	-2	-2	-1	17
Ålefjær	1	0	0	0	-2	-1	1	0	0	-1	0	0	-2
Tveit	5	2	0	-2	-3	1	4	0	-4	-2	-2	1	0
Hånes	0	-6	-3	-8	-11	-1	-2	-6	-5	0	-2	-3	-47
Indre Randesund	38	9	1	-3	-11	17	38	1	-8	-2	-1	2	80
Ytre Randesund	10	3	0	0	-5	4	13	1	1	1	0	-2	27
I alt	73	42	31	52	102	42	71	46	8	2	-3	-3	463

Ser vi først på summen av alle aldersgruppene innenfor de ulike bydelene, viser flere en negativ gjennomsnittlig nettoinnflytting noe som betyr at flere har flyttet ut av bydelen enn inn i den. Slettheia, Hellemyr og Hånes skiller seg ut med sterkest negativ nettoinnflytting. De bydelene som har høyest nettoinnflytting er Kvadraturen/Eg, Lund/Sødal,

Kongsgård/Gimlekollen og Indre Randesund. For aldersgruppen 0-6 år ser man at det har vært størst innflytting til Vågsbygd, Kongsgård/Gimlekollen, og Indre Randesund, mens det har vært størst utflytting på Grim, Kvadraturen/Eg og Lund/Sødal. De samme trendene gjelder også for personer mellom 30-39 år, som er alderen de fleste velger å få barn.

Fordeling av barn mellom 0-5 år og mellom 6-15 år i de ulike bydelene vises i tabellen nedenfor. Tallene er fra 2012 og er hentet fra Kristiansand Kommune sine nettsider under Kristiansandsstatistikken. Det bor flest barn under 5 år i Vågsbygd, Lund/Sødal og Indre Randesund. Færrest er det på Tinnheia, Kvadraturen/Eg, Strai, Mosby, Ytre Randesund og Ålefjær. Det er lignende andeler for barn mellom 6 og 15 år.

Tabell 6: Antall barn 0-15 år i de ulike bydelene i 2012 (Kristiansand Kommune, 2014b)

	2012	0-5år	Prosent	6-15år	Prosent
Flekkerøy		309	5 %	550	5,27 %
Ytre Vågsbygd		427	7 %	903	8,65 %
Midtre Vågsbygd		644	10 %	1055	10,10 %
Slettheia		376	6 %	569	5,45 %
Hellemyr		383	5 %	657	6,29 %
Tinnheia		203	3 %	313	3,00 %
Grim		386	6 %	500	4,79 %
Kvadraturen/ Eg		230	3 %	210	2,01 %
Lund/ Sødal		724	11 %	1006	9,64 %
Kongsgård/ Gimlekollen		465	7 %	771	7,38 %
Strai		178	3 %	306	2,93 %
Mosby		173	3 %	270	2,59 %
Justvik		346	6 %	341	3,27 %
Ålefjær		30	0 %	56	0,54 %
Tveit		224	4 %	408	3,91 %
Hånes		351	5 %	591	5,66 %
Indre Randesund		875	14 %	1531	14,66 %
Ytre Randesund		222	3 %	394	3,77 %
Uoppgitt		3	0 %	10	0,10 %
Sum		6 549	100 %	10 441	100 %

2.3 Barnehager

I følge foreldreportalen på Kristiansand kommune sine nettsider har Kristiansand 107 barnehager per 01.01.2014 hvorav 78 er private og 29 er kommunale. Tabell 7 viser en oversikt over antall barnehager og barnehageplasser fordelt på bydeler i Kristiansand.

Tabell 7: Antall barnehager og barnehageplasser 2014 (Kristiansand Kommune, 2014d)

Bydel	Antall barnehager	Antall plasser	Prosent av plasser
Kvadraturen/Eg/Odderøya	7	282	5,5 %
Grim	6	307	6,0 %
Strai	4	121	2,4 %
Mosby	3	81	1,6 %
Midtre Vågsbygd	11	532	10,4 %
Ytre Vågsbygd	8	348	6,8 %
Flekkerøy	2	170	3,3 %
Slettheia	5	210	4,1 %
Hellemyr	4	238	4,7 %
Tinnheia	4	91	1,8 %
Lund	12	675	13,2 %
Gimlekollen	7	395	7,7 %
Justvik/Ålefjær/Erkleiv	4	274	5,4 %
Hånes	8	346	6,8 %
Indre Randesund	10	418	8,2 %
Ytre Randesund	7	495	9,7 %
Tveit	5	122	2,4 %
Sum	107	5105	100,0 %

Som tabellen viser er det flest barnehageplasser i Vågsbygd, Lund og Randesund.

Sammenligner vi med tabell 6 ser vi at disse bydelene også har høyest andel av barn under 5 år. Bydelene som har færrest barnehager er Mosby, Strai, Tveit, Flekkerøya og Tinnheia. De fire førstnevnte er de bydelene som ligger lengst unna sentrum, og samtidig også har den minste andelen av barn. Tinnheia ligger ganske nærme byen, men her bor det en relativt liten andel av barn.

I og med at antall barn og barnehageplasser i en bydel sammenfaller såpass bra kan dette tyde på at folk har barnehageplass i nærheten av bostedet sitt.

3. Teori

I dette kapitlet beskriver jeg innledningsvis en del sentrale kjennetegn ved boligmarkedet før jeg går videre med å presentere en modell som beskriver hvordan boligpriser varierer med nærhet til sentrum. Deretter vil jeg gjøre rede for den hedonistiske prisfunksjonen som tar utgangspunkt i hvordan etterspørselen etter forskjellige attributter ved boligen er med på å forklare variasjonen i boligpris. Til slutt formulerer jeg seks hypoteser basert på teorien som er gjennomgått.

3.1 Kjennetegn ved boligmarkedet

En bolig er både et konsumgode og et kapitalobjekt med en lang levetid. Det vil si at når man kjøper en bolig kjøpes boligjenestene som produseres av boligen samtidig som man investerer kapital i boligen (NOU, 2002).

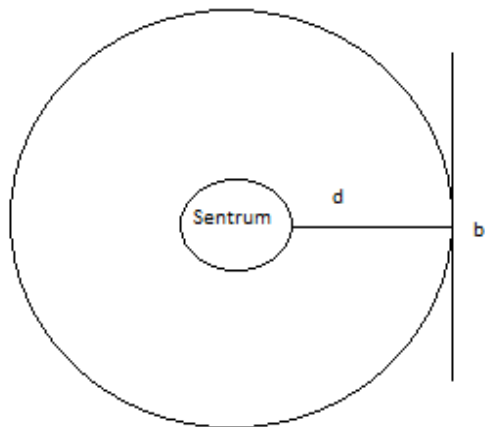
Markedet for tomtearealer og boliger er et produkt differensiert marked i motsetning til råvaremarkedet hvor for eksempel mais og korn er homogene goder. Boliger som har den samme prisen kan ha ulike attributter ved seg. Og selv om boliger kan være helt like, vil aldri beliggenheten være den samme da de ikke kan oppta samme plass (DiPasquale & Wheaton, 1996). Boliger er således et *heterogent* gode.

Boliger skiller seg fra andre goder spesielt fordi de er *immobile*. Det nødvendiggjør at husholdningene flytter på seg, noe som kan medføre økte transaksjonskostnader i form av flyttekostnader i tillegg til meglerkostnader og dokumentavgift. Som et alternativ til å flytte kan huseiere renovere boligen og dermed øke sitt boligkonsum (Smith, Rosen, & Fallis, 1988).

Tilbudet av boliger endrer seg lite fra år til år da nybygging utgjør kun 1% av den samlede boligmassen. En antar derfor at det samlede tilbudet av boliger er gitt på kort sikt, uavhengig av prisen. Boligmassen vil øke hvis nybyggingen er større enn avgangen. På lang sikt vil nybygging bidra til å dempe boligprisene, men siden tilbudet av boliger endres langsomt, må det være forhold på etterspørselssiden som forklarer større endringer i pris (NOU, 2002).

3.2 Alonso- Muth- Mills- modellen

Alonso, Muth og Mills tar for seg det urbane tomtemarkedet hvor de forklarer forskjellen i boligprisene med avstanden til sentrum. I denne klassiske modellen presentert av DiPasquale og Wheaton (1996) viser de hvordan boligprisene øker jo nærmere bysentrum man bor. Forutsetningene for dette er en forenklet fremstilling av virkeligheten der det kun er ett sentrum hvor alle jobber. Husholdningene er identiske og antall pendlere per husholdning er konstant. Det er derfor attraktivt for alle å bo nærme sentrum for å spare både tid og transportkostnader. Dette gjelder selv om vi tar høyde for beliggenhet i forhold til andre fasiliteter som kjøpesenter, skole, transportruter, hav og lignende (DiPasquale & Wheaton, 1996). Videre forutsetter vi at byen har en gitt bygningsstruktur bestemt av historisk bygging. Den monosentriske byen kan illustreres som en sirkel med sentrum i midten:



Figur 2: Sirkulær monosentrisk by (Robertsen, 2013)

Arbeiderne pendler inn til sentrum fra hjemmet langs en rett linje (d) til transportkostnad k per km. Alle boligene er identiske og har en husleie $R(d)$ som varierer med avstand til sentrum. Husleietjenester produseres ved hjelp av et fast antall tomteareal q per hus og en fast byggekostnad (c). Boligtettheten er dermed $1/q$. Inntekten (y) til husholdningene kan bli brukt på pendling, husleie og andre goder (x). Tomteareal allokeres til høyeste pris, da de med høyest betalingsvillighet leier husene. Siden tilbudet av areal på et avgrenset området er uelastisk og etterspørselen er elastisk så er det sistnevnte som bestemmer prisen. Så lenge husholdningene er identiske vil forbruket av andre goder (x^0) være konstant uansett beliggenhet ($x = x^0$). Når markedet er i likevekt synker derfor husleia i takt med økningen i transportkostnader når man beveger seg ut fra bysentrum. Forskjellen i husleie tilsvarende med

andre ord eksakt transportkostnadene, og husholdningene har ingen insentiv til å flytte. Markedet er nå i likevekt. Vi kan uttrykke husleia som følgende ligning;

$$R(d) = y - kd - x^0 \quad (3.1)$$

Når man befinner seg i bysentrum ($d = 0$) og man ikke har noen transportkostnader vil husleia være $R(d) = y - x^0$. Når man beveger seg utover fra bykjernen vil husleia synke i takt med økte transportkostnader. På bygrensen (b) vil husleia være lavest. Utenfor bygrensa er jordbruk den alternative arealbruken og avkastningen per mål er r^a . Med fast boligtetthet vil tomteleia være $r^a q$. På bygrensa ($d = b$) vil husleia være sammensatt av både tomteleie samt den annualiserte kostnaden ved å konstruere et hus (byggekostnaden):

$$R(b) = r^a q + c \quad (3.2)$$

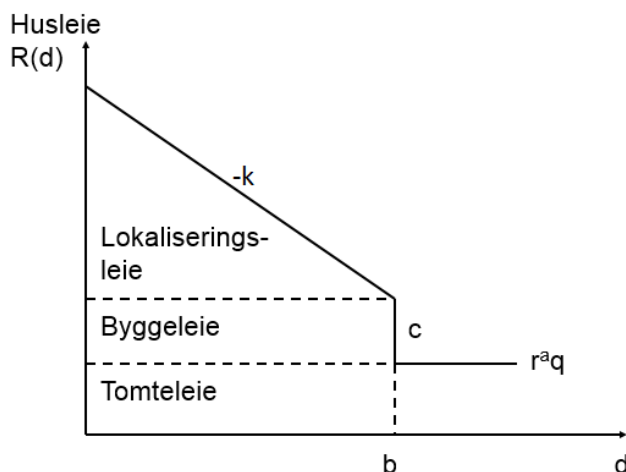
Setter vi $R(d) = R(b)$ og løser med hensyn på x^0 finner vi det alle husholdningene på bygrensa kan konsumere av andre goder:

$$x^0 = y - kb - (r^a q + c) \quad (3.3)$$

Ved å sette dette uttrykket inn i ligning (3.1) vil vi finne at husleiegradienten i avstand d fra sentrum er:

$$R(d) = (r^a q + c) + k(b - d) \quad (3.4)$$

Figuren nedenfor illustrerer husleiegradienten $R(d)$:



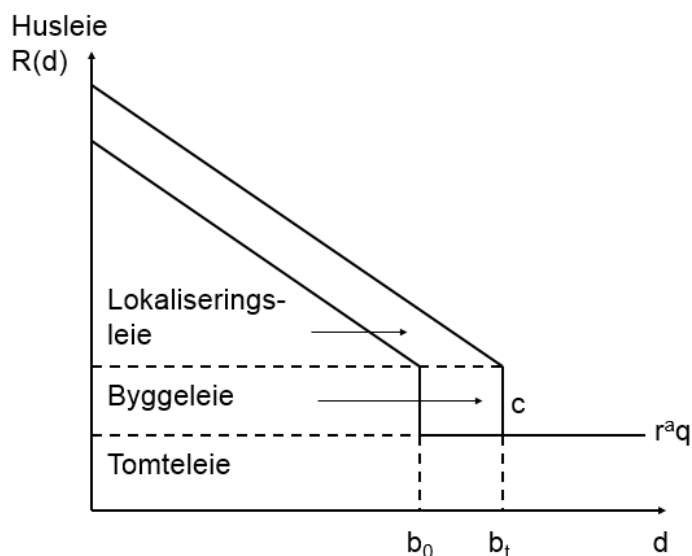
Figur 3: Husleiegradienten (DiPasquale & Wheaton, 1996)

Husleien er gitt langs den vertikale aksen og avstand fra sentrum langs den horisontale aksen. Sentrum er representert der de to aksene møtes, altså i origo. Tomteleie og byggeleie er konstant langs alle beliggenheter innenfor og på bygrensa. Utenfor bygrensa er det kun jordleie. Vi finner helningen til husleiegradienten ved å derivere uttrykket over med hensyn på d :

$$\frac{\partial R(d)}{\partial d} = -k \quad (3.5)$$

Da får vi en lokaliseringsleie som avtar eksakt med avstanden til sentrum med økningen i transportkostnader k . Vi ser altså at det eneste som medfører endring i husleie er lokalisering i forhold til sentrum.

Over tid vil bygrensa utvides som følge av befolkningsvekst. Dette kan illustreres ved at bygrensa skifter fra b_0 til b_t som vist i figur 4:



Figur 4: Virkninger av befolkningsvekst (DiPasquale & Wheaton, 1996)

Når bygrensen utvides vil vi få høyere hus- og lokaliseringsleie for alle lokaliseringer innenfor bygrensa. Den prosentvise økningen i husleie vil være størst i byens utkant og minst i bykjernen. Videre vil prisveksten i husleie på et bestemt sted avta med tiden etter hvert som byen utvides. Vi har at husleia på et hvert tidspunkt er;

$$R_t(d) = (r^a q + c) + k(b_t - d) \quad d \leq b_t \quad (3.6)$$

Ved utledning av eiendomsprisen forventer vi derfor at veksten i husleie skal fortsette inn fremtiden. Vi antar at kalkulasjonsrenta er i og at vi har uendelig tidshorisont ($t \rightarrow \infty$).

Eiendomsprisen $P_t(d)$ er den kapitaliserte verdien av husleia på et tidspunkt t ;

$$P_t(d) = NV(R_t(d)) = \frac{r^a q}{i} + \frac{c}{i} + \frac{k(b_t - d)}{i} - \frac{k b_t g}{i(i - g)} \quad d < b_t, \quad i > g \quad (3.7)$$

Det første leddet tilsvareer nåverdien av all fremtidig landbruksavkastning. Deretter har vi nåverdien av byggeleie (c/i). Det tredje leddet er nåverdien av sparte transportkostnader også kalt lokaliseringsverdien. Det siste leddet er den forventede veksten i lokaliseringsverdi ettersom bygrensen utvides med en vekstrate g (DiPasquale & Wheaton, 1996).

3.3 Den hedonistiske prisfunksjonen

I tillegg til avstand til sentrum er det en hel rekke andre egenskaper som er av verdi for husholdningene. Vi har attributter som er knyttet til selve boligen som for eksempel størrelse, antall bad og soverom, og attributter som er knyttet til lokaliseringen, for eksempel støy, nabolag, nærhet til butikk, industri, skole og transport. Tanken bak hedonistisk metode er at prisen på en bolig er avhengig av ulike egenskaper/attributter ved boligen og hvordan en vektlegger disse.

Jeg presenterer her en oversiktlig fremstilling av den hedonistiske metoden skrevet av Liv Osland i 2001. Den hedonistiske metoden er svært ofte brukt i studier om boligmarkedet. Metoden er utviklet av Sherwin Rosen i 1974, på bakgrunn av Kelvin J. Lancaster (1966) sin ide om at varer kan være sammensatt av nyttebærende attributter. I denne modellen betraktes en bolig som en vektor bestående av n objektive målte attributter (Osland, 2001). Sammensetningen av attributter kan skrives som:

$$Z = (Z_1, Z_2, \dots, Z_n) \quad (3.8)$$

Det som bestemmer totalprisen på bolig er etterspørselen og tilbudet etter enkeltattributter. Det vil si prisene på attributtene, også kalt hedonistisk pris. Man observerer disse prisene indirekte gjennom boligens totale pris. Markedsprisen for en bolig er dermed en funksjon av dens sammensetning av attributter, kalt den hedonistiske prisfunksjonen (Robertsen & Theisen, 2010):

$$P(Z) = P(Z_1, Z_2, \dots, Z_n) \quad (3.9)$$

Videre skal jeg utlede den hedonistiske prisfunksjonen som et resultat av samspillet mellom etterspørernes «budfunksjoner» og tilbydernes «offerfunksjon». I den sammenheng antas det at det finnes et stort antall boliger i markedet slik at valget mellom ulike kombinasjoner av attributter er kontinuerlig videre i modellen. Det forutsettes ingen asymmetrisk informasjon i markedet. Det vil si at hver enkelt aktør har full informasjon om priser og attributter for boliger. Videre har hver enkelt aktør ikke noen innflytelse på markedsforhold og priser. Vi har et effektivt marked hvor tilpasning skjer friksjonsfritt og ser derfor helt bort fra flyttekostnader, søke- og transaksjonskostnader (Osland, 2001).

3.3.1 Likevekt på etterspørselssiden av markedet

Det antas at hver husholdning kun kjøper én bolig til konsum og maksimerer nytten gitt budsjettbetingelsen:

Maksimer:

$$U_j = (Z, X, \alpha_j) \quad (3.9)$$

Gitt:

$$Y_j = X + P(Z) \quad (3.10)$$

Nytten (U) består av en vektor av de forskjellige attributtene (Z), en vektor av alle andre konsumgoder enn bolig (X), og en vektor av parametere som karakteriserer preferansene (α) for husholdning (j). Det antas at nyttefunksjonen er strengt konkav. Husholdningene har en gitt inntekt (Y) som begrenser konsumet deres, og består av bolig og andre goder. Prisen på andre goder settes lik 1.

Første- og andreordensderiverte av prisfunksjonen $P(Z)$ finnes, men har ubestemt fortegn. Når man maksimerer nytten får man følgende uttrykk:

$$\frac{\frac{\partial U_j}{\partial Z_i}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_i} \quad (3.11)$$

Den venstre siden av uttrykket angir den marginale substitusjonsrate mellom Z og X . Det vil si hvor mye du er villig til å gi opp av Z for å få mer av X gitt at du ligger på samme nyttenivå. Dette uttrykket skal være lik den partiellderiverte av prisfunksjonen med hensyn til de respektive boligattributtene og angir hvor mye prisen endrer seg når mengden av attributt i endrer seg. Dette kalles for den hedonistiske prisen for et attributt i og tilsvarer helningen til prisfunksjonen for optimal mengde av Z_i (Osland, 2001).

Den maksimale betalingsvilligheten for ulike sammensetninger av attributtvektorer definerer *budfunksjonen*, og er med på å forklare etterspørselssiden i markedslivekten.

Budfunksjonen kan utledes ved å sette inn de optimale verdiene for Z^* og X^* , slik at $Y_j = X^* + P(Z^*)$. Uttrykket gjøres om:

$$X^* = Y_j - P(Z^*) \quad (3.12)$$

Dette uttrykket settes inn i nyttefunksjonen i ligning (3.9) og gir:

$$U_j = U(Z^*, Y_j - P(Z^*), \alpha_j) = U_j^* \quad (3.13)$$

Når nytten holdes konstant lik U^* og man antar at inntekten er gitt er det rimelig å forutsette at den maksimale betalingsvilligheten θ er lik den prisen man faktisk betaler $P(Z^*)$ (Osland, 2001). Nyttefunksjonen blir da:

$$U_j^* = U(Z, Y_j - \theta_j, \alpha_j) \quad (3.14)$$

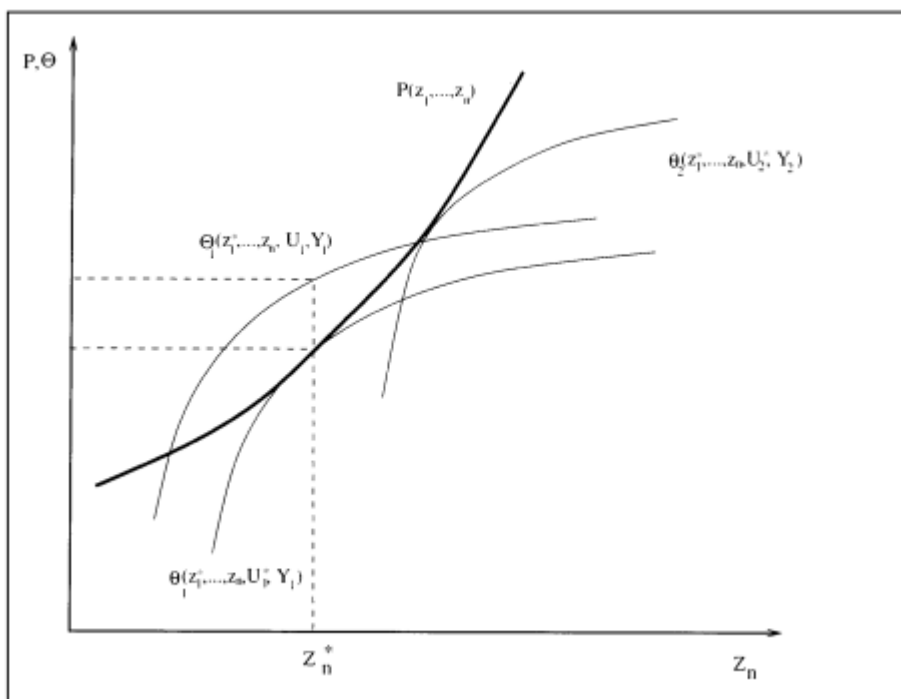
Funksjonen ovenfor uttrykker maksimal betalingsvilje for andre sammensetninger av boligattributioner enn den optimale. Ved andre sammensetninger beregnes en subjektiv pris hvor inntekten nøyaktig brukes opp og husholdningene forblir på det optimale nyttenivået. Siden budfunksjonen vil variere med inntekt- og nyttenivå kan den uttrykkes ved:

$$\theta = \theta(Z, Y_j, U_j, \alpha_j) \quad (3.15)$$

Maksimal betalingsvillighet for en partiell økning i et boligattribution uttrykkes ved implisitt derivasjon i følgende uttrykk:

$$\frac{\partial \theta_j}{\partial Z_i} = \frac{\frac{\partial U_j}{\partial Z_i}}{\frac{\partial U_j}{\partial X}} > 0 \quad i = 1 \dots n \quad (3.16)$$

Så lenge nyttefunksjonen er strengt konkav kan vi ved dobbeltderivering finne at den er positiv men avtagende for partielle økninger i boligattributioner ($\frac{\partial^2 \theta_j}{\partial Z_i^2} < 0$). Budfunksjonen gir et sett av indifferenskurver til hvert nyttenivå (Osland, 2001). Grafisk vil budfunksjonen se slik ut:



Figur 5: Husholdningenes budfunksjoner (Osland, 2001)

I figuren over måles prisen langs den vertikale aksene, og attributtmengde langs den horisontale. Det antas at husholdningen er optimalt tilpasset i alle attributt bortsett fra Z_n . Nyttensnivået stiger ved bevegelse nedover i diagrammet, og maksimeres der sammensetningen av boligattributter gjør at man kommer på den laveste oppnåelige budkurven (Θ). Det er preferansene (α) som gjør at husholdningene tilpasser seg forskjellig, og har forskjellige budfunksjoner. Noen husholdninger ønsker for eksempel en større bolig, andre ønsker seg en mindre. Tenker man seg at Z er boligareal, angir Θ_2 i figur 5 en husholdning som etterspør mer boligareal enn husholdning 1 (Θ_1).

Prisfunksjonen $P(Z)$ er gitt eksogent og er en konveks kurve som viser hvordan den hedonistiske prisen stiger ved en partiell økning i et attributt. Likevekten finnes der $P(Z)$ og Θ tangerer. Likevektbetingelsen er gitt ved å sette ligning 3.11 lik 3.16:

$$\frac{\partial \theta_j}{\partial Z_n} = \frac{\frac{\partial U_j}{\partial Z_n}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_n} \quad j \dots, m \quad (3.17)$$

Denne ligningen forteller oss at marginal betalingsvillighet er lik den marginale prisen på attributtet. Dette gir lik helning i optimum for de to kurvene. Videre er det en forutsetning for

nyttmaksimum at betalingsvilligheten er lik det laveste beløpet man må betale for en bolig med den optimale sammensetningen av attributter: $\theta(Z^*, Y_j, U_j^*, \alpha_j) = P(Z)$, hvor $P(Z)$ er det minste beløpet husholdningene må betale på markedet (Osland, 2001).

3.3.2 Likevekt på tilbudssiden av markedet

Bedrifter produserer boliger og er derfor tilbyderne i markedet. De vil tilpasse seg slik at profitten maksimeres ved å endre antall produserte enheter eller sammensetningen av attributter. For å tilfredsstille kontinuerlig variasjon i attributter antas det at det finnes mange små bedrifter og at hver bedrift spesialiserer seg og produserer kun én boligtype med en gitt sammensetning av attributter. Profittfunksjonen til hver bedrift defineres som:

$$\pi = M \times P(Z) - C(M, Z, \beta) \quad (3.17)$$

Det først leddet av ligningen over er en inntektsfunksjon som består av antall boliger M multiplisert med den hedonistiske prisfunksjonen $P(Z)$. Prisfunksjonen oppfattes av den enkelte bedrift som gitt og uavhengig av antall boliger som bedriften produserer. Deretter trekkes kostnadsfunksjonen (C) fra inntekten for å finne profitten. Kostnadsfunksjonen er en funksjon av antall boliger (M), attributter (Z) og en vektor av skiftparameteren β . Sistnevnte kan være faktorpriser eller produksjonsteknologi for den enkelte bedrift. Kostnadsfunksjonen antas å være en konveks stigende funksjon av antall boliger, og grensekostnadene i produksjon av attributter er positive og ikke-avtagende.

Førsteordens betingelsene for maksimal profitt er:

$$\frac{\partial P}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} \quad i = 1, \dots, n \quad (3.18)$$

Venstre del av ligningen viser den implisitte prisen for et gitt attributt ved at vi deriverer prisen på boligen med hensyn på attributt i . For at maksimal profitt skal finne sted må denne være lik grensekostnaden ($\partial C/\partial Z_i$) per bolig (M) ved en partiell økning i mengden boligattributter (Osland, 2001).

Videre bør bedriften produsere et antall boliger slik at grenseinntekten (prisen på en bolig) er lik grensekostnaden i produksjon av boliger:

$$P(Z) = \frac{\partial C}{\partial M} \quad (3.19)$$

For å sikre andreordensbetingelsene for maksimum må man i tillegg til å anta en konveks kostnadsfunksjon forutsette at $\frac{\partial^2 C}{\partial Z_i^2} > \frac{\partial^2 p}{\partial Z_i^2}$ i det relevante området for maksimum (Osland, 2001).

