

Hva betyr avstand til E18 for boligprisene i Askim?

Stine Karterud

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 innestår for de metoder som er anvendt og de konklusjoner som er trukket.

Forord

Denne oppgaven er gjennomført som siste del av masterutdanningen i økonomi og administrasjon, med spesialisering innen finansiell økonomi, ved Universitetet i Agder. Oppgaven er obligatorisk og utgjør 30 studiepoeng og har som hovedmål å lære studenten å gjennomføre et vitenskapelig arbeid.

Arbeidet med oppgaven har vært en veldig spennende og lærerik prosess. Datainnsamlingen tok mye tid, men til gjengjeld fikk jeg et realistisk datamateriale som kan belyse hva avstand til E18 betyr for boligprisene i Askim.

Jeg vil benytte anledningen til å takke min veileder, Anne Wenche Emblem, for god veiledning og raske tilbakemeldinger underveis.

Stine Karterud
Kristiansand, 15. mai 2015

Sammendrag

Formålet med denne masteroppgaven er å undersøke hvilken betydning avstand til ny E18 har for boligprisene i Askim. For å belyse dette blir det for hver enkelt bolig målt kjøreavstand i meter til nærmeste påkjørsel på motorveien. Det antas at økt kjøreavstand til påkjørsel vil redusere boligprisene ved at E18 gir bedret fremkommelighet til større byer hvor mange har sine arbeidsplasser. På en annen side kan det være en negativ faktor å bo nærme E18 på grunn av økt støy og forurensing. På bakgrunn av dette blir det også målt luftavstand i meter til E18 fra hver enkelt bolig.

For å besvare oppgaven på best mulig måte er det samlet inn informasjon om solgte boliger i Askim fra år 2003-2014. Kjøreavstand i meter til nærmeste påkjørsel er inkludert som en kontinuerlig variabel i analysen, mens avstand i meter målt i luftavstand blir kategorisert inn i ulike dummy-variabler. Det endelige datasettet består av 3.546 observasjoner og blir analysert i tre ulike modeller. På bakgrunn av analysen utført i kapittel 6 blir den semi-logaritmiske modellen valgt, den forklarer 77,50 prosent av variasjonen i boligprisene.

Den semi-logaritmiske modellen viser en signifikant negativ sammenheng mellom boligpriser og kjøreavstand i meter til nærmeste påkjørsel på motorveien. For hver meter økning i kjøreavstand til nærmeste påkjørsel vil boligprisen bli redusert, dette innebærer at konklusjonen samsvarer med antagelsen. Med samme modell er det også utført en regresjon som kategoriserer avstanden til motorvei målt i luftlinje inn i ulike dummy-variabler for å kunne se om det er negativt å bo nærme motorveien målt i luftlinje i forhold til å bo lenger unna. Analysen viser at boliger som ligger innenfor 500 meter og 501-1000 meter fra motorveien i luftlinje har en lavere pris enn boliger som ligger over 1000 meter unna. Det som er uventet er at boligene som ligger mellom 501-1000 meter unna motorveien i luftlinje har den laveste prisen.

Innholdsfortegnelse

Forord	ii
Sammendrag	iii
Innholdsfortegnelse	iv
Figurliste	vi
Tabelloversikt	vii
Vedleggsliste	viii
1. Innledning	1
2. Bakgrunn	3
2.1 Askim	3
2.2 Befolkningsvekst	5
2.3 Europavei E18	9
3. Teori	12
3.1 Hva bestemmer prisen på en bolig?	12
3.2 Alonso – Muth – Mills – modellen	13
3.3 Hedonistisk prisfunksjon	16
3.4 Eksternaliteter	25
3.5 Hypoteser	26
4. Økonometrisk modell	28
4.1 Introduksjon av relevante regresjonsmodeller	28
4.1.1 Enkel og multipl regressjon	28
4.1.2 Logaritmisk funksjonsform	29
4.2 Valg av uavhengige variabler	30
4.3 De klassiske forutsetningene for OLS	33
4.4 Modellens forklaringskraft	35
4.5 Hypotese testing	36
4.6 Spesifisering av økonometrisk modell	38
5. Innhenting, bearbeiding og beskrivelse av datamaterialet	40
5.1 Innsamling av datamaterialet	40
5.2 Rensing og kontroll av data	43
5.3 Beskrivelse av datamaterialet	44
5.4 Korrelasjon	55

6. Analyse	57
6.1 Gjennomgang av regresjonsmodellene	57
6.1.1 Multipl lineær regresjon	57
6.1.2 Multipl lineær regresjon definert med avstands-dummyer	63
6.1.3 Dobbel-logaritmisk regresjon.....	64
6.1.4 Dobbel-logaritmisk regresjon definert med avstands-dummyer.....	70
6.1.5 Semi-logaritmisk regresjon	71
6.1.6 Semi-logaritmisk regresjon definert med avstands-dummyer	76
6.2 Valg av modell.....	77
6.3 Hypotesetesting	77
7. Nærmere drøfting av hovedproblemstilling	80
7.1 Tolkning av resultatene	80
7.2 Undersøkelse av verdiendring	81
7.3 Kjente begrensninger i analysen.....	84
8. Konklusjon.....	85
Litteraturliste	87

Figurliste

Figur 1: Kart over Østfold (Østfold fylkeskommune, 2012).....	3
Figur 2: Kart over Askim inkludert tettbebyggelser.....	4
Figur 3: Nyetablerte bedrifter i Askim, tall fra (Statistisk sentralbyrå, 2014e).....	5
Figur 4: Befolkningstetthet i Oslo-regionen 1. Januar 2010 (Høydahl, 2010).....	6
Figur 5: Gjennomsnittlig kvadratmeterpris (Statistisk sentralbyrå, 2014d).....	6
Figur 6: Nettoinnflytting til Askim fra år 2000-2014 (Statistisk sentralbyrå, 2015).....	8
Figur 7: Innbyggertall i Askim, sortert på alder (Statistisk sentralbyrå, 2014c).....	9
Figur 8: E18 Østfold (Statens vegvesen, 2014b).....	9
Figur 9: E18 gjennom Askim.....	11
Figur 10: Monosentrisk by.....	13
Figur 11: komponenter av husleien (DiPasquale & Wheaton, 1996).....	15
Figur 12: Husleiegradient i en voksende by (DiPasquale & Wheaton, 1996).....	15
Figur 13: Husholdningers budfunksjoner (Osland, 2001).....	19
Figur 14: Produsentenes offerfunksjon (Osland, 2001).....	23
Figur 15: Markedslikevekt (Osland, 2001).....	24
Figur 16: Solgte boliger med hensyn på salgspris, 2003-2014.....	47
Figur 17: Gjennomsnittlig boligpriser fra 2003-2014.....	48
Figur 18: Solgte boliger med hensyn på fellesgjeld, 2003-2014.....	49
Figur 19: Solgte boliger med hensyn på boareal, 2003-2014.....	51
Figur 20: Solgte boliger med hensyn på alder, 2003-2014.....	52
Figur 21: Solgte boliger med hensyn på avstand til sentrum, 2003-2014.....	54
Figur 22: Solgte boliger med hensyn på avstand til påkjørsel motorvei, 2003-2014.....	54
Figur 23: Residualplot, multippel lineær regresjon.....	60
Figur 24: Normalfordeling i feilleddet, multippel lineær regresjon.....	63
Figur 25: Residualplot, dobbelt-logaritmisk regresjon.....	67
Figur 26: Normalfordeling i feilleddet, dobbelt-logaritmisk regresjon.....	70
Figur 27: Residualplot, semi-logaritmisk regresjon.....	73
Figur 28: Normalfordeling i feilleddet, semi-logaritmisk regresjon.....	76
Figur 29: Støykart over Askim (Statens vegvesen, 2013).....	81
Figur 30: Boligpris basert på kjøreavstand til motorvei.....	82
Figur 31: Boligpris basert på luftavstand til motorvei.....	83

Tabelloversikt

Tabell 1: Boligprisøkning for 2003-2014 (Statistisk sentralbyrå, 2014b).....	7
Tabell 2: Nettoinnflytting til Askim fra år 2000-2014 (Statistisk sentralbyrå, 2015).....	7
Tabell 3: Lineær regresjon av boa.....	43
Tabell 4: Oversikt over utelatte variabler.....	44
Tabell 5: kategoriske variabler i analysen.....	45
Tabell 6: kontinuerlige variabler i analysen.....	46
Tabell 7: Gjennomsnittlig salgspris for boliger solgt i Askim, sortert på år.....	48
Tabell 8: Gjennomsnittlig fellesgjeld for boliger solgt i Askim, sortert på år.....	50
Tabell 9: Gjennomsnittlig boareal for boliger solgt i Askim, sortert på år.....	51
Tabell 10: Gjennomsnittlig alder for solgte boliger i Askim, sortert på år.....	53
Tabell 11: Korrelasjonsmatrise.....	55
Tabell 12: Multipl linear regresjon.....	58
Tabell 13: VIF-test, multipl linear regresjon.....	59
Tabell 14: Breusch – Pagan test, multipl linear regresjon.....	61
Tabell 15: Robuste standardfeil, multipl linear regresjon.....	62
Tabell 16: Multipl linear regresjon, avstandsdummyer.....	64
Tabell 17: Dobbel-logaritmisk regresjon.....	65
Tabell 18: VIF-test, dobbel-logaritmisk regresjon.....	66
Tabell 19: Breusch – Pagan test, dobbel-logaritmisk regresjon.....	68
Tabell 20: robuste standardfeil, dobbel-logaritmisk regresjon.....	69
Tabell 21: Dobbel-logaritmisk regresjon, avstandsdummyer.....	70
Tabell 22: Semi-logaritmisk regresjon.....	71
Tabell 23: VIF-test, semi-logaritmisk regresjon.....	72
Tabell 24: Breusch – Pagan test, semi-logaritmisk regresjon.....	74
Tabell 25: robuste standardfeil, semi-logaritmisk regresjon.....	75
Tabell 26: Semi-logaritmisk regresjon, avstandsdummyer.....	76
Tabell 27: Verdiendring med avstand som kontinuerlig variabel.....	82
Tabell 28: Verdiendring med avstands-dummyer.....	83

Vedleggsliste

Vedlegg 1: Postnummer med avstand til sentrum	91
Vedlegg 2: Gjennomsnittlig salgpris etter postnummer, 2003-2014	92
Vedlegg 3: Gjennomsnittlig fellesgjeld etter postnummer, 2003-2014	93
Vedlegg 4: Gjennomsnittlig boareal etter postnummer, 2003-2014	94
Vedlegg 5: Gjennomsnittlig alder etter postnummer, 2003-2014	95
Vedlegg 6: Fullstendig korrelasjonsmatrise	96
Vedlegg 7: Fullstendig multipl linear regresjon med avstandsdummyer	98
Vedlegg 8: Fullstendig dobbelt-logaritmisk regresjon med avstandsdummyer	99
Vedlegg 9: Fullstendig semi-logaritmisk regresjon med avstandsdummyer	100
Vedlegg 10: kommandoer i STATA	101

1. Innledning

Når en bolig skal kjøpes må ulike aspekter som boareal, avstand til sentrum, boligtype og eierform vurderes nøye. I den senere tid har beliggenhet, ikke bare til kommunesentrum, men til større byer blitt en veldig viktig faktor ved at pendling opptar mye tid. Både teori og empiri viser at vei og fremkommelighet er viktig og det kan tenkes at det er positivt for en liten by å få bedret fremkommelighet til større kommuner som har flere arbeidsplasser. Ved at en motorvei produserer mye støy og forurensing kan den på en annen side ha negativ innvirkning på boligprisen. Ut i fra disse ulike implikasjonene for boligpris vil det være interessant å se hvilken effekt avstand til E18 har på boligprisene i Askim.

Problemstillingen i denne oppgaven er:

Hva betyr avstand til E18 for boligprisene i Askim?

Ved at oppgaven kaster lys over implikasjoner veitrafikken kan ha for boligområder vil en rekke aktører være interessert i resultatene. Foretak som driver med eiendomsutvikling og boligbygging kan dra nytte av analysen ved plassering av nybygg. Det offentlige, derunder lignende pendlerkommuner, kan benytte analysen til å planlegge hvor fremtidige boligområder skal utvikles og hvor fremtidige motorveier skal plasseres.

Opgaven innledes med litt bakgrunnsinformasjon om Askim, befolkningsveksten i dette området hvor fokus er rettet mot bosted og arbeidssted. Deretter vil utbyggingen av E18 bli belyst.

Videre blir relevant teori presentert som gir en grunnleggende innføring i både teori og begreper som videre vil bli benyttet i oppgaven. De to modellene som blir gjennomgått er Alonso-Muth-Mills-modellen og den hedonistiske prisfunksjonen. Hypoteser basert på modellene blir utformet avslutningsvis.

I neste kapittel vil den økonometriske modellen bli gjennomgått, som forklarer valgt metode. Deretter vil fremgangsmåten for datainnsamlingen bli belyst og det vil bli en kort beskrivelse av datamaterialet.

I analysedelen vil tre regresjonsmodeller bli testet og analysert. For hver regresjonsmodell vil det bli utført to regresjoner. Den første har avstand til motorvei som en kontinuerlig variabel, hvor avstanden er definert som kjøreavstand i meter fra hver enkelt bolig til nærmeste påkjørsel på motorveien. Den andre har kategorisert avstand til motorvei som dummy-variabler, denne avstanden er målt i meter i luftavstand fra hver enkelt bolig til motorveien. Grunnen til at det er delt opp er at en bolig kan ligge nærme veien målt i luftavstand slik at støy og forurensing blir et problem. Til slutt blir den beste modellen valgt, hypotesene blir testet og en konklusjon blir trukket.