Offerfunksjonen er definert ved $\phi = (Z, \pi, \beta)$ og angir den minste prisen produsentene er villig til å akseptere for å kunne tilby boliger med forskjellige attributter, til et konstant profittnivå og gitt det optimale antall bolig som produseres. Offerfunksjonen utledes ved å ta utgangspunkt i de optimale verdiene av attributter Z^* , boliger M^* og profittnivå π^* som innsatt i profittfunksjonen gir (Osland, 2001):

$$\pi^* = M^* \times P(Z^*) - C(M^*, Z^*, \beta) \quad (3.20)$$

Når profittnivået holdes konstant kan vi sette offerfunksjonen inn for den hedonistiske prisfunksjonen. Profittfunksjonen uttrykkes da som:

$$\pi^* = M^* \times \phi(Z, \pi, \beta) - C(M^*, Z^*, \beta) \quad (3.21)$$

Deretter deriveres ligningen over med hensyn på M og Z_i og man kan finne førsteordensbetingelsene:

$$\phi(Z^*, \pi^*, \beta) = \frac{\partial C}{\partial M} \quad (3.22)$$

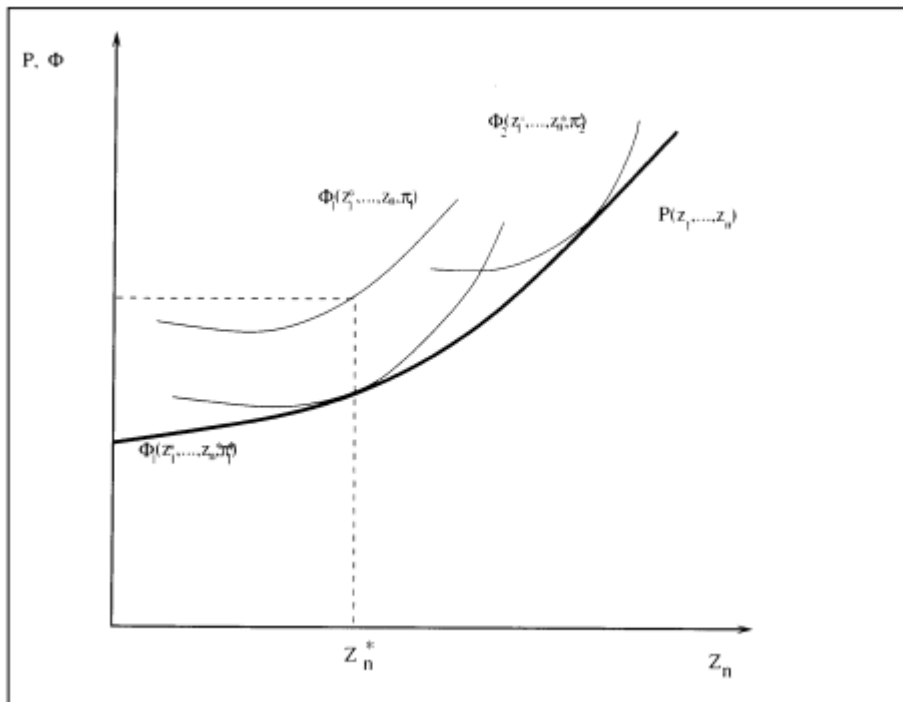
$$\frac{\partial \phi}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} \quad i = 1, \dots, n \quad (3.23)$$

Ligning (3.22) viser at grensekostnaden for å produsere en bolig til må være lik den laveste prisen produsentene er villig til å akseptere (offerfunksjonen). Løses denne med hensyn på M

og setter man uttrykket inn i ligning (3.21) elimineres M , og profittfunksjonen definerer dermed implisitt en relasjon mellom offerprisen og boligattributter (Osland, 2001):

$$\phi = \phi(Z, \pi^*, \beta) \quad (3.24)$$

Offerfunksjonen kan også fremstilles grafisk i figur 6 hvor optimal tilpasning finner sted der offerfunksjonen tangerer den eksogent gitte prisfunksjonen:



Figur 6: Produsentenes offerfunksjon (Osland, 2001)

Offerkurvene er gitt ved et sett isoprofitkurver i grafen over. Skiftparameteren (β) representerer ulik mengde tilbud av attributt (Z_n). En produsent med offerfunksjon Φ_2 tilbyr mer av Z_n enn det en produsent med offerfunksjon Φ_1 gjør. Kurvene er konvekse og profittnivået stiger ved bevegelse opp i diagrammet, $\frac{\partial \phi}{\partial \pi} > 0$.

Likevekt på tilbudssiden er gitt ved førsteordenbetingelsene i ligning (3.18) og (3.23) når $\Phi(Z^*, \pi^*, \beta) = P(Z^*)$:

$$\frac{\partial \phi}{\partial Z_n} = \frac{\frac{\partial C}{\partial Z_n}}{M} = \frac{\partial P}{\partial Z_n} \quad (3.25)$$

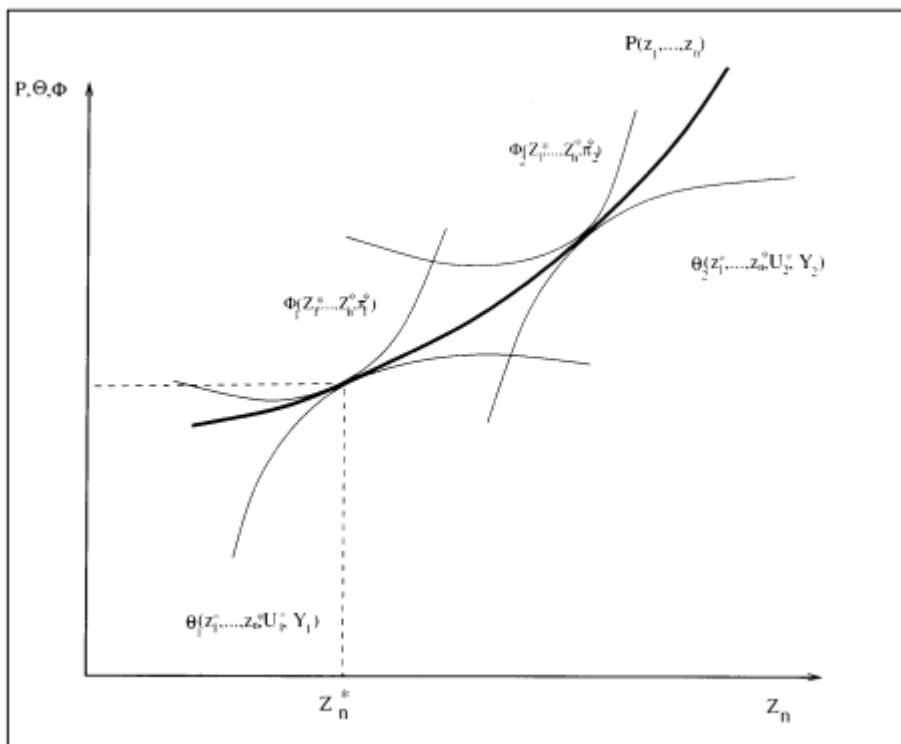
I likevekt vil altså offerprisen være lik den eksogent gitte prisfunksjonen (Osland, 2001).

3.3.3 Markedslikevekt

Når husholdningenes budfunksjoner og produsentenes offerfunksjoner tangerer hverandre har vi markedslikevekt:

$$\frac{\partial \theta}{\partial Z_i} = \frac{\partial P}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} = \frac{\partial \phi}{\partial Z_i} \quad (3.26)$$

Den hedonistiske prisfunksjonen er omtales av Osland (2001) som en «omhylling av både konsumentenes budfunksjoner og produsentenes offerfunksjoner». Markedslikevekten illustreres i figuren nedenfor.



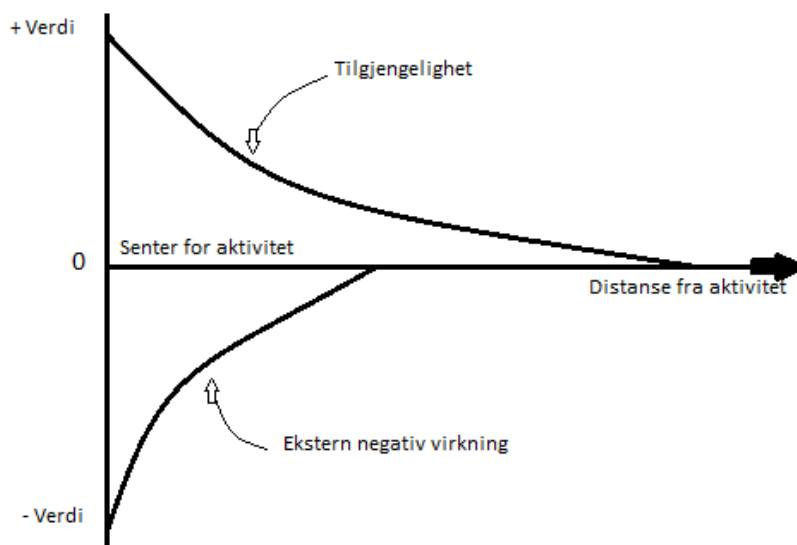
Figur 7: Markedslikevekt (Osland, 2001)

3.4 Eksternaliteter

Det har i den senere tid blitt et økt fokus på eksternaliteter og deres effekt på boligpriser. Eksternaliteter kan være handlinger som påvirker andre enten positivt eller negativt (DiPasquale & Wheaton, 1996). Li og Brown (1980) viste i en studie at det både kan gi en positiv og en negativ effekt på boligpriser ved å bo nær motorvei, industri og kommersielle områder. Bor man for eksempel nær en motorvei har man kanskje kortere reisevei til jobb og større tilgjengelighet til andre plasser, men samtidig bringer dette med seg økt støynivå og forurenset luft rundt boligen. Dette forventes å påvirke boligprisene ved at boliger i nærheten av en motorvei har lavere verdi enn boliger lengere unna (Li & Brown, 1980).

I en nyere studie utført av Rosiers, Lagana og Thériault (2001) fant de at boligprisene synker med avstand fra skole, og at størrelsen på skolen påvirker boligprisene negativt i Quebec, Canada. De fant også at den optimale boavstanden fra skolen er mellom 300 og 500 meter, ca. 9-15 minutter gåavstand. De forklarer den positive effekten med at husholdningene sparer transportkostnader, og den negative effekten med støy og mye trafikk i tilknytning til skolen. (Des Rosiers, Lagana, & Thériault, 2001).

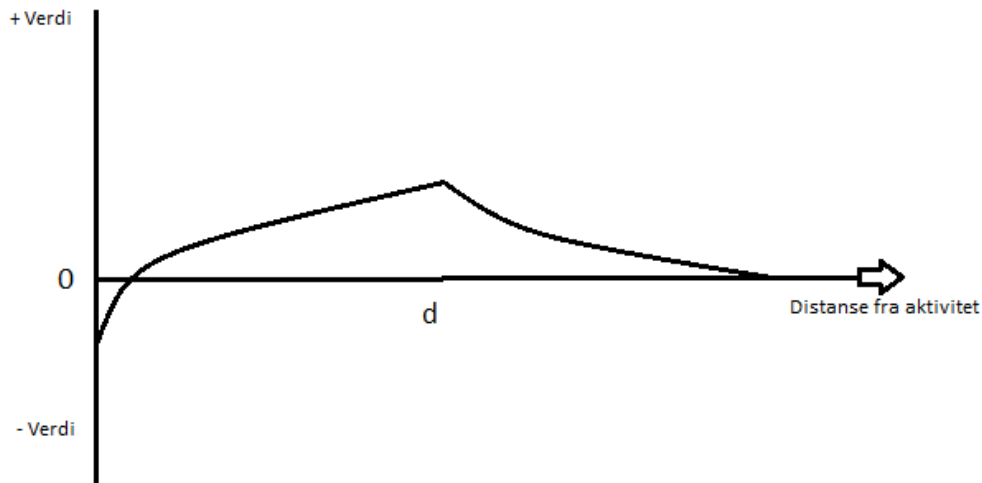
Dette kan vises grafisk i Figur 8 hvor vi ser at verdien av boligen består av en positiv og negativ komponent ved avstand til en slik fasilitet/aktivitet. Tilgjengelighetskurven og den negative eksterne virkningen synker avtakende med avstand fra aktiviteten.



Figur 8: Positiv og negativ effekt på boligpriser (Li & Brown, 1980)

I følge Li og Brown (1980) var studien deres den første som estimerte både den positive og negative effekten av en eksternalitet. Som vi ser av figuren over, så vil vi forvente at den negative eksterne virkningen opphøre raskere enn den positive eksterne virkningen.

Nettoeffekten kan vises i figur 9 under:



Figur 9: Nettoeffekten av eksternaliteter på boligpriser (Li & Brown, 1980)

Grafen i figuren over viser at det vil gi en negativ effekt på boligprisen å bo tett inntil aktiviteten. Med økt distanse fra eksternaliteten øker derimot verdien på boligen helt til punkt d hvor verdien er høyest. Deretter vil boligprisene synke avtakende med avstand fra aktiviteten.

3.5 Hypoteser

Basert på teorien jeg har gjennomgått og problemstillingen min vil jeg nå formulere følgende seks hypoteser.

Hovedhypoteser: Avstand til barnehage

Hypotese 1:

Alonso-modellen i delkapittel 3.2 viser at boligprisen synker med avstand fra sentrum på grunn av økte transportkostnader. Under forutsetning om at små barn blir transportert til og fra bosted og barnehagen benytter jeg meg av denne teorien, og forventer at boligprisene synker med avstand fra barnehage. Jeg antar i tillegg at prisen vil synke avtakende, dvs. at prisen vil synke mindre og mindre ved en økning i avstand fra barnehage.

H1: Det er en negativ sammenheng mellom avstand til barnehage og boligpris

Hypotese 2:

Videre ønsker jeg å undersøke om det vil gi en negativ effekt på prisen å bo veldig nærme en barnehage på grunn at økt trafikk og støy.

H2: Pris på bolig som er lokalisert veldig nærme en barnehage vil være lavere enn pris på bolig som ligger litt lengere unna en barnehage

Kontrollhypoteser:

De følgende bi-hypotesene inkluderes for å redusere risikoen for at jeg tildeler forklaringskraft til variabler som egentlig ikke kan forklare variasjonen i prisen.

Hypotese 3: Avstand til sentrum

Attributter knyttet til hvilket område man bor i forventes å påvirke prisen da det kan være forskjeller mellom enkelte områder. Dette kan for eksempel være nærhet til sjø, natur og kollektiv transport, kommunale tjenestetilbud eller forekomst av kriminalitet i et område. Alonso-modellen forteller oss at boligprisene vil øke jo nærmere sentrum man beveger seg, selv når man kontrollerer for andre lokaliseringsattributter.

H3: Det er en negativ sammenheng mellom avstand til sentrum og boligpris

Hypotese 4: Boareal

Størrelsen på boligen forventes å virke inn på verdien. En økning i boareal er forbundet med en økende, men avtakende nytte. Det vil si at man kan forvente synkende marginal betalingsvillighet for hver partielle økning i kvadratmeter.

H4: Det er en positiv sammenheng mellom boareal og boligpris

Hypotese 5: Sammenheng mellom alder på bolig og boligpris

Alder på bolig er med på å forklare kvaliteten. Jo eldre en bolig er jo større sannsynlighet er det for vedlikeholdsbehov og oppgraderinger. Jeg antar derfor at alderen på bolig har negativ innvirkning på prisen. I følge Robertsen og Theisen (2009) er prisnedgangen størst de første 20-30 årene etter bygging for deretter å avta. De argumenterer med at dette skyldes at boliger ofte renoveres etter 20-30 år, slik at prisforskjellen mellom en 25 år gammel bolig og en 50 år gammel bolig vil være mindre enn en ny bolig og en 25 år gammel bolig.

H5: Det er en negativ sammenheng mellom alder på bolig og boligpris

Hypotese 6: Sammenheng mellom fellesgjeld og boligpris

I en tidligere undersøkelse av boligmarkedet i Kristiansand er det ved en lineær modell funnet at omsetningsprisen på bolig faller med 0,89 kroner når fellesgjelden øker med én krone (Robertsen & Theisen, 2009). Det er derfor interessant å se om jeg finner lignende resultater i min undersøkelse.

H6: Det er en negativ sammenheng mellom fellesgjeld og boligpris

4. Økonometrisk metode

Regresjonsanalyse er en statistisk metode som forsøker å forklare variasjonen i den avhengige variabelen som funksjonen av forandringer i et sett med uavhengige variabler (Studenmund, 2006). I analyser om boligmarkedet anvendes ofte den hedonistiske modellen, som jeg har gjennomgått i kapittel 3. Det er altså denne modellen jeg benytter meg av når jeg nå skal estimere en prisfunksjon som forklarer forholdet mellom boligprisen og en rekke attributter. Vi har fra teorien at prisen på en bolig er en funksjon av de ulike attributtene ved boligen: $P(Z) = P(Z_1, Z_2, \dots, Z_n)$. Videre vi jeg benytte symbolene $Y(X) = Y(X_1, X_2, \dots, X_n)$, hvor Y representerer den avhengige variabelen «boligpris», og X representerer den uavhengige variabelen «attributt».

4.1 Valg av funksjonsform

4.1.1 Enkel og multippel regresjonsmodell

Den enkleste formen for regresjonsmodell er den hvor vi bare har en uavhengig variabel, X :

$$Y = \alpha + \beta_1 X + \epsilon_i$$

Y svarer til den avhengige variabelen som i min oppgave vil være boligpris. Konstantleddet (α) angir verdien av Y når den uavhengige variabelen er null. Betaen indikerer hvor mye Y vil øke dersom X øker med én enhet, også kjent som stigningstallet til den uavhengige variabelen. Restleddet ϵ er lagt til regresjonsligningen og viser den variasjonen i Y som ikke kan forklares med den inkluderte uavhengige variabelen. Selv om man inkluderer flere uavhengige variabler i en regresjonsmodell er det ikke sikkert at variasjonen i den avhengige vil bli fullstendig forklart. Dette kan komme av feil i målinger, feil valg av funksjonsform eller tilfeldigheter. Det er faktisk nesten helt umulig å unngå en form for error i analysen, så vi kan forvente at restleddet ikke vil være lik null (Studenmund, 2006).

I min oppgave har jeg flere enn én uavhengig variabel, og må derfor bruke multippel lineær regresjonsmodell hvor k angir antall uavhengige variabler og i antall observasjoner:

$$Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 D_{1i} + \dots + \beta_k X_{ki} + \epsilon_i, \quad (i = 1, 2, 3, \dots, N)$$

Hver enkelt beta viser hvor mye den avhengige variabelen endrer seg når den tilhørende variabelen endres med én enhet gitt at alle de andre uavhengige variablene holdes konstant. Dummyvariabelen (D) har verdien 1 dersom boligen innehar en definert egenskap, og verdien 0 dersom den ikke har det. Den tilhørende betaen forteller hvor mye Y endrer seg dersom det vi undersøker har den bestemte egenskapen. Variabler som ikke er med i modellen blir ikke holdt konstant når vi estimerer betaene.

Siden vi ikke kan observere de virkelige verdiene av koeffisientene til modellen, må vi kalkulere de ved hjelp av datasettet. Den estimerte regresjonsligningen ser slik ut:

$$\hat{Y} = \alpha + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 D_{1i} + \dots + \hat{\beta}_k X_{ki} + e_i$$

Her representerer «beta-hattene» de estimerte verdiene og e_i residuale til modellen. Jo nærmere \hat{Y} er den virkelige verdien av Y , jo bedre passer modellen vår. Forskjellen mellom disse to er definert som residuale:

$$e_i = Y - \hat{Y}_i$$

Den mest vanlige og enkleste måten å estimere en regresjonsmodell på er ved bruk av en metode kalt Minste Kvadratets Metode eller «Ordinary Least Squares» (OLS) som anslår betaene ved å velge de $\hat{\beta}$ som minimerer kvadratsummen av residualene.

Jo mindre residuale er, desto bedre passer den estimerte ligningen (\hat{Y}) de observerte verdiene (Y). Forskjellen mellom residualet e og feilleddet ϵ er verdt å merke seg. Mens sistnevnte er forskjellen mellom den observerte Y og den virkelige Y , er residuale forskjellen mellom den observerte Y og den estimerte \hat{Y} . Feilleddet som også kalles for error kan aldri bli observert, mens residuale blir kalkulert ved en regresjon (Studenmund, 2006).

4.1.2 Logaritmiske funksjonsformer

Forholdet mellom den avhengige og uavhengige variabelen trenger ikke nødvendigvis å være lineært. Konsekvensene av å bruke gal funksjonsform kan gi feilaktige resultater, så derfor er det viktig å ha en formening om hvordan forholdet er mellom variablene slik at vi kan omkode variablene.

Dobbeltlogaritmisk

I en dobbeltlogaritmisk funksjonsform tar man den naturlige logaritmen (ln) til den avhengige variabelen og ln til de uavhengige variablene. Dette vil gi en lineær form på koeffisientene, mens variablene vil være ikke-lineære.

$$\ln Y = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \beta_3 D_{1i} + \epsilon$$

Dersom en uavhengig variabel øker med 1 prosent når de andre uavhengige variablene holdes konstant, så vil den tilhørende β -koeffisienten fortelle hvor mange prosent den avhengige variabelen vil endre seg med. Siden dummyvariabler kan ha verdien null, tar man ikke ln av disse. Dummyvariabelens beta forteller hvor mange prosent Y endrer seg dersom variabelen inntreffer (dvs. lik 1).

Semilogaritmisk

Den semilogaritmiske funksjonsformen er en variant av den dobbeltlogaritmiske formen hvor man tar den naturlige logaritmen av kun den avhengige variabelen Y .

$$\ln Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 D_{1i} + \epsilon_i$$

Betaverdiene gir endringen i *prosent* på prisen når X øker med én *enhet*. Dummyvariabelens beta forteller hvor mange prosent Y endrer seg dersom variabelen inntreffer (dvs. lik 1).

4.2 Forutsetninger for OLS

De klassiske forutsetningene 1-6 må bli møtt for at OLS estimatorene skal bli best mulig (Studenmund, 2006):

1. Regresjonsmodellen er lineær, korrekt spesifisert og har et feilledd.

Antagelsen om en lineær modell bør ikke bety at teorien bak er det. Ikke-lineære sammenhenger kan gjøres om til en lineær regresjon ved omkoding som vist tidligere. Med korrekt spesifisering menes at vi bruker riktig funksjonsform og at vi har minst mulig utelatte variabler.

2. Feilleddet har et gjennomsnitt på null.

Når vi tar hensyn til hele populasjonens feilledd så skal gjennomsnittet til hele populasjonen være lik null. Sannsynligheten for å oppnå dette øker med et større utvalg.

3. De uavhengige variablene korrelerer ikke med feilleddet.

Det antas at de observerte verdiene av de uavhengige variablene bestemmes uavhengig av verdien på feilleddet. Hvis ikke ville X også forklart en andel av variasjonen i Y som egentlig ville vært forklart av feilleddet.

4. Feilleddets observasjoner er ikke korrelert med hverandre.

Hvis for eksempel det at en observasjons feilledd er positiv bidrar til å øke sannsynligheten for at feilleddet til en annen observasjon også er positiv, er de to observasjonene av feilleddet positivt korrelert. Hvis det finnes en slik systematisk korrelasjon mellom to observasjoner vil det være vanskelig å få et riktig bilde av koeffisientenes feilledd.

5. Feilleddet er homoskedastisk.

Det vil si at variansen eller spredningen er konstant for hver observasjon av feilleddet. Brudd på denne forutsetningen kalles heteroskedastisitet, og kan enklest ses dersom et plott av residualene viser en tydelig vifteform.

6. Ingen perfekt multikollinearitet mellom de uavhengige variablene.

Det vil si at ingen uavhengig variabel er en perfekt lineær funksjon av en annen forklaringsvariabel. Perfekt kollinearitet mellom to variabler innebærer at den relative

endringen i den ene følges av den samme endringen i den andre variabelen. Det betyr at de to faktisk måles det samme. Løsningen vil være å droppe en av disse variablene.

7. Feilleddet er normalfordelt.

Normalfordeling innebærer at observasjonene skal være klokke-formet, symmetrisk og omtrent alle verdiene skal være ± 3 standardavvik fra gjennomsnittet. Dette er ikke et krav for OLS, men hypotesetester som F-test og t-test bygger på at feilleddene er normalfordelte og vil trolig ikke være egnet hvis de ikke er det.

Gauss-Markov teoremet sier at dersom de klassiske modellforutsetningene 1-6 er tilstede, vil de estimerte betaene beregnet ved OLS være de estimatorene som har minst varians blant alle mulige forventningsrette, lineære estimatorer (BLUE= «Best Linear Unbiased Estimator»). Med forventningsrett menes at $\hat{\beta}$ har en forventet verdi lik den samme verdien av β og kan oppnås ved et stort antall observasjoner, små residualer og stor spredning i X-variabelen (Studenmund, 2006).

4.3 Modellens forklaringskraft

Modellens forklaringskraft forteller oss noe om kvaliteten på modellen, og gir oss mulighet til å sammenligne forskjellige datasett, funksjonsformer og kombinasjoner av uavhengige variabler. Forklaringskraften, R^2 som viser andelen av variansen i Y forklart av regresjonslinjen, finner vi ved følgende formel (Studenmund, 2006):

$$R^2 = \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS} = 1 - \frac{\sum e_i^2}{\sum (Y_i - \bar{Y})^2} \quad 0 \leq R^2 \leq 1$$

Vi har at $TSS = ESS + RSS$:

$$TSS = \sum_{i=1}^N (Y_i - \bar{Y})^2,$$
$$ESS = \sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2$$
$$RSS = \sum_{i=1}^N e_i^2$$

«Explained sum of squares» (ESS) måler variasjonen forklart av den estimerte regresjonslinjen mens «residual sum of squares» (RSS) er den delen som forklares utenfor modellen. Da vil «total sum of squares» (TSS) måle den totale variasjonen forklart av regresjonen.

Jo nærmere tallet 1 R^2 ligger, desto bedre passer den estimerte regresjonslinjen observasjonene. Er R^2 nær null viser det at modellen ikke forklarer verdiene til Y bedre enn et enkelt gjennomsnitt av Y, og at modellen derfor er dårlig.

En ulempe ved R^2 er at ved å inkludere flere uavhengige variabler vil aldri forklaringskraften kunne svekkes. Den vil faktisk øke selv om variabelen(e) ikke er med på å forklare variasjonen i Y, med unntak av når koeffisienten til den nye variabelen er null. Når vi inkluderer en variabel, antall observasjoner (N) er konstant og antall estimerte koeffisienter (K) øker med én, vil modellen få en økt forklaringskraft fordi vi ikke tar hensyn til at antall frihetsgrader synker med én. Tar vi imidlertid hensyn til dette vil \bar{R}^2 gi et mer riktig bilde fordi den kun vil øke dersom den ekstra forklaringskraften er sterkere enn tapet av frihetsgrader.

$$\bar{R}^2 = 1 - \frac{\sum e_i^2 / (N - K - 1)}{\sum (Y_i - \bar{Y})^2 / (N - 1)}$$

R^2 erstattes derfor i mange tilfeller av \bar{R}^2 fordi dette gir et riktigere mål på forklaringskraften til modellen. Dette er imidlertid kun ett mål på kvaliteten til modellen og man burde se dette i sammenheng med andre kvalitetsmål slik som tidligere studier, forventede fortegn til beta-koeffisientene og statistisk signifikans.

4.4 Hypotesetesting

Før man går i gang med selve analysen er det hensiktsmessig å formulere noen hypoteser som skal testes. Nullhypotesen (H_0) er et utsagn som skal si noe om hva man ikke forventer, alternativhypotesen (H_A) skal si noe om de resultatene man forventer. For å vite om nullhypotesen kan forkastes eller ikke, må kritisk verdi, også kalt signifikansnivået, bestemmes. Den kritiske verdien sier noe om sannsynligheten for å forkaste nullhypotesen når den er sann (Type I error). Sannsynligheten for å akseptere en falsk nullhypotese kalles type II error. Jo lavere signifikansnivå man velger, mindre sannsynlig er det at man gjør en type I feil, men det øker sannsynligheten for å gjøre en type II feil. Studenmund (2006) anbefaler en kritisk verdi på 5% dersom konsekvensene av å gjøre en type II feil ikke er for store. Hvis parameterens signifikansnivå er mindre enn 5% kan man da med 95% sannsynlighet forkaste H_0 .

T-test er vanlig å benytte for å teste hypoteser som omhandler de enkelte koeffisientene. Her finner vi kritisk t-verdi fra en tabell og bestemmes på grunnlag av om vi har en ensidig eller tosidig test, spesifisert signifikansnivå og antall frihetsgrader. Antall frihetsgrader er $(N - K - 1)$ hvor N er antall observasjoner og K er antall koeffisienter med fratrekk fra konstantleddet. Dersom den kalkulerede t-verdien er større (i absolutt verdi) enn den kritiske t-verdien og har det samme fortegnet som alternativhypotesen kan nullhypotesen forkastes.

En alternativ tilnærming til t-testen er *p-verdi* som gir oss den laveste signifikante verdien vi kan forkaste nullhypotesen ved. P-verdien er mellom 0 og 1 og jo lavere verdien er jo større sannsynlighet er det for at vi kan forkaste nullhypotesen.

Hvis p-verdien er mindre enn signifikansnivået og hvis regresjonskoeffisienten har det samme fortegnet som alternativhypotesen kan vi altså forkaste nullhypotesen (Studenmund, 2006).

5. Innhenting, bearbeiding og beskrivelse av datamaterialet

5.1 Innsamling av datamateriale

Det første jeg gjorde var å skaffe meg en oversikt over alle barnehager i Kristiansand kommune. På kommunens nettsider fant jeg at det er totalt 107 antall barnehager i drift i dag, sammen med informasjon om adresse og antall plasser. Siden jeg er interessert i om avstand til barnehage påvirker boligprisene i Kristiansand var det viktig å sikre seg nok data i ulike avstander fra barnehage. Jeg valgte derfor å se på antall boligsalg de tre siste årene, nærmere bestemt boligsalg mellom 14. februar 2011 og 14. februar 2014. Dette medførte at jeg også var nødt til å finne ut av hvilke barnehager som har åpnet, stengt og byttet lokalisering de tre siste årene. Dette viste seg å være en prosess som tok mye tid. På finn.no sin kartfunksjon kunne jeg plassere barnehagene på kartet. Kartet er ikke oppdatert, så det dukket opp barnehager som ikke lenger er i drift samtidig som det manglet en del nyåpnede barnehager. For å finne ut av hvilke barnehager som har åpnet de tre siste årene sendte jeg e-post til alle barnehagene, og de jeg ikke hørte noe fra ringte jeg til, eller sendte en ny e-post. Barnehager som er lagt ned var det ikke så lett å få noe informasjon om annet enn at jeg kunne se på finn.kart.no at det har vært barnehager på plasser hvor det ikke er nå lenger. I tillegg til internettsøk var jeg i kontakt med både kommunen og fylkesmannen uten at de kunne gi meg noe klart svar. I de tilfellene hvor jeg ikke klarte å finne noe informasjon om når nedleggelse av barnehage har funnet sted har jeg derfor valgt å se bort fra gjeldene observasjoner. En oversikt over hvilke barnehager som er med i undersøkelsen finnes i vedlegg 1 og 2. Jeg gjør en forenkling og betrakter alle barnehagene som like. Jeg har altså ikke skilt mellom størrelsen på barnehagene, eller andre kjennetegn som for eksempel familiebarnehage, bedriftsbarnehage, privat eller kommunal barnehage.

Informasjonen om boligomsetningene er sekundærdata hentet fra Eiendomsverdi AS sin database på nettet. I følge nettsidene deres er Eiendomsverdi «et selskap som overvåker og registrerer aktivitet og utvikling i de norsk eiendomsmarkedene» (Eiendomsverdi, 2014). Jeg sorterte utvalget mitt på postnumre i Kristiansand og datamaterialet bestod av totalt 6972 observasjoner før datarensing. All data er behandlet i Stata, og kommandoene fra «Do-filen» er samlet i vedlegg 18. Fra eiendomsverdi finner jeg i tillegg til pris en rekke andre detaljer jeg har nytte av:

- Adresse
- Eierform
- Boligtype
- Boareal
- Bruttoareal
- Salgsdato
- Fellesgjeld
- Byggeår

Avstand til barnehage:

Avstand fra bolig til barnehage er målt ved luftavstand i meter til *nærmeste* barnehage for hvert eneste boligsalg med hver sin unike adresse. Dette var en svært tidkrevende prosess siden jeg har 5609 observasjoner. Målingen er som sagt gjort ved bruk av finn.no sin kartfunksjon. I tillegg til å ha avstand til barnehage som en kontinuerlig variabel, har jeg valgt å kategorisere avstanden inn i 0-50 meter, 51-100 meter, 101-500 meter, 501-1000 meter og over 1000 meter som dummyvariabler. Jeg definerer «veldig nære» som innen 50 meter, mens mellom 51 og 100 er «nære», 101-500 meter og 501-1000 meter er «innen gåavstand», mens over 1000 meter er «kjøreavstand». Denne kategoriseringen håper jeg vil si meg noe om en mulig negativ ekstern effekt av å bo veldig nærme en barnehage. Jeg valgte luftavstand fordi jeg håper det vil gi et mer riktig bilde av hva som er nært og hva som ikke er nært. For eksempel ligger noen boliger helt inntil en barnehage.

Avstand til sentrum:

Jeg har gitt alle boliger med samme postnummer lik avstand til sentrum. For å avgrense postnumrene har jeg brukt en nettside laget av Erik Bolstad hvor det er tydelig markert på kartet hvor de ulike postnumrene ligger (Bolstad, 2014). Jeg valgte et punkt midt i området til et postnummer som en tilnærming til avstanden for alle boligene innenfor tilhørende postnummer (gjennomsnitt). Avstanden er målt i antall kilometer kjøreavstand til Markens gate 25 ved hjelp av finn.no sitt kart. Denne adressen ligger i postnummer 4611, og alle boligene som er solgt innenfor dette postnummeret har jeg satt avstand til sentrum lik 0,1. En liste over avstandene til de ulike postnumrene finnes i vedlegg 8.

Eierform:

Eierformen antas å påvirke prisen på bolig. Eierform deles inn i selveier, borettslag og aksjeleilighet. Disse variablene er kategoriske og må derfor gjøres om til dummy-variabler i analysen.

Boligtype:

Boligene er delt inn i enebolig, tomannsbolig, rekkehus og leilighet og er omkodet til dummy-variabler.

Boareal:

Boareal, eller primærrrom er målt fra innerveggene og utgjør de viktigste rommene i boligen, som soverom, bad, stue og kjøkken. Dette vil være en kontinuerlig variabel i analysen min.

Bruttoareal:

Bruttoareal er arealet av hele boligen målt fra yttersiden. Siden det er mange observasjoner som mangler denne opplysningen velger jeg heller å bruke variabelen boareal. I de tilfeller hvor jeg har bruttoareal oppgitt, men ikke boareal, bruker jeg bruttoareal til å estimere boarealet ved en regresjon. Dette gjøres i neste delkapittel.

Fellesgjeld:

Fellesgjeld tas med som en kontinuerlig variabel.

Alder:

Alder er en kontinuerlig variabel beregnet ved å ta salgsdato minus byggeår.

Området:

Lokaliseringsattributter fanges opp av postnummerdummyer.

5.2 Datarensing

I datasettet mitt manglet det en del opplysninger om enkelte observasjoner som gjorde at jeg ikke kunne bruke disse i analysen min. Jeg har som nevnt tidligere utelatt de observasjonene hvor jeg ikke visste når den nærmeste barnehagen stengte. Dette er viktig for å minske risikoen for å gjøre systematiske feil i observasjonene. En observasjonsfeil kan defineres som forskjellen mellom den rapporterte verdien og den virkelige verdien (Zikmund, Babin, Carr, & Griffin, 2010). Videre har jeg utelatt alle observasjoner som manglet byggeår, pris og adresse i tillegg til observasjoner hvor både boareal og bruttoareal på bolig mangler eller er lik null. Ved tilfeller der informasjon om boareal mangler, men bruttoareal er oppgitt har jeg estimert boareal ved hjelp av lineær regresjon:

$$boa = \alpha + \beta * bta + \epsilon,$$

hvor *boa* er boareal, *bta* er bruttoareal, α er konstantleddet og ϵ er feilleddet. β representerer regresjonskoeffisienten til bruttoarealet og forklarer styrken på forholdet mellom boareal og bruttoareal. Tabellen under viser regresjonsanalysen:

Tabell 8: Regresjon boligareal og bruttoareal

Source	SS	df	MS	Number of obs = 3961		
Model	10040215.1	1	10040215.1	F(1, 3959) =	44022.74	
Residual	902924.535	3959	228.068839	Prob > F =	0.0000	
Total	10943139.7	3960	2763.41911	R-squared =	0.9175	
				Adj R-squared =	0.9175	
				Root MSE =	15.102	

boa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
bta	.7156465	.0034108	209.82	0.000	.7089594	.7223337
_cons	14.86128	.5023733	29.58	0.000	13.87634	15.84621

Som vi ser er konstantleddet 14.86128 mens den estimerte regresjonskoeffisienten til bruttoarealet er 0.7156. Dette ga meg regresjonsligningen $boa = 14.86128 + (0.7156465 * bta)$, og har hjulpet meg med å estimere boarealet til 22 observasjoner der det mangler¹.

¹ Boarealet er beregnet på bakgrunn av 3961 observasjoner der bruttoarealet er oppgitt.

Under følger en oversikt over antall observasjoner som er utelatt grunnet manglende data.

Tabell 9: Utelatte observasjoner

Antall observasjoner	6972
Mangler byggeår	-60
Mangler pris	-45
Mangler boa og bta	-1192
Mangler adresse	-2
Mangler opplysning om stenging	-64
Gjenværende observasjoner	5609

5.3 Beskrivelse av datamaterialet

Som en innledning til analysen presenterer jeg her datamaterialet mitt, og gir en kort deskriptiv beskrivelse av dataene. Tabell 10 viser en oversikt over de kontinuerlige variablene, mens tabell 11 viser de kategoriske variablene som har dummyverdier 0 og 1.

Tabell 10: Kontinuerlige variabler

Variable	Obs	Mean	Std. Dev.	Min	Max
prisen	5609	2526074	1239147	50000	1.13e+07
fgjeld	5609	89572.38	254821.5	0	2835000
alder	5609	40.05687	27.53583	0	311
BOA	5609	104.2893	52.20413	16	418
avstandbhg	5609	325.7755	270.8529	0	4500
sentrum	5609	5.694117	3.918426	.1	18.9

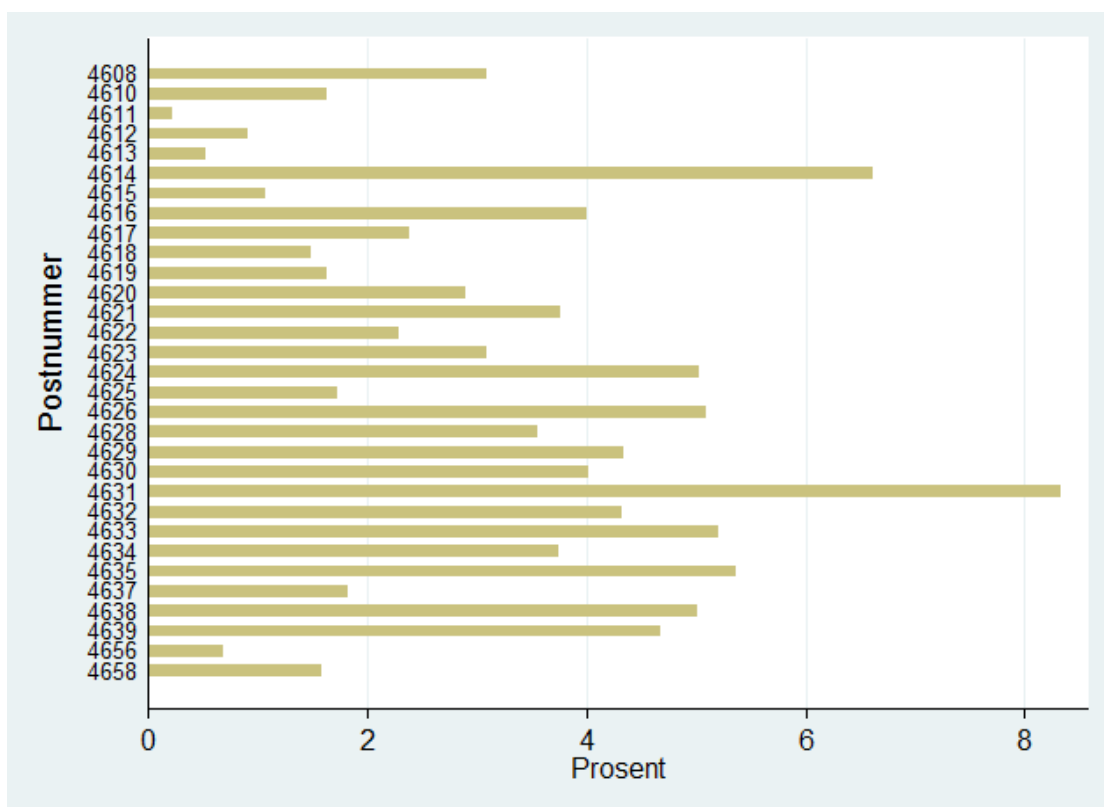
Tabell 11: Kategoriske variabler

Variable	Obs	Mean	Std. Dev.	Min	Max
borettslag	5609	.285969	.451915	0	1
selveier	5609	.714031	.451915	0	1
enebolig	5609	.2495989	.4328194	0	1
tomannsbolig	5609	.0960956	.2947485	0	1
leilighet	5609	.5286147	.499225	0	1
rekkehus	5609	.1256909	.3315302	0	1
P4608	5609	.0308433	.1729084	0	1
P4610	5609	.0162239	.126347	0	1
P4611	5609	.0021394	.0462085	0	1
P4612	5609	.0090925	.0949287	0	1
P4613	5609	.0051703	.0717248	0	1
P4614	5609	.0661437	.2485553	0	1
P4615	5609	.0106971	.1028813	0	1
P4616	5609	.0399358	.1958259	0	1
P4617	5609	.0238902	.1527206	0	1
P4618	5609	.0147976	.120753	0	1
P4619	5609	.0162239	.126347	0	1
P4620	5609	.0288822	.1674902	0	1
P4621	5609	.0376181	.1902878	0	1
P4622	5609	.0228205	.1493441	0	1
P4623	5609	.0308433	.1729084	0	1
P4624	5609	.0502763	.2185341	0	1
P4625	5609	.0172936	.1303748	0	1
P4626	5609	.0509895	.2199959	0	1
P4628	5609	.0354787	.1850029	0	1
P4629	5609	.0433232	.2036019	0	1
P4630	5609	.0401141	.1962443	0	1
P4631	5609	.083259	.276298	0	1
P4632	5609	.0431449	.2032014	0	1
P4633	5609	.0520592	.2221662	0	1
P4634	5609	.0374398	.1898539	0	1
P4635	5609	.0536638	.2253731	0	1
P4637	5609	.0181851	.1336321	0	1
P4638	5609	.0500981	.2181667	0	1
P4639	5609	.0467106	.2110372	0	1
P4656	5609	.0067748	.0820374	0	1
P4658	5609	.0158674	.1249735	0	1
innenfor50	5609	.0492066	.2163185	0	1
innenfor100	5609	.0779105	.2680546	0	1
innenfor500	5609	.7206276	.4487309	0	1
innenfor1000	5609	.1306828	.3370832	0	1

Tabellene viser antall observasjoner, gjennomsnittet, standardavviket og minste og største verdi av observasjonene. Gjennomsnittet er summen av alle verdiene til en variabel delt på antall observasjoner. Standardavviket forteller noe om spredningen på verdiene av observasjonene, det vil si hvor mye verdiene avviker fra gjennomsnittet. Minimums- og maksimumsverdiene gir oss et bilde på variasjonsbredden i observasjonene. Når min og maks-verdiene ligger nær hverandre, vil de også ligge nær gjennomsnittet og variabelen har et lite standardavvik. I motsatt tilfelle vil de ligge langt fra gjennomsnittet og vi vil ha et større standardavvik.

Postnummer

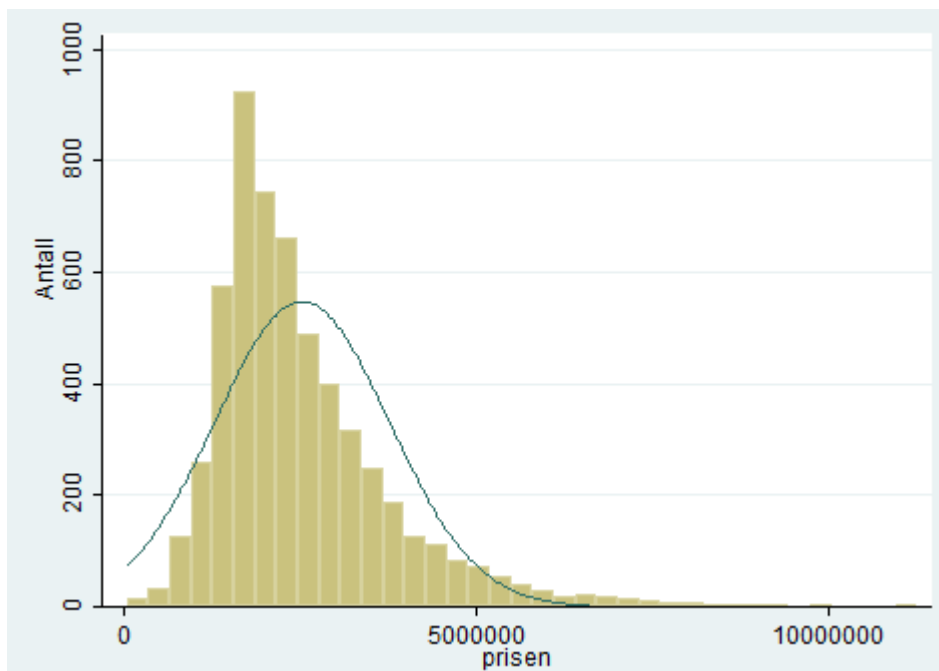
Utvalget mitt er sortert på postnumre. Fordelingen av solgte boliger innenfor de ulike postnumrene finnes i figuren nedenfor. Som vi ser av histogrammet under så ligger de fleste av observasjonene i postnummer 4631 og 4614. I vedlegg 3 har jeg lagt ved en oversikt over antall solgte boliger mellom 2011-2014 etter postnummer.



Figur 10: Oversikt over solgte boliger mellom 2011-2014 etter postnummer

Pris

Gjennomsnittsprisen på boliger i utvalget mitt er 2 526 074 kr. Den minste verdien er på 50 000 og er lav fordi jeg ikke har tatt hensyn til fellesgjelden på 735 000. Den største verdien er på 11.250.000 og er uten fellesgjeld. Det store spranget mellom den minste og største verdien gjør at standardavviket blir høyt. Klokkeformen rundt histogrammet i figuren nedenfor viser hvordan variabelen ville sett ut hvis den hadde vært perfekt normalfordelt (Midtbø, 2012). I dette tilfellet ser den litt skjev ut da gjennomsnittet er nærmere minimumsverdien enn maksimumsverdien.

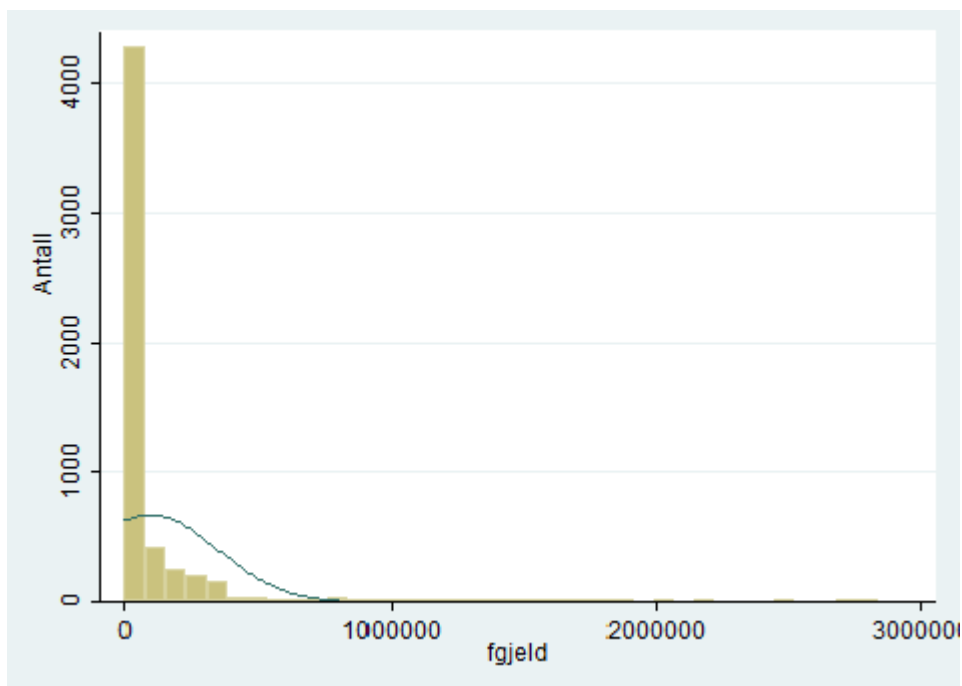


Figur 11: Gjennomsnittspris solgte boliger, 2011-2014

En oversikt over gjennomsnittsprisen i de ulike postnumrene finnes i vedlegg 4 hvor de er rangert lav-høy. Noen av de laveste gjennomsnittsprisene i utvalget mitt er blant annet postnummer 4617, 4629, 4614, 4619, 4624 og 4626. Noen av de høyeste gjennomsnittsprisene finnes i postnummer 4637, 4610, 4625, 4656 og 4633.

Fellesgjeld

Gjennomsnittlig fellesgjeld er 89 572, med et standardavvik på 1 239 147. Minste verdi er null, og største verdi 2 835 000. Grunnen til at gjennomsnittet er såpass lavt er fordi kun 1721 observasjoner har fellesgjeld, og gjennomsnittet er regnet ut fra alle 5609 de observasjonene.

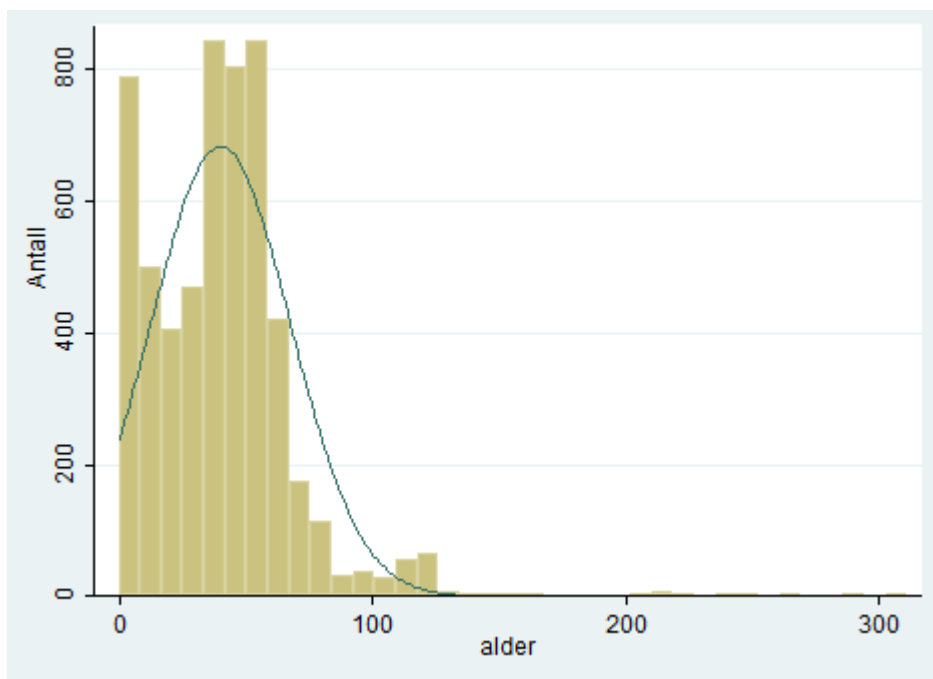


Figur 12: Gjennomsnittlig fellesgjeld solgte boliger, 2011-2014

En oversikt over gjennomsnittlig fellesgjeld i de ulike postnumrene ligger i vedlegg 5. En del av postnumrene med høy gjennomsnittlig fellesgjeld ser ut til å være postnumre med lav gjennomsnittlig boligpris. Samtidig har en del av postnumrene med lav gjennomsnittlig fellesgjeld høy gjennomsnittlig boligpris. Siden fellesgjeld og pris ikke er summert er det vanskelig å sammenligne det gjennomsnittlige prisnivået i de ulike postnumrene.

Alder

Gjennomsnittlig alder på boliger i utvalget mitt er 40 år. Minste verdi er null og den eldste boligen er 311 år. Som vi ser av histogrammet under er de fleste boligene i utvalget mindre enn 60 år og ganske få boliger som er over 100 år. Standardavviket ligger i gjennomsnitt 27.53 år unna gjennomsnittet og skyldes nok enkelte høye observasjoner.

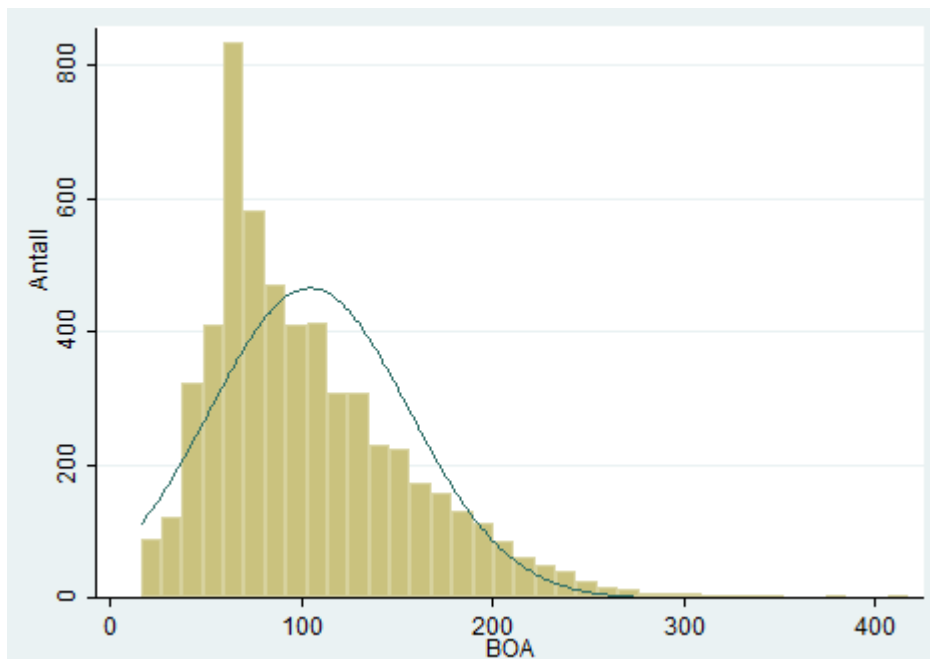


Figur 13: Gjennomsnittlig alder solgte boliger, 2011-2014

En oversikt over gjennomsnittlig alder på boliger i de ulike postnumrene finnes i vedlegg 6. De stedene i utvalget mitt som har størst gjennomsnittlig alder på boligene er Kvadraturen, Grim og Lund. Lavest gjennomsnittsalder på boligene i utvalget mitt er på Ytre Randesund, Indre Randesund og Justvik/Ålefjær.

Boareal

Den minste observasjonen er en liten bolig på 16 kvadratmeter, og den største er en bolig på hele 418 kvadratmeter. Gjennomsnittsverdien er 104,28 og jeg ser at utvalget mitt består av flere mindre boliger enn større boliger, noe som gir meg et litt skjevt utvalg. Men allikevel ligger de fleste av observasjonene rundt gjennomsnittet. Standardavviket forteller oss at i gjennomsnitt ligger observasjonene 52.2 m² fra gjennomsnittet.

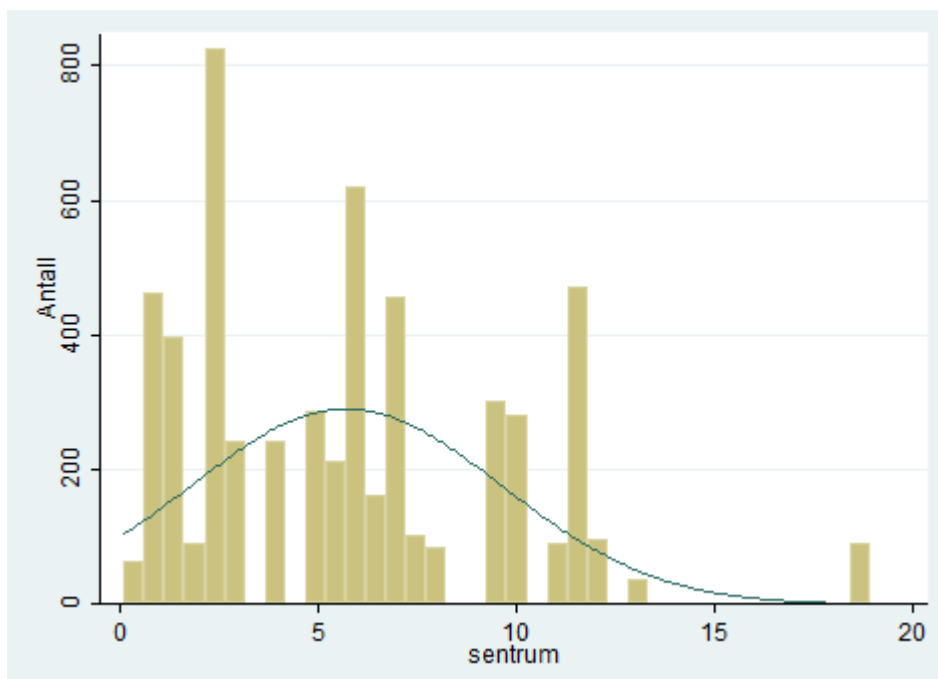


Figur 14: Gjennomsnittlig boareal solgte boliger, 2011-2014

En oversikt over gjennomsnittlig boareal i de ulike postnumrene finnes i vedlegg 7. Gjennomsnittlig boareal i utvalget er minst i postnumre som tilhører Kvadraturen/Eg og Grim, og størst i Indre Randesund, Tveit, Flekkerøya, og Hamresanden.

Avstand til sentrum

Gjennomsnittsavstanden til sentrum er på 5,69 km, hvor minimumsverdien er 0,1 km og maksimumsverdien er 18,9 km. Av histogrammet ser man at det er flere observasjoner som ligger nærmere sentrum enn langt unna. Standardavviket er på 3,91 km unna gjennomsnittet.



Figur 15: Gjennomsnittsavstand til sentrum

En oversikt over avstanden for hver postnummer finnes i vedlegg 8.