2. Bakgrunn

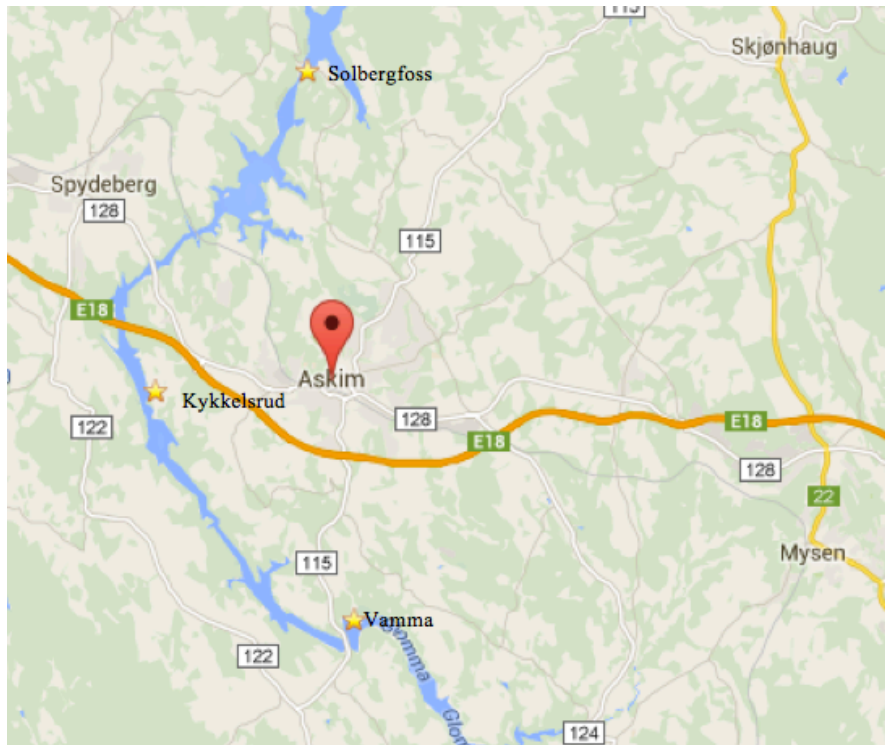
2.1 Askim

Askim ligger i Østfold Kommune og har i følge Statistisk sentralbyrå et innbyggertall på 15.430 personer per 1. Januar 2014 (Statistisk sentralbyrå, 2014a). Du finner Askim øst for Glomma med et areal på 69 km². I 1837 ble Askim kommune opprettet av det kommunale selvstyret, det har ikke vært noen endringer i grensene siden den tid. Et viktig veiskille kom i 1996, da kommunestyret vedtok at Askim skulle ha bystatus. Som sett på kartet nedenfor grenser kommunen i vest mot Spydeberg, i nordøst mot Trøgstad, i sørøst mot Eidsberg og i sør mot Skiptvet.



Figur 1: Kart over Østfold (Østfold fylkeskommune, 2012)

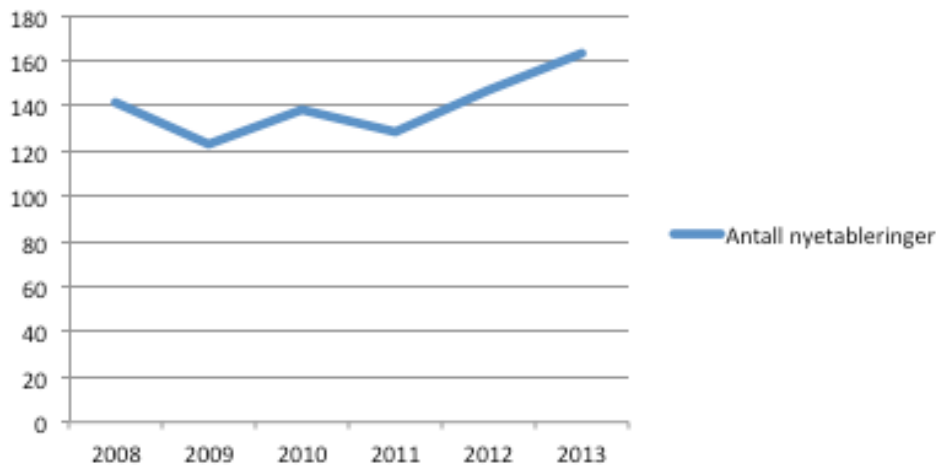
39 prosent av kommunen er dyrket jord og 40 prosent er skog. Dette betyr at de fleste innbyggerne er lokalisert i det som kalles administrasjonssenteret Askim (Askimbyen), som for øvrig er Østfolds fjerde største tettsted. Mindretallet av befolkningen er videre fordelt på tettbebyggelser som Solbergfoss, Vamma og Kykkelsrud. Disse er merket med en gul stjerne i kartet nedenfor.



Figur 2: Kart over Askim inkludert tettbebyggelser

Askim har historisk sett vært en industrikommune. Askim Gummivarefabrikk ble grunnlagt i 1920 og ble ansett som en hjørnesteinsbedrift. Viking Dekk AS som utgjorde den siste delen av Askim Gummivarefabrikk, ble nedlagt i 1991. I 2013 utgjorde industrien kun 9 prosent av kommunens arbeidsplasser, hvor Glava AS og Askim Frukt- og Bærpresseri AS sto for tyngden av næringens sysselsetting. Askim har de senere år blitt et handels- og servicesentrum for hele indre Østfold (Thorsnæs, 2014).

Næringslivsutviklingen i Askim ser ut til å gå riktig vei. I en analyse utført i samarbeid mellom analyseselskapet Dun & Bradstreet og landets fremste økonomiske magasin Kapital har det blitt utviklet en statistisk modell som rangerer norske byers evne til å skape et bærekraftig næringsliv. I 2009 kom Askim på 9. plass i denne analysen. (Askim kommune, 2009). Figur 3 nedenfor viser at det relativt jevnt har blitt skapt nye foretak i Askim i perioden 2009-2013. Tendensen de to siste årene har vært stigende.



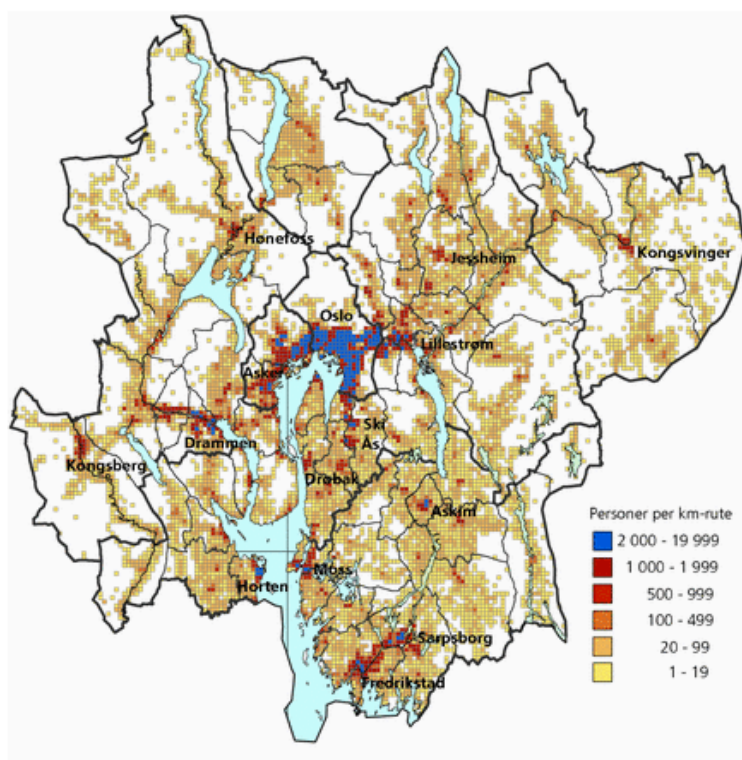
Figur 3: Nyetablerte bedrifter i Askim, tall fra (Statistisk sentralbyrå, 2014e)

Til tross for at Askim ser ut til å ha et bærekraftig næringsliv, har en stor andel av innbyggerne jobb utenfor kommunen. I 2011 var det totalt 7.129 sysselsatte personer bosatt i Askim. Av disse var det kun 3.227 som faktisk arbeider innenfor bostedskommunen. Det vil si at hele 54,7 prosent av de sysselsatte som er bosatt i Askim Pendler, hvor den største andelen pendler til et annet fylke (Statistisk sentralbyrå, 2011c). Etter videre undersøkelse kan det fastslås at 42,1 prosent av denne utpendlingen var til Ski- og Oslo kommune (Statistisk sentralbyrå, 2013b).

2.2 Befolkningsvekst

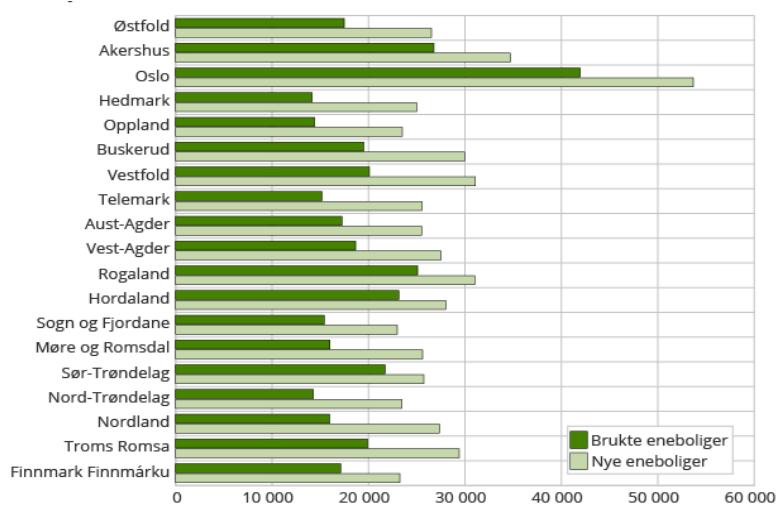
Over 50 prosent av landets befolkningsvekst fra år 2000-2010 har skjedd innenfor en radius på 10 mil fra Oslo. Denne befolkningsveksten inkluderer innvandringsoverskudd, nettoinnflytting fra resten av landet og nesten halvparten av Norges fødselsoverskudd. Hvis rådhuset i Oslo er utgangspunktet så utgjør kommunene som ligger 10 mil herifra kun 6 prosent av landets areal men de innehar 35 prosent av landets befolkning. Dette tilsier at befolkningstettheten i disse områdene er relativt stor i forhold til resten av landet (Høydahl, 2010).

Fra figur 4 nedenfor kan vi se at befolkningstettheten er størst i området mellom Oslo, Asker, Ski og Lillestrøm. Dette er områder som ligger rundt 2-3 mil i fra Oslos rådhus i følge gulesiders karttjeneste. Områder som ligger lenger unna har også en relativt stor befolkningstetthet, et eksempel er Askim som ligger omtrent 5 mil fra Oslo.



Figur 4: Befolkningstetthet i Oslo-regionen 1. Januar 2010 (Høydahl, 2010)

Befolkningstettheten i Oslo gjør at etterspørselen etter boliger øker, noe som fører til at prisene også skyter i været. Det er derfor mange som velger å være mer mobile når det gjelder valg av bosted. Fra figuren nedenfor kan vi lese at prisene på både nye og brukte eneboliger er betraktelig mye høyere i Oslo enn alle andre fylker i 2013. Det kan derfor tenkes at husstander med barn, eller husstander som foretrekker større plass, hage, eller lignende velger å bosette seg for eksempel i Østfold selv om de arbeider i Oslo.



Figur 5: Gjennomsnittlig kvadratmeterpris (Statistisk sentralbyrå, 2014d)

Fra år 2003-2014 har det vært en kraftig økning i boligprisene over hele landet. Tabell 1 nedenfor viser prosentvise endringer i gjennomsnittlig kvadratmeterpris for eneboliger, småhus og blokkleiligheter både i Askim og Oslo. Av tabellen fremkommer det at det er relativt store forskjeller i boligprisutviklingen i Oslo i forhold til Askim når vi ser på alle årene under ett. For eneboliger, småhus og blokkleiligheter har økningen i pris vært generelt større i Oslo enn i Askim med noen unntak. Fra år 2013-2014 har boligprisveksten vært mye høyere for eneboliger og blokkleiligheter i Askim i forhold til Oslo.