Boligtype

Tabellen under viser at halvparten av utvalget mitt består av leiligheter, en fjerdedel består av eneboliger. Videre har jeg kodet enebolig=1, tomannsbolig=2, rekkehus=3 og leilighet=4.

Tabell 12: Solgte boliger etter boligtype, 2011-2014

Boligtype	Antall	Prosent
Enebolig (1)	1400	25%
Tomannsbolig (2)	539	9,6%
Rekkehus (3)	705	12,6%
Leilighet (4)	2965	52,9%
Total	5609	100,0%

I vedlegg 9 finnes det en oversikt over antall boligtyper solgt innenfor de ulike postnumrene.

Eierform

I tabellen under vises fordelingen av de ulike eierformene i utvalget mitt. Da aksjeleilighet og borettslag begge er andelsleiligheter, og det ikke eksisterer vesentlige forskjeller mellom disse

to har jeg valgt å videre behandle aksjeleilighet som borettslag. Selveier er den eierformen som dominerer utvalget mitt.

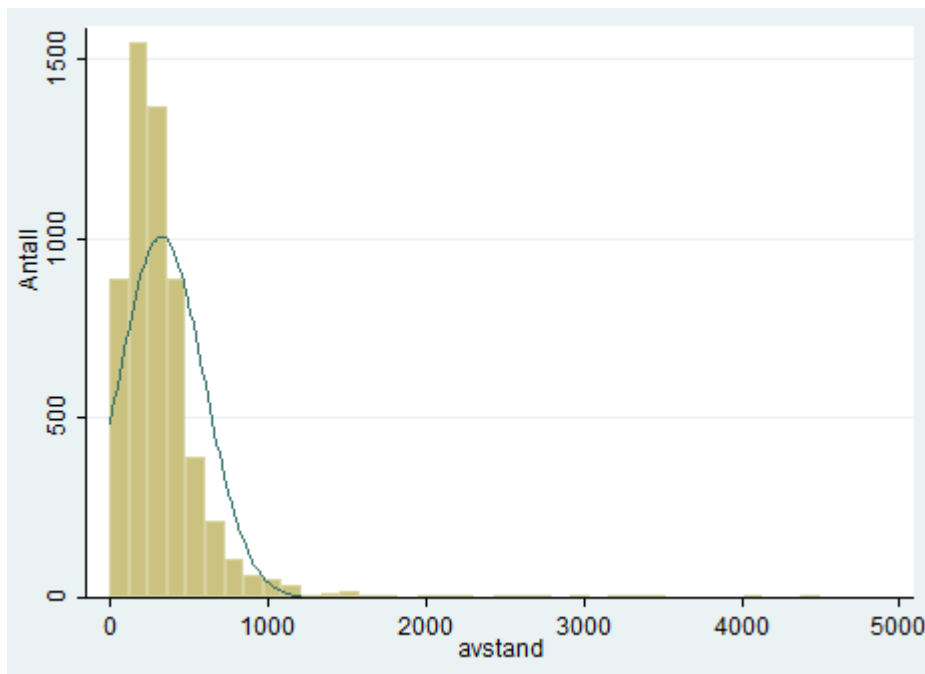
Tabell 13: Solgte boliger etter eierform, 2011-2014

Eierform	Antall	Prosent
Selveier (1)	4005	71,4%
Borettslag (2)	1546	27,6%
Aksjeleilighet (3)	58	1,0%
Totalt	5609	100,0%

En oversikt over eierform i de ulike postnumrene er i vedlegg 10.

Avstand til nærmeste barnehage

Den gjennomsnittlige avstanden til nærmeste barnehage er 325,77 meter, hvor den nærmeste barnehagen ligger i samme bygg som boligen og barnehagen lengst unna ligger 4500 meter fra boligen. Histogrammet viser at jeg har flere boliger som ligger nærme en barnehage enn boliger som ligger langt unna en barnehage. Standardavviket er på 270,85 meter.



Figur 16: Gjennomsnittsavstand til barnehage, 2011-2014

En oversikt over gjennomsnittsavstanden til barnehage i de ulike postnumrene finnes i vedlegg 11.

Kategorisering av avstand til barnehage

For å se om det finnes en forskjell mellom ulike avstander til barnehage og prisen på bolig har jeg kategorisert avstanden. Tabellen under viser kategorisering, omkoding, tilhørende variabler i analysen samt antall og prosent. Som vi ser er det flest observasjoner av boliger som ligger mellom 100 og 1000 meter unna en barnehage. Det er ganske liten andel boliger som ligger innen 50 meter fra en barnehage og enda lavere andel boliger som ligger over 1 km unna.

Tabell 14: Solgte boliger etter avstandskategorier

Grenser	Variabel	Antall observasjoner	Prosent
Under 50 meter (1)	Innenfor50	276	4,9%
Mellom 51-100 meter (2)	Innenfor100	437	7,8%
Mellom 101-500meter (3)	Innenfor500	4042	72,06%
Mellom 501-1000 (4)	Innenfor1000	733	13,07%
Over 1000 meter (5)	Over1000	121	2,16%
Totalt		5609	100,0%

En oversikt over fordelingen av avstandskategorier innenfor de ulike postnumrene finnes i vedlegg 12.

6. Analyse

6.1 Korrelasjon

Perfekt multikollinearitet mellom variabler fører til brudd på forutsetning nr. 6 for OLS. Korrelasjonskoeffisienten, r , måler den lineære sammenhengen mellom to kontinuerlige variabler (Midtbø, 2012). Absoluttverdien til r er mellom +1 og -1 og fortegnet indikerer retningen til korrelasjonen mellom de to variablene. Jo nærmere r er ± 1 , jo sterkere er korrelasjonen. Dersom $r=0$ er to variabler ukorrelerte, og de samvarierer ikke. En positiv korrelasjon betyr at dersom en variabel øker vil den andre også øke. Hvis den ene synker så vil den andre også synke. En negativ korrelasjon betyr at hvis den ene variabelen øker, så synker den andre og omvendt.

I korrelasjonsmatrisen har jeg både med både den kontinuerlige variabelen «avstand til barnehage», og avstandsdummyene. Disse skal dog ikke brukes samtidig som forklaringsvariabler i den videre analysen. Jeg har utelatt postnumrene i korrelasjonsmatrisen som gjengis under på grunn av plassmangel. Fullstendig matrise i vedlegg 13.

Tabell 15: Korrelasjonsmatrise

	prisen	BOA	alder	fgjeld	selveier	borettslag	enebolig	
prisen	1.0000							
BOA	0.7185*	1.0000						
alder	-0.0595*	-0.0287*	1.0000					
fgjeld	-0.3256*	-0.1957*	-0.1651*	1.0000				
selveier	0.4393*	0.3949*	-0.0158	-0.5348*	1.0000			
borettslag	-0.4393*	-0.3949*	0.0158	0.5348*	-1.0000*	1.0000		
enebolig	0.5117*	0.6807*	0.0351*	-0.1994*	0.3568*	-0.3568*	1.0000	
tomannsbolig	0.0366*	0.0986*	0.0133	-0.0938*	0.1863*	-0.1863*	-0.1880*	
rekkehus	-0.0286*	0.0637*	-0.0017	-0.0628*	0.0769*	-0.0769*	-0.2187*	
leilighet	-0.4463*	-0.6906*	-0.0371*	0.2700*	-0.4704*	0.4704*	-0.6107*	
innenfor50	-0.0185	-0.0833*	0.0115	-0.0198	-0.0366*	0.0366*	-0.0588*	
innenfor100	0.0023	-0.0144	0.0447*	0.0003	-0.0118	0.0118	-0.0216	
innenfor500	-0.0177	-0.0667*	0.0101	0.0930*	-0.0766*	0.0766*	-0.0917*	
innenfor1000	0.0416*	0.1307*	-0.0567*	-0.0992*	0.1365*	-0.1365*	0.1394*	
over1000	-0.0187	0.0533*	0.0009	-0.0282*	-0.0038	0.0038	0.0873*	
avstandbhg	0.0423*	0.1685*	-0.0473*	-0.0767*	0.1348*	-0.1348*	0.2007*	
sentrum	0.0981*	0.3495*	-0.3849*	0.0052	0.1380*	-0.1380*	0.3223*	
		tomann~g	rekkehus	leilig~t	innen~50	inne~100	inne~500	inn~1000
tomannsbolig	1.0000							
rekkehus	-0.1236*	1.0000						
leilighet	-0.3453*	-0.4015*	1.0000					
innenfor50	-0.0322*	-0.0266*	0.0877*	1.0000				
innenfor100	-0.0361*	0.0242	0.0240	-0.0661*	1.0000			
innenfor500	0.0197	-0.0396*	0.0942*	-0.3654*	-0.4668*	1.0000		
innenfor1000	0.0279*	0.0540*	-0.1732*	-0.0882*	-0.1127*	-0.6227*	1.0000	
over1000	-0.0109	-0.0082	-0.0638*	-0.0338*	-0.0432*	-0.2385*	-0.0576*	1.0000
avstandbhg	0.0423*	0.0098	-0.2054*	-0.2452*	-0.2635*	-0.2947*	0.4856*	0.0576*
sentrum	0.0329*	0.0717*	-0.3465*	-0.0646*	-0.0425*	-0.1729*	0.2558*	0.0576*
		over1000	avstan~g	sentrum				
over1000	1.0000							
avstandbhg	0.6348*	1.0000						
sentrum	0.1152*	0.2898*	1.0000					

I korrelasjonsmatrisen indikerer stjernene at koeffisienten er signifikant på 5%, det betyr at man med 95% sannsynlighet kan fastslå samvariasjon mellom to variabler. Som vi ser er boareal og pris sterkt positivt korrelerte og betyr at disse to måler mye av det samme. Fellesgjeld og alder er som forventet negativt korrelert med prisen. Det som er overraskende er at koeffisienten til «avstand til barnehage» og «avstand til sentrum» er positivt korrelert med prisen. Jeg skal se nærmere på dette i analysen. Når det gjelder korrelasjonen mellom de forskjellige uavhengige variablene kan man se at borettslag og selveier er perfekt negativt

korrelerte fordi disse er dummyvariabler hvor den ene utelukker den andre. Jeg må derfor utelate den ene variabelen i den videre analysen. Det samme må jeg gjøre med minst ett postnummer og en boligtype. Enebolig har en sterk positiv korrelasjon med boareal, og leilighet en sterk negativ korrelasjon med boareal. Sentrum er negativt korrelert med alder, noe som indikerer at alder synker med avstand til sentrum.

6.2 Valg av funksjonsform

Forutsetningen om linearitet står sentralt i OLS. Har vi en ikke-lineær sammenheng kan feil funksjonsform ha store konsekvenser for koeffisientene i analysen. I dette delkapittelet skal jeg teste en lineær, dobbeltlogaritmisk og semilogaritmisk modell, før jeg velger den som best passer dataene mine.

6.2.1 Multippel lineær regresjon

For å unngå multikollinearitet ble selveier, leilighet, postnummer 4611 og 4612 utelatt av analysen. Jeg har tidligere definert sentrum som Markensgate 25, som ligger i postnummer 4611. Jeg velger i tillegg å utelate 4612 fordi en sentral del av Markensgate ligger i dette postnummeret. Basisboligen min vil derfor være en selveierleilighet i postnummer 4611 og 4612 med boreal lik null.

Da jeg først kjørte regresjonsanalysen hadde jeg med postnummerdummyene (med unntak av 4611 og 4612) i tillegg til avstand til sentrum. Da fikk jeg en forklaringskraft på 70,39%, men ingen av de nevnte variablene ble signifikante på 5%-nivå. Jeg gjorde også en VIF-test (Variance Inflation Factor-test) hvor hver forklaringsvariabel vurderes mot alle de andre forklaringsvariablene samtidig. En tommelfingerregel sier at VIF-verdier over 10 gir multikollinearitetsproblemer. Testen viste høye verdier for variabel «avstand til sentrum» og postnummerdummyene med en gjennomsnittlig VIF på 22388,85 og den høyeste enkeltverdien på 222985,03. Dette betyr problemer med multikollinearitet mellom postnumrene og avstand til sentrum.

Da jeg utelot avstand til sentrum ble alle utenom fire postnummerdummyer signifikante og forklaringskraften økte til 70,40%. På grunnlag av dette har jeg derfor valgt å kun gå videre

med postnummerdummyene da dette også vil fungere som et mål på avstand til sentrum i tillegg til andre lokaliserings-attributter.

Regresjonsanalysen for den lineære modellen vises i tabellen under. Ligningen med de inkluderte variablene gjengis slik:

$$P = \beta_0 + \beta_1 BOA + \beta_2 alder + \beta_3 fgjeld + \beta_4 borettslag + \beta_5 enebolig + \beta_6 rekkehus \\ + \beta_7 tomannsbolig + \beta_8 P4608 + \dots + \beta_{36} P4658 + \beta_{37} avstandbhg$$

Tabell 16: Lineær multipel regresjon

Source	SS	df	MS	Number of obs = 5609		
Model	6.0788e+15	37	1.6429e+14	F(37, 5571) =	361.46	
Residual	2.5322e+15	5571	4.5453e+11	Prob > F	= 0.0000	
				R-squared	= 0.7059	
				Adj R-squared	= 0.7040	
Total	8.6110e+15	5608	1.5355e+12	Root MSE	= 6.7e+05	

prisen	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	17134.14	276.5588	61.95	0.000	16591.98	17676.31
alder	-7034.856	387.3623	-18.16	0.000	-7794.238	-6275.475
fgjeld	-.8532699	.0459033	-18.59	0.000	-.9432583	-.7632815
borettslag	-221228.5	29486.68	-7.50	0.000	-279033.9	-163423.1
enebolig	117385.1	37652.69	3.12	0.002	43571.11	191199
rekkehus	-184587.3	33014.25	-5.59	0.000	-249308.1	-119866.5
tomannsbolig	-135855.6	37388.22	-3.63	0.000	-209151.1	-62560.05
P4608	430036	99902.63	4.30	0.000	234187.9	625884.1
P4610	788777.5	112004.2	7.04	0.000	569205.5	1008349
P4613	-564487.2	151885.7	-3.72	0.000	-862242.4	-266732
P4614	-212652.9	92333.9	-2.30	0.021	-393663.3	-31642.43
P4615	-218346.8	123424.8	-1.77	0.077	-460307.5	23613.87
P4616	-373905.5	97511.28	-3.83	0.000	-565065.6	-182745.4
P4617	-484063.8	103804.1	-4.66	0.000	-687560.2	-280567.3
P4618	-1017894	114016.7	-8.93	0.000	-1241411	-794376.6
P4619	-1309078	112166.1	-11.67	0.000	-1528967	-1089189
P4620	-798238.2	101544.5	-7.86	0.000	-997304.9	-599171.4
P4621	-467027.5	98263.06	-4.75	0.000	-659661.4	-274393.6
P4622	-798620.8	105104.9	-7.60	0.000	-1004667	-592574.2
P4623	-576234.4	100912.3	-5.71	0.000	-774061.9	-378406.9
P4624	-992339.1	95656.51	-10.37	0.000	-1179863	-804815.1
P4625	-455042.3	111481.7	-4.08	0.000	-673589.9	-236494.7
P4626	-1054634	95316.97	-11.06	0.000	-1241493	-867775.8
P4628	-950668	99137.11	-9.59	0.000	-1145015	-756320.6
P4629	-670666.1	96998.4	-6.91	0.000	-860820.8	-480511.4
P4630	127517.7	96820.37	1.32	0.188	-62287.95	317323.4
P4631	288025.6	91847.45	3.14	0.002	107968.7	468082.4
P4632	110843.6	96525.57	1.15	0.251	-78384.17	300071.3
P4633	-151454	95252.04	-1.59	0.112	-338185.2	35277.11
P4634	-639870.5	99844.87	-6.41	0.000	-835605.4	-444135.7
P4635	-621656.1	95338.53	-6.52	0.000	-808556.8	-434755.4
P4637	-453725.6	110478.8	-4.11	0.000	-670307.2	-237144.1
P4638	-631499.1	95633.7	-6.60	0.000	-818978.5	-444019.8
P4639	-306575.4	96970.85	-3.16	0.002	-496676	-116474.7
P4656	-479748.7	140457.7	-3.42	0.001	-755100.7	-204396.8
P4658	-1288636	113629.1	-11.34	0.000	-1511393	-1065879
avstandbhg	-313.5105	38.32291	-8.18	0.000	-388.6383	-238.3826
_cons	1685971	90779.41	18.57	0.000	1508008	1863934

Den lineære modellen har et konstantledd på 1685971 som indikerer prisen på basisboligen² med et boareal på 0 km². Betaen til de kontinuerlige variablene viser hvor mye prisen endrer seg ved en økning i variabelen med én enhet. Denne modellen viser at boligprisen synker med 313,51 kroner per meter avstand fra nærmeste barnehage. Dummyvariablene viser hvor mye prisen vil endre seg dersom de inntreffer. Er det for eksempel en enebolig forventes prisen å øke med 117385 kroner.

VIF-testen viste en gjennomsnittlig VIF på 3,55 og den høyeste verdien lå på 7,95 og tilhørte postnummer 4631. Det er følgelig ingen problemer med multikollinearitet.

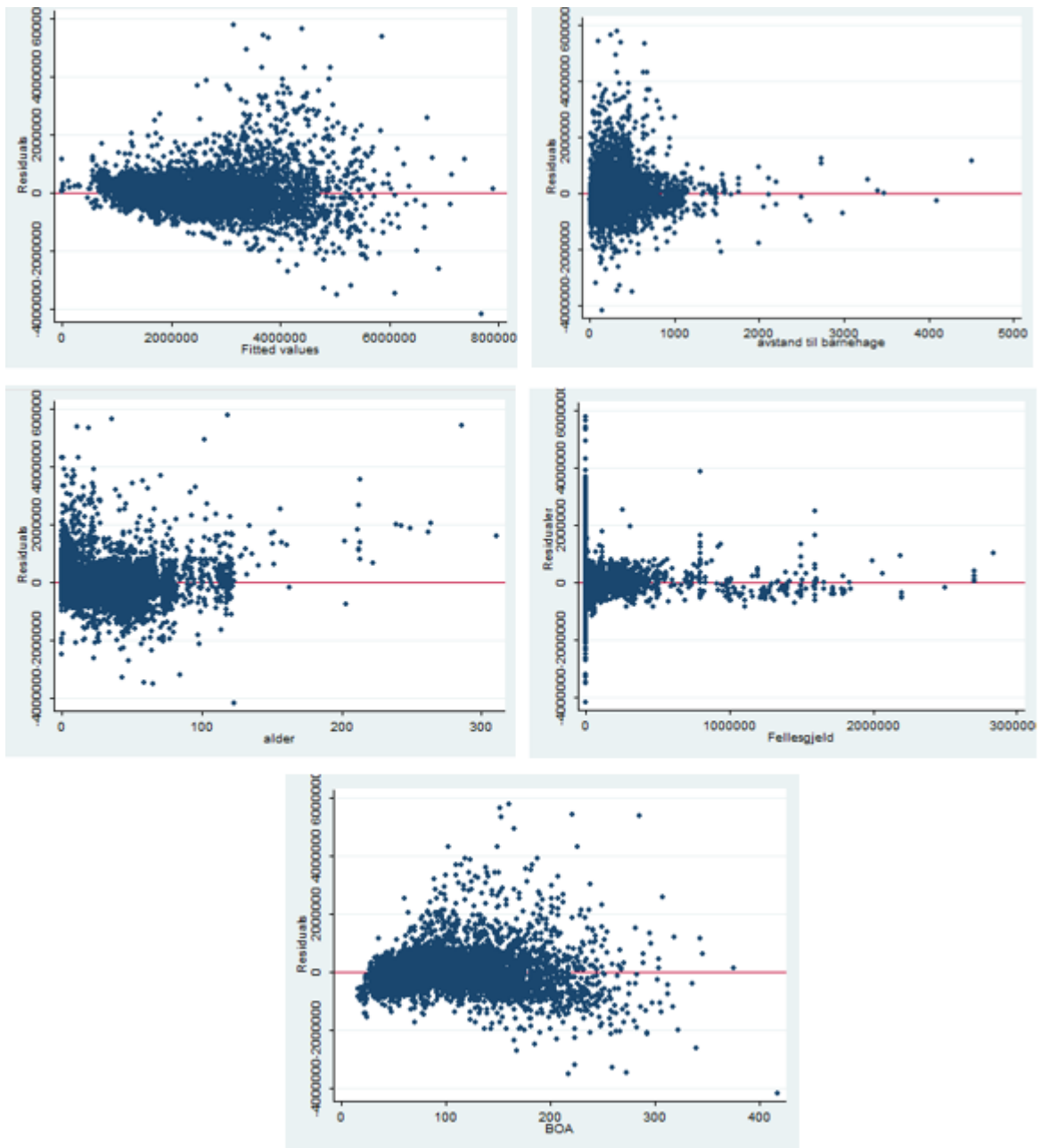
Tabell 17: Variance Inflation Factor (VIF)- test

Variable	VIF	1/VIF
P4631	7.95	0.125853
P4614	6.50	0.153881
P4635	5.70	0.175555
P4633	5.53	0.180987
P4626	5.43	0.184325
P4624	5.39	0.185475
P4638	5.37	0.186189
P4639	5.17	0.193532
P4629	4.81	0.207807
P4632	4.75	0.210676
P4616	4.50	0.222281
P4630	4.45	0.224505
P4634	4.43	0.225560
P4621	4.31	0.231820
P4628	4.15	0.240948
P4623	3.76	0.266215
P4608	3.68	0.271623
P4620	3.57	0.280196
enebolig	3.28	0.305174
P4617	3.10	0.322499
P4622	3.04	0.328951
P4637	2.69	0.371855
P4625	2.61	0.383671
BOA	2.57	0.388838
P4658	2.49	0.401918
P4619	2.48	0.403553
P4610	2.47	0.404720
P4618	2.34	0.427583
borettslag	2.19	0.456444
P4615	1.99	0.502662
fgjeld	1.69	0.592370
P4656	1.64	0.610434
tomannsbolig	1.50	0.667391
rekkehus	1.48	0.676556
P4613	1.46	0.682937
alder	1.40	0.712397
avstandbhg	1.33	0.752262
Mean VIF	3.55	

² Basisboligen her er en selveierleilighet i postnummer 4611 og 4612 med boareal lik null.

Selv om modellen viser en høy forklaringskraft, vil jeg kjøre diagnostikk på den for å se om den holder på regresjonsforutsetningene til OLS om homoskedastisitet, normalfordelte restledd og linearitet.

Heteroskedastisitet er brudd på forutsetning nr.5 som sier at variansen eller spredningen er konstant for hver observasjon av feilleddet (Studenmund, 2006). Dette kan sees ved at residualplottene viser en vifteform. Hvis plottene er jevnt spredt rundt regresjonslinjen har man ikke problemer med heteroskedastisitet. Figuren under viser hele modellen sitt residualplott, samt plottet til de uavhengige variablene fellesgjeld, boareal, alder og avstand til barnehage.



Figur 17: Residualplott, multipl linear regresjon

Residualplottene over viser at modellen og samtlige variabler har heteroskedastisitet. Jeg gjør en Breusch-Pagan test for heteroskedastisitet i Stata som gir følgende resultater:

Tabell 18: Breusch- Pagan test, multipel lineær regresjon

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of prisen	
chi2(1)	= 2876.95
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fgjeld	
chi2(1)	= 104.73
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: BOA	
chi2(1)	= 1842.14
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: alder	
chi2(1)	= 81.98
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: avstandbhg	
chi2(1)	= 4.53
Prob > chi2	= 0.0333

Testen viser at nullhypotesen om lik varians kan forkastes for variablene fellesgjeld, alder, boareal og for hele modellen. I samtlige tilfeller indikeres problemer med heteroskedastisitet. For avstand til barnehage derimot kan nullhypotesen forkastes på 5%-signifikansnivå, men ikke på 1%-signifikansnivå.

Det finnes to former for heteroskedastisitet. Den ene er ren heteroskedastisitet og den andre er sammensatt heteroskedastisitet. Førstnevnte kommer av ulik spredning, men rett funksjonsform. Den sammensatte kommer av ulik spredning i en situasjon der man har feil spesifikasjon av modellen som følge av utelatte variabler eller feil funksjonsform på modellen. Konsekvensene

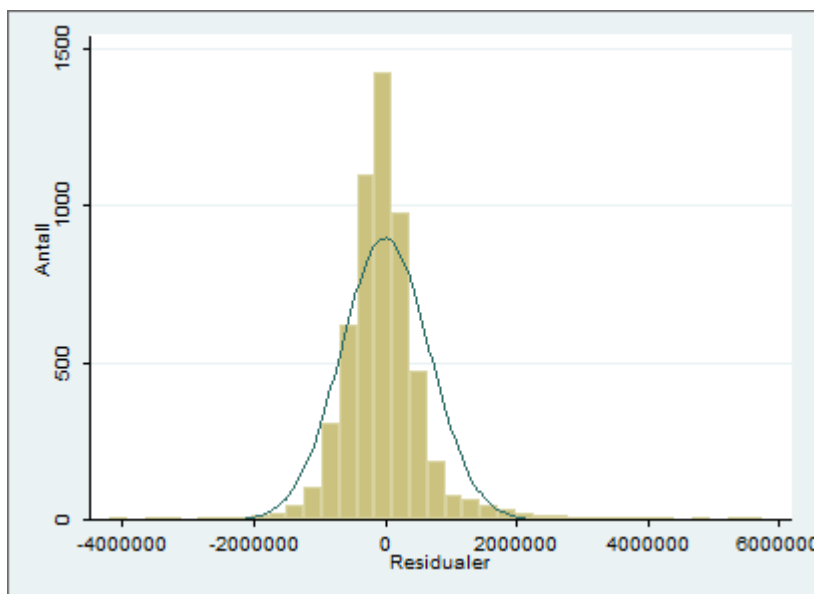
av heteroskedastisitet vil være at variansen ikke minimeres, slik at modellen ikke lenger vil være «BLUE». I tillegg kan estimert standardavvik til koeffisientene bli skjeve, slik at t-verdiene kan bli for høye og P-verdiene for lave slik at man dermed feilaktig forkaster nullhypotesen. Ren heteroskedastisitet vil imidlertid ikke ha noen konsekvenser for beta-estimatene.

Stata gir meg korrigerede (robuste) standardfeil som tar hensyn til at det foreligger heteroskedastisitet uavhengig om det er ren eller sammensatt heteroskedastisitet. De robuste standardfeilene vises i tabellen under. Som man kan se er fortsatt variablene i modellen signifikante. For å spare plass vises ikke postnumrene her, selv om de er med i analysen. Kun tre av postnumrene er ikke-signifikante på 5%-nivå.

Tabell 19: Robuste standardfeil, multipl linear regresjon

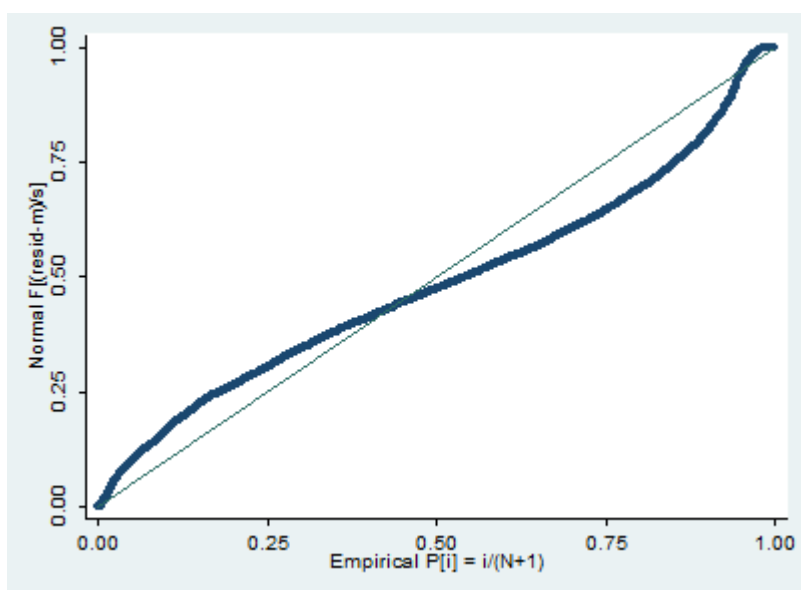
prisen	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
BOA	17134.14	456.5687	37.53	0.000	16239.09	18029.2
alder	-7034.856	764.4701	-9.20	0.000	-8533.516	-5536.197
fgjeld	-.8532699	.03835	-22.25	0.000	-.9284508	-.778089
borettslag	-221228.5	22015.3	-10.05	0.000	-264387	-178069.9
enebolig	117385.1	45826.87	2.56	0.010	27546.51	207223.6
tomannsbolig	-135855.6	35334.23	-3.84	0.000	-205124.4	-66586.68
rekkehus	-184587.3	27530.19	-6.70	0.000	-238557.2	-130617.4
avstandbhg	-313.5105	42.85904	-7.31	0.000	-397.5309	-229.4901
_cons	1685971	96437.6	17.48	0.000	1496916	1875026

I følge OLS-forutsetning nr. 7 skal feilleddene være normalfordelte. Som vi ser av grafen under ser den symmetrisk ut, og har et gjennomsnitt rundt null. Når man har et så stort utvalg som jeg har vil det trolig ikke være et stort problem med å få normalfordeling.



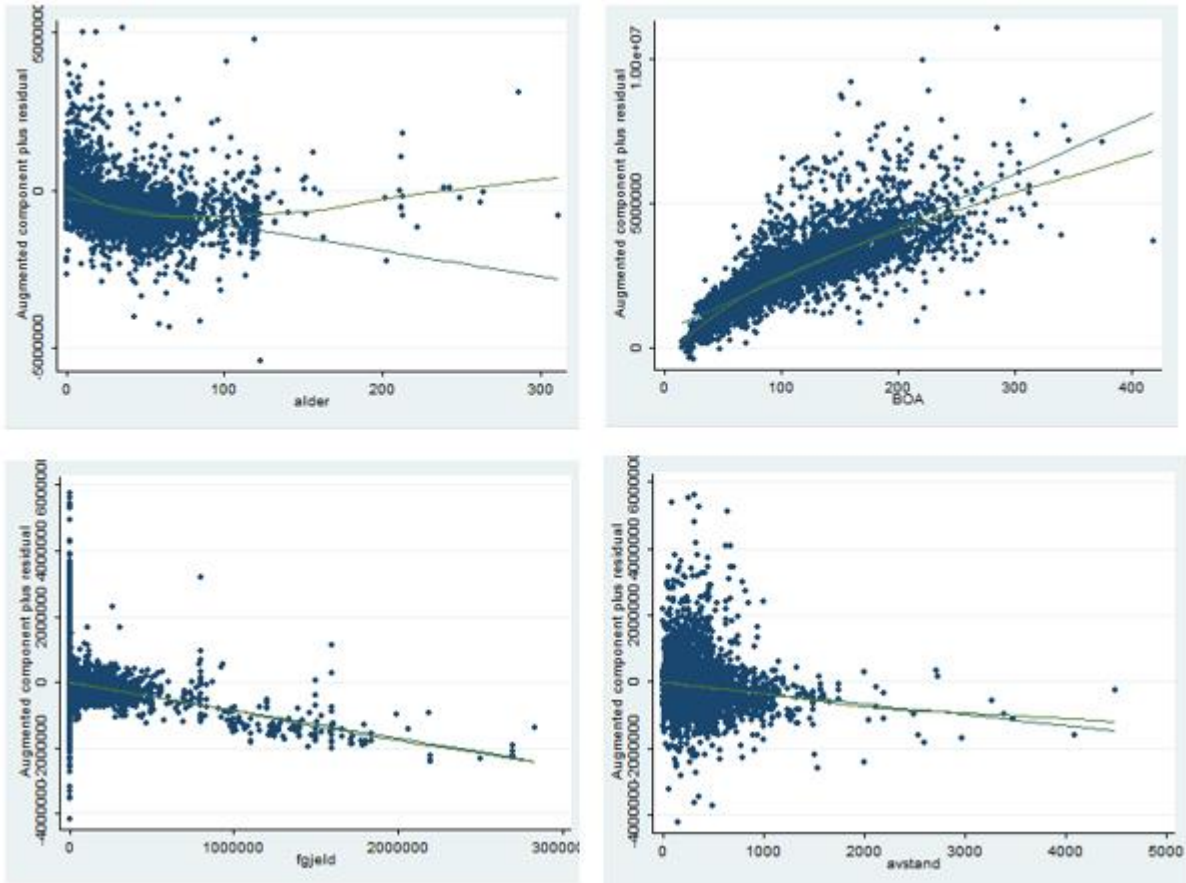
Figur 18: Normalfordeling residualer, multippel lineær regresjon

Figuren under viser at modellen ikke er perfekt normalfordelt, da residualene helst skal ligge nøyaktig på 45-graderslinjen.



Figur 19: Normalskråplott, multippel lineær regresjon

Nå vil jeg se nærmere på om forutsetningen om linearitet holder. Figuren under viser prikkdiagrammatrisene til de kontinuerlige variablene for en lineær modell. Den grønne kurven kalles for «lowess-kurven» og en stor avstand mellom den og den lineære indikerer ikke-linearitet (Midtbø, 2012).



Figur 20: Prikkdiagrammatrise, lineær modell

Kurven for variablene alder og boareal indikerer en tydelig ikke-linearitet. Kurven for fellesgjeld ligger så godt som på den lineære kurven og tyder derfor på korrekt spesifisering for denne variabelen. For «avstand til barnehage» er det en liten avstand mellom kurvene, og det kan derfor indikere en ikke-lineær sammenheng.

6.2.2 Multipel lineær regresjon med avstandsdummyer

Siden jeg er interessert i å se om det finnes en forskjell mellom ulike avstander til barnehage gjør jeg en egen estimering av dette. For å spare plass er ikke postnumrene med de videre fremstillingene, selv om den er med i selve analysen. I denne modellen har jeg tatt ut variabelen «over 1000 meter» for å unngå multikollinearitet, og vil derfor bruke denne som sammenligningsgrunnlag. Basisboligen min her vil altså være en selveierleilighet i postnummer 4611 eller 4612, med boreal lik null og over 1000 meter til nærmeste barnehage. Regresjonsanalysen vises i tabell 20, hvor ligningen til modellen gjengis slik:

$$\begin{aligned}
P = & \beta_0 + \beta_1 BOA + \beta_2 alder + \beta_3 fgjeld + \beta_4 borettslag + \beta_5 enebolig \\
& + \beta_6 tomannsbolig + \beta_7 rekkehus + \beta_8 P4608 + \dots + \beta_{36} P4658 \\
& + \beta_{37} innenfor50 + \beta_{38} innenfor100 + \beta_{39} innenfor500 \\
& + \beta_{40} innenfor1000
\end{aligned}$$

Tabell 20: Multipl linear regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 5609		
Model	6.0785e+15	40	1.5196e+14	F(40, 5568) =	334.10	
Residual	2.5325e+15	5568	4.5484e+11	Prob > F =	0.0000	
Total	8.6110e+15	5608	1.5355e+12	R-squared =	0.7059	
				Adj R-squared =	0.7038	
				Root MSE =	6.7e+05	

prisen	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	17120.74	277.0743	61.79	0.000	16577.57	17663.92
alder	-7114.777	387.6571	-18.35	0.000	-7874.736	-6354.818
fgjeld	-.8677918	.0460155	-18.86	0.000	-.9580001	-.7775835
borettslag	-206920.4	29612.28	-6.99	0.000	-264972	-148868.8
enebolig	115499	37657.11	3.07	0.002	41676.38	189321.6
tomannsbolig	-138856.1	37399.04	-3.71	0.000	-212172.8	-65539.39
rekkehus	-182797	33057.44	-5.53	0.000	-247602.5	-117991.5
innenfor50	442404.8	79031.06	5.60	0.000	287473.1	597336.6
innenfor100	525927.2	73813.97	7.13	0.000	381223	670631.4
innenfor500	446727.5	66829.13	6.68	0.000	315716.3	577738.7
innenfor1000	310854.3	68707.34	4.52	0.000	176161.1	445547.5
_cons	1129495	111779.9	10.10	0.000	910362.3	1348627

Som resultat av at jeg bruker dummyvariabler på avstand til barnehage mister jeg noe forklaringskraft. Denne modellen viser at prisen på en bolig som ligger under 50 meter fra en barnehage er 442404,8 kroner høyere enn prisen på en bolig som ligger over 1000 meter unna en barnehage. For en bolig som ligger mellom 51 og 100 meter unna en barnehage er prisen 525927,2 kroner høyere enn for en bolig som ligger over 1000 meter unna en barnehage. Deretter ser man at en prisen på en bolig som ligger mellom 101 og 500 meter unna en barnehage er 446727,5 kroner høyere enn prisen for en bolig som ligger over 1000 meter unna en barnehage. Til er prisen 310854,3 kroner høyere for en bolig som ligger mellom 501 og 1000 meter unna en barnehage enn prisen for bolig som har over 1000 meter avstand til barnehage.

6.2.3 Dobbellogaritmisk regresjon

I den dobbeltlogaritmiske funksjonsformen tar jeg ln til prisen og de uavhengige variablene utenom dummy-variablene. Dobbellogaritmisk funksjonsform kan bare bli brukt når variablene har positive verdier. Derfor vil ikke dummy-variabler bli logaritmisk omkodet, da disse kun har verdi 0 og 1. For kontinuerlige variabler som inneholder verdien null, slik som fellesgjeld, alder og avstand til barnehage, har jeg lagt til verdien 1. Den dobbeltlogaritmiske modellen vises i tabellen under, hvor regresjonslinjen gjengis slik:

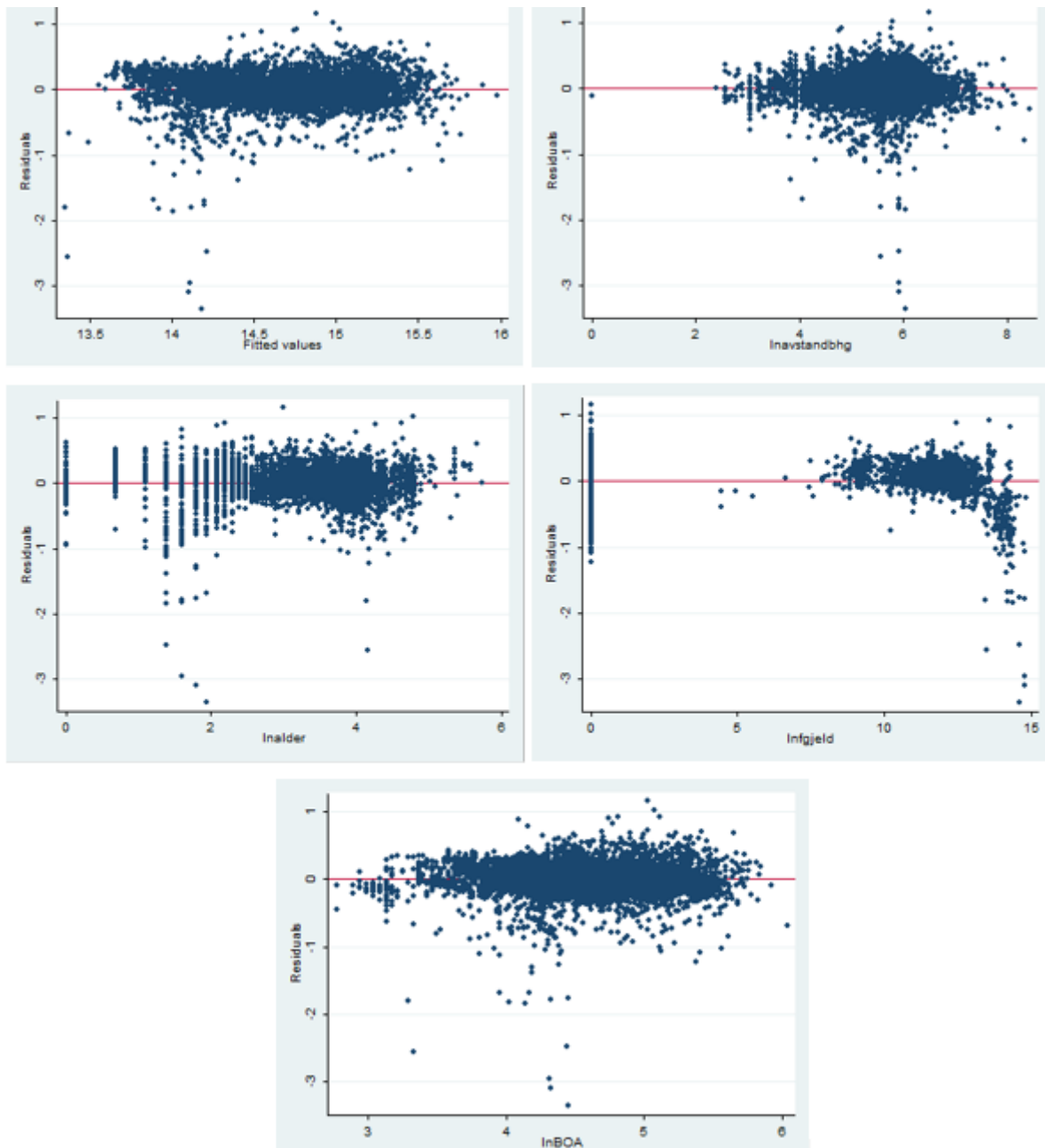
$$\begin{aligned} \ln P = & \beta_0 + \beta_1 \ln BOA + \beta_2 \ln alder + \beta_3 \ln fgjeld + \beta_4 \text{borettslag} + \beta_5 \text{enebolig} \\ & + \beta_6 \text{tomannsbolig} + \beta_7 \text{rekkehus} + \beta_8 P4608 + \dots + \beta_{36} P4658 \\ & + \beta_{37} \ln avstandbhg \end{aligned}$$

Tabell 21: Dobbellogaritmisk regresjon

Source	SS	df	MS	Number of obs = 5609		
Model	918.502894	37	24.8244025	F(37, 5571) = 411.50		
Residual	336.080304	5571	.060326746	Prob > F = 0.0000		
Total	1254.5832	5608	.223713124	R-squared = 0.7321		
				Adj R-squared = 0.7303		
				Root MSE = .24562		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6813301	.0103732	65.68	0.000	.6609946	.7016656
lnalder	-.0619838	.0040789	-15.20	0.000	-.0699801	-.0539875
lnfgjeld	-.0215539	.0013574	-15.88	0.000	-.024215	-.0188929
borettslag	-.0425678	.01651	-2.58	0.010	-.0749338	-.0102018
enebolig	.0253083	.0134451	1.88	0.060	-.0010493	.0516659
tomannsbolig	-.0758699	.0139585	-5.44	0.000	-.1032339	-.0485058
rekkehus	-.0881616	.0124409	-7.09	0.000	-.1125506	-.0637726
lnavstandbhg	-.0300941	.0048331	-6.23	0.000	-.0395688	-.0206193
_cons	12.07163	.0602539	200.35	0.000	11.95351	12.18975

Den dobbeltlogaritmiske modellen viser at dersom avstand til barnehage øker med 1%, så synker prisen med 3%. Denne modellen forklarer 73,03% av variasjonen i prisen, men jeg kan ikke utelukkende vurdere om den dobbeltlogaritmiske modellen er bedre enn den lineære fordi den avhengige variabelen er forskjellige i de to modellene. Jeg skal også her se på restleddet, teste for heteroskedastisitet og undersøke linearitetsforutsetningene.

I residualplottet for modellen nedenfor ser man at observasjonene er ganske tett samlet langs linjen, men at det er noen avvikende plott spesielt for lave predikerte verdier.



Figur 21: Residualplott, dobbeltlogaritmisk regresjon

For variabelen fellesgjeld ser det ut som at variansen i residuale øker etter hvert som de predikerte verdiene blir større. For de andre variablene er det ikke så lett å se et mønster, men en Breusch-Pagan-test viser at jeg kan forkaste nullhypotesen om konstant varians for variablene lnpris, lnalder, lnBOA og lnfgjeld. Det vil si at det er heteroskedastisitet i alle disse variablene. For den naturlige logaritmen av avstand til barnehage derimot er homoskedastisk på 1% signifikansnivå.

Tabell 22: Breusch-Pagan test, dobbeltlogaritmisk regresjon

```
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: lnnavstandbhg

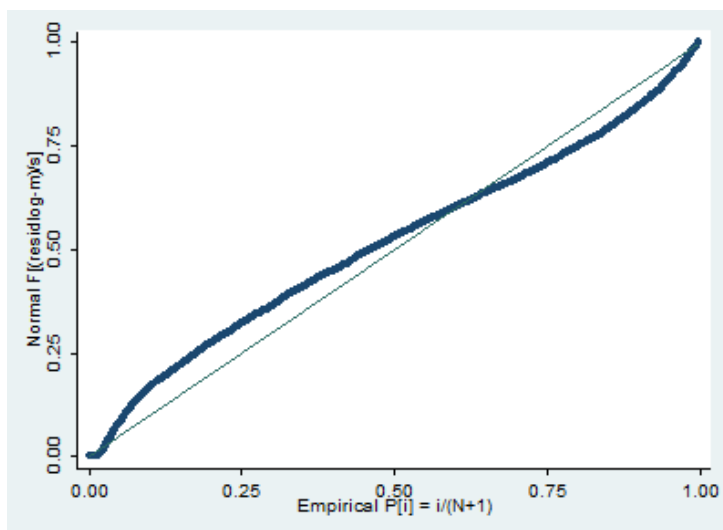
chi2(1)      =      6.31
Prob > chi2  =      0.0120
```

Siden modellen indikerer problemer med heteroskedastisitet kjører også en regresjon som gir robuste standardfeil på den dobbeltlogaritmiske modellen som vist i tabellen under. Alle variablene vist her er signifikante på 5%-nivå. Av postnumrene er det åtte som ikke er signifikante.

Tabell 23: Robuste standardfeil, dobbeltlogaritmisk regresjon

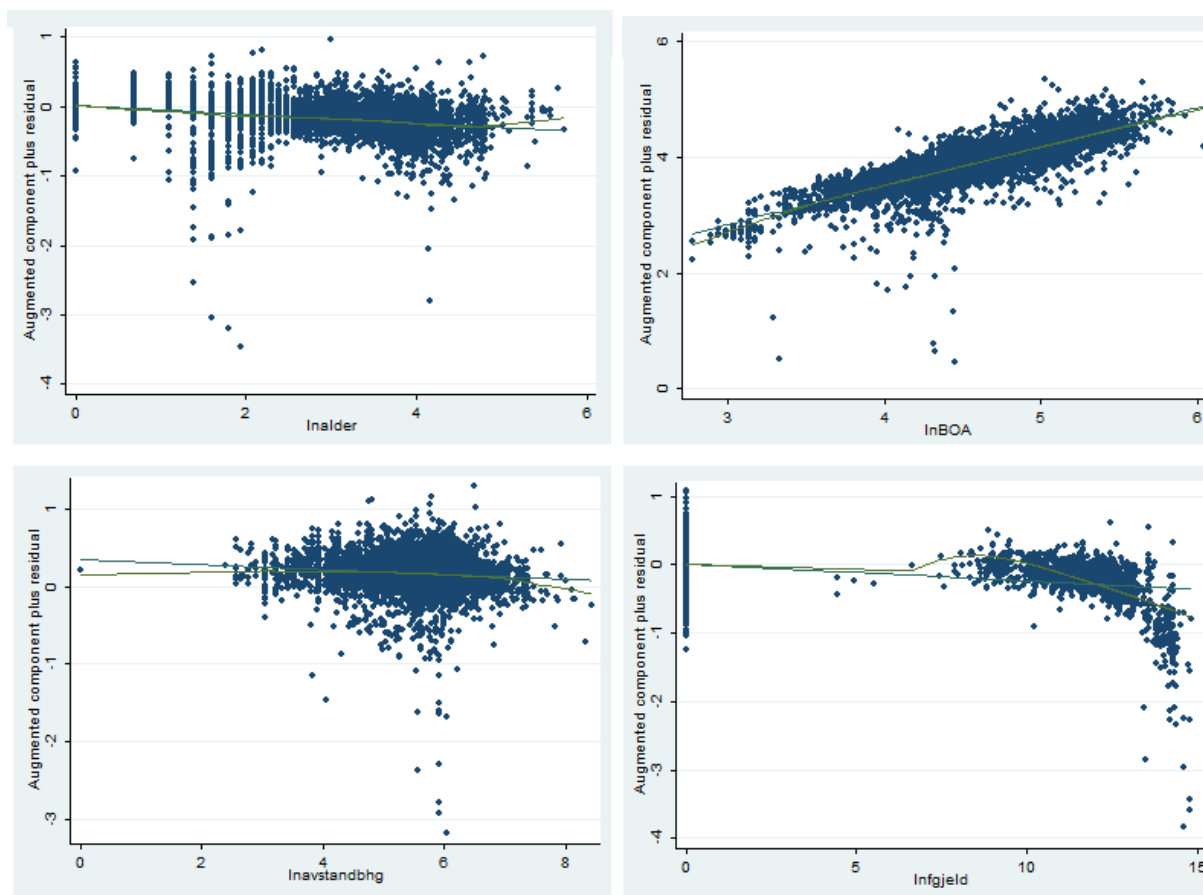
lnpris	Robust					[95% Conf. Interval]	
	Coef.	Std. Err.	t	P> t			
lnBOA	.6813301	.0113594	59.98	0.000	.6590611	.703599	
lnalder	-.0619838	.0048853	-12.69	0.000	-.0715608	-.0524068	
lnfgjeld	-.0215539	.0014437	-14.93	0.000	-.0243841	-.0187238	
borettslag	-.0425678	.0135099	-3.15	0.002	-.0690524	-.0160831	
enebolig	.0253083	.0128512	1.97	0.049	.0001149	.0505017	
tomannsbolig	-.0758699	.0118069	-6.43	0.000	-.099016	-.0527237	
rekkehus	-.0881616	.0103515	-8.52	0.000	-.1084545	-.0678687	
lnnavstandbhg	-.0300941	.0045971	-6.55	0.000	-.0391062	-.0210819	
_cons	12.07163	.0634927	190.13	0.000	11.94716	12.1961	

Forutsetningen om normalfordelte restledd ser ut til å være bedre ivaretatt i denne modellen enn i den lineære, da vi kan se av grafen under at kurven ligger noe nærmere den lineære linjen.



Figur 22: Normalskråplott, dobbeltlogaritmisk regresjon

Figuren nedenfor viser lowess-kurven til variablene i den dobbeltlogaritmiske modellen. Det kan se ut som at linearitetsforutsetningen er dårligere ivaretatt her enn i den lineære modellen for avstand til barnehage og fellesgjeld. For boareal og alder derimot vil denne funksjonsformen passe bedre.



Figur 23: Prikkdiagrammatrise, dobbeltlogaritmisk modell

6.2.4 Dobbellogaritmisk regresjon med avstandsdummyer

Regresjonsmodellen nedenfor viser den dobbeltlogaritmiske funksjonsformen med avstand delt inn i kategorier. Modellen forklarer 73% av variasjonen i prisen. Regresjonsanalysen vises under, hvor regresjonsligningen gjengis slik:

$$\begin{aligned} \ln P = & \beta_0 + \beta_1 \ln BOA + \beta_2 \ln alder + \beta_3 \ln fgjeld + \beta_4 \text{borettslag} + \beta_5 \text{enebolig} \\ & + \beta_6 \text{tomannsbolig} + \beta_7 \text{rekkehus} + \beta_8 P4608 + \dots + \beta_{36} P468 \\ & + \beta_{37} \text{innenfor50} + \beta_{38} \text{innenfor100} + \beta_{39} \text{innenfor500} \\ & + \beta_{40} \text{innenfor1000} \end{aligned}$$

Tabell 24: Dobbellogaritmisk regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 5609		
Model	918.201782	40	22.9550445	F(40, 5568) =	379.97	
Residual	336.381416	5568	.060413329	Prob > F =	0.0000	
Total	1254.5832	5608	.223713124	R-squared =	0.7319	
				Adj R-squared =	0.7300	
				Root MSE =	.24579	

lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6785775	.0104042	65.22	0.000	.6581813	.6989737
lnalder	-.061286	.0040946	-14.97	0.000	-.0693129	-.0532591
lnfgjeld	-.022026	.001361	-16.18	0.000	-.0246942	-.0193579
borettslag	-.0348055	.0165651	-2.10	0.036	-.0672796	-.0023313
enebolig	.0262101	.0134735	1.95	0.052	-.0002032	.0526234
tomannsbolig	-.0769046	.0139683	-5.51	0.000	-.104288	-.0495212
rekkehus	-.0878822	.0124606	-7.05	0.000	-.1123099	-.0634546
innenfor50	.1309923	.0288724	4.54	0.000	.0743911	.1875936
innenfor100	.1338311	.0269489	4.97	0.000	.0810007	.1866614
innenfor500	.1058237	.0244007	4.34	0.000	.0579888	.1536587
innenfor1000	.0733339	.0250672	2.93	0.003	.0241924	.1224754
_cons	11.79622	.0600005	196.60	0.000	11.67859	11.91384

Den dobbeltlogaritmiske modellen viser at hvis man befinner seg innenfor 50 meter fra en barnehage, så vil prisen på boligen være 13,09% høyere enn hvis man hadde vært over 1000 meter fra en barnehage. Hvis boligen befinner seg mellom 51 og 100 meter unna en barnehage vil prisen være 13,38% høyere enn hvis man hadde vært over 1000 meter fra en barnehage. Befinner boligen seg mellom 101 og 500 meter unna en barnehage vil prisen være 10,58% høyere enn hvis den hadde ligget over 1000 meter unna en barnehage. Til slutt vil prisen på en bolig som ligger mellom 501 og 1000 meter unna en barnehage være 7,33% lavere enn en bolig som ligger over 1000 meter unna.

6.2.5 Semilogaritmisk regresjon

I den semilogaritmiske regresjonsmodellen tar jeg kun den naturlige logaritmen av prisen, og lar de uavhengige variablene være uendret. Jeg får da en forklaringskraft på 73,90% som er noe høyere sammenlignet med forklaringskraften til den dobbeltlogaritmiske med avstand som kontinuerlig variabel. Regresjonsanalysen vises i tabell 25, hvor den tilhørende ligningen gjengis slik:

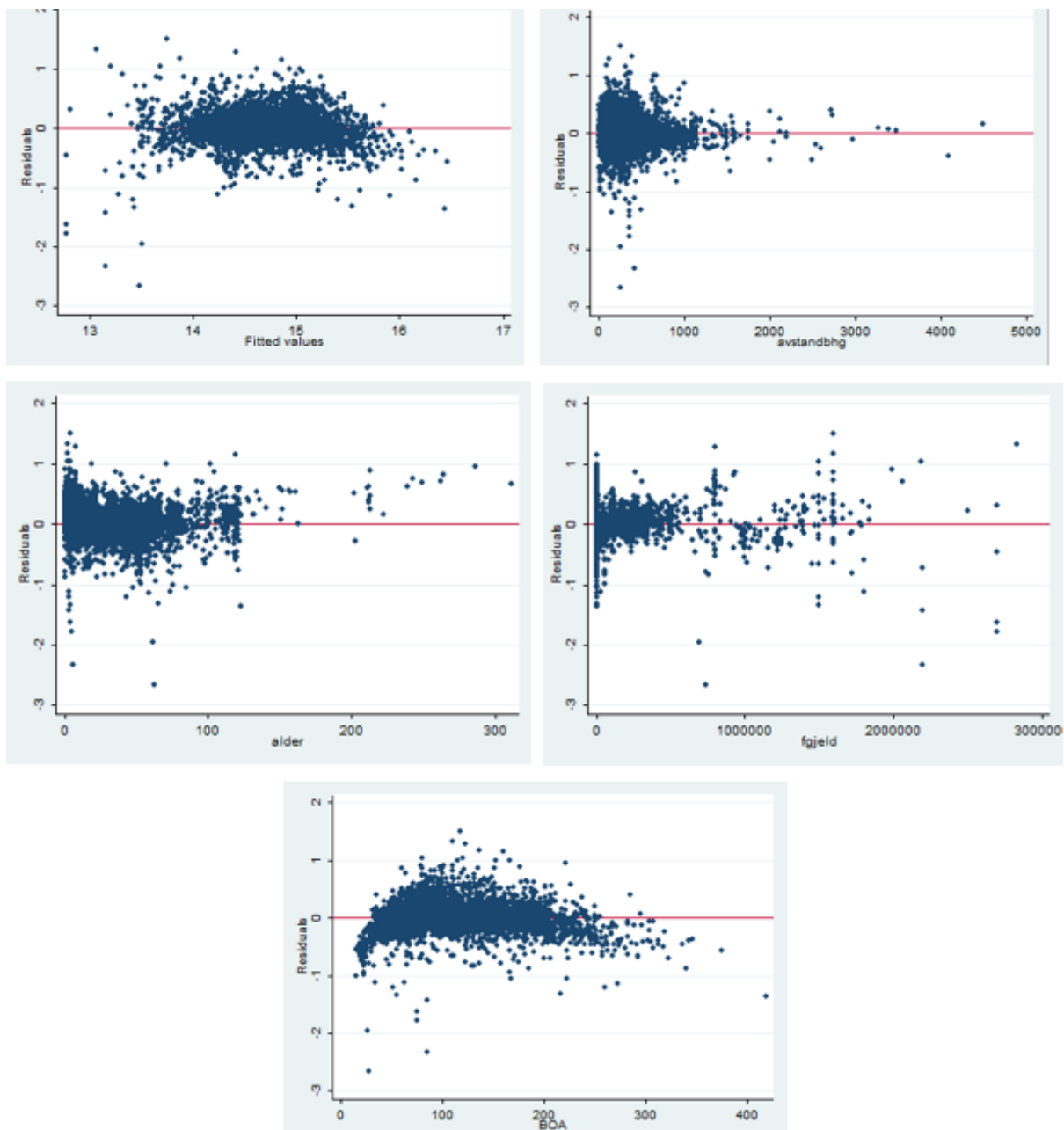
$$\begin{aligned} \ln P = & \beta_0 + \beta_1 BOA + \beta_2 alder + \beta_3 fgjeld + \beta_4 borettslag + \beta_5 enebolig \\ & + \beta_6 tomannsbolig + \beta_7 rekkehus + \beta_8 P4608 + \dots \beta_{36} P458 \\ & + \beta_{37} avstandbhg \end{aligned}$$

Tabell 25: Semilogaritmisk regresjon

Source	SS	df	MS	Number of obs = 5609		
Model	929.341212	37	25.11733	F(37, 5571) = 430.23		
Residual	325.241986	5571	.058381258	Prob > F = 0.0000		
Total	1254.5832	5608	.223713124	R-squared = 0.7408		
				Adj R-squared = 0.7390		
				Root MSE = .24162		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0059018	.0000991	59.54	0.000	.0057075	.0060961
alder	-.0024415	.0001388	-17.59	0.000	-.0027137	-.0021694
fgjeld	-6.49e-07	1.65e-08	-39.43	0.000	-6.81e-07	-6.16e-07
borettslag	-.0678226	.0105677	-6.42	0.000	-.0885395	-.0471057
enebolig	.0511522	.0134944	3.79	0.000	.024698	.0776064
tomannsbolig	.0101311	.0133996	0.76	0.450	-.0161373	.0363995
rekkehus	-.0064697	.011832	-0.55	0.585	-.029665	.0167256
avstandbhg	-.0001093	.0000137	-7.96	0.000	-.0001362	-.0000824
_cons	14.38485	.0325345	442.14	0.000	14.32107	14.44863

Modellen viser at hvis avstand til barnehage øker med en meter, så vil prisen synke med 0,01093%.