Tabell 1: Boligprisøkning for 2003-2014 (Statistisk sentralbyrå, 2014b)

	2003-2004	2005-2006	2007-2008	2009-2010	2011-2012	2013-2014	2003-2014
Enebolig i Askim	1,8 %	6,3 %	0,1 %	1,6 %	-2,2 %	10,5 %	67,9 %
Småhus i Askim	11,7 %	3,6 %	10,6 %	4,2 %	13,6 %	-1,1 %	90,5 %
Blokkleilighet i Askim	8,4 %	3,5 %	6,1 %	9,5 %	14,3 %	10,1 %	97,9 %
Enebolig i Oslo	8,4 %	10,2 %	1,8 %	8,2 %	6,6 %	1,6 %	110,7 %
Småhus i Oslo	10,8 %	15,0 %	0,1 %	11,7 %	6,4 %	2,0 %	133,9 %
Blokkleilighet i Oslo	11,8 %	16,6 %	-6,0 %	8,0 %	8,1 %	2,0 %	118,7 %

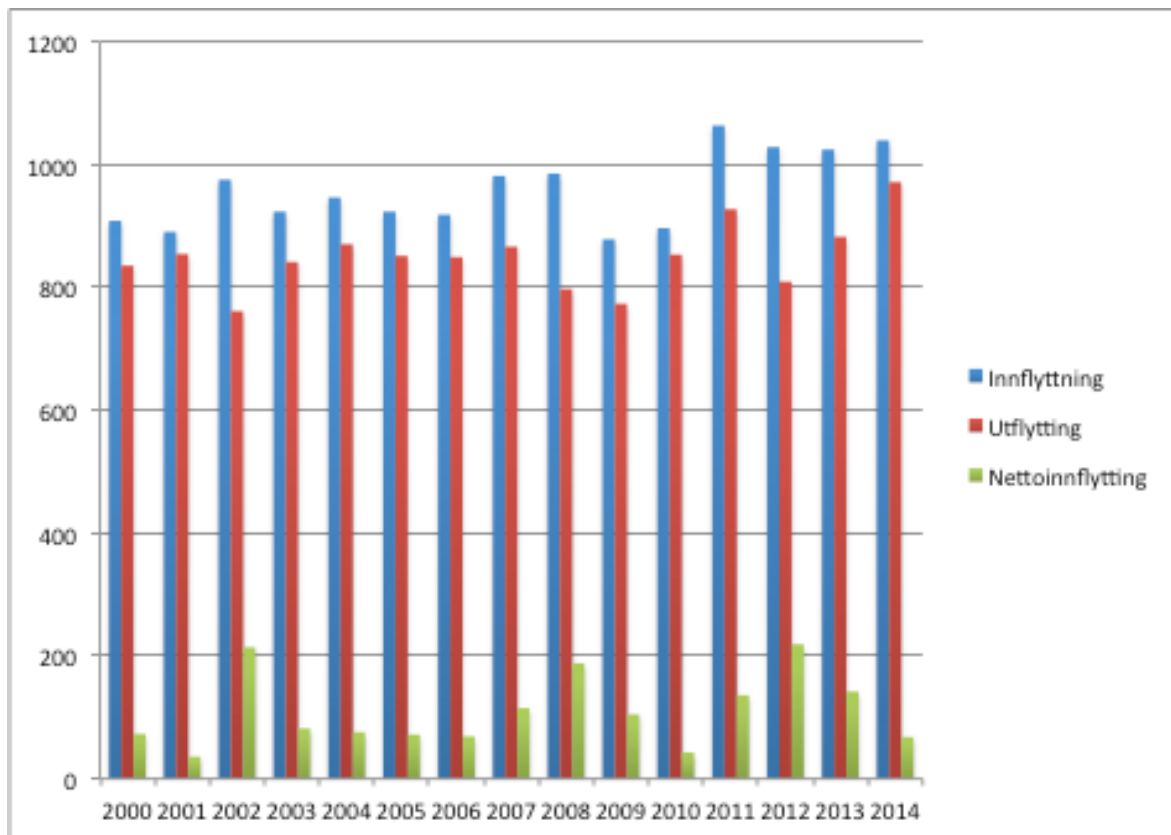
Med denne prisutviklingen i friskt minne kan det være interessant å se på hvordan bosetningsmønsteret i Askim har vært. Tall fra Statistisk sentralbyrå kan bekrefte at det er mange som bosetter seg utenfor Oslo selv om de er sysselsatt i Oslo. I 2011 var det totalt 425.299 personer som var sysselsatt i Oslo, av disse var det 163.108 personer som pendlet inn til hovedstaden fra et annet fylke, dette tilsvarer 37,5 prosent innpendling (Statistisk Sentralbyrå, 2011b). I forrige avsnitt ble det informert om at 42,1prosent av de som pendler fra Askim arbeider i Ski eller Oslo kommune. Selv om dette kun utgjør en brøkdel av alle som pendler inn til Oslo hver dag vil det være interessant å se på tall som viser antall innflyttinger og utflyttinger fra Askim gjennom tidene. Tabellen nedenfor viser innflyttinger, utflyttinger og nettoinnflyttinger fra år 2000-2014.

Tabell 2: Nettoinnflytting til Askim fra år 2000-2014 (Statistisk sentralbyrå, 2015)

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014
Innflyttinger	908	890	975	923	946	923	918	981	985	878	896	1063	1028	1024	1039
Utflyttinger	835	854	761	841	870	851	849	866	797	773	853	927	809	882	971
Nettoinnflytting	73	36	214	82	76	72	69	115	188	105	43	136	219	142	68

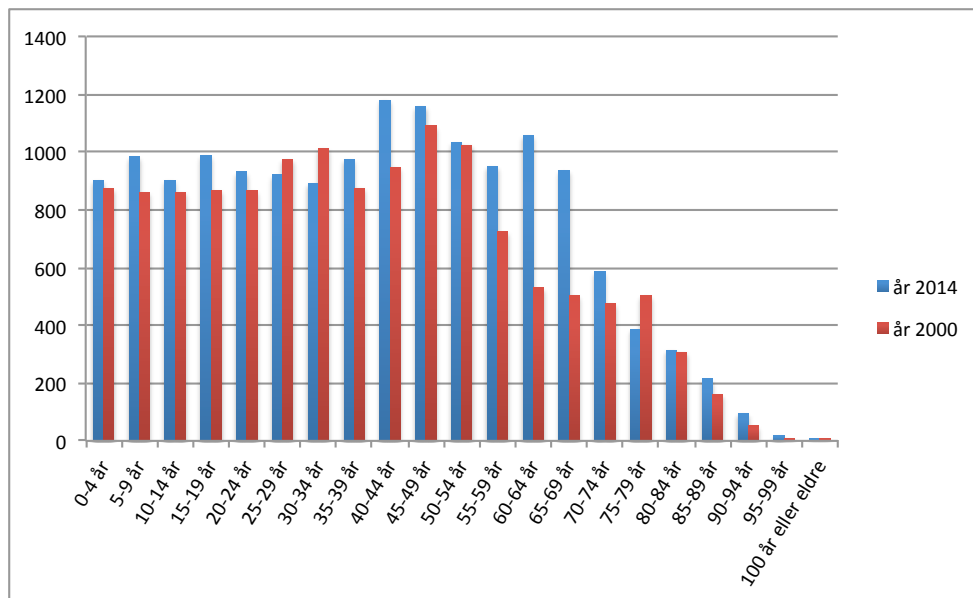
Tallene indikerer at innflyttingen de siste 15 årene har vært relativt jevnt stigende med unntak av noen år. De siste fire årene har antall innflyttinger vært over 1000 personer, noe som er

relativt mange i forhold til tidligere år. Antall utflyttinger har variert en del, men ligget stabilt på mellom 761-971 personer per år. Nettoinnflyttingen, som er innflyttinger minus utflyttinger, har vært positiv de siste 15 årene. I 2002 var nettoinnflyttingen relativt stor i forhold til de nærliggende årene, innflyttingen var rimelig høy og utflyttingen relativt lav. Fra 2007-2013, med unntak av 2010, kan det vises at nettoinnflyttingen har vært høy. Figuren nedenfor viser dette mer visuelt.



Figur 6: Nettoinnflytting til Askim fra år 2000-2014 (Statistisk sentralbyrå, 2015)

Det finnes ingen målinger på hvilke aldersgrupper som flytter inn og ut. Som en indikator kan det være interessant å se på hvilke aldersgrupper som bodde i Askim i år 2000 i forhold til i år 2014. En kan anta at det er en naturlig økning i alle aldersgrupper basert på at det har vært en befolkningsvekst. Fra figur 7 nedenfor kan vi se at veksten klart har vært størst i gruppene 40-44 år og 55-69 år. Det har vært en nedgang i aldersgruppen 25-34 år. Ut i fra dette er det vanskelig å konkludere hvem som faktisk flytter til og fra Askim, men det kan tenkes at barnefamilier som eksempelvis trenger mer plass velger å bo et annet sted enn der de arbeider.



Figur 7: Innbyggertall i Askim, sortert på alder (Statistisk sentralbyrå, 2014c)

2.3 Europavei E18

Den 24. februar i år 2000 ble «Østfoldpakka» vedtatt i Stortinget, en ny E18 gjennom Østfold skal bygges. Den planlagte strekningen vil utgjøre rundt 70 km og vil strekke seg fra Riksgrensen, gjennom Marker-, Eidsberg-, Trøgstad-, Askim-, Spydeberg-, Hobøl-, Ski- og Ås kommune. Vegen vil bli bygget i 9 forskjellige traseer. Kartet nedenfor viser hvilken rekkefølge de ulike trassene skal bygges ved hjelp av nummerering. Heltrukken gul linje indikerer at traseen er ferdigstilt og er i bruk, stiplet gul linje viser at byggingen har startet, mens rød heltrukken linje viser at vegen ikke er påbegynt.



Figur 8: E18 Østfold (Statens vegvesen, 2014b)

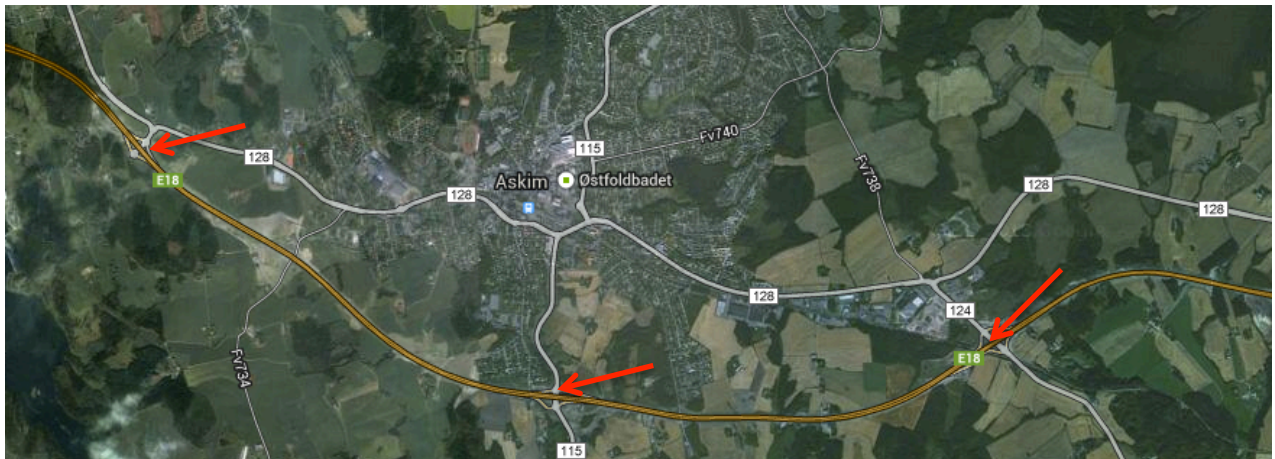
En rask og oversiktlig fremstilling som viser årstall for de ulike traseene er nødvendig her:

1. Marker grense – Melleby ble ferdigstilt i oktober 1998
2. Ørje – Eidsberg grense ble ferdigstilt oktober 2003
3. Sekkelsten – Krosby ble åpnet for trafikk i 2005
4. Momarken – Sekkelsten ble åpnet for trafikk i 2007
5. Krosby – Knapstad og åpnet for trafikk i November 2010
6. Melleby – Momarken ble åpnet for trafikk i Oktober 2014
7. Knapstad – Retvet er forventet ferdig høsten 2016
8. Riksgrensen – Ørje er ikke påbegynt
9. Retvet – Vinterbro er ikke påbegynt

Formålet med Østfoldpakken er at den skal «Bedre fremkommeligheten og skape en sikrere E18 gjennom Østfold ved å legge denne utenom tettstedene» (Statens vegvesen, 2014a).

Google Maps informerer at den fysiske kjøreavstanden ikke har blitt redusert dersom du benytter den nye motorveien. Den gamle veien er 28,8 km lang og uten trafikk vil det ta 32 minutter fra Askim til Ski, med ny E18 tar det 28 minutter men avstanden er 29,5 km. Fremkommeligheten har blitt bedret ved at det er mindre risiko for kø med en firefelts motorvei, grunnet høyere kvalitet på veien er sikkerheten også mer ivaretatt.

Ved at trafikksikkerheten og fremkommeligheten har blitt bedret vil det være interessant å se på hvordan boligprisene i Askim reflekterer dette. Til tross for at motorveien går utenfor sentrum, viser satellittbildet nedenfor at endel boligområder ikke ligger uforstyrret av trafikken. Askim består av endel jordbruk slik at det finnes mange flate områder hvor lyd kan reise relativt langt. Det kan dermed tenkes at motorveien har en positiv effekt på boligprisene ved at den gir økt fremkommelighet, men at den også drar med seg noen negative ringvirkninger som støy og forurensing. I følge helsedirektoratet har avstanden til støykilden liten betydning da det er støyopplevelsen og desibelnivået som avgjør om noe kan klassifiseres som nabostøy. Topografien kan som i dette tilfellet være relativt flat slik at støyen bæres over lengre avstander (Gersh, 2013). De røde pilene i satellittbilde nedenfor viser hvor de forskjellige påkjørslene til motorveien ligger.



Figur 9: E18 gjennom Askim

3. Teori

I dette kapittelet vil det først bli en kort diskusjon om hva som bestemmer prisen på en bolig. Deretter vil det bli en gjennomgang av to modeller som forklarer boligprisen på hver sin måte. Alonso-Muth-Mills-modellen forklarer hvordan boligprisene varierer med avstand til sentrum, mens den hedonistiske prisfunksjonen som sier noe om hvordan forskjellige attributter ved boligen kan være med å forklare boligprisen. Avslutningsvis vil det bli formulert hypoteser som senere i oppgaven skal bli testet.