Figuren nedenfor viser residualplottet for modellen og for alle de kontinuerlige variablene. Det tyder på heteroskedastisitet i modellen da variasjonen i residualene blir trangere ved økte predikerte verdier. Det samme gjelder for variablene avstand til barnehage, alder og boareal. For fellesgjeld øker variasjonen i residualene med økte predikerte verdier.



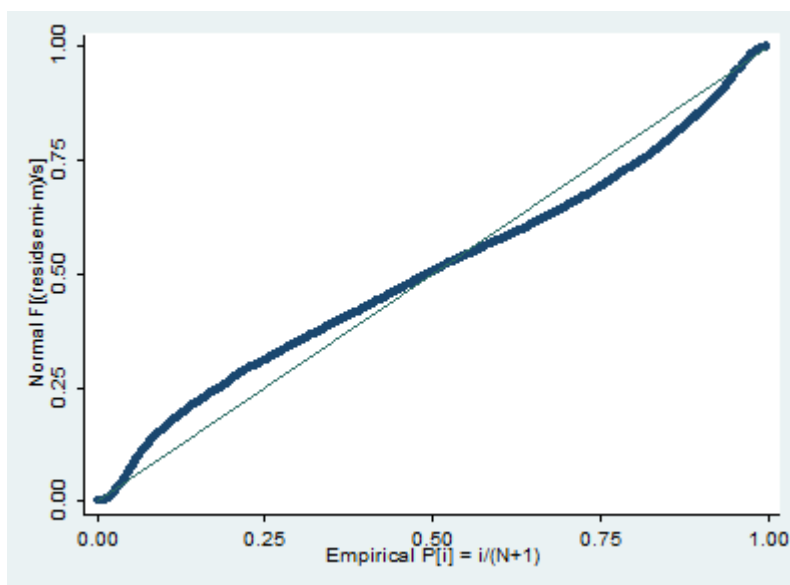
Figur 24: Residualplott, semilogaritmisk regresjon

En BP-test viser at jeg har rett i mine antakelser; ingen av variablene har konstant varians i residuale. Jeg kjører en regresjonsanalyse med robuste standardfeil for å ta hensyn til heteroskedastisitet i den semilogaritmiske modellen, og får ingen store endringer i p-verdi her heller. Seks av postnumrene er ikke signifikante på 5%-nivå.

Tabell 26: Robuste standardfeil, semilogaritmisk regresjon

lnpris	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
BOA	.0059018	.0001432	41.21	0.000	.005621	.0061825
alder	-.0024415	.0002288	-10.67	0.000	-.0028901	-.0019929
fgjeld	-6.49e-07	4.38e-08	-14.83	0.000	-7.34e-07	-5.63e-07
borettslag	-.0678226	.0125973	-5.38	0.000	-.0925181	-.043127
enebolig	.0511522	.014443	3.54	0.000	.0228382	.0794661
tomannsbolig	.0101311	.011697	0.87	0.386	-.0127996	.0330618
rekkehus	-.0064697	.0095526	-0.68	0.498	-.0251965	.0122571
avstandbhg	-.0001093	.0000135	-8.07	0.000	-.0001358	-.0000827
_cons	14.38485	.0432641	332.49	0.000	14.30004	14.46967

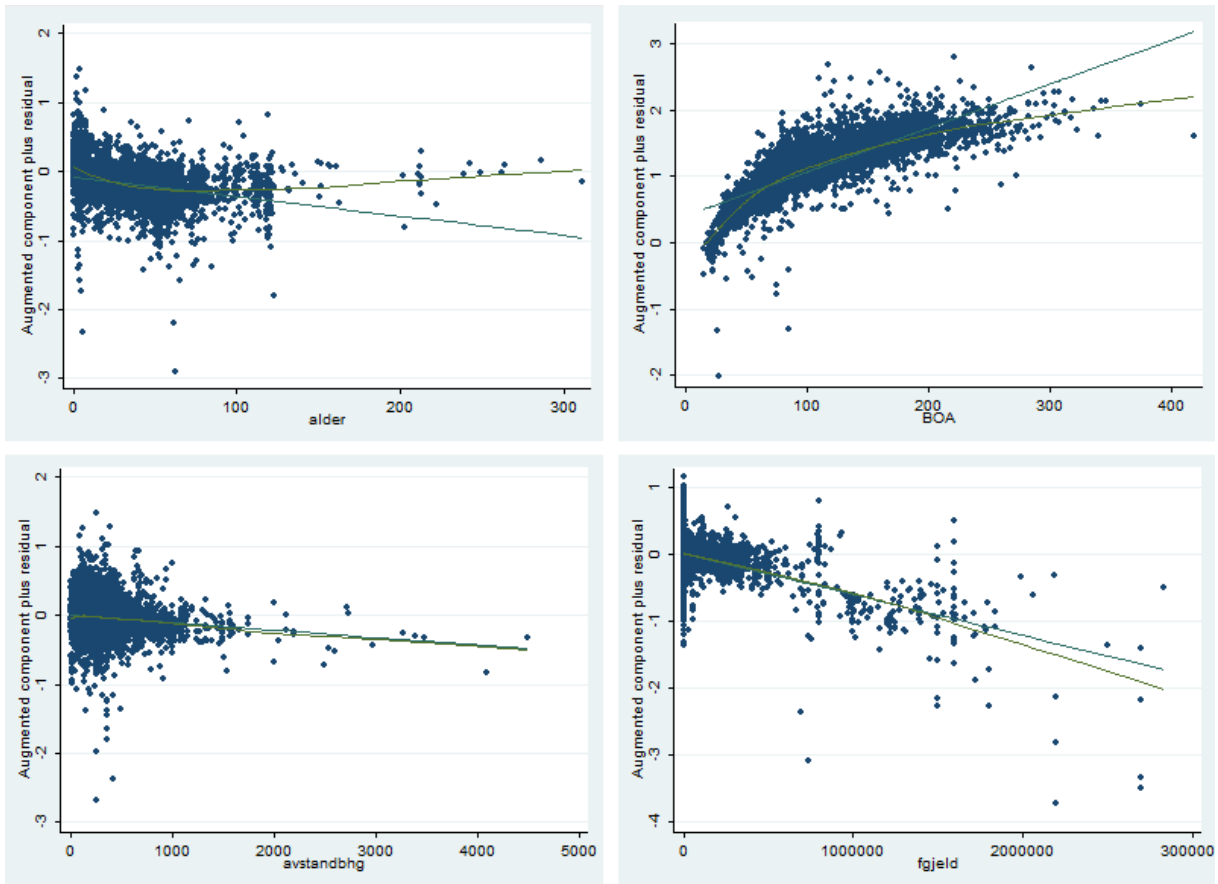
Figuren under viser at restleddene er tilnærmet normalfordelt da de predikerte verdiene ligger nær den lineære 45-graders linjen. Avvikene over linjen er omtrent like store som avvikene over og skjæringspunktet ligger ca. på 0,5. Det vil følgelig være naturlig å anta normalfordeling her.



Figur 25: Normalskråplott, semilogaritmisk regresjon

I figuren nedenfor vises lowess-kurvene for variablene i den semilogaritmiske modellen. Siden det er stor avstand mellom den lineære kurven og den grønne lowess-kurven for boareal, fellesgjeld og alder vil ikke denne modellen være den beste for disse variablene. For avstand til barnehage derimot gir denne funksjonsformen den beste spesifiseringen av alle

modellene jeg har testet.



Figur 26: Prikkdiagrammatrise, semilogaritmisk modell

6.2.6 Semilogaritmisk regresjon med avstandsdummyer

Denne modellen viser at de uavhengige variablene forklarer 73,88% av variasjonen i prisen. Dette er noe bedre enn den dobbeltlogaritmiske modellen som viste en forklaringskraft på 73% med kategorisering av avstanden. Regresjonsanalysen vises i tabellen under, hvor regresjonsligningen ser slik ut:

$$\begin{aligned} \ln P = & \beta_0 + \beta_1 BOA + \beta_2 alder + \beta_3 fgjeld + \beta_4 borettslag + \beta_5 enebolig \\ & + \beta_6 tomannsbolig + \beta_7 rekkehus + \beta_8 P4608 + \dots + \beta_{36} \\ & + \beta_{37} innenfor50 + \beta_{38} innenfor100 + \beta_{39} innenfor500 \\ & + \beta_{40} innenfor1000 \end{aligned}$$

Tabell 27: Semilogaritmisk regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 5609		
Model	929.227622	40	23.2306906	F(40, 5568) = 397.56		
Residual	325.355575	5568	.058433113	Prob > F = 0.0000		
Total	1254.5832	5608	.223713124	R-squared = 0.7407		
				Adj R-squared = 0.7388		
				Root MSE = .24173		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0058883	.0000993	59.29	0.000	.0056936	.006083
alder	-.0024672	.0001389	-17.76	0.000	-.0027396	-.0021948
fgjeld	-6.55e-07	1.65e-08	-39.70	0.000	-6.87e-07	-6.22e-07
borettslag	-.0622835	.0106138	-5.87	0.000	-.0830908	-.0414762
enebolig	.0506023	.0134973	3.75	0.000	.0241423	.0770624
tomannsbolig	.0086997	.0134048	0.65	0.516	-.017579	.0349783
rekkehus	-.0059729	.0118487	-0.50	0.614	-.029201	.0172551
innenfor50	.1272144	.0283269	4.49	0.000	.0716826	.1827461
innenfor100	.172622	.0264569	6.52	0.000	.1207561	.2244879
innenfor500	.1569905	.0239534	6.55	0.000	.1100326	.2039485
innenfor1000	.1095975	.0246266	4.45	0.000	.0613198	.1578752
_cons	14.19022	.040065	354.18	0.000	14.11168	14.26876

Man kan se av regresjonen at for en bolig innenfor 50 meter fra en barnehage er prisen 12,72% høyere enn for en bolig 1000 meter unna. For en bolig mellom 51 og 100 meter unna en barnehage er prisen 17,26% høyere enn en for bolig over 1000 meter unna. Deretter vil prisen på en bolig som ligger mellom 101-500 meter unna en barnehage være 15,69% høyere enn en bolig som ligger over 1000 meter unna en barnehage. Til slutt vil en bolig som ligger mellom 501 og 1000 meter unna en barnehage ha en pris som er 10,95% høyere enn en bolig som er over 1000 meter unna.

6.2.7 Valg av modell

Når modell skal velges gjøres dette på bakgrunn av egnethet med datasettet, restleddets forutsetninger og forklaringskraften. Jeg har tidligere antatt ikke-linearitet for variabelen alder, boareal og avstand til barnehage. Variabelen jeg er interessert i å undersøke er først og fremst avstand til barnehage. Prikkdiagrammatrisen i figur 20 viste en svak konveks lowess- kurve for avstand til barnehage. Dette kan tale for en ikke-lineær sammenheng og jeg velger derfor å gå videre med en logaritmisk funksjonsform. Den semilogaritmiske regresjonsmodellen fikk en høyere forklaringskraft enn den dobbeltlogaritmiske, samtidig som den har bedre forutsetninger for normalfordelte restledd. Prikkdiagrammatrisen for

«avstand til barnehage» i den semilogaritmiske modellen viste også at forutsetningen om linearitet er best ivaretatt her. På bakgrunn av dette velger jeg å gå videre med den semilogaritmiske modellen i hypotesetestingen. Fullstendig fremstilling av de semilogaritmiske regresjonsmodellene finnes i vedlegg 14 og 15. De samme modellene med robuste standardfeil finnes i vedlegg 16 og 17.

6.3 Testing av hypotesene

Ved testing av hypotesene velger jeg å bruke P-verdien som angir det laveste nivået nullhypotesen kan forkastes ved. Når jeg tolker resultatene går jeg ut fra den semilogaritmiske regresjonsmodellen med avstand som kontinuerlig variabel bortsett fra i hypotese 2 hvor jeg tester avstandsdummyene.

Hypotese 1: Avstand til barnehage

$H_0 =$ Det er ikke en negativ sammenheng mellom avstand til barnehage og boligpris

$H_A =$ Det er en negativ sammenheng mellom avstand til barnehage og boligpris

Koeffisienten til variabelen «avstandbhg» forteller oss at boligprisene synker med 0,01093% for hver meter avstand fra barnehage. Siden koeffisienten har en P-verdi på 0,000 kan jeg med 99% sikkerhet konkludere med at det er en negativ sammenheng mellom avstand til barnehage og boligpris, og forkaster dermed nullhypotesen.

Hypotese 2: Avstand til barnehage- dummy

$H_0 =$ Prisen på en bolig som ligger veldig nærme en barnehage vil være høyere enn en bolig som ligger litt lengere unna

$H_A =$ Prisen på en bolig som ligger veldig nærme en barnehage vil være lavere enn en bolig som ligger litt lenger unna

Koeffisienten til boliger innenfor 50 meter fra nærmeste barnehage har en P-verdi på 0,000. Variabelen er derfor signifikant på et 1%-nivå og jeg kan med 99% sikkerhet fastslå at nærhet til barnehage innenfor den gitte grensen har betydning for boligprisen. Modellen viser at en

bolig innenfor 50 meter fra nærmeste barnehage vil ha en pris som er 12,72% høyere enn en bolig som ligger over 1000 meter fra nærmeste barnehage.

Koeffisienten til boliger mellom 51-100 meter unna nærmeste barnehage har en P-verdi på 0,000. Variabelen er derfor signifikant på et 1%-nivå og jeg kan med 99% sikkerhet fastslå at nærhet til barnehage innenfor den gitte grensen har betydning for boligprisen. Modellen viser at en bolig mellom 51-100 meter fra nærmeste barnehage vil ha en pris som er 17,26% høyere enn en bolig som ligger over 1000 meter fra nærmeste barnehage.

Koeffisienten til boliger mellom 101-500 meter unna nærmeste barnehage har en P-verdi på 0,000. Variabelen er derfor signifikant, og jeg kan med 99% sikkerhet konkludere med at nærhet til barnehage innenfor den gitte grensen har betydning for boligprisen. Modellen viser at en bolig mellom 101-500 meter fra nærmeste barnehage vil ha en pris som er 15,69% høyere enn en bolig som ligger over 1000 meter fra nærmeste barnehage.

Koeffisienten til boliger mellom 501-1000 meter unna nærmeste barnehage har en P-verdi på 0,000. Variabelen er derfor signifikant på et 1%-nivå og jeg kan med 99% sikkerhet fastslå at nærhet til barnehage innenfor den gitte grensen har betydning for boligprisen. Modellen viser at en bolig mellom 501-1000 meter fra nærmeste barnehage vil ha en pris som er 10,95% høyere enn en bolig som ligger over 1000 meter fra nærmeste barnehage.

Koeffisientene til innenfor 50 meter, 51-100 meter 101-500 og 501-1000 meter er henholdsvis 12,72%, 17,26% og 15,69% og 10,95%. Økningen i betaverdien fra innenfor 50 meter til 51-100 meter fører til at jeg kan forkaste nullhypotesen og konkludere med at prise på bolig vil være lavere innenfor 50 meter fra en barnehage enn pris på bolig som ligger mellom 51-100 meter og 101-500 meter unna en barnehage.

Hypotese 3: Avstand til sentrum

H_0 = Det er ikke en sammenheng mellom avstand til sentrum og boligpris

H_A = Det er en sammenheng mellom avstand til sentrum og boligpris

Som forventet er de fleste av koeffisientene for postnumrene negative, noe som betyr at prisene er høyere i sentrum enn utenfor. Kun fem postnumre er positive. Postnummer 4608 og

4610 ligger veldig nær sentrum og havet, noe som kan forklare hvorfor disse er positive. Videre er postnummer 4630, 4631 og 4632 også positive, men disse ligger på Lund som er et veldig attraktivt boligområde. Siden postnummerdummyene fanger opp andre lokaliseringattributter i tillegg til avstand til sentrum er det vanskelig å finne om prisen blir lavere jo lenger unna sentrum du kommer. Attraktiviteten til et område spiller inn som en viktig faktor her. Tre postnumre er ikke signifikante, ett postnummer er signifikant på 2%-nivå og to er signifikante på 5%-nivå. De resterende 23 postnumrene er signifikante på 1%-nivå. Modellen med robuste standardfeil i vedlegg 16 viser at seks av postnumrene ikke er signifikante på 5% nivå, men at resten er signifikante på 1%-nivå. Jeg kan dermed forkaste nullhypotesen og konkludere med at det er en sammenheng mellom avstand til sentrum og boligpris.

Hypotese 4: Boareal

H₀ = Det er ikke en positiv sammenheng mellom boareal og boligpris

H_A = Det er en positiv sammenheng mellom boareal og boligpris

Koeffisienten til boarealet viser at når boarealet øker med en kvadratmeter, så vil prisen øke med 0,59%. Den har en P-verdi på 0,000 og jeg kan med 99% sikkerhet forkaste nullhypotesen og konkludere med at det er en positiv sammenheng mellom boareal og boligpris.

Hypotese 5: Alder

H₀ = Det er ikke en negativ sammenheng mellom alder på bolig og boligpris

H_A = Det er en negativ sammenheng mellom alder på bolig og boligpris

Koeffisienten til alder viser at når alder på bolig øker med ett år, så vil prisen synke med 0,24%. Den har en P-verdi på 0,000 og er signifikant på 1%-nivå. Jeg kan forkaste nullhypotesen og konkludere med at det er en negativ sammenheng mellom alder på bolig og prisen.

Hypotese 6: Fellesgjeld

H_0 = Det er ikke en negativ sammenheng mellom fellesgjeld og boligpris

H_A = Det er en negativ sammenheng mellom fellesgjeld og boligpris

Koeffisienten til fellesgjeld viser at når fellesgjelden øker med en krone, så vil prisen synke med -0,0000649%. Den har en P-verdi på 0,000 og er derfor signifikant på 1%-nivå. Jeg forkaster nullhypotesen og konkluderer med at det er en negativ sammenheng mellom fellesgjeld og boligpris.

Jeg har ikke laget hypoteser om boligtype og eierform, men fra analysen kan man se at eierform har betydning for boligprisene da «borettslag» er signifikant på 1%-nivå.

Koeffisienten forteller oss at prisen på en borettslagsbolig vil være -6,78% lavere enn en selveierbolig. Dette kan være fordi borettslagsleiligheter har en innskuddspris (fellesgjeld) som ikke er regnet med i prisen. Videre viser analysen at prisen for en enebolig vil være 5,11% høyere enn prisen for en leilighet. Variablene «tomannsbolig» og «rekkehus» er ikke signifikante hver for seg, men en kommando «testparm» i Stata viser at boligtypene til sammen er signifikante. Boligtype har med andre ord betydning for boligprisen.

7. Nærmere drøfting av hovedproblemstillingene

Hypotesetestingen i forrige kapittel viste at nærhet til barnehage har både en positiv og negativ effekt på boligprisene, da verdien på boligen stiger med nærhet til barnehage for å så synke ved lokalisering tett inntil en barnehage. I dette kapitlet lager jeg stiliserte eksempler på hvordan boligprisen endrer seg dersom jeg endrer avstand til barnehage.

7.1 Undersøkelse av verdiendring

Jeg tar utgangspunkt i basisboligen som er en selveierleilighet lokalisert i sentrum (postnummer 4611 og 4612). Jeg setter boareal lik 60 kvadratmeter, fellesgjeld lik null, alder lik null og avstand til barnehage lik null. Prisen for denne boligen vil være:

$$P = e^{14,38485+(0,0059018*60)} = 2\,517\,956$$

Når jeg nå øker avstanden til nærmeste barnehage med 100 meter vil prisen være:

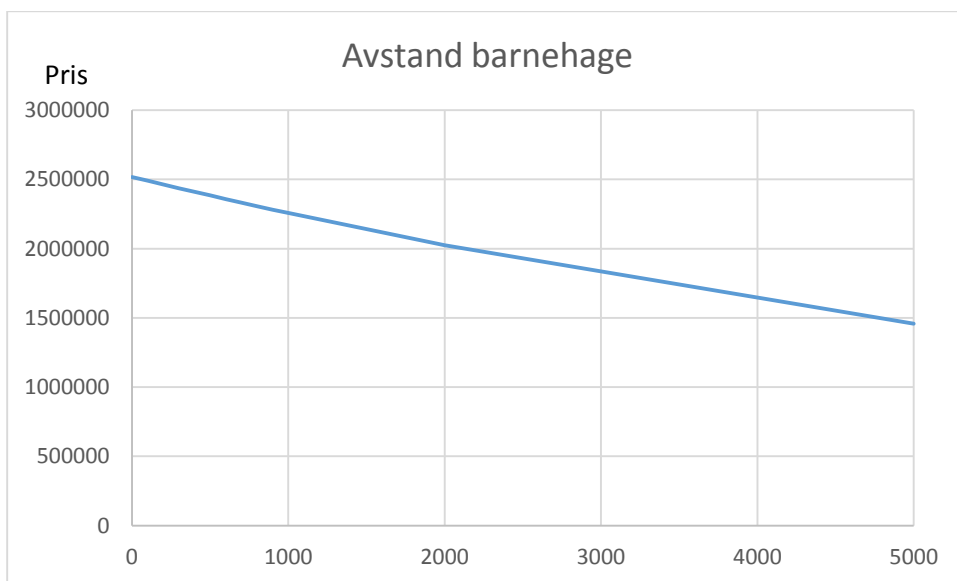
$$P = e^{14,38485+(0,0059018*60)+(-0,0001093*100)} = 2\,490\,585$$

Prisen har sunket med 27 371 kroner. I tabellen nedenfor har jeg estimert verdien på basisboligen for ulike avstander.

Tabell 28: Undersøkelse av verdiendring

Avstand i meter	Pris	Endring
0	2517956	0
100	2490585	-27371
200	2463511	-27074
300	2436731	-26780
400	2410243	-26488
500	2384042	-26201
600	2358127	-25915
700	2332493	-25634
800	2307137	-25356
900	2282058	-25079
1000	2257251	-24807

Tabellen demonstrerer at prisen synker avtakende ved økning i avstand til barnehage. Jeg velger i tillegg og vise grafen på dette i figuren under hvor den vertikale aksene er pris og den horisontale er antall meter.



Figur 27: Avstand til barnehage

Videre ønsker jeg å undersøke verdiendringen ved ulike avstandsdummyer. Dersom nærmeste barnehage er under 50 meter unna vil basisboligen ha prisen:

$$P = e^{14,19022+(0,0058883*60)+(0,1272144)} = 2\,351\,896$$

Dersom nærmeste barnehage er mellom 51 og 100 meter unna vil basisboligen ha prisen:

$$P = e^{14,19022+(0,0058883*60)+(0,172622)} = 2\,461\,152$$

Dersom nærmeste barnehage er mellom 101 og 500 meter unna vil basisboligen ha prisen:

$$P = e^{14,19022+(0,0058883*60)+(0,1569905)} = 2\,422\,979$$

Dersom nærmeste barnehage er mellom 501 og 1000 meter unna vil basisboligen ha prisen:

$$P = e^{14,19022+(0,0058883*60)+(0,1095975)} = 2\,310\,826$$

Dersom nærmeste barnehage er over 1000 meter unna vil basisboligen ha prisen:

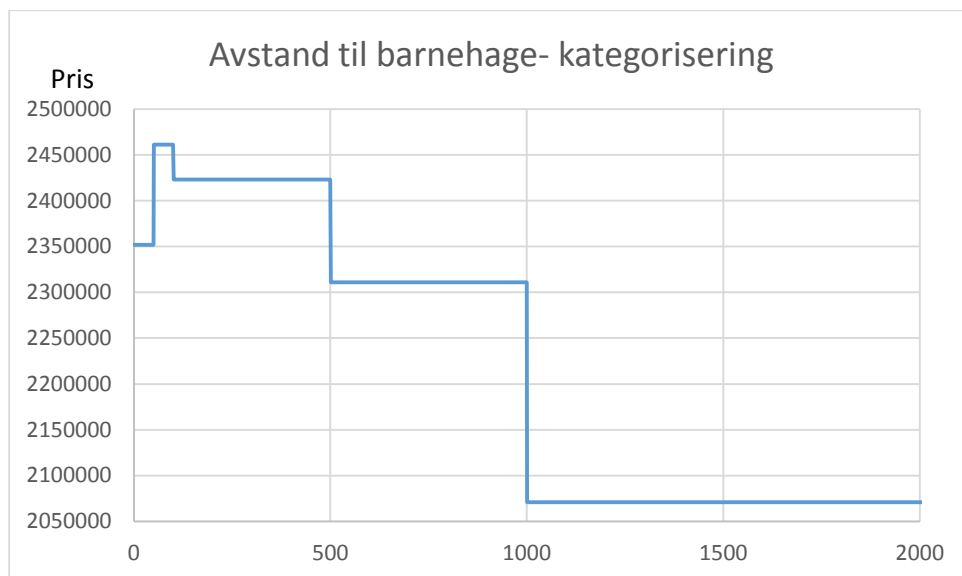
$$P = e^{14,19022+(0,0058883*60)} = 2\,070\,950$$

Tabellen under viser en oversikt over endringen i verdi når jeg endrer de ulike avstandene:

Tabell 29: Undersøkelse av verdiendring, kategorisering av avstand

Avstand meter	Pris	Endring
0-50 (1)	2351896	0
51-100 (2)	2461152	109256
101-500 (3)	2422979	-38173
501-1000 (4)	2310826	-112153
Over 1000 (5)	2070950	-239876

Ligger basisboligen mellom 51 og 100 meter unna en barnehage vil verdien på boligen være 109 256 kroner høyere enn hvis den hadde ligget innenfor 50 meter fra en barnehage. Verdien på boligen synker deretter med 38 173 kroner dersom den hadde ligget mellom 101 og 500 meter unna, og den synker ytterligere med 112 153 kroner dersom den hadde ligget mellom 501 og 1000 meter unna en barnehage. Til slutt synker verdien på boligen med 239 876 kroner dersom den hadde ligget over 1000 meter fra en barnehage. Den optimale avstanden vil ligge mellom 51 og 100 meter siden det er innenfor dette intervallet prisen er høyest. Dette illustreres i figuren under:



Figur 28: Avstand til barnehage- kategorisering

7.2 Kritiske vurderinger av analysen

Noen av barnehagene ligger ved siden av skoler, bedrifter, sykehjem eller nær andre fasiliteter. Dette er ikke tatt hensyn til i analysen, og kan dermed gi en forsterket (positiv eller negativ) sammenheng med boligprisen som egentlig ikke skyldes nærhet til barnehage i seg selv. Modellen min er nokså forenklet, og det hadde vært en fordel og kontrollert for flere variabler som forventes å påvirke boligprisen. Dette er imidlertid omfattende da dette krever god innsikt i nærmiljøene i Kristiansand.

8. Konklusjon

Formålet med denne oppgaven var å undersøke om det er en sammenheng mellom avstand til barnehage og boligpriser i Kristiansand. Resultatet av analysen viser at det er en negativ sammenheng mellom avstand til barnehage og boligpriser. Den semilogaritmiske modellen ble valgt da den passer datasettet mitt best. Den viste at boligprisene synker med 0,01093% når avstand til barnehage øker med én meter. En dummy-kategorisering av avstand til barnehage viser at prisen på en bolig innenfor 50 meter fra en barnehage er 12,72% høyere enn en bolig som ligger over 1000 meter fra en barnehage. Deretter er prisen på en bolig som ligger mellom 51 og 100 meter unna en barnehage 17,26% høyere, og prisen på en bolig som er mellom 101 og 500 meter unna er 16,69% høyere enn en bolig som ligger over 1000 meter unna. Til slutt vil en bolig som ligger mellom 501 og 1000 meter unna en barnehage ha en pris som er 10,95% høyere enn en bolig som ligger over 1000 meter unna. Med andre ord viser dette at boligprisen øker når man beveger seg fra under 50 meter til mellom 51-100 meter unna en barnehage, for å deretter avta. Den lave prisen helt inntil barnehagen kan skyldes at støy og trafikk gjør det uattraktivt å bo så nærme. For eiendomsutviklere eller boligbyggere kan dette bety at man burde være forsiktige med å bygge boliger innen 50 meter fra en eksisterende barnehage, og at den optimale avstanden ville vært mellom 51 og 100 meter unna siden det er der boligprisene vil være høyest. Hvis det er slik at småbarnsforeldre velger å bosette seg i eneboliger, mens enslige, par uten barn og eldre bosetter seg i leiligheter, kan dette bety at hvis det skal bygges noe nær en barnehage vil eneboliger muligens være det beste valget siden småbarnsfamilier er de som etterspør nærhet til barnehage.

Analysen viser også at avstand til sentrum er med på å forklare variasjonen i prisen da samtlige postnummerdummyer er signifikante. Negative koeffisienter indikerer at prisen er lavere enn i sentrum. Da andre lokaliseringsattributter enn avstand til sentrum også fanges opp av postnumrene gir ikke koeffisientene et bilde av at boligprisene synker jo lenger vekk fra sentrum man kommer. Resultatene viser at alle postnumrene med unntak av noen få attraktive områder, er negative og betyr at boligprisene er høyere i sentrum enn utenfor.

Jeg har også testet andre variabler som kan være med på å forklare boligprisen, og finner at prisen øker med 0,59% for hver økning i kvadratmeter, prisen synker med -0,24% når alderen på bolig øker med ett år og at prisen synker med -0,0000649% for hver kroners økning i

fellesgjeld. Kort oppsummert finner jeg signifikante sammenhenger mellom boligpris og avstand til barnehage, avstand til sentrum, boareal, alder på bolig, fellesgjeld, boligtype og eierform.

Som forslag til videre forskning kunne det vært interessant og undersøkt om antall barnehageplasser ved den nærmeste barnehagen har noe å si for boligprisen. En stor barnehage med mange plasser kan antas å gi en sterkere sammenheng enn det jeg har funnet, samtidig som en liten barnehage kanskje har en svakere eller ingen sammenheng i det hele tatt.

Et annet forslag er å undersøke om man finner lignende resultater i andre byer. Kristiansand er en nokså tettbebygd kommune med mange barnehager innen nær avstand fra boliger, som kan gjøre at etterspørselen etter nærhet til barnehage kanskje ikke er så høy som den muligens ville vært i en kommune hvor det var færre barnehager og lengere avstander mellom boliger og barnehager. Samtidig lurere jeg på om den negative effekten ved å ha en barnehage som nærmeste nabo ville forsvunnet i en slik kommune, da fordelene med å være nær kanskje vil veie opp for trafikken og støyen forbundet med dette.

9. Litteraturhenvisninger

- Bolstad, E. (2014). Postnummer i Vest-Agder med koorinatar. Hentet fra <http://www.erikbolstad.no/geo/noreg/postnummer/vest-agder/>.
- Des Rosiers, F., Lagana, A., & Thériault, M. (2001). Size and proximity effects of primary schools on surrounding house values. *Journal of Property Research*, 18(2), 149-168.
- DiPasquale, D., & Wheaton, W. C. (1996). *Urban Economics and Real Estate Markets*. New Jersey, USA: Prentice- Hall.
- Eiendomsverdi. (2014). Eiendomsinformasjon satt i system. Hentet 14.02.14 fra www.eiendomsverdi.no.
- Kristiansand Kommune. (2014a). Bydelskart. Hentet 26.04.2014 fra <http://www.kristiansand.kommune.no/Documents/Kommuneplan%202011-2022/Vedtatt%20kommuneplan%202011-2022/Kart/OversiktUtsnittBydeler.pdf>.
- Kristiansand Kommune. (2014b). Delområder. Hentet 28.04.2014 fra <http://www.kristiansand.kommune.no/no/Om-Kristiansand/Kristiansandstatistikken/Befolkning-1/Befolkning/>.
- Kristiansand Kommune. (2014c). Flyttematrise 2012. Hentet 28.04.2014 fra <http://www.kristiansand.kommune.no/no/Om-Kristiansand/Kristiansandstatistikken/Befolkning-1/Flytting/>.
- Kristiansand Kommune. (2014d). Kristiansand- A flourishing city. Hentet 26.04.2014 fra <http://www.kristiansand.kommune.no/no/Om-Kristiansand/About-Kristiansand/>.
- Li, M. M., & Brown, H. J. (1980). Micro-Neighborhood Externalities and Hedonic Housing Prices. *Land Economics*, 56(2), 125-141.
- Midtbø, T. (2012). *STATA- en entusiastisk innføring*. Oslo: Universitetsforlaget.
- NOU. (2002). *Boligmarkedene og boligpolitikken*. Oslo: Statensforvaltningstjeneste, Informasjonsforvaltning.
- Osland, L. (2001). Den hedonistiske metoden og estimering av attributtpriser. *Norsk Økonomisk Tidsskrift*, 115, 1-22.
- Robertsen, K. (2013). [Forelesningsnotater BE-409, Eiendomsøkonomi].
- Robertsen, K., & Theisen, T. (2009). The Impact of Financial Arrangements and Institutional Form on Housing Prices. *Journal of Real Estate Finance and Economics*, 42(3), 371-392.
- Robertsen, K., & Theisen, T. (2010). Boligmarkedet i Kristiansand *Økonomi og tid: 18 essays i Pufendorf-tradisjon: festskrift til Professor Arild Sæther på 70-årsdagen, 8.august 2010* (pp. 243-260). Bergen: Fagbokforlaget.
- Smith, L. B., Rosen, K. T., & Fallis, G. (1988). Recent Developements in Economic Models of Housing Markets. *Journal of Economic Literature*, 26(1), 29-64.
- Statistisk Sentralbyrå. (2011a). Beboede boliger etter bygningstype, fylke, kommune og bydel. Prosent. Hentet 13.02.14 fra <http://www.ssb.no/befolkning/statistikker/fobbolig/hvert-10-aar/2013-02-26?fane=tabell&sort=nummer&tabell=101959>.
- Statistisk Sentralbyrå. (2011b). Privathusholdninger, etter eierstatus, fylke, kommune og bydel. 2001 og 2011. Prosent. Hentet 28.04.2014 fra <http://www.ssb.no/a/kortnavn/fobhushold/tab-2012-12-18-13.html>.
- Statistisk Sentralbyrå. (2011c). Privathusholdninger, etter husholdningstype, fylke, kommune og bydel. 2011. Prosent. Hentet 13.02.2014 fra <http://www.ssb.no/a/kortnavn/fobhushold/tab-2012-12-18-08.html>.

- Statistisk Sentralbyrå. (2013a). Berekna folkemengd ved årsskiftet, 15 desember 2013. Hentet 13.02.2014 fra <http://www.ssb.no/befolkning/statistikker/folkber>.
- Statistisk Sentralbyrå. (2013b). Innenlandske flyttinger. Hentet 18.02.2014 fra <https://www.ssb.no/statistikbanken/SelectVarVal/Define.asp?MainTable=FlyttInnUtNetto&KortNavnWeb=flytting&PLanguage=0&checked=true>.
- Studenmund, A. H. (2006). *Using Econometrics* (5 ed.): Pearson Education.
- Zikmund, W. G., Babin, B. J., Carr, J. C., & Griffin, M. (2010). *Business Research Methods*: South-Western, Cengage Learning.

Vedlegg

Vedlegg 1: Barnehager i drift per 01.01.2014

Barnehage	Område
Odderøya barnehage	Kvadraturen/Eg/Odderøya
Revehiet naturbarnehage	Kvadraturen/Eg/Odderøya
Kløvergården barnehage	Kvadraturen/Eg/Odderøya
Sløyden barnehage	Kvadraturen/Eg/Odderøya
Fergefjellet barnehage	Kvadraturen/Eg/Odderøya
Alvejordet steinerbarnehage	Kvadraturen/Eg/Odderøya
Trollhaugen barnehage	Kvadraturen/Eg/Odderøya
Hellinga barnehage	Tinnheia
Stasjonshaven barnehage	Grim
Wergerlandsveien barnehage	Grim
Ravnedalen barnehage	Grim
Læringsverkstedet avd. Grim	Grim
Møllestua barnehage	Grim
Eventyrskogen barnehage	Strai
Gangdalen familiebarnehage	Strai
Strai barnehage	Strai
Fantebakken familiebarnehage	Mosby
Høietun barnehage	Mosby
Mosby barnehage	Mosby
Mosby Oppvekstsenter, avd. Barnehage	Mosby
Vågsbygd FUS bhage avd. Karuss	Mitre Vågsbygd
Vågsbygd familiebarnehage	Mitre Vågsbygd
Vågsbygds FUS bhage avd. Lumber	Mitre Vågsbygd
Karuss barnehage	Mitre Vågsbygd
Fiskåtangen barnehage	Mitre Vågsbygd
Eplehagen barnehage	Mitre Vågsbygd
Auglandstjønn barnehage	Mitre Vågsbygd
Skårungen barnehage	Mitre Vågsbygd
Rognelia familiebarnehage	Mitre Vågsbygd
Rådyrveien familiebarnehage	Voie/Møvig
Læringsverkstedet avd Bråvannsåsen	Voie/Møvig
Voie barnehage	Voie/Møvig
Ternevig barnehage AS	Voie/Møvig
Kanutten familiebarnehage	Voie/Møvig
Mummidalen barnehage A/S	Voie/Møvig
Andungen barnehage	Voie/Møvig
Voietun barnehage	Voie/Møvig
Vågsbygd FUS bhage avd. Voie	Voie/Møvig
Ravneheia barnehage	Voie/Møvig
Taremareskogen barnehage	Flekkerøy
Flekkerøya barnehage	Flekkerøy

Bokfinken familiebarnehage	Slettheia
Linerla barnehage	Slettheia
Nedre Slettheia Fam.bhg	Slettheia
Slettheia familiebhg	Slettheia
Øvre Slettheia barnehage	Slettheia
Fidjemoen familiebarnehage	Hellemyr
Hellemyr barnehage	Hellemyr
Skomagergada barnehage	Hellemyr
Solkollen barnehage avd. Hellemyr	Hellemyr
Tinnstua barnehage	Tinnheia
Lykkeliten familiebarnehage	Tinnheia
Lille Brum fam.barnehage	Tinnheia
Knøttebo familiebarnehage	Tinnheia
Hokus Pokus barnehage	Lund
Blåmann Steinerbarnehage	Lund
Kongsgård flyktningsbarnehage	Lund
Læringsverkstedet Kuholmen	Lund
Lillebo familiebarnehage	Lund
Lund Barnehage A/S	Lund
Læringsverkstedet avd lund	Lund
Solsikken familiebarnehage	Lund
Stiftegården familiebarnehage	Lund
Bamsebo barnehage	Lund
Barnas rom og familiebarnehage	Lund
Roligheden gård barnehage	Lund
Barnas rom og familiebhg Gimlekollen	Gimlekollen
Bergtorasvei barnehage	Gimlekollen
Jegersberg barnehage	Gimlekollen
Linntjønn FUS barnehage	Gimlekollen
Læringsverkstedet bhg avd Gimlekollen	Gimlekollen
Nemo familiebarnehage	Gimlekollen
Presteheia barnehage	Gimlekollen
Havlimyra barnehage	Justvik
Læringsverkstedet avd. Eidet	Justvik
Kalvehagen familiebarnehage	Justvik
Erkleiv familiebarnehage	Ålefjær/Erkleiv
Berkhuskollen familiebarnehage	Hånes
Hvetekornet familiebarnehage	Hånes
Jordbærveien barnehage	Hånes
Rakkerungan barnehage	Hånes
Solkollen barnehage Hånes	Hånes
Timenes gårdsbarnehage	Hånes
Sørlandsparken barnehage	Hånes
Søm barnehage	Søm
Askeladden barnehage	Søm

Barnas hage Ronatoppen fam.bhg	Søm
Liantjønn barnehage	Søm
Portveien familiebarnehage	Søm
Snikkedalen familiebarnehage	Søm
Rødhette FUS barnehage	Søm
Solkollen barnehage Søm	Søm
Tittelitue familiebarnehage	Søm
Veslefrikk barnehage	Søm
Espira Dvergsnes barnehage	Søm
Blekkulf barnehage	Randesund
Fidje barnehage	Randesund
Maurtua barnehage	Randesund
Maritippen barnehage	Søm/dvergsnes
Rakkerungan gårdsbarnehage	Randesund
Tømmerstien familiebarnehage	Randesund
Hamretun barnehage	Tveit
Brattvolltoppen familiebarnehage	Tveit
Læringsverkstedet avd. Tveit	Tveit
Mølledammen Gård familiebarnehage	Tveit
Refshalen Gård barnehage	Tveit
Trollstua barnehage	Tveit

Vedlegg 2: Barnehager stengt de siste tre årene

Barnehage
Tobienborg familiebarnehage
Trollstua familiebarnehage
Vabua barnehage
Lundskråga barnehage
Eidet barnehage
Oddernes og Kongsgård barnehage
Fredheim familiebarnehage
Lille Brum familiebarnehage
Andøysløyfen barnehage

Vedlegg 3: Solgte boliger i de ulike postnumrene og bydelene, 2011-2014

Postnummer	Antall observasjoner	Bydel
4608	173	Kvadraturen
4610	91	Kvadraturen
4611	12	Kvadraturen
4612	51	Kvadraturen
4613	29	Kvadraturen
4614	371	Kvadraturen
4615	60	Eg
4616	224	Grim
4617	134	Grim
4618	83	Strai
4619	91	Mosby
4620	162	Midtre Vågsbygd
4621	211	Midtre Vågsbygd
4622	128	Ytre Vågsbygd
4623	173	Ytre Vågsbygd
4624	282	Ytre Vågsbygd
4625	97	Flekkerøya
4626	286	Slettheia
4628	199	Hellemyr
4629	243	Tinnheia
4630	225	Lund/Sødal
4631	467	Lund/Sødal
4632	242	Lund/Sødal
4633	292	Gimlekollen
4634	210	Justvik/Ålefjær
4635	301	Hånes
4637	102	Indre Randesund (Søm)
4638	281	Indre Randesund (Søm)
4639	262	Ytre Randesund
4656	38	Hamresanden (v/Hånes)
4658	89	Tveit

Vedlegg 4: Gjennomsnittlig boligpris etter postnummer, 2011-2014

Postnummer	Obs	Gjennomsnitt	Standardavvik	Min.	Maks.
4617	134	1874888	576317,5	950000	4500000
4629	243	1892037	599948,7	670000	4000000
4614	371	1908663	945244,7	425000	6900000
4619	91	1960110	485571,2	975000	3350000
4624	282	2004504	813128,9	650000	5350000
4626	286	2014757	829295,4	750000	5950000
4612	51	2020833	1207370	650000	6700000
4621	211	2094905	1040912	920000	7250000
4615	60	2117167	724780,7	1350000	5500000
4616	224	2189688	935091,8	1000000	6100000
4622	128	2215492	868770	50000	4775000
4620	162	2368719	726108,2	950000	5200000
4618	83	2387229	642091,8	1050000	3900000
4635	301	2396130	1074195	1050000	11300000
4611	12	2430000	844522,5	1700000	3950000
4628	199	2526070	886565,1	540000	4720000
4632	242	2589161	1243264	600000	8500000
4634	210	2625019	1042671	890000	5800000
4608	173	2651214	1630741	850000	10000000
4613	29	2695345	1404491	1350000	6000000
4638	281	2747847	879127,1	750000	6050000
4658	89	2759551	687262,6	1300000	4500000
4631	467	2834733	1378818	650000	8775000
4623	173	2873020	1494536	270000	9200000
4639	262	2958340	1776796	50000	9100000
4630	225	3168276	1378558	1200000	9100000
4633	292	3288048	1519380	450000	9250000
4656	38	3334145	946926,4	1250000	7050000
4625	97	3441289	1521509	1490000	8900000
4610	91	3510632	1819805	1215000	7900000
4637	102	3601422	883006,3	1990000	6350000

Vedlegg 5: Gjennomsnittlig fellesgjeld etter postnummer, 2011-2014

Fellesgjeld	Obs.	Gjennomsnitt	Standardavvik	Min.	Maks.
4611	12	0	0	0	0
4613	29	0	0	0	0
4618	83	0	0	0	0
4625	97	0	0	0	0
4637	102	0	0	0	0
4656	38	0	0	0	0
4658	89	0	0	0	0
4610	91	3077	13153,55	0	77718
4608	173	5843,48	15950,98	0	89750
4635	301	10132,3	129440,7	0	378300
4634	210	15531,39	29141,66	0	103373
4620	162	19870,02	101517,1	0	568000
4619	91	23951,95	88796,23	0	445621
4638	281	35475,58	207193,8	0	1605288
4630	225	37463,87	99569,82	0	465675
4623	173	41172,57	158050,4	0	1720000
4626	286	41906,55	67852,49	0	473647
4631	467	57265,18	84308,33	0	426830
4617	134	63375,5	101938,4	0	337533
4628	199	72427,83	280333,1	0	1500000
4621	211	89479,19	147432,7	0	800000
4632	242	90577,23	115736,9	0	383570
4616	224	90975,44	138353	0	518817
4622	128	93086,43	177962	0	792647
4633	292	93633,15	287174,8	0	1500000
4615	60	99788,85	67751,3	0	230856
4624	282	122999,7	186890,5	0	1008063
4614	371	172609,2	409145,9	0	2187500
4629	243	205285,4	219421,7	0	936830
4612	51	330441,2	689454,2	0	2835000
4639	262	392857,2	691183,4	0	2700000

Vedlegg 6: Gjennomsnittlig alder på bolig etter postnummer, 2011-2014

Postnummer	Observasjoner	Gjennomsnitt	Standardavvik	Min.	Maks.
4639	262	11,748	13,538	0	129
4638	281	20,672	13,554	1	117
4634	210	21,723	20,737	0	131
4618	83	26,373	26,528	0	108
4628	199	28,07	14,717	1	62
4635	301	29,604	9,257	2	49
4633	292	30,349	19,589	0	132
4625	97	31,876	45,41	0	243
4623	173	32,479	20,596	0	141
4619	91	32,604	25,506	1	120
4624	282	34,088	12,541	1	76
4621	211	37,27	18,302	0	82
4656	38	38,473	16,819	1	56
4622	128	39,039	17,324	0	75
4658	89	39,82	29,773	1	152
4637	102	40,48	9,545	2	56
4626	286	42,062	6,01	14	56
4612	51	43,019	59,006	2	213
4620	162	43,753	19,17	0	78
4629	243	44,802	8,747	6	54
4614	371	49,824	45,452	0	311
4610	91	51,12	40,669	7	121
4608	173	52,421	43,713	0	123
4617	134	52,679	15,601	7	93
4631	467	52,792	21,701	0	136
4630	225	54,666	34,044	0	286
4615	60	55,933	1,614	53	59
4613	29	56,275	43,72	1	222
4632	242	56,603	13,594	7	93
4611	12	61,083	43,292	7	112
4616	224	64,991	23,749	11	123

Vedlegg 7: Gjennomsnittlig boareal etter postnummer, 2011-2014

Boareal	Observasjoner	Gjennomsnitt	Standardavvik	Min.	Maks.
4608	173	60,479	37,492	16	375
4614	371	62,397	35,639	19	295
4612	51	67,742	33,948	27	190
4611	12	71	30,283	38	132
4615	60	81,4	35,793	53	226
4617	134	81,611	30,856	40	298
4610	91	83,004	36,655	22	210
4621	211	84,367	44,778	33	296
4632	242	85,061	44,214	16	343
4631	467	86,196	43,941	22	266
4616	224	91,863	50,915	40	418
4629	243	93,884	31,155	29	200
4624	282	107,479	46,909	34	240
4626	286	107,583	50,456	40	304
4635	301	107,867	39,503	47	285
4622	128	111,479	44,389	27	230
4630	225	112,155	53,068	29	273
4638	281	116,062	48,798	40	282
4620	162	116,24	43,291	42	293
4619	91	116,725	42,97	50	257
4618	83	118,53	53,254	39	273
4634	210	118,575	38,803	52	257
4613	29	120,069	76,357	48	305
4623	173	125,118	53,134	49	289
4639	262	125,919	59,425	45	336
4628	199	126,908	49,844	30	307
4633	292	128,626	60,968	37	319
4656	38	149,315	41,714	70	227
4625	97	150,134	66,708	52	346
4658	89	159,405	52,749	58	291
4637	102	161,002	44,692	79	312

Vedlegg 8: Avstand til sentrum fra de ulike postnumrene

Postnummer	Avstand sentrum km
4611	0,1
4612	0,3
4614	0,7
4610	0,8
4616	1,4
4608	1,5
4615	1,7
4613	1,9
4631	2,3
4617	2,5
4630	2,6
4632	2,8
4629	3,9
4626	4,7
4621	5,4
4628	5,8
4633	5,8
4622	5,9
4620	6,2
4623	6,9
4624	6,9
4637	7,5
4618	8,2
4635	9,5
4638	10,1
4619	11,2
4634	11,5
4639	11,7
4625	12
4656	12,9
4658	18,9

Vedlegg 9: Boligtyper solgt mellom 2011-2014 etter postnummer

Postnummer	Enebolig	Leilighet	Rekkehus	Tomanns.	Total
4608	1	171	0	1	173
4610	1	87	0	3	91
4611	0	12	0	0	12
4612	3	45	0	3	51
4613	11	14	1	3	29
4614	23	328	4	16	371
4615	4	44	11	1	60
4616	33	137	25	29	224
4617	4	84	27	19	134
4618	33	31	5	14	83
4619	50	23	9	9	91
4620	41	34	36	51	162
4621	29	157	6	19	211
4622	25	43	32	28	128
4623	85	45	26	17	173
4624	79	129	34	40	282
4625	68	16	1	12	97
4626	78	140	46	22	286
4628	81	63	26	29	199
4629	37	139	58	9	243
4630	59	119	5	42	225
4631	47	353	45	22	467
4632	15	162	46	19	242
4633	115	109	32	36	292
4634	71	53	74	12	210
4635	69	159	51	22	301
4637	65	1	28	8	102
4638	63	162	48	8	281
4639	106	101	29	26	262
4656	27	3	0	8	38
4658	77	1	0	11	89
Total	1400	2965	705	539	5609

Vedlegg 10: Eierform på boliger solgt 2011-2014 etter postnummer

Postnummer	Aksjeleilighet	Borettslag	Selveier	Total
4608	0	13	160	173
4610	0	0	91	91
4611	0	0	12	12
4612	0	21	30	51
4613	0	0	29	29
4614	4	115	252	371
4615	0	46	14	60
4616	0	89	135	224
4617	0	36	98	134
4618	0	0	83	83
4619	0	9	82	91
4620	0	6	156	162
4621	0	126	85	211
4622	0	37	91	128
4623	0	19	154	173
4624	0	123	159	282
4625	0	0	97	97
4626	2	74	210	286
4628	0	24	175	199
4629	0	156	87	243
4630	0	22	203	225
4631	44	189	234	467
4632	7	93	142	242
4633	1	62	229	292
4634	0	49	161	210
4635	0	149	152	301
4637	0	0	102	102
4638	0	14	267	281
4639	0	74	188	262
4656	0	0	38	38
4658	0	0	89	89
Total	58	1546	4005	5609

Vedlegg 11: Gjennomsnittlig avstand til barnehage, 2011-2014 etter postnummer

Postnummer	Observasjoner	Gjennomsnitt	Standardavvik	Min.	Maks.
4610	91	66,61538	53,88502	12	311
4616	224	168,25	67,64733	33	456
4632	242	190,5537	92,86973	13	496
4626	286	195,7203	113,7287	16	535
4614	371	205,5741	80,9434	12	355
4629	243	217,9218	132,6864	25	688
4615	60	219	114,2191	63	455
4631	467	233,3769	117,7704	20	577
4624	282	257,7128	137,1458	48	1000
4621	211	311,3507	118,7921	16	645
4635	301	312,2359	213,1986	26	960
4638	281	317,2313	186,4353	17	715
4628	199	324,1156	317,1807	20	2120
4623	173	348,2197	152,804	45	830
4608	173	352,6647	142,6148	199	712
4622	128	354,1016	177,9072	15	740
4613	29	356,069	96,03122	129	550
4639	262	356,6679	352,2923	70	2980
4637	102	376,9314	173,9492	41	710
4633	292	377,9452	236,0983	20	1000
4611	12	385	45,55117	266	425
4620	162	394,4815	128,4402	89	680
4630	225	414,8444	242,6968	50	1560
4612	51	417	53,67756	350	595
4618	83	442,9277	312,2754	59	1120
4658	89	554,0225	525,0923	0	3280
4619	91	567,4176	232,043	14	1130
4617	134	600,6716	435,3352	13	1760
4625	97	642,3814	416,1779	35	1550
4634	210	644,1857	621,8374	30	4500
4656	38	684,5789	268,4297	252	1350

Vedlegg 12: Avstandskategorier etter postnummer

Postnummer	Innenfor 50	Innenfor 100	Innenfor 500	Innenfor 1000	Over 1000	Total
4608	0	0	146	27	0	173
4610	45	26	20	0	0	91
4611	0	0	12	0	0	12
4612	0	0	44	7	0	51
4613	0	0	26	3	0	29
4614	20	33	318	0	0	371
4615	0	8	52	0	0	60
4616	4	31	189	0	0	224
4617	1	9	58	42	24	134
4618	0	15	35	29	4	83
4619	1	1	34	52	3	91
4620	0	1	124	37	0	162
4621	2	3	196	10	0	211
4622	4	7	90	27	0	128
4623	1	7	137	28	0	173
4624	3	16	245	18	0	282
4625	2	5	37	34	19	97
4626	25	55	204	2	0	286
4628	14	21	146	11	7	199
4629	19	28	180	16	0	243
4630	1	7	172	31	14	225
4631	40	37	389	1	0	467
4632	24	22	196	0	0	242
4633	18	15	195	64	0	292
4634	4	5	92	77	32	210
4635	22	35	188	56	0	301
4637	1	7	68	26	0	102
4638	19	18	180	64	0	281
4639	0	20	214	19	9	262
4656	0	0	13	22	3	38
4658	6	5	42	30	6	89
Total	276	437	4,042	733	121	5,609

sentrum	-0.0698	-0.3392	-0.1060	-0.2235	-0.1275	0.0784	0.1805	0.0223	-0.0148	0.0080	0.0549	0.0708	0.2135	-0.0588
	P4628	P4629	P4630	P4631	P4632	P4633	P4634	P4635	P4637	P4638	P4639	P4656	P4658	innen-50
P4628	1.0000													
P4629	-0.0408	1.0000												
P4630	-0.0392	-0.0435	1.0000											
P4631	-0.0578	-0.0641	-0.0616	1.0000										
P4632	-0.0407	-0.0452	-0.0434	-0.0640	1.0000									
P4633	-0.0449	-0.0499	-0.0479	-0.0706	-0.0498	1.0000								
P4634	-0.0378	-0.0420	-0.0403	-0.0594	-0.0419	-0.0462	1.0000							
P4635	-0.0457	-0.0507	-0.0487	-0.0718	-0.0506	-0.0558	-0.0470	1.0000						
P4637	-0.0261	-0.0290	-0.0278	-0.0410	-0.0289	-0.0319	-0.0268	-0.0324	1.0000					
P4638	-0.0440	-0.0489	-0.0469	-0.0692	-0.0488	-0.0538	-0.0453	-0.0547	-0.0313	1.0000				
P4639	-0.0425	-0.0471	-0.0453	-0.0667	-0.0470	-0.0519	-0.0437	-0.0527	-0.0301	-0.0508	1.0000			
P4656	-0.0158	-0.0176	-0.0169	-0.0249	-0.0175	-0.0194	-0.0163	-0.0197	-0.0112	-0.0190	-0.0183	1.0000		
P4658	-0.0244	-0.0270	-0.0260	-0.0383	-0.0270	-0.0298	-0.0250	-0.0302	-0.0173	-0.0292	-0.0281	-0.0105	1.0000	
innenfor50	0.0187	0.0285	-0.0423	0.0508	0.0491	0.0135	-0.0275	0.0263	-0.0248	0.0195	-0.0504	-0.0188	0.0107	1.0000
innenfor100	0.0198	0.0296	-0.0357	0.0015	0.0103	-0.0232	-0.0398	0.0341	-0.0047	-0.0119	-0.0013	-0.0240	-0.0103	-0.0661
innenfor500	0.0056	0.0095	0.0200	0.0755	0.0423	-0.0276	-0.1242	-0.0510	-0.0164	-0.0410	0.0474	-0.0697	-0.0704	-0.3654
innenfor1000	-0.0429	-0.0409	0.0043	-0.1149	-0.0823	0.0615	0.1381	0.0391	0.0502	0.0661	-0.0382	0.1098	0.0778	-0.0882
over1000	0.0180	-0.0316	0.0572	-0.0447	-0.0315	-0.0348	0.1776	-0.0354	-0.0202	-0.0341	0.0195	0.0326	0.0401	-0.0338
avstandbhg	-0.0012	-0.0847	0.0672	-0.1028	-0.1060	0.0451	0.2319	-0.0119	0.0257	-0.0072	0.0252	0.1094	0.1070	-0.2452
sentrum	0.0052	-0.0974	-0.1614	-0.2611	-0.1569	0.0063	0.2922	0.2374	0.0627	0.2582	0.3393	0.1519	0.4280	-0.0646

	inne-100	inne-500	inn-1000	over1000	avstan-g	sentrum
innenfor100	1.0000					
innenfor500	-0.4668	1.0000				
innenfor1000	-0.1127	-0.6227	1.0000			
over1000	-0.0432	-0.2385	-0.0576	1.0000		
avstandbhg	-0.2635	-0.2947	0.4856	0.6348	1.0000	
sentrum	-0.0425	-0.1729	0.2558	0.1152	0.2898	1.0000