3.1 Hva bestemmer prisen på en bolig?

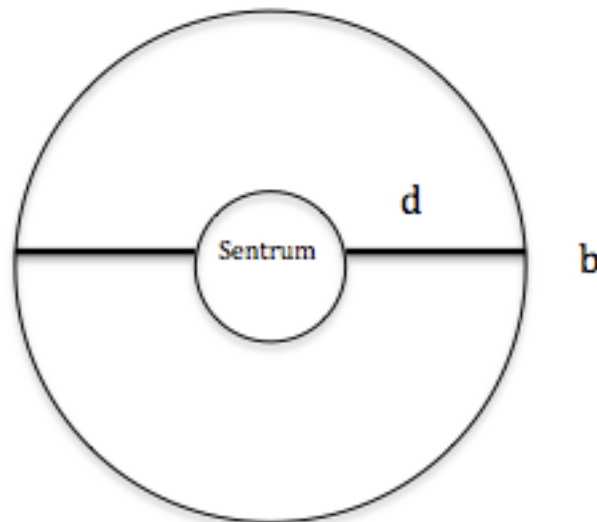
Prisutviklingen på boliger har en stor påvirkning på velstanden til norske boligeiere, ved at hele 80 prosent av husholdningene eier en bolig (Statistisk sentralbyrå, 2007). Det er mye som kan forårsake endringer i etterspørselen og prisen etter boliger. Her kan for eksempel befolkningsendring, inntektsnivå, tilgang til kreditt, rentenivå, skatt og forventinger til fremtiden nevnes som viktige faktorer. Noen av disse faktorene kan fort endre seg; et godt eksempel på dette er når bankene setter ned boliglånsrenten stimulerer dette til høyere boligpriser. Dette kan sirkulere ved at høyere boligpriser øker lånebehovet. Den høye boligprisen sikrer bankene mot tap ved at de tar pant i objektet som har en økt verdi, og når husholdningene får høyere lån vil de by opp prisen på boligene (Anundsen & Jansen, 2013).

To boliger kan aldri være helt identiske ved at de aldri kan oppta samme plass. Boligmarkedet skiller seg således ut i fra andre markeder ved at hver enkelt bolig er et heterogent og immobilt gode. To boliger som fremstår som identiske ved at de har like attributter kan ha to forskjellige priser ved at den ene for eksempel opptar en mer attraktiv tomt. Når dette er sagt kan vi også ha tilfeller av at to helt ulike boliger blir solgt for den samme prisen.

Tilbudet av nye boliger endrer seg relativt lite i løpet av et år (cirka 1 prosent), det antas dermed at tilbudet er gitt og dermed uelastisk på kort sikt. Etterspørselen etter boliger er utledet ved å sortere alle i markedet ut i fra betalingsvilje. På kort sikt er det dermed etterspørselssiden som bestemmer boligprisen (Emblem, 2014).

3.2 Alonso – Muth – Mills – modellen

Alonso-Muth-Mills-modellen tar for seg det urbane tomtemarkedet og forklarer forskjellen i boligprisene ved hjelp av avstand til sentrum. Denne modellen som er utledet av DiPasquale og Wheaton (1996) sier at boligprisene blir høyere jo nærmere sentrum du bor. For å illustrere dette må det tas noen forutsetninger som forenkler virkeligheten. Vi har en monosentrisk by som innebærer at alle arbeidsplasser er lokalisert i sentrum. Byen er formet som en sirkel hvor alle husholdninger pendler en rett linje inn til sentrum. Byggningsstrukturen er gitt basert på historisk bygging, slik at det er ikke mulig å erstatte tomter med høyere bygninger. Alle tomtene er differensierte ved at hver enkelt tomt er unik. Tilbud på et bestemt avgrenset sted er gitt, mens etterspørselen er elastisk og bestemmer dermed prisen. Dette innebærer at husholdninger med høyest betalingsvillighet vil okkupere boligen.



Figur 10: Monosentrisk by

I denne modellen er alle husholdningene identiske slik at antall pendlere per husholdning er gitt. Inntekten (y), som er eksogent gitt, blir brukt på transportkostnader (kd), husleie $R(d)$, samt annet konsum (x^0). Transportkostnadene fremkommer ved at arbeiderne pendler til og fra sentrum langs en rett linje (d) til en transportkostnad k per km. Husleien $R(d)$ varierer med avstand til sentrum. Husleietjenester er gitt ved et fast antall tomteareal q per hus og en fast byggekostnad (c). Boligtettheten blir dermed $1/q$. Siden alle husholdningene er identiske vil konsumet av andre goder (x^0) være konstant uansett hvilken lokalisering husholdningen har

($x = x^0$). Forskjeller i husleie vil dermed tilsvare forskjeller i pendlekostnader, husleien vil synke med økningen i avstand fra sentrum. Markedet er nå i likevekt og husleie-likningen kan uttrykkes slik:

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

Når du bor i sentrum så vil avstanden til sentrum være lik null ($d=0$), slik at transportkostnadene forsvinner. Husleien vil dermed være høyest her og husleie-likningen blir komprimert til $R(0) = y - x^0$. Ved bevegelse ut i fra sentrum ser vi at husleien reduseres ettersom vi får økte transportkostnader. Husleien vil være lavest på bygrensen (b) siden her er transportkostnadene høyest ved at pendleravstanden til sentrum er størst. Utenfor bygrensen er jordbruk alternativ bruk av areal og avkastninger per mål er r^a . Med fast arealbruk kan hver tomt leies ut til $r^a q$. Husleien på bygrensen ($d=b$) vil dermed være leie av tomt $r^a q$ samt annuiteten av byggekostandene som er leie av selve bygget (c). Dette kan uttrykkes slik:

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

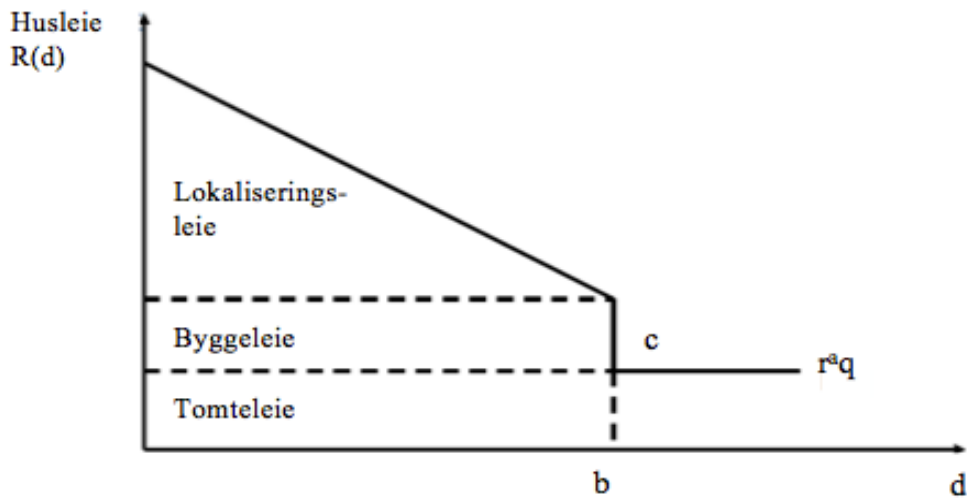
Husholdningenes forbruk på annet konsum får vi ved å sette $R(b) = R(d)$ og $d=b$ og løse med hensyn på x^0 . Husholdningene på bygrensen vil dermed konsumere andre goder basert på denne likningen:

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

Ved å sette dette uttrykket inn i likning (3.1) vil vi få at husleien uansett lokalisering vil være summen av tomteleie, byggeleie og sparte pendlekostander ettersom hvor langt fra sentrum boligen er plassert.

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

Disse komponentene som husleien består av kan fremstilles grafisk som vist i figuren nedenfor. Lokaliseringsrenten ($k(b-d)$) er den komponenten som varierer med avstand til sentrum, mens byggeleien (c) og tomteleien ($r^a q$) er konstant uavhengig av hvor boligen er plassert. X-aksen viser avstand fra sentrum (d), mens Y-aksen viser husleien $R(d)$.



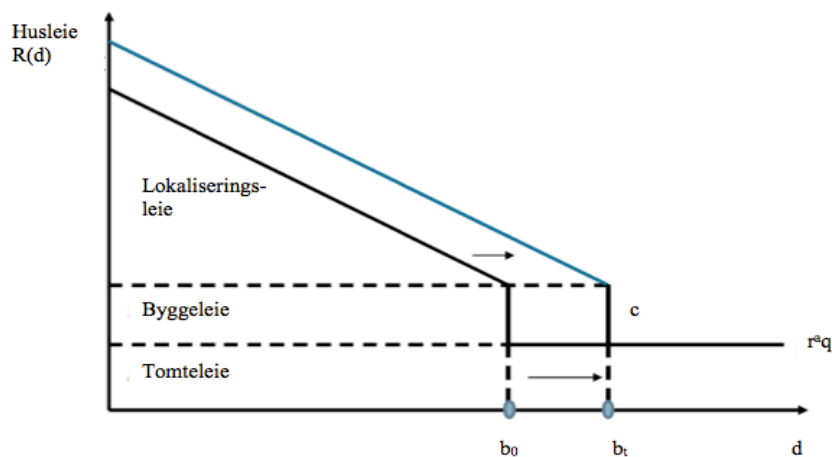
Figur 11: komponenter av husleien (DiPasquale & Wheaton, 1996)

Husleiegradienten som er helningen på lokaliseringsleien får vi ved å derivere likning (3.4) med hensyn på d :

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

Dette forteller oss at lokaliseringsleien vil være fallende ettersom avstanden mellom bygrensen og sentrum ($b-d$) blir lavere. Jo nærmere bygrensen boligen ligger, desto lavere lokaliseringsleie må du betale.

Byer vokser og utvides etter hvert som befolkningen øker. Dette kan vises ved at bygrensen skifter fra b_0 til b_t som vist i figuren nedenfor.



Figur 12: Husleiegradient i en voksende by (DiPasquale & Wheaton, 1996)

En utvidelse fra b_0 til b_t vil føre til at husleien øker for alle innenfor bygrensen, ved at lokalisierungsleien øker. Den prosentvise økningen vil være størst i utkanten av byen og minst i sentrum. Ved hjelp av variabelen tid (t), kan vi uttrykke husleien på ethvert tidspunkt:

$$R_t(d) = (r^a q + c) + k(b_t - d), \quad \text{hvor } d \leq b_t \quad (3.6)$$

I virkeligheten er det ikke alltid slik at alle byer danner en sirkel rundt bysentrum. Variabelen v viser hvilken topografi byen har. Dersom det er mulig å bosette seg i en 360 graders sirkel rundt bysentrum vil $v = 1$, mens $v < 1$ tilsier at bosetningen rundt bysentrum ikke er en hel sirkel som følge av for eksempel sjø. Når antall husholdninger (n) og tomtestørrelsen per bolig (q) inkluderes vil vi kunne utforme uttrykket $b = \sqrt{\frac{nq}{\pi v}}$, som sier oss hvor bygrensen vil gå. Bygrensen vil øke dersom antall husholdninger og tomtestørrelsen per bolig øker eller når sirkelen rundt bysentrum avtar.

3.3 Hedonistisk prisfunksjon

Alonso – Muth – Mills – modellen gir ikke et realistisk bilde av boligprisene ved at det kun er avstand til sentrum bestemmer prisene. Den hedonistiske prisfunksjonen utviklet av Rosen (1974) inkluderer andre attributter ved boligene slik at boligprisene vil variere i flere dimensjoner enn kun avstanden til sentrum (Osland, 2001). I følge Osland er denne metoden mye brukt i ulike studier av boligmarkedet, ved at den tar for seg boliger som heterogene goder med ulike egenskaper/attributter.

Ved et boligkjøp avhenger prisen normalt av flere attributter, dette kan eksempelvis være størrelse, kvalitet, antall rom, hage, garasje, utsikt, støy og så videre. I denne oppgaven vil «nærhet til motorvei» være en sentral attributt i analysen. En boligs attributter kan uttrykkes ved hjelp av en vektor som består av n ulike attributter (Osland, 2001). Dette kan uttrykkes slik:

$$Z = (Z_1, \dots, Z_n) \quad (3.7)$$

Markedsprisen på boligen, som blir betraktet som et heterogent gode, blir utledet som en funksjon av godets sammensetning av attributter. Denne funksjonen blir gjerne kalt for den hedonistiske prisfunksjonen (Robertsen & Theisen, 2010):

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

Rosen (1974) mener at suksessen bak modellen er at den innehar markedsmekanismer mellom kjøperne og selgerne av de heterogene boligene (Robertsen & Theisen, 2010). Som vi kommer inn på nedenfor vil etterspørsel bli formulert i «budfunksjoner» og tilbud i «offerfunksjoner» for hvert enkelt attributt.

3.3.1 Likevekt på etterspørselssiden av markedet

På etterspørselssiden finner vi husholdningene. Nytten (U) for husholdning (j) vil bli basert på bakgrunn av boligens attributter (Z), andre konsumvarer enn boligen (X) og husholdningens preferanser (α_j). Det går ut i fra at nyttefunksjonen er stengt konkav.

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

Hver husholdning vil ha en ikke-lineær budsjettrestriksjon som setter en begrensning på nytten. Teorien antar at hver husholdning kun kjøper en bolig, og at boligen brukes til eget konsum.

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

Y_j angir inntekt målt i enheter av X for husholdning j og $P(Z)$ er den hedonistiske prisfunksjonen. Prisen på andre goder settes her lik 1. Teorien bygger på at den første- og andrederiverte av den hedonistiske prisfunksjonen $P(Z)$ finnes, men med ubestemt fortegn. Formen antas å være slik at en entydig indre løsning på nyttemaksimeringsproblemet finnes. Husholdningene tilpasser seg slik at nytten maksimeres gitt budsjettrestriksjonen. Vi finner optimum ved å partiellderivere nyttefunksjonen (3.9) med hensyn på Z_i og X , uttrykket vi får vil se slik ut:

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

Venstre side av likning (3.11) viser den marginale substitusjonsraten mellom Z_i og X , mens høyre side viser den partiellderiverte hedonistiske prisfunksjonen med hensyn på de ulike attributtene. I optimum skal dermed den marginale substitusjonsraten mellom Z_i og X være lik den partiellderiverte hedonistiske prisfunksjonen med hensyn på de ulike attributtene.

Budfunksjonen

Definisjon av budfunksjonen er at den viser maksimal betalingsvillighet for ulike hustyper eller sammensetninger av attributtvektorer, når nyttenivået og inntekten holdes konstant (Osland, 2001). Alonso (1964) utredet grunnlaget for budfunksjonen i forbindelse med tomteareal. Med utgangspunkt i de optimale verdiene for boligvektoren Z^* og konsum av andre varer X^* , kan budfunksjonen utledes slik:

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

Setter vi dette inn i nyttefunksjonen får vi dette uttrykket:

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

Videre lar vi inntekten være gitt og nyttenivået være konstant lik U^* . Den maksimale betalingsvilligheten θ_j vil da være lik den prisen man faktisk betaler $P(Z^*)$. Nyttefunksjonen kan således uttrykkes slik:

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

Dette uttrykket viser maksimal betalingsvillighet ved andre attributtsammensetninger enn hva som er det optimale. Disse ulike kombinasjonene vil oppfattes som likeverdige ved at det beregnes en subjektiv pris som gjør at inntekten brukes opp slik at husholdningene fortsetter

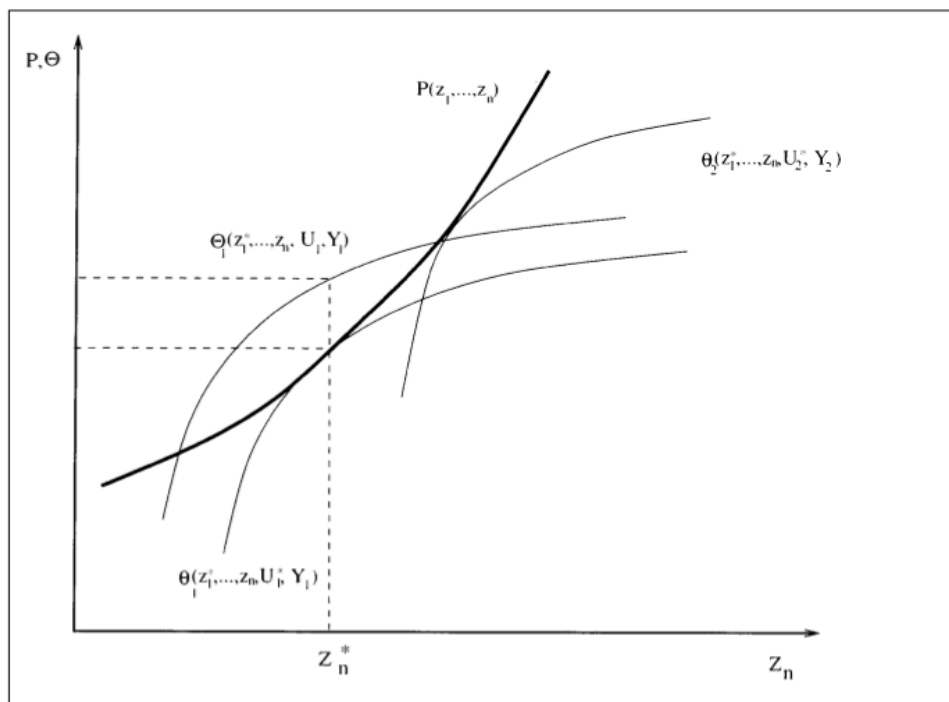
å være på det samme optimale nyttenivået. På grunn av variasjoner i inntekt- og nyttenivå, kan budfunksjonen mer generelt uttrykkes slik:

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

Ved implisitt derivasjon kan maksimal betalingsvillighet for en partiell økning i et boligattributt uttrykkes slik:

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

Så lenge nyttefunksjonen er strengt konkav kan vi dobbelt derivere og vise at $\frac{\partial^2 \theta_j}{\partial Z_i^2} < 0$. Dette innebærer at betalingsvilligheten er positiv men avtagende ved partielle økninger i boligattributter. Budfunksjonen gir et sett av indifferenskurver til hvert nyttenivå, dette kan vises grafisk:



Figur 13: Husholdningers budfunksjoner (Osland, 2001)

Figuren ovenfor viser kroner langs den vertikale aksen og ulike boligattributter langs den horisontale aksen. Det antas at konsumenten er optimalt tilpasset i alle attributter bortsett fra Z_n som i dette eksempelet er boligareal (Osland, 2001). Nyttensnivået øker ved bevegelse nedover i diagrammet, og nyttemaksimeringen finner sted der sammensetningen av boligattributter er slik at man kommer på den lavest mulige budkurven (θ). Hver husholdning har ulike preferanser (α) og dermed også ulike nyttefunksjoner og budfunksjoner. Dette vises i figur 13 ved at husholdningen med budfunksjonen (θ_2) har preferanser for et større boligareal i forhold til husholdningen med budfunksjonen (θ_1).

Den eksogent gitte hedonistiske prisfunksjonen vises ved hjelp av den konvekse kurven i figuren ovenfor, den stiger ved en partiell økning i boligareal. Likevekt for konsumentene således være der prisfunksjonen tangerer med den lavest oppnåelige budfunksjonen. Likevekten er gitt ved å sette likning 3.11 lik 3.16:

$$\frac{\partial \theta_j}{\partial Z_n} = \frac{\frac{\partial U_j}{\partial Z_n}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_n} \quad j=1, \dots, m \quad (3.17)$$

Tolkningen av dette er at i nyttemaksimeringen vil den marginale betalingsvilligheten være lik den implisitte prisen på attributtet. Helningen på de to kurvene må være identiske. I tillegg til dette krever likevekt at $\theta(Z^*, Y_j, U_j, \alpha_j) = P(Z)$, som tilsier at $P(Z)$ er det minste beløpet husholdningene må betale på markedet for en bolig med attributtvektoren Z .

3.3.2 Likevekt på tilbudssiden av markedet

Som vi nettopp har sett tilpasser etterspørselssiden seg slik at nytten maksimeres. På den andre siden har vi mange små bedrifter som er tilbyderne av boliger i markedet, de vil tilpasse seg slik at profitten maksimeres. Det antas at hver bedrift spesialiserer på og produserer en boligtype med en gitt sammensetning av attributter. Profittfunksjonen til hver enkelt bedrift kan defineres slik:

$$\Pi = M \times P(Z) - C(M, Z, \beta) \quad (3.18)$$

Profittfunksjonen er gitt ved inntekter minus kostnader. Inntektsfunksjonen er ikke-lineær og er definert ved antall boliger M multiplisert med den hedonistiske prisfunksjonen $P(Z)$.

Prisfunksjonen er gitt og er uavhengig av antall boliger som blir produsert.

Kostnadsfunksjonen (C) er en funksjon av antall boliger (M), attributter (Z) og en vektor av skiftparametere (β). Hvor skiftparameterne for eksempel representerer faktorpriser eller produksjonsteknologi for den enkelte bedrift. Kostnadsfunksjonen vil være en konveks stigende funksjon av antall boliger som blir produsert og grensekostnadene i produksjon av attributter antas å være positive og ikke-avtagende. Ved at det finnes mange små tilbydere vil det være kontinuerlig variasjon i attributter ved boligene som tilbys. Førsteordens betingelse for maksimal profitt er gitt ved:

$$\frac{\partial P}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} \quad i=1, \dots, n \quad (3.19)$$

Denne betingelsen sier at den implisitte prisen for attributt i er lik grensekostnaden ($\partial C / \partial Z_i$) per bolig (M) med en partiell økning i mengden boligattributter. Videre kan det vises at bedriften bør produsere et antall boliger slik at grenseinntekt, prisen på boligen, er lik grensekostnader i produksjon av boliger (Osland, 2001):

$$P(Z) = \frac{\partial C}{\partial M} \quad (3.20)$$

For å sikre andreordensbetingelse for maksimum er det ikke tilstrekkelig å kun anta at

kostnadsfunksjonen er konveks, man må i tillegg forutsette at vi har $\frac{\partial^2 C}{\partial Z_i^2} > \frac{\partial^2 P}{\partial Z_i^2}$ i det

relevante området for maksimum.

Offerfunksjonen

Offerfunksjonen $\Phi(Z, \pi, \beta)$ er en viktig komponent på tilbudssiden av markedet. Det er den minste prisen produsentene er villig til å akseptere for å tilby boliger med ulike attributter, til et konstant profittnivå og gitt et optimalt antall boliger som produseres (Osland, 2001). I utledningen av offerfunksjonen tas det dermed utgangspunkt i de optimale verdiene med hensyn på ulike attributter Z^* , tilbud av boliger M^* og profitt π^* . Dette gir denne profittfunksjonen:

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

Dersom profittnivået holdes konstant, kan profittfunksjonen uttrykkes slik (Rosen, 1974):

$$\pi = M\Phi(Z^*, \pi^*, \beta) - C(M^*, Z^*, \beta) \quad (3.22)$$

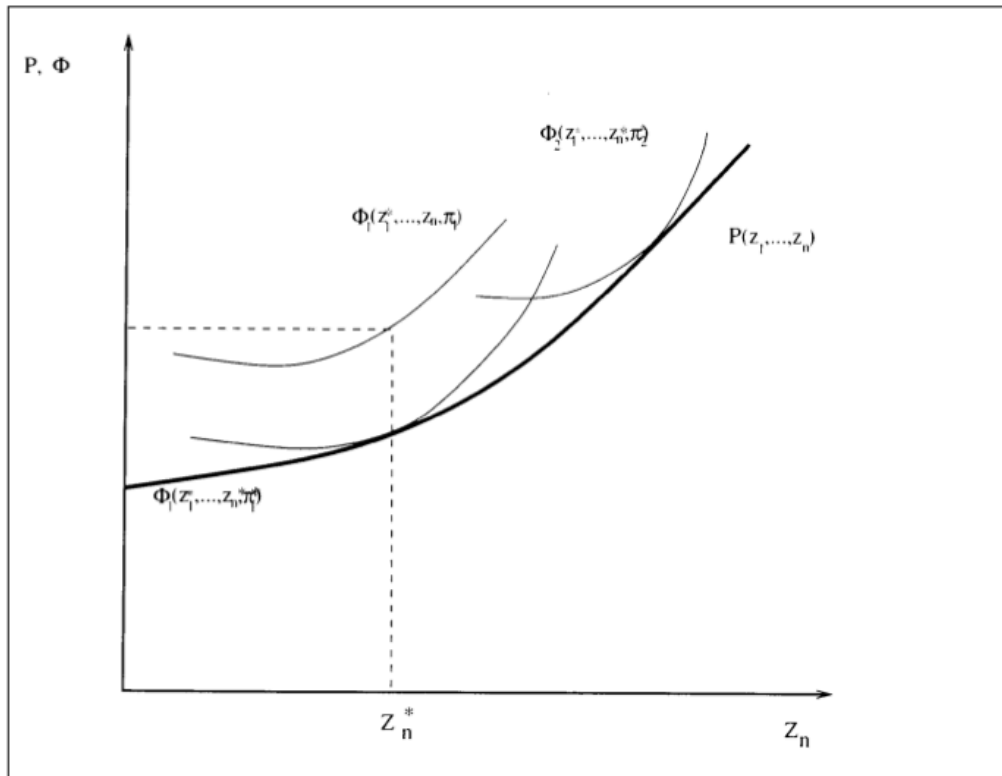
Ved å derivere denne profittfunksjonen med hensyn på M og Z_i kan førsteordensbetingelsene bli funnet:

$$\Phi(Z^*, \pi^*, \beta) = \frac{\partial C}{\partial M} \quad (3.23)$$

$$\frac{\partial \Phi}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} \quad i=1, \dots, n \quad (3.24)$$

Av likning 3.23 kan vi se at grensekostnaden ved å produsere en bolig skal være lik den minste prisen som produsentene er villige til å akseptere. Dersom vi løser denne med hensyn på M og setter uttrykket inn i likning 3.22 vil M elimineres, og profittfunksjonen definerer dermed implisitt relasjon mellom offerpriser og boligattributter (Osland, 2001):

$$\Phi = \Phi(Z, \pi^*, \beta) \quad (3.25)$$



Figur 14: Produsentenes offerfunksjon (Osland, 2001)

Figuren ovenfor viser at optimal tilpasning for hver enkelt produsent vil være der hvor offerfunksjonen tangerer med den eksogent gitte prisfunksjonen. Offerkurvene er konvekse og profittnivået øker ved bevegelse oppover i diagrammet ved at $\frac{\partial \Phi}{\partial \pi} > 0$. Skiftparameteren (β) har innvirkning på hvor mye produsenten tilbyr av attributtet Z_n . En produsent med offerfunksjon Φ_1 vil tilby relativt mindre boligareal enn en produsent med offerfunksjon Φ_2 .