Vedlegg 14: Fullstendig semilogaritmisk regresjonsmodell

Source	SS	df	MS	Number of obs = 5609		
Model	929.341212	37	25.11733	F(37, 5571) = 430.23		
Residual	325.241986	5571	.058381258	Prob > F = 0.0000		
Total	1254.5832	5608	.223713124	R-squared = 0.7408		
				Adj R-squared = 0.7390		
				Root MSE = .24162		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0059018	.0000991	59.54	0.000	.0057075	.0060961
alder	-.0024415	.0001388	-17.59	0.000	-.0027137	-.0021694
fgjeld	-6.49e-07	1.65e-08	-39.43	0.000	-6.81e-07	-6.16e-07
borettslag	-.0678226	.0105677	-6.42	0.000	-.0885395	-.0471057
enebolig	.0511522	.0134944	3.79	0.000	.024698	.0776064
tomannsbolig	.0101311	.0133996	0.76	0.450	-.0161373	.0363995
rekkehus	-.0064697	.011832	-0.55	0.585	-.029665	.0167256
P4608	.0761735	.0358041	2.13	0.033	.0059834	.1463635
P4610	.2014572	.0401412	5.02	0.000	.1227648	.2801497
P4613	-.2473302	.0544344	-4.54	0.000	-.3540428	-.1406176
P4614	-.1217964	.0330916	-3.68	0.000	-.1866688	-.0569241
P4615	-.0653311	.0442342	-1.48	0.140	-.1520475	.0213852
P4616	-.1474957	.0349471	-4.22	0.000	-.2160056	-.0789858
P4617	-.2078928	.0372024	-5.59	0.000	-.2808239	-.1349616
P4618	-.3464604	.0408625	-8.48	0.000	-.4265668	-.266354
P4619	-.4812849	.0401992	-11.97	0.000	-.560091	-.4024787
P4620	-.2856688	.0363925	-7.85	0.000	-.3570124	-.2143252
P4621	-.2026588	.0352165	-5.75	0.000	-.2716969	-.1336207
P4622	-.3423689	.0376686	-9.09	0.000	-.416214	-.2685238
P4623	-.2401692	.036166	-6.64	0.000	-.3110686	-.1692698
P4624	-.3791377	.0342824	-11.06	0.000	-.4463445	-.3119309
P4625	-.1932711	.039954	-4.84	0.000	-.2715965	-.1149458
P4626	-.4246413	.0341607	-12.43	0.000	-.4916096	-.3576731
P4628	-.325707	.0355298	-9.17	0.000	-.3953592	-.2560548
P4629	-.2301649	.0347633	-6.62	0.000	-.2983145	-.1620153
P4630	.0337754	.0346995	0.97	0.330	-.0342492	.1017999
P4631	.0831128	.0329172	2.52	0.012	.0185822	.1476434
P4632	.0346649	.0345938	1.00	0.316	-.0331525	.1024823
P4633	-.0780096	.0341374	-2.29	0.022	-.1449322	-.0110869
P4634	-.2408733	.0357834	-6.73	0.000	-.3110227	-.1707238
P4635	-.2068196	.0341684	-6.05	0.000	-.273803	-.1398362
P4637	-.1603049	.0395945	-4.05	0.000	-.2379256	-.0826842
P4638	-.191807	.0342742	-5.60	0.000	-.2589978	-.1246162
P4639	-.1876417	.0347534	-5.40	0.000	-.2557719	-.1195114
P4656	-.1528209	.0503387	-3.04	0.002	-.2515044	-.0541374
P4658	-.4166777	.0407236	-10.23	0.000	-.4965118	-.3368436
avstandbhg	-.0001093	.0000137	-7.96	0.000	-.0001362	-.0000824
_cons	14.38485	.0325345	442.14	0.000	14.32107	14.44863

Vedlegg 15: Fullstendig semilogaritmisk regresjonsmodell med avstandsdummyer

Source	SS	df	MS	Number of obs = 5609		
Model	929.227622	40	23.2306906	F(40, 5568) = 397.56		
Residual	325.355575	5568	.058433113	Prob > F = 0.0000		
Total	1254.5832	5608	.223713124	R-squared = 0.7407		
				Adj R-squared = 0.7388		
				Root MSE = .24173		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0058883	.0000993	59.29	0.000	.0056936	.006083
alder	-.0024672	.0001389	-17.76	0.000	-.0027396	-.0021948
fgjeld	-6.55e-07	1.65e-08	-39.70	0.000	-6.87e-07	-6.22e-07
borettslag	-.0622835	.0106138	-5.87	0.000	-.0830908	-.0414762
enebolig	.0506023	.0134973	3.75	0.000	.0241423	.0770624
tomannsbolig	.0086997	.0134048	0.65	0.516	-.017579	.0349783
rekkehus	-.0059729	.0118487	-0.50	0.614	-.029201	.0172551
P4608	.0844629	.035813	2.36	0.018	.0142554	.1546704
P4610	.2445781	.0408824	5.98	0.000	.1644326	.3247236
P4613	-.2402985	.0544482	-4.41	0.000	-.3470381	-.1335588
P4614	-.1049173	.0330292	-3.18	0.001	-.1696675	-.0401672
P4615	-.0548423	.0442159	-1.24	0.215	-.1415227	.0318381
P4616	-.1284017	.0348411	-3.69	0.000	-.1967038	-.0600997
P4617	-.1922791	.0374758	-5.13	0.000	-.2657463	-.1188119
P4618	-.3332213	.0410562	-8.12	0.000	-.4137075	-.2527352
P4619	-.4708384	.0404555	-11.64	0.000	-.5501469	-.3915298
P4620	-.2773264	.0364159	-7.62	0.000	-.3487157	-.2059371
P4621	-.1972143	.0352135	-5.60	0.000	-.2662464	-.1281821
P4622	-.3315854	.0376965	-8.80	0.000	-.4054853	-.2576855
P4623	-.2308754	.0361675	-6.38	0.000	-.3017778	-.1599729
P4624	-.3662504	.0342356	-10.70	0.000	-.4333656	-.2991352
P4625	-.175264	.040256	-4.35	0.000	-.2541816	-.0963465
P4626	-.406969	.0341707	-11.91	0.000	-.473957	-.339981
P4628	-.3123284	.0355867	-8.78	0.000	-.3820922	-.2425646
P4629	-.21256	.0347275	-6.12	0.000	-.2806395	-.1444806
P4630	.0450261	.0347463	1.30	0.195	-.0230903	.1131425
P4631	.0968404	.0328959	2.94	0.003	.0323516	.1613292
P4632	.053929	.0345413	1.56	0.119	-.0137854	.1216433
P4633	-.0680253	.0341918	-1.99	0.047	-.1350547	-.000996
P4634	-.2310864	.0359708	-6.42	0.000	-.3016031	-.1605697
P4635	-.1938887	.0342232	-5.67	0.000	-.2609795	-.1267979
P4637	-.1490248	.0396302	-3.76	0.000	-.2267155	-.0713341
P4638	-.1749151	.034324	-5.10	0.000	-.2422036	-.1076266
P4639	-.1781985	.0347915	-5.12	0.000	-.2464033	-.1099936
P4656	-.1464709	.0505237	-2.90	0.004	-.245517	-.0474248
P4658	-.4080739	.0408437	-9.99	0.000	-.4881436	-.3280043
innenfor50	.1272144	.0283269	4.49	0.000	.0716826	.1827461
innenfor100	.172622	.0264569	6.52	0.000	.1207561	.2244879
innenfor500	.1569905	.0239534	6.55	0.000	.1100326	.2039485
innenfor1000	.1095975	.0246266	4.45	0.000	.0613198	.1578752
_cons	14.19022	.040065	354.18	0.000	14.11168	14.26876

Vedlegg 16: Robuste standarfeil, semilogaritmisk regresjon med avstand som kontinuerlig variabel

Linear regression		Number of obs = 5609				
		F(37, 5571) = 376.34				
		Prob > F = 0.0000				
		R-squared = 0.7408				
		Root MSE = .24162				
lnpris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0059018	.0001432	41.21	0.000	.005621	.0061825
alder	-.0024415	.0002288	-10.67	0.000	-.0028901	-.0019929
fgjeld	-6.49e-07	4.38e-08	-14.83	0.000	-7.34e-07	-5.63e-07
borettslag	-.0678226	.0125973	-5.38	0.000	-.0925181	-.043127
enebolig	.0511522	.014443	3.54	0.000	.0228382	.0794661
tomannsbolig	.0101311	.011697	0.87	0.386	-.0127996	.0330618
rekkehus	-.0064697	.0095526	-0.68	0.498	-.0251965	.0122571
P4608	.0761735	.0491274	1.55	0.121	-.0201353	.1724822
P4610	.2014572	.0550755	3.66	0.000	.0934877	.3094268
P4613	-.2473302	.0635086	-3.89	0.000	-.3718319	-.1228285
P4614	-.1217964	.0444708	-2.74	0.006	-.2089765	-.0346164
P4615	-.0653311	.0461127	-1.42	0.157	-.1557301	.0250678
P4616	-.1474957	.044022	-3.35	0.001	-.233796	-.0611954
P4617	-.2078928	.0441641	-4.71	0.000	-.2944715	-.121314
P4618	-.3464604	.0479257	-7.23	0.000	-.4404135	-.2525073
P4619	-.4812849	.0469498	-10.25	0.000	-.5733248	-.3892449
P4620	-.2856688	.0447125	-6.39	0.000	-.3733228	-.1980147
P4621	-.2026588	.0444577	-4.56	0.000	-.2898133	-.1155043
P4622	-.3423689	.0540016	-6.34	0.000	-.4482331	-.2365048
P4623	-.2401692	.0459797	-5.22	0.000	-.3303074	-.150031
P4624	-.3791377	.0429047	-8.84	0.000	-.4632475	-.2950278
P4625	-.1932711	.053162	-3.64	0.000	-.2974893	-.0890529
P4626	-.4246413	.0434256	-9.78	0.000	-.5097725	-.3395101
P4628	-.325707	.0446986	-7.29	0.000	-.4133337	-.2380803
P4629	-.2301649	.0438336	-5.25	0.000	-.316096	-.1442339
P4630	.0337754	.0456005	0.74	0.459	-.0556195	.1231702
P4631	.0831128	.0436626	1.90	0.057	-.0024829	.1687085
P4632	.0346649	.0443501	0.78	0.434	-.0522786	.1216084
P4633	-.0780096	.0450145	-1.73	0.083	-.1662556	.0102365
P4634	-.2408733	.0453075	-5.32	0.000	-.3296936	-.1520529
P4635	-.2068196	.0433288	-4.77	0.000	-.2917609	-.1218783
P4637	-.1603049	.045648	-3.51	0.000	-.2497928	-.0708171
P4638	-.191807	.0433629	-4.42	0.000	-.2768152	-.1067988
P4639	-.1876417	.0493572	-3.80	0.000	-.284401	-.0908823
P4656	-.1528209	.0555286	-2.75	0.006	-.2616786	-.0439632
P4658	-.4166777	.0521165	-8.00	0.000	-.5188464	-.314509
avstandbhg	-.0001093	.0000135	-8.07	0.000	-.0001358	-.0000827
_cons	14.38485	.0432641	332.49	0.000	14.30004	14.46967

Vedlegg 17: Robuste standardfeil, semilogaritmisk regresjon med avstandsdummyer

Linear regression						Number of obs = 5609	
						F(40, 5568) = 350.42	
						Prob > F = 0.0000	
						R-squared = 0.7407	
						Root MSE = .24173	
lnpris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]		
BOA	.0058883	.0001425	41.32	0.000	.005609	.0061677	
alder	-.0024672	.0002292	-10.76	0.000	-.0029165	-.0020179	
fgjeld	-6.55e-07	4.38e-08	-14.95	0.000	-7.41e-07	-5.69e-07	
borettslag	-.0622835	.0125832	-4.95	0.000	-.0869515	-.0376155	
enebolig	.0506023	.0143984	3.51	0.000	.0223758	.0788288	
tomannsbolig	.0086997	.0117056	0.74	0.457	-.014248	.0316473	
rekkehus	-.0059729	.0095734	-0.62	0.533	-.0247405	.0127947	
P4608	.0844629	.0493081	1.71	0.087	-.0122002	.181126	
P4610	.2445781	.0561519	4.36	0.000	.1344985	.3546577	
P4613	-.2402985	.0632446	-3.80	0.000	-.3642825	-.1163145	
P4614	-.1049173	.0447245	-2.35	0.019	-.1925948	-.0172399	
P4615	-.0548423	.0461059	-1.19	0.234	-.1452278	.0355432	
P4616	-.1284017	.0441975	-2.91	0.004	-.2150462	-.0417573	
P4617	-.1922791	.044733	-4.30	0.000	-.2799733	-.1045849	
P4618	-.3332213	.0485069	-6.87	0.000	-.4283137	-.2381289	
P4619	-.4708384	.047334	-9.95	0.000	-.5636315	-.3780452	
P4620	-.2773264	.0449409	-6.17	0.000	-.3654281	-.1892247	
P4621	-.1972143	.0446824	-4.41	0.000	-.2848092	-.1096193	
P4622	-.3315854	.0542772	-6.11	0.000	-.43799	-.2251808	
P4623	-.2308754	.0462504	-4.99	0.000	-.3215442	-.1402066	
P4624	-.3662504	.0431254	-8.49	0.000	-.450793	-.2817078	
P4625	-.175264	.0539812	-3.25	0.001	-.2810883	-.0694397	
P4626	-.406969	.0437235	-9.31	0.000	-.4926842	-.3212539	
P4628	-.3123284	.0450553	-6.93	0.000	-.4006543	-.2240025	
P4629	-.21256	.0440732	-4.82	0.000	-.2989607	-.1261594	
P4630	.0450261	.0459104	0.98	0.327	-.0449761	.1350284	
P4631	.0968404	.0437038	2.22	0.027	.0111638	.182517	
P4632	.053929	.0446218	1.21	0.227	-.0335471	.141405	
P4633	-.0680253	.0453686	-1.50	0.134	-.1569654	.0209148	
P4634	-.2310864	.0459755	-5.03	0.000	-.3212163	-.1409565	
P4635	-.1938887	.0436348	-4.44	0.000	-.27943	-.1083475	
P4637	-.1490248	.045898	-3.25	0.001	-.2390028	-.0590468	
P4638	-.1749151	.0436735	-4.01	0.000	-.2605323	-.0892979	
P4639	-.1781985	.0496647	-3.59	0.000	-.2755608	-.0808362	
P4656	-.1464709	.0559121	-2.62	0.009	-.2560805	-.0368613	
P4658	-.4080739	.0523481	-7.80	0.000	-.5106967	-.3054511	
innenfor50	.1272144	.0268417	4.74	0.000	.0745941	.1798346	
innenfor100	.172622	.02286	7.55	0.000	.1278076	.2174364	
innenfor500	.1569905	.0203535	7.71	0.000	.1170897	.1968914	
innenfor1000	.1095975	.0206201	5.32	0.000	.069174	.150021	
_cons	14.19022	.04852	292.46	0.000	14.0951	14.28534	

Vedlegg 18: Kommandoer i STATA

```
cd "C:\Users\HeleneSofie\Documents\Masteroppgave\data"
```

***Setter inn txt-fil:**

```
insheet using "txt\0813.txt", clear  
insheet using "txt\1416.txt", clear  
insheet using "txt\1722.txt", clear  
insheet using "txt\2329.txt", clear  
insheet using "txt\3034.txt", clear  
insheet using "txt\353738.txt", clear  
insheet using "txt\395658.txt", clear
```

Lagrer alle txt-filene som dta-filer

***Henter inn dta-filene og setter de sammen**

```
use "dta\0813.dta"  
append using "dta\1416.dta"  
append using "dta\1722.dta"  
append using "dta\2329.dta"  
append using "dta\3034.dta"  
append using "dta\353738.dta"  
append using "dta\395658.dta"
```

```
rename avstand avstandbhg
```

***Renser dataene**

```
drop regdato  
drop omshast  
drop prisant  
drop pris  
drop fellesgjeld  
drop m2prom  
drop megler
```

```
drop if missing(bygger)
drop if prisen==0
drop if missing(boa)& missing(bta)
drop if missing(avstandbhg)
```

***Estimerer BOA ved hjelp av bta**

```
reg boa bta
generate BOA=boa
replace BOA=(14.86128+(0.7156465*bta)) if boa==.
```

***Finner alder på boligen**

```
generate alder=0
replace alder=(2011-bygger) if salgsr==2011
replace alder=(2012-bygger) if salgsr==2012
replace alder=(2013-bygger) if salgsr==2013
replace alder=(2014-bygger) if salgsr==2014
```

***Oppretter dummyer for boligtype**

```
generate enebolig=0
generate tomannsbolig=0
generate rekkehus=0
generate leilighet=0
replace enebolig=1 if boligtype=="Enebolig"
replace tomannsbolig=1 if boligtype=="Tomannsbolig"
replace rekkehus=1 if boligtype=="Rekkehus"
replace leilighet=1 if boligtype=="Leilighet"
```

***Koder boligtypene**

```
generate Boligtype=0
replace Boligtype=1 if boligtype=="Enebolig"
replace Boligtype=2 if boligtype=="Tomannsbolig"
replace Boligtype=3 if boligtype=="Rekkehus"
replace Boligtype=4 if boligtype=="Leilighet"
```

***Oppretter dummyer for eierform**

generate selveier=0

generate borettslag=0

replace selveier=1 if eierform=="Selveier"

replace borettslag=1 if eierform=="Borettslag"

replace borettslag=1 if eierform=="Aksjeleilighet"

***Koder eierform**

generate Eierform=0

replace Eierform=1 if eierform=="Selveier"

replace Eierform=2 if eierform=="Borettslag"

replace Eierform=3 if eierform=="Aksjeleilighet"

***Oppretter postnummerdummyer**

generate P4608=0

generate P4610=0

generate P4611=0

generate P4612=0

generate P4613=0

generate P4614=0

generate P4615=0

generate P4616=0

generate P4617=0

generate P4618=0

generate P4619=0

generate P4620=0

generate P4621=0

generate P4622=0

generate P4623=0

generate P4624=0

generate P4625=0

generate P4626=0

generate P4628=0
generate P4629=0
generate P4630=0
generate P4631=0
generate P4632=0
generate P4633=0
generate P4634=0
generate P4635=0
generate P4637=0
generate P4638=0
generate P4639=0
generate P4656=0
generate P4658=0
replace P4608=1 if postnummer==4608
replace P4610=1 if postnummer==4610
replace P4611=1 if postnummer==4611
replace P4612=1 if postnummer==4612
replace P4613=1 if postnummer==4613
replace P4614=1 if postnummer==4614
replace P4615=1 if postnummer==4615
replace P4616=1 if postnummer==4616
replace P4617=1 if postnummer==4617
replace P4618=1 if postnummer==4618
replace P4619=1 if postnummer==4619
replace P4620=1 if postnummer==4620
replace P4621=1 if postnummer==4621
replace P4622=1 if postnummer==4622
replace P4623=1 if postnummer==4623
replace P4624=1 if postnummer==4624
replace P4625=1 if postnummer==4625
replace P4626=1 if postnummer==4626
replace P4628=1 if postnummer==4628
replace P4629=1 if postnummer==4629
replace P4630=1 if postnummer==4630


```
replace P4631=1 if postnummer==4631
replace P4632=1 if postnummer==4632
replace P4633=1 if postnummer==4633
replace P4634=1 if postnummer==4634
replace P4635=1 if postnummer==4635
replace P4637=1 if postnummer==4637
replace P4638=1 if postnummer==4638
replace P4639=1 if postnummer==4639
replace P4656=1 if postnummer==4656
replace P4658=1 if postnummer==4658
```

***Oppretter grenser (dummyer)**

```
generate innenfor50=0
generate innenfor100=0
generate innenfor500=0
generate innenfor1000=0
generate over1000=0
```

```
replace innenfor50=1 if avstandbhg<=50
replace innenfor100=1 if avstandbhg>50 & avstandbhg<=100
replace innenfor500=1 if avstandbhg>100 & avstandbhg<=500
replace innenfor1000=1 if avstandbhg>500 & avstandbhg<=1000
replace over1000=1 if avstandbhg>1000
```

***Oppretter koder på grenser**

```
generate grenser=0
replace grenser=1 if innenfor50==1
replace grenser=2 if innenfor100==1
replace grenser=3 if innenfor500=1
replace grenser=4 if innenfor1000==1
replace grenser=5 if over1000==1
```

***Oppretter avstand til sentrum**

```
generate sentrum=0
```

replace sentrum=1.5 if postnummer==4608
replace sentrum=0.1 if postnummer==4611
replace sentrum=0.8 if postnummer==4610
replace sentrum=0.3 if postnummer==4612
replace sentrum=1.9 if postnummer==4613
replace sentrum=0.7 if postnummer==4614
replace sentrum=1.7 if postnummer==4615
replace sentrum=1.4 if postnummer==4616
replace sentrum=2.5 if postnummer==4617
replace sentrum=8.2 if postnummer==4618
replace sentrum=11.2 if postnummer==4619
replace sentrum=6.2 if postnummer==4620
replace sentrum=5.4 if postnummer==4621
replace sentrum=5.9 if postnummer==4622
replace sentrum=6.9 if postnummer==4623
replace sentrum=6.9 if postnummer==4624
replace sentrum=12 if postnummer==4625
replace sentrum=4.7 if postnummer==4626
replace sentrum=5.8 if postnummer==4628
replace sentrum=3.9 if postnummer==4629
replace sentrum=2.6 if postnummer==4630
replace sentrum=2.3 if postnummer==4631
replace sentrum=2.8 if postnummer==4632
replace sentrum=5.8 if postnummer==4633
replace sentrum=11.5 if postnummer==4634
replace sentrum=9.6 if postnummer==4635
replace sentrum=7.5 if postnummer==4637
replace sentrum=10.1 if postnummer==4638
replace sentrum=11.7 if postnummer==4639
replace sentrum=12.9 if postnummer==4656
replace sentrum=18.9 if postnummer==4658

***Oversikter**

sum prisen fgjeld alder BOA avstandbhg sentrum

sum enebolig tomannsbolig rekkehus leilighet selveier borettslag innenfor50 innenfor100
innenfor500 innenfor1000 over1000

histogram prisen, norm frequency

sum pris if postnummer==4608

sum pris if postnummer==4610

sum pris if postnummer==4611

....

Sum pris if postnummer==4658

histogram fgjeld, norm frequency

sum fgjeld if postnummer==4608

sum fgjeld if postnummer==4610

...

Sum fgjeld if postnummer==4658

histogram alder, norm frequency

sum alder if postnummer==4608

sum alder if postnummer==4610

....

Sum alder if postnummer==4658

histogram BOA, norm frequency

sum BOA if postnummer==4608

sum BOA if postnummer==4610

...

Sum BOA if postnummer==4658

histogram avstandbhg, norm frequency

sum avstandbhg if postnummer==4608

sum avstandbhg if postnummer==4610

...

Sum avstandbhg if postnummer==4658

histogram sentrum, norm frequency

sum sentrum if postnummer==4608

sum sentrum if postnummer==4610

...

Sum sentrum if postnummer==4658

histogram Boligtype, frequency

tab2 postnummer Boligtype

tabulate Eierform

tab2 postnummer eierform

catplot postnummer, percent

***Korrelasjonsmatrise**

corr prisen BOA alder fgjeld selveier borettslag enebolig tomannsbolig rekkehus leilighet
P4608 P4610 P4611 P4612 P4613 P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621
P4622 P4623 P4624 P4625 P4626 P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635
P4637 P4638 P4639 P4656 P4658 innenfor50 innenfor100 innenfor500 innenfor1000
over1000 avstandbhg sentrum

pwcorr prisen BOA alder fgjeld selveier borettslag enebolig tomannsbolig rekkehus leilighet
innenfor50 innenfor100 innenfor500 innenfor1000 over1000 avstandbhg sentrum, listwise
star(5)

***Lineær regresjon med avstand sentrum, postnummer og avstand barnehage**

reg pris BOA alder fgjeld borettslag enebolig rekkehus tomannsbolig P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626

P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
avstandbhg sentrum

***Test for multikollinearitet**

vif

***Lineær regresjon med postnummer og avstand barnehage**

reg prisen BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626
P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
avstandbhg

***Sjekker for heteroskedastisitet**

rvfplot, yline(0)
rvpplot fgjeld, yline(0)
rvpplot alder, yline(0)
rvpplot BOA, yline(0)
rvpplot avstandbhg, yline(0)

***BP-test**

hetttest
hetttest fgjeld
hetttest BOA
hetttest alder
hetttest avstandbhg

***Test for multikollinearitet**

vif

***Predikerer**

predict pred

***Residualaler**

predict resid, residual

***Sjekker om restleddet er normalfordelt**

histogram resid, norm

pnorm res

***Undersøker linearitetsforutsetningen**

acprplot BOA, lowess

acprplot fgjeld, lowess

acprplot alder, lowess

acprplot avstandbhg, lowess

***Lineær regresjon med kategorisering av avstand**

reg prisen BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626
P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
innenfor50 innenfor100 innenfor50 innenfor1000

***Logaritmisk regresjonsmodell**

generate lnpris=ln(prisen)

generate lnalder=ln(alder+1)

generate lnfgjeld=ln(fgjeld+1)

generate lnBOA=ln(BOA)

generate lnavstandbhg=ln(avstandbhg+1)

***Dobbeltlogaritmisk regresjon med kontinuerlig avstand**

reg lnpris lnBOA lnalder lnfgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610
P4613 P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625
P4626 P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656
P4658 lnavstandbhg

***Sjekker for heteroskedastisitet**

rvfplot, yline(0)

rvpplot lnavstandbhg, yline(0)

rvpplot lnalder, yline(0)

rvpplot lnfgjeld, yline(0)

rvpplot lnBOA, yline(0)

***Breusch-Pagan test for heteroskedastisitet**

hettest

hettest lnavstandbhg

hettest lnalder

hettest lnfgjeld

hettest lnBOA

***Residualer**

predict residlog, residual

***Sjekker normalfordeling i residualene**

pnorm residlog

***Sjekker linearitetsforutsetningene dobbeltlogaritmisk modell**

acprplot lnBOA, lowess

acprplot lnfgjeld, lowess

acprplot lnalder, lowess

acprplot lnavstandbhg, lowess

***Dobbeltlogaritmisk regresjon med kategorisering av avstand**

reg lnpris lnBOA lnalder lnfgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610
P4613 P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625
P4626 P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656
P4658 innenfor50 innenfor100 innenfor500 innenfor1000

***Semilogaritmisk regresjon med kontinuerlig avstand**

reg lnpris BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626
P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
avstandbhg

***Sjekker for heteroskedastisitet**

rvfplot, yline(0)

rvpplot fgjeld, yline(0)

rvpplot alder, yline(0)

rvpplot BOA, yline(0)

rvpplot avstandbhg, yline(0)

***BP-test**

hettest

hettest fgjeld

hettest BOA

hettest alder

hettest avstandbhg

***Residualer**

predict residsemi, residuals

***Sjekker normalfordeling i residualene**

pnorm residsemi

***Sjekker linearitetsforutsetningene semilogaritmisk modell**

acprplot BOA, lowess

acprplot fgjeld, lowess

acprplot alder, lowess

acprplot avstandbhg, lowess

***Robuste standardfeil semilogaritmisk modell**

reg lnpris BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613

P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626

P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658

avstandbhg, robust

***Semilogaritmisk regresjon med kategorisering av avstanden**

reg lnpris BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626
P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
innenfor50 innenfor100 innenfor500 innenfor1000

***Robuste standardfeil semilogaritmisk modell med kategorisering av avstanden**

reg lnpris BOA alder fgjeld borettslag enebolig tomannsbolig rekkehus P4608 P4610 P4613
P4614 P4615 P4616 P4617 P4618 P4619 P4620 P4621 P4622 P4623 P4624 P4625 P4626
P4628 P4629 P4630 P4631 P4632 P4633 P4634 P4635 P4637 P4638 P4639 P4656 P4658
innenfor50 innenfor100 innenfor500 innenfor1000, robust

***Sjekker om boligtypene er signifikante**

testparm enebolig tomannsbolig rekkehus