For å oppnå likevekt på tilbudssiden må offerkurvene til hver produsent tagere den eksogent gitte prisfunksjonen, dette finner vi fra førsteordensbetingelsene i ligning (3.19) og (3.24):

$$\frac{\partial \Phi}{\partial Z_n} = \frac{\frac{\partial C}{\partial Z_n}}{M} = \frac{\partial P}{\partial Z_n} \quad (3.26)$$

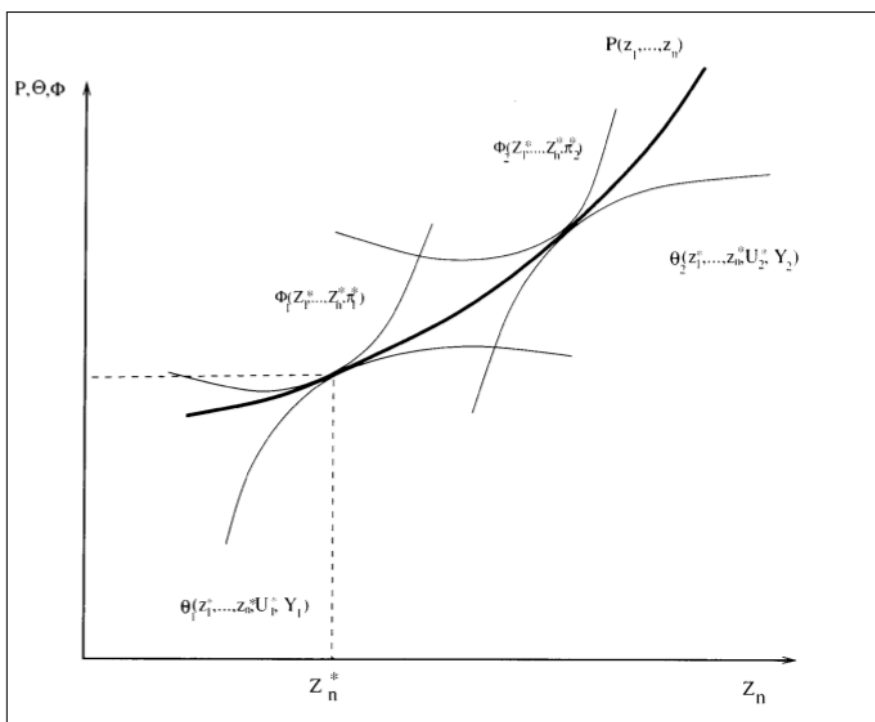
I likevekt kreves det også at offerprisen skal være lik den eksogent gitte prisfunksjonen $\Phi(Z^*, \pi^*, \beta) = P(Z^*)$ (Osland, 2001).

3.3.3 Markedslikevekt

Markedslikevekt oppstår ved at husholdningenes budfunksjon tangerer produsentenes offerfunksjon:

$$\frac{\partial \theta}{\partial Z_i} = \frac{\partial P}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} = \frac{\partial \Phi}{\partial Z_i} \quad (3.27)$$

Den hedonistiske prisfunksjonen inkluderer således både konsumentenes budfunksjoner og produsentenes offerfunksjoner. Markedslikevekten er illustrert i figuren nedenfor:



Figur 15: Markedslikevekt (Osland, 2001)

3.4 Eksternaliteter

Gjennom tidene har eksternaliteter fått en mer sentral rolle når vi ser på boligpriser. Kvaliteten på det urbane livet er påvirket av generell estetikk av det fysiske miljøet, dette innebærer hvordan utbyggingen i samfunnet ser ut samt utforming og bruk av offentlige rom. Eksternaliteter kan ses på som handlinger som enten har en positiv eller negativ innvirkning på andre (DiPasquale & Wheaton, 1996).

Det er blitt gjennomført mange studier som tar for seg hvordan boligprisene reagerer på investeringer i infrastruktur. Flyplasser og tunge industrianlegg er kjent for å produsere støy og forurensning, siden tidlig 1970-tallet har økonomer studert hvilken virkning slike anlegg har på nærliggende boliger. Generelle funn tyder på at boligene som blir utsatt for dårligere luftkvalitet kan redusere boligprisene med rundt 15 prosent (Nelson, 1978). På samme måte har støy fra flyplasser vist å redusere boligprisene med 5-10 prosent (Mieszkowski & Saper, 1978). Motorveier kan på to forskjellige måter også gi positive og negative eksterne virkninger i nærliggende områder. De kan øke eiendomsverdiene ved at transportnettverket gir økt tilgjengelighet til for eksempel sentrum og arbeidsplasser, de kan også være med å redusere eiendomsverdiene ved at umiddelbar nærhet til motorveier kan føre til økt støy og luftforurensning (DiPasquale & Wheaton, 1996).

Ulike undersøkelser i forskjellige land støtter ideen om at ikke-eksistens av stillhet er en av de høyest rangerte miljøproblemene vi står ovenfor i verden. I Sverige er veitrafikk en av de vanligste kildene til støy og omkring 1,6 millioner mennesker blir berørt. Det her i denne sammenheng blitt gjennomført en studie som estimerer en fordelsvurdering av en reduksjon av trafikkstøy som for eksempel kan være nyttig i ulike typer investeringsanalyser av vei- og boligutbygging (Wilhelmsson, 2000). Det som gjør denne studien interessant er at den sikter inn på problemstillingen ved at den inkluderer støy-variabelen i den hedonistiske prisfunksjonen. Resultatene av den empiriske analysen indikerer at det er en total rabatt på 30 prosent på hus som ligger på et sted med mye støy, sammenlignet med et hus som ligger i et rolig område.

3.5 Hypoteser

I dette avsnittet blir det formulert hypoteser som er basert gjennomgått teori og problemstilling.

Hypotese 1: Kjøreavstand til nærmeste påkjørsel på motorveien

Alonso-Muth-Mills-modellen som er beskrevet i avsnitt 3.2 tar utgangspunkt i at prisen på en bolig blir bestemt ut i fra sparte transportkostnader. Redusert tidsbruk og reisekostnader blir reflektert i betalingsvilligheten. Dette utgangspunktet blir adoptert til hypotese 1, hvor det vil bli undersøkt om boliger som har kortere kjøreavstand målt i meter til nærmeste påkjørsel har en høyere pris enn de som må kjøre lenger for å komme inn på motorveien.

H_0^1 : Kjøreavstand til påkjørsel motorvei målt i meter har ingen betydning for boligprisen

H_A^1 : Kjøreavstand til påkjørsel motorvei målt i meter har betydning for boligprisen

Hypotese 2: Negativ effekt å bo veldig nærme motorveien målt i luftavstand

Motorveien kan ha både negativ og positiv effekt på boligprisen. Det kan tenkes at dersom en bolig er plassert nærme motorveien målt i luftavstand vil det være mye støy og luftforurensing. Under denne hypotesen vil det bli undersøkt om det å bo nærme motorveien målt i luftavstand vil gi en negativ effekt på boligprisen.

H_0^2 : Prisen på en bolig som er lokalisert nærme motorveien målt i luftavstand vil være høyere enn en bolig som ligger lenger unna.

H_A^2 : Prisen på en bolig som er lokalisert nærme motorveien målt i luftavstand vil være lavere enn en bolig som ligger lenger unna.

Disse to hypotesene vil være hovedhypoteser. I tillegg vil det nedenfor bli presentert en kontrollhypotese for å redusere risikoen for feil, ved at det er flere faktorer enn nærhet til motorvei som forklarer boligprisen.

Hypotese 3: Sammenheng mellom boareal og boligpris

Størrelsen på en bolig antas å ha en innvirkning på boligprisen. En stor bolig koster vanligvis mer enn en liten bolig ved at den for eksempel kan dekke flere behov. Antagelsen her vil være at vi har en økende, men avtagende nytte.

H_0^3 : Det er ingen sammenheng mellom boareal og boligpris

H_A^3 : Det er sammenheng mellom boareal og boligpris

4. Økonometrisk modell

Regresjons analyse er en statistisk teknikk som forsøker å forklare bevegelser i en avhengig variabel som en funksjon av bevegelser i et sett av andre uavhengige-, eller forklarede variabler (Studenmund, 2011). Denne type analyse blir blant annet brukt til å forklare boligpriser, og den hedonistiske prisfunksjonen vil dermed kunne bli estimert på bakgrunn av en regresjonsanalyse. Boligprisen er den avhengige variabelen som blir estimert på bakgrunn av et sett uavhengige variabler som er attributtene ved boligen. Fra teorien har vi at prisfunksjonen på en bolig kan bli uttrykt som en funksjon av de ulike attributtene n $P(Z) = P(Z_1, Z_2, \dots, Z_n)$. Videre i oppgaven og i analysen vil dette erstattes med $Y(X) = Y(X_1, X_2, \dots, X_n)$, hvor Y representerer den avhengige variabelen og X representerer de uavhengige variablene.

4.1 Introduksjon av relevante regresjonsmodeller

4.1.1 Enkel og multipel regresjon

Den enkle regresjonsmodellen inneholder kun en uavhengig variabel og den kan uttrykkes slik:

$$Y = \beta_0 + \beta X_i + \varepsilon_i$$

Hvor den avhengige variabelen Y indikerer boligprisen. Estimatet av skjæringskoeffisienten β_0 blir tolket som boligprisen dersom den uavhengige variabelen X har en verdi lik null, dette konstantleddet viser skjæringspunktet på den vertikale aksene. Helningskoeffisienten β_1 kan tolkes som at dersom X øker med en enhet vil Y være forventet, alt annet konstant, å øke med β_1 enheter (Brooks, 2008).

Ved at en bolig høyst sannsynlig inneholder mer enn en attributt kan man benytte en multipel regresjonsmodell som kan formuleres slik:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

Hvor p er antall uavhengige variabler som er inkludert i modellen. Som i den enkle modellen angir β_p også her endringen i den avhengige variabelen dersom X_p endres med en enhet når alle andre variabler holdes konstant. Med denne informasjonen kan vi konkludere med at jo høyere verdi vi får på β_p , jo større innvirkning har den uavhengige variabelen på den avhengige variabelen.

I praksis er det ikke mulig å observere verdiene til de ulike koeffisientene i modellen, slik at disse må bli estimert på bakgrunn av innsamlet informasjon i et datasett. De estimerte

koeffisientene i er merket med « $\hat{\beta}_p$ » og regresjonsmodellen kan nå uttrykkes slik:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} + \hat{\beta}_2 X_{2i} + \dots + \hat{\beta}_p X_{pi} + \varepsilon_i$$

Avstanden mellom de estimerte boligprisene (\hat{Y}_i) og de observerte boligprisene (Y_i) bestemmer hvor god den estimerte ligningen er. Denne avstanden beregnes gjennom residuale ε_i , dette feilleddet kan defineres på denne måten:

$$\varepsilon_i = Y_i - \hat{Y}_i$$

Den mest brukte metoden for å skaffe disse anslagene er Minste Kvadraters Metode eller «Ordinary Least Squares» (OLS). Denne estimeringsteknikken beregner $\hat{\beta}$ slik at kvadratsummen av residualene blir minimert (Studenmund, 2011).

4.1.2 Logaritmisk funksjonsform

I virkeligheten er det ofte slik at ulike sammenhenger ikke er lineære. Logaritmer gjør det mulig å uttrykke endringer i variabler som prosentvise endringer. Dette kan være nyttig ved at det finnes mange ulike relasjoner som på en naturlig måte kan uttrykkes i prosent (Stock & Watson, 2012). Valg av feil funksjonsform kan få store konsekvenser, og det er derfor viktig å se på hvilket forhold de ulike variablene har til hverandre. Dobbel-logaritmisk og semi-logaritmisk er de to hovedgruppene som vil få en kort gjennomgang.

Dobbelt-logaritmisk

Denne formen er ikke-lineær i variablene, samtidig som den er lineær i koeffisientene. I den dobbelt-logaritmiske funksjonsformen tar man den naturlige logaritmen av den avhengige variabelen og av de uavhengige variablene (Studenmund, 2011):

$$\ln Y_i = \beta_0 + \beta_1 \ln X_{1i} + \beta_2 \ln X_{2i} + \dots + \beta_p \ln X_{pi} + \varepsilon_i$$

Tolkningen her blir hvis den uavhengige variabelen X_1 øker med en prosent, når alle de andre uavhengige variablene holdes konstant, vil den avhengige variabelen endres med β_1 prosent. I denne funksjonsformen er alle de uavhengige variablene omkodet ved hjelp av de naturlige logaritmene. Unntaket er dummyvariabler, som vanligvis har verdien 0 eller 1 avhengig om en bestemt tilstand holder. En slik variabel med en verdi lik 0 vil ikke kunne påvirke boligprisen, mens en variabel med verdi lik 1 vil endre boligprisen med prosenten den tilhørende koeffisienten har.

Semi-logaritmisk

Denne formen er en variant av den dobbelt-logaritmiske funksjonsformen hvor noen, men ikke alle variablene er uttrykt ved hjelp av naturlige logaritmer. Et eksempel kan være at en tar den naturlige logaritmen av kun den avhengige variabelen Y (Studenmund, 2011):

$$\ln Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \varepsilon_i$$

Her vil endringen i boligprisene bli gitt prosentvis; Dersom X øker med en *enhet*, vil betaverdien gi den *prosentvise* endringen i prisen, når alt annet holdes konstant.

Dummyvariablenes betaverdi vil indikere prosentendringen i Y dersom en tilstand holder (dvs. dummyvariabel lik 1).

4.2 Valg av uavhengige variabler

For å få en korrekt estimert regresjonslikning er det tre komponenter som må være riktig spesifisert; korrekt valg av de uavhengige variablene, riktig funksjonsform og riktig form på det stokastiske feilleddet (Studenmund, 2011). Det faktum at forskere kan bestemme hvilke uavhengige variabler som skal inkluderes i estimeringen kan bli sett på som både en styrke og

en svakhet. Styrken ligger i det faktum at ligningene kan formuleres slik at de passer individuelle behov. Svakheten ligger i at forskerne kan anslå mange ulike spesifikasjoner helt til deres poeng er bevist, selv om mange av de andre resultatene motbeviste det.

En viktig uavhengig variabel kan bli utelatt i analysen ved at den for eksempel blir glemt eller at det ikke er mulig å samle inn eller observere data. Når dette skjer vil den estimerte ligningen bli feil. Utelatelse av en relevant forklaringsvariabel, som for eksempel boareal fra estimering av boligprisene, hindrer deg i å få et anslag på koeffisienten til boareal og fører også vanligvis til skjevhet i de estimerte koeffisientene til variablene i ligningen.

Anta at den riktige regresjonsmodellen forklarer boligprisen ved hjelp av to uavhengige variabler:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \quad (4.1)$$

Hvor ε_i viser det stokastiske feilleddet. Hvis X_2 utelates fra estimeringen vil likningen se slik ut:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i^* \quad (4.2)$$

ε_i^* vil nå være lik:

$$\varepsilon_i^* = \varepsilon_i + \beta_2 X_{2i} \quad (4.3)$$

Fra ligning (4.3) kan det vises at feilleddet inneholder virkningene fra eventuelle utelatte variabler. Dersom vi ser på ligning (4.2) og (4.3) i sammenheng kan det virke som om vi får objektive anslag for β_0 og β_1 selv om vi har utelatt X_2 fra ligningen. Uheldigvis er ikke dette tilfellet ved at de inkluderte koeffisientene nesten alltid plukker opp noe av effekten av den utelatte variabelen som gjør at de vil endre seg og skape skjevheter. For å unngå denne skjevheten må X_1 og X_2 være perfekt ukorrelert (Studenmund, 2011).

Dersom en irrelevant variabel blir inkludert i analysen vil ikke dette skape skjevhet, men variansen til de beregnede koeffisientene for variablene som er inkludert vil øke. Dersom man

antar at den virkelige regresjonsmodellen forklarer boligprisen ved hjelp av en variabel kan den spesifiseres slik:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i \quad (4.4)$$

Dersom en irrelevant variabel inkluderes i analysen vil regresjonsmodellen inkludere en ekstra uavhengig variabel:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i^{**} \quad (4.5)$$

Feilleddet fra denne modellen kan uttrykkes:

$$\varepsilon_i^{**} = \varepsilon_i - \beta_2 X_{2i} \quad (4.6)$$

Hvis den virkelige koeffisienten av den irrelevante variabelen er lik null ($\beta_2=0$) vil ikke denne skape skjevhet i de andre koeffisientene i regresjonsanalysen. Inkluderingen av den irrelevante variabelen vil øke variansen av de beregnede koeffisientene, denne økningen i variansen kan ha en tendens til å redusere den absolutte størrelsen av t-verdiene.

Inkluderingen vil også vanligvis redusere \bar{R}^2 , dette kommer vi nærmere inn på under avsnittet som forteller om modellens forklaringskraft.

Det er blitt utviklet fire forskjellige punkter som gir veiledning om en uavhengig variabel skal inkluderes i estimeringen eller ikke (Studenmund, 2011):

1. Teori: er variabelens plass i ligningen entydig og høres inkluderingen teoretisk riktig ut?
2. T-Test: er variabelens estimerte koeffisient signifikant i den forventede retningen?
3. \bar{R}^2 : Blir den totale tilpasning av ligningen (justert for frihetsgrader) bedre når variabelen inkluderes i ligningen?
4. Skjevhet: skjer det en vesentlig endring i koeffisientene til de andre uavhengige variablene når den aktuelle variabelen inkluderes i ligningen?

Hvis alle de fire vilkårene ovenfor er oppfylt kan det konkluderes med at variabelen hører til i ligningen. Dersom ingen av vilkårene er oppfylt kan man trygt utelate variabelen fra regresjonen. I mange tilfeller vil det være slik at ikke alle fire vilkårene stemmer, økonometrikeren må i disse tilfellene bruke skjønn og ikke stole på at en variabel skal inkluderes kun fordi for eksempel \bar{R}^2 øker. Teori, ikke statistisk passform, bør være det viktigste kriteriet for inkluderingen av en forklarende variabel, ved at man risikerer å produsere uriktige resultater.

4.3 De klassiske forutsetningene for OLS

De klassiske forutsetningene må være oppfylt for at OLS estimatorene skal være de beste tilgjengelige estimatene for regresjonsmodeller. Hvis en eller flere av forutsetningene ikke holder kan det være at andre estimeringsteknikker kan gi bedre estimater enn OLS. En viktig del av regresjonsanalysen blir dermed å avgjøre om de klassiske forutsetningene holder for den aktuelle ligningen (Studenmund, 2011). De 7 klassiske forutsetningene er som følger:

1. Regresjonsmodellen er lineær, den er korrekt spesifisert og har et additivt feilledd. Forutsetningen om at modellen er lineær krever ikke at den underliggende teorien er lineær ved at for eksempel en eksponentiell funksjon kan bli omgjort til en lineær modell ved hjelp av naturlige logaritmer. Videre må modellen være korrekt spesifisert ved at den har riktig funksjonsform og at det ikke skal være utelatt noen relevante forklaringsvariabler. Til slutt forventes det at et stokastisk feilledd er lagt til ligningen.
2. Gjennomsnittet på feilleddet er lik null. Dersom man har et lite utvalg med få observasjoner er det lite sannsynlig av gjennomsnittet er nøyaktig lik null. Når størrelsen på utvalget går mot uendelig vil gjennomsnittet nærme på feilleddet gå mot null.
3. Alle forklaringsvariablene er ukorrelert med feilleddet. Hvis for eksempel feilleddet og en forklaringsvariabel X ble positivt korrelert, så ville den estimerte koeffisienten trolig være høyere enn den ellers ville vært. OLS metoden vil feilaktig tilskrive variasjonen i Y forårsaket av ε til X .

4. Observasjoner av feilleddet er ikke korrelert med hverandre. Observasjonene av feilleddet er dermed uavhengig av hverandre. Hvis det oppstår en systematisk sammenheng mellom to ulike observasjoner så vil det være vanskeligere for OLS å få nøyaktige beregninger av standardavvik for koeffisientene.
5. Feilleddet har en konstant varians. Dette kalles «homoscedasticity» og innebærer at vi har konstant varians for ulike observasjoner av feilleddet. Dersom konstant varians ikke forekommer vil OLS generere feilaktige estimater av standardavviket av koeffisientene.
6. Det oppstår ikke perfekt multikollinearitet mellom de forklarende variablene. Dette innebærer at ingen forklarende variabel er en perfekt lineær funksjon av noen annen forklarende variabel (er). Dersom dette oppstår vil OLS metoden ikke kunne skille de ulike variablene fra hverandre.
7. Feilleddet er normalfordelt. Denne forutsetningen om normalitet er ikke nødvendig for OLS estimeringen, men tross for at denne er valgfri finnes det to gode grunner til å ta den med:
 - a. Feilleddet kan betraktes som summen av et antall mindre påvirkninger eller feil. Ettersom antallet av disse påvirkningene blir større har fordelingen av feilleddet en tendens til å nærme seg normalfordeling.
 - b. Med mindre feilleddet er normalfordelt vil ikke t-statistikken og F-statistikken være helt aktuell, med mindre utvalget som testes er veldig stort.

Gitt at de klassiske forutsetningene 1-6 holder har Gauss-Markov teoremet blitt innfridd. Dette teoremet sier at «OLS er BLUE» som står for at estimeringsmetoden er «Best Linear Unbiased Estimator». Dette forteller at koeffisientene, altså betaverdiene, estimert av OLS vil være de estimatorene med minst varians blant alle lineære estimater (Studenmund, 2011).

4.4 Modellens forklaringskraft

Modellens forklaringskraft kan fortelle oss noe om kvaliteten på regresjonen, den gir også mulighet til å sammenligne ulike modeller som har forskjellige datasett eller ulike kombinasjoner av uavhengige variabler. Vi kan aldri være helt sikre på det faktum at en modell representerer sannheten mer enn en annen modell. Modellens forklaringskraft er kun en komponent i dette valget, en kan dermed ikke la seg bli altfor påvirket av det. R^2 , som er uttrykket for forklaringskraften, måler den prosentvise endringen i variansen i Y som er forklart av regresjonslikningen (Studenmund, 2011). Denne formelen kan uttrykkes slik:

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

Hvor dekomponeringen av den totale variansen er lik $TSS = ESS + RSS$, hvor:

$$TSS = \sum_{i=1}^N (Y_i - \bar{Y})^2$$

$$ESS = \sum_{i=1}^N (\hat{Y}_i - \bar{Y})^2$$

$$RSS = \sum_{i=1}^N e_i^2$$

TSS «Total sum of squares» måler den totale variasjonen som er forklart av regresjonen. Denne består av de to komponentene «Explained sum of squares» som måler variasjonen innenfor den estimerte modellen og «Residual sum of squares» som måler variasjonen som ikke er forklart i den estimerte modellen. En verdi av R^2 nær en viser at modellen har en utmerket passform, mens en verdi nær null viser at verdiene av Y_i bedre kunne vært forklart ved hjelp av utvalgets gjennomsnitt. Ettersom OLS velger koeffisient estimater som minimerer RSS, vil modellen gi den høyest mulige R^2 .

Et stort problem med R^2 er at dersom en ny uavhengig variabel legges til i regresjonen vil dette aldri kunne senke R^2 . Den vil dermed være liten hjelp dersom det skal avgjøres om en ekstra uavhengig variabel forbedrer vår evne til å forklare den avhengige variabelen. På grunn

av dette problemet har det blitt utviklet et annet uttrykk som måler modellens forklaringskraft. Dette uttrykket er R^2 justert for frihetsgrader:

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

\bar{R}^2 måler dermed den prosentvise endringen i variansen i Y som er forklart av regresjonslikningen justert for frihetsgrader. Dette uttrykket kan øke, synke eller forbli uendret dersom en ny uavhengig variabel blir lagt til i regresjonslikningen. Utfallet vil avhenge av om den nye tilsetningen oppveier for tap av frihetsgrader. Det er viktig å understreke at dette kun er et enkelt mål på forklaringskraften til regresjonsmodellen. Det vil være avgjørende å inkludere andre faktorer når det skal vurderes hvilke variabler som skal tas med, dette ble diskutert i avsnitt 4.2 ovenfor.

4.5 Hypotese testing

Det første steget i hypotese testing er å formulere hypotesene som skal bli testet.

Nullhypotesen (H_0) er typisk en uttalelse som sier noe om hva man *ikke* forventer, mens den alternative hypotesen (H_A) vanligvis sier noe om hva man forventer. Når det skal avgjøres om nullhypotesen skal forkastes kreves det at vi setter en signifikantgrense, som forteller sannsynligheten for at vi konkluderer feil. Det er veldig vanlig å bruke en signifikantgrense på 5 prosent, men både 10- og 1 prosent er mye brukt (Brooks, 2008). Det er nesten umulig å bevise at en gitt hypotese er riktig. Selv om vi ikke kan bevise at en gitt teori er riktig ved å benytte hypotesetesting, kan vi ofte avvise nullhypotesen med en viss grad av signifikans.

Ved at regresjonskoeffisientene bare er estimater av de virkelige parameterne, vil det være urealistisk å tro at konklusjoner som trekkes fra regresjonsanalysen alltid vil være korrekte.

Det er to typer feil som kan forekomme i hypotese testing (Studenmund, 2011):

- Type I feil: vi forkaster en sann nullhypotese
- Type II feil: vi forkaster ikke en falsk nullhypotese

Disse feilene kan begrenses ved at det alltid vil velges en kritisk grense. Dersom en parameters signifikantnivå er mindre enn 5 prosent kan vi forkaste nullhypotesen med 95 prosent sikkerhet, dersom vi velger en signifikantgrense på 5 prosent.

T-test

T-test er den mest benyttede metoden for å undersøke hypoteser som omhandler de enkelte koeffisientene i regresjonen. Denne testen er passende å bruke når det stokastiske feilleddet er normalfordelt og når variansen til fordelingen må estimeres. Siden dette vanligvis er tilfellet, har denne hypotesetestings metoden blitt veldig vanlig i praksis. T-tester utføres vanligvis kun på helnings-koeffisienter og den aktuelle formen på modellen for koeffisient k er (Studenmund, 2011):

$$t_k = \frac{(\hat{\beta}_k - \beta_{H_0})}{SE(\hat{\beta}_k)}, \quad (k=1,2,\dots,k)$$

hvor:

$\hat{\beta}_k$ = den estimerte regresjons koeffisienten av variabel k

β_{H_0} = grenseverdien som følger av nullhypotesen for β_k , denne er vanligvis lik null

$SE(\hat{\beta}_k)$ = det estimerte standardavviket av $\hat{\beta}_k$

For å bestemme om vi skal forkaste nullhypotesen eller ikke basert på den beregnede t-verdien benyttes det en kritisk t-verdi. Den kritiske t-verdien, t_c , er valgt ut i fra en t-tabell. Denne verdien bestemmes ut i fra om vi har en ensidig eller tosidig test, hvilket signifikant nivå vi velger og antall frihetsgrader. Hvor antall frihetsgrader er definert som antall observasjoner (N) minus antall koeffisienter (k) og et konstantledd ($N - k - 1$). Dersom den absolutte verdien av den kalkulerede t-verdien $|t_k|$ er større enn den kritiske t-verdien og har det samme fortegnet som den alternative hypotesen, kan vi forkaste nullhypotesen.

P-verdier er en alternativ tilnærming til t-testen. P-verdien forteller det laveste nivået av signifikans som vi kan forkaste nullhypotesen på. Alle p-verdier er gitt i sannsynlighet, slik at de vil alltid være mellom null og en. En lav p-verdi vil skape tvil om nullhypotesen, for å kunne forkaste H_0 må vi dermed ha en lav p-verdi.

F-test

T-test blir brukt til å teste hypoteser som kun inneholder en koeffisient, dersom man vil teste flere koeffisienter samtidig kan en F-test benyttes. Slike hypoteser kan for eksempel ofte se slik ut:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_A: \beta_p \neq 0$$

Den generelle F-testen kan uttrykkes slik (Studenmund, 2011):

$$F = \frac{\frac{(RSS_M - RSS)}{M}}{\frac{RSS}{(N - k - 1)}}$$

Hvor:

RSS_M = «residual sum of squares» fra den begrensede ligningen

RSS = «residual sum of squares» fra den opprinnelige ligningen

M = antall begrensninger som er lagt på ligningen (vanligvis lik antall β som er eliminert fra den opprinnelige ligningen)

$(N-k-1)$ = Antall frihetsgrader i den opprinnelige ligningen

For å avgjøre om nullhypotesen i dette tilfellet skal forkastes eller ikke er beslutningsregelen som følger; dersom den beregnede F-verdien er større enn den aktuelle kritiske F-verdien forkaster vi H_0 .

4.6 Spesifisering av økonometrisk modell

Den økonometriske modellen som vil bli benyttet i denne oppgaven vil bli bestemt ved hjelp av analysen som skal gjennomføres i kapittel 6. Den regresjonsmodellen som innfrir de klassiske forutsetningene for OLS og som best beskriver dataene vil bli valgt. Som nevnt tidligere vil det for hver modell bli gjennomført to regresjoner, en hvor kjøreavstand målt i meter til nærmeste påkjørsel er med som en kontinuerlig variabel og en hvor luftavstand til motorvei målt i meter er inkludert som kategoriserte dummy-variabler.

Den avhengige variabelen i analysen er boligpris. De uavhengige variablene er valgt med utgangspunkt i problemstilling og teori. Følgende forklaringsvariabler som kan være med i den første analysen er kjøreavstand målt i meter til nærmeste påkjørsel på motorvei, boareal, fellesgjeld, år boligen ble solgt, alder på bolig, boligtype, eierform, postnummer og avstand til kommunesentrum. I analyse nummer 2 vil kjøreavstand til nærmeste påkjørsel bli erstattet med luftavstand til motorveien. Grunnen til at analyse 2 blir inkludert er for å undersøke om det er negativt å bo for nærme motorveien ved at det innebærer en del støy og økt forurensing.

Hvordan funksjonsformen ser ut vet vi ikke enda, men for de to analysene vil prisen henholdsvis være en funksjon av følgende variabler:

$$P = f(\text{avstandpåkjørselmotorvei, boareal, fellesgjeld, år boligen ble solgt, alder, boligtype, eierform, postnummer, avstand til sentrum})$$
$$P = f(\text{luftavstandtilmotorvei, boareal, fellesgjeld, år boligen ble solgt, alder, boligtype, eierform, postnummer, avstand til sentrum})$$

5. Innhenting, bearbeiding og beskrivelse av datamaterialet

5.1 Innsamling av datamaterialet

For å kunne besvare problemstillingen på best mulig måte er det behov for å ha et bredt datagrunnlag i forhold til hvilke boliger som er solgt i perioden og hvilke attributter disse boligene innehar. For å få tilgang til et slikt datagrunnlag har jeg brukt Eiendomsverdi AS sine databaser. På Eiendomsverdi AS sine nettsider kan vi lese at de er «et selskap som overvåker og registrerer aktivitet og utvikling i de norske eiendomsmarkedene» (Eiendomsverdi, 2015). Ved tilgang til deres databaser kan en finne informasjon om alle landets solgte eiendommer.

Under avsnittet «Europavei E18» ble det informert om at «Østfoldpakka» ble vedtatt i stortinget i år 2000. Motorveien som går gjennom Askim ble ferdigstilt i år 2005. Ved at det er små forandringer i boligprisene på kort sikt, er det derfor innhentet observasjoner fra to år før motorstrekket ble ferdigstilt og frem til år 2014. Det innsamlede datamaterialet omfatter perioden 1. januar 2003 til 31. desember 2014.

Informasjon om boligene solgt i denne perioden er hentet fra Eiendomsverdi AS sine omsetningsrapporter. Utvalget er basert på postnumre i Askim og det komplette datamaterialet består av 4297 observasjoner før eventuelle korrigeringer. Det er innhentet informasjon om følgende variabler:

- Adresse
- Eierform
- Boligtype
- Boareal / P-rom
- Bruttoareal
- Salgsdato
- Salgspris
- Fellesgjeld
- Byggeår

Videre vil det bli en nærmere omtale om de ulike attributtene som skal inkluderes i analysen.

Avstand til påkjørsel motorvei:

Avstand til påkjørsel motorvei er målt i kjøreavstand i meter fra boligens adresse til nærmeste påkjørsel på motorveien. Denne målingen var en tidkrevende prosess ved at det var 4297 solgte boliger i perioden. Avstandsmålingen er gjort ved hjelp av Google Maps hvor den korteste veien konstant har blitt valgt. Dette vil være inkludert som en uavhengig kontinuerlig variabel i analysen.

Avstand til motorvei målt i luftlinje:

Det er også gjort manuelle målinger som viser luftavstand i meter fra hver enkelt bolig til motorveien. Grunnen til at disse målingene er inkludert er at en bolig kan være lokalisert nærme motorveien i luftlinje, men må kanskje kjøre flere kilometer for å faktisk benytte veien. Det kan tenkes at dette vil kunne ha en negativ effekt på boligprisen ved at det blir mer støy og forurensing. Luftlinjemålingen er gjort ved hjelp av Finn.no sin karttjeneste. Avstand til motorvei i luftlinje vil bli lagt inn som kategoriserte dummy-variabler i analysen for å kunne si noe om det er negativt for boligprisen å bo nærme motorveien i forhold til de som ligger lenger unna. Kategoriseringen av dummy-variablene er basert på støymålinger utført av Statens Vegvesen langs den nye E18 (Statens vegvesen, 2006). Denne målingen gir grunnlag for å anta at boliger som ligger innenfor en radius av 500 meter i luftavstand fra motorveien vil være utsatt for endel støy og forurensing. Ved at Askim er en relativt flat by med en del jordbruk vil det være rimelig å anta at eksternalitetene knyttet til motorveien også påvirker boliger som ligger mellom 501-1000 meter unna motorveien i luftlinje. Boligene som ligger over 1000 meter fra motorveien målt i luftlinje antas å være upåvirket.

Avstand til sentrum:

Når det gjelder avstand til sentrum har alle boliger med likt postnummer fått lik avstand til sentrum. Erik Bolstad har utviklet en nettside som naturlig plasserer et postnummer i midten av det aktuelle området, slik at vi får en naturlig gjennomsnittsavstand til sentrum for alle postnumrene (Bolstad, 2014). Avstanden er målt i antall kilometers kjøreavstand til Storgata 9 1830 som ligger midt i Askim sentrum. Alle boligene som da ligger innenfor postnummer 1830 har fått en avstand på 0,1 kilometer. Alle postnumrene har blitt dobbeltsjekket ved hjelp av Postens karttjeneste som viser områdene postnumrene dekker. Postnummer 1814 var feilplassert, videre i analysen har dette blitt rettet. En liste over de ulike postnumrene og deres avstand til sentrum er vist i vedlegg 1.

Eierform:

Eierformen kan ha en effekt på boligprisen. I Eiendomsverdi sin database har vi at eierformen kan deles inn i selveier, borettslag eller aksjeleilighet. Ved at disse variablene må omkodes vil de bli tatt med som dummy-variabler i analysen.

Boligtype:

I databasen skilles det mellom enebolig, rekkehus, leilighet og tomannsbolig. Disse variablene vil på samme måte som eierform bli representert ved hjelp av dummy-variabler.

Boareal / P-rom:

Boareal/primærrom er det innvendige arealet minus boder, garasje, terrasser felles trapperom og lignende (Boligareal, 2010). Dette vil være inkludert som en kontinuerlig variabel i analysen.

Bruttoareal:

Bruttoareal er målt fra ytterveggenes ytterside. På grunn av at det er flere boliger i datasettet som mangler bruttoareal i forhold til boareal vil sistnevnte bli benyttet i analysen. Bruttoareal vil bli benyttet for å estimere boareal i tilfeller hvor boareal ikke er oppgitt, men bruttoareal er tilgjengelig.

Fellesgjeld:

Fellesgjeld vil naturlig ha en innvirkning på boligprisen og blir tatt med i analysen som en kontinuerlig variabel.

Byggeår:

Alder på boligen vil bli tatt med som en uavhengig kontinuerlig variabel. Alderen blir definert som salgsår minus byggeår for å finne alderen på boligen når den ble solgt.

Salgspris:

Salgspris er den prisen som boligen faktisk blir solgt for. Det er den prisen kjøperen er villig til å gi for det bestemte settet av attributter og selgeren er villig til å akseptere. Ved at det er salgsprisen som skal estimeres, blir denne tatt med som den avhengige variabelen i analysen.

5.2 Rensing og kontroll av data

Ved innsamling av data vil det alltid være en mulighet for feil ved at noen variabler kan bli utelatt eller feilregistrert. For at analysen skal være så troverdig som mulig er det avgjørende at dataene er korrekte. Det er derfor viktig å korrigere variabler som er mangelfulle og feilregistrerte. Analyseprogrammet STATA er blitt brukt og alle kommandoer er lagt i vedlegg 10.

For å minimalisere feil i datasettet har jeg valgt å utelate de observasjonene som mangler byggeår, salgspris, kjøreavstand til nærmeste påkjørsel målt i meter samt luftavstand til motorvei målt i meter. I situasjoner hvor både boareal og bruttoareal ikke er oppgitt er disse også utelatt. Dersom informasjon om boareal ikke er tilgjengelig men bruttoareal er oppgitt, har boarealet blitt estimert på bakgrunn av lineær regresjon:

$$Y = \beta_0 + \beta \times bta + \varepsilon$$

hvor:

Y= boa

β_0 = konstantleddet

β = helningskoeffisienten

bta = bruttoareal

ε = feilleddet

Tabell 3: Lineær regresjon av boa

Source	SS	df	MS	Number of obs = 3547		
Model	7605734.04	1	7605734.04	F(1, 3545) =	7907.35	
Residual	3409780.04	3545	961.856147	Prob > F =	0.0000	
Total	11015514.1	3546	3106.46195	R-squared =	0.6905	
				Adj R-squared =	0.6904	
				Root MSE =	31.014	
boa	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
bta	.5500932	.0061862	88.92	0.000	.5379644	.562222
_cons	38.77722	.9779229	39.65	0.000	36.85987	40.69457

Denne regresjonsligningen vil estimere boarealet på bakgrunn av bruttoarealet for observasjoner som mangler informasjon om boarealet:

$BOA = 38.77722 + (0.5500932 * bta)$. Dette indikerer at boarealet vil være lik konstantleddet på 38.78 kvadratmeter pluss omtrent 55 prosent av bruttoarealet. På bakgrunn av denne likningen ble boarealet til 15 boliger estimert, slik at disse kan være med i den videre analysen.

Den laveste salgsprisen som er registrert i perioden er 25 000 kr. Ut i fra en sjekk på finn.no er det informert om at det er fellesgjeld på denne boligen, men størrelsen er ikke oppgitt. Denne observasjonen vil derfor bli utelatt i analysen. Naturlig nok var det noen observasjoner som hadde mangelfulle opplysninger. Etter korrigering av dataene sitter vi igjen med 3 546 Observasjoner:

Tabell 4: Oversikt over utelatte variabler

Totale observasjoner	4297
Mangler pris	-214
Mangler byggeår	-68
Mangler boa og bta	-460
Mangler avstand til motorvei i luftlinje	-8
Manglet str. på fellesgjeld	-1
Gjenstående observasjoner	3546

5.3 Beskrivelse av datamaterialet

Problemstillingen i denne oppgaven er deskriptiv ved at formålet er å beskrive en situasjon på et bestemt område. Det deskriptive designet i denne situasjonen innebærer å se sammenhenger mellom flere variabler som forklarer boligprisen i Askim (Gripsrud, 2010).

Tabell 5 viser de kategoriske variablene som er omtalt som dummy-variabler som har verdier på enten 0 eller 1. Tabell 6 viser en forenklet oversikt over de kontinuerlige variablene i analysen. Ut i fra disse tabellene får vi informasjon om antall observasjoner, gjennomsnitt, standardavvik samt minste og største observasjon av de enkelte variablene. Gjennomsnittet er som kjent summen av alle verdiene til en variabel dividert med antall observasjoner.

Standardavviket forteller noe om observasjonene vil være i tråd med gjennomsnittet, dersom

en variabel har et høyt standardavvik indikerer dette at variasjonen vil være stor rundt snittverdien. Minimums- og maksimumsverdiene gir et bilde av spredningen i observasjonene ved at vi får informasjon om den laveste- og den høyeste observerte verdien.

Tabell 5: kategoriske variabler i analysen

Variable	Obs	Mean	Std. Dev.	Min	Max
enebolig	3546	.4035533	.4906791	0	1
tomannsbolig	3546	.1209814	.3261516	0	1
rekkehus	3546	.0688099	.2531663	0	1
leilighet	3546	.4066554	.4912788	0	1
selveier	3546	.7986464	.4010682	0	1
borettslag	3546	.2013536	.4010682	0	1
postnr1807	3546	.1395939	.3466141	0	1
postnr1808	3546	.0544275	.2268914	0	1
postnr1809	3546	.1401579	.3471998	0	1
postnr1811	3546	.1099831	.3129128	0	1
postnr1812	3546	.0738861	.2616223	0	1
postnr1813	3546	.07304	.2602389	0	1
postnr1814	3546	.0614777	.2402384	0	1
postnr1815	3546	.036943	.1886486	0	1
postnr1830	3546	.0265087	.1606652	0	1
postnr1831	3546	.0806543	.2723418	0	1
postnr1832	3546	.2033277	.402531	0	1
innenfor500	3546	.1037789	.3050165	0	1
innenfor1000	3546	.2081218	.4060217	0	1
over1000	3546	.6880993	.4633349	0	1
aar03	3546	.0868584	.2816672	0	1
aar04	3546	.0792442	.2701576	0	1
aar05	3546	.07304	.2602389	0	1
aar06	3546	.0682459	.252203	0	1
aar07	3546	.0879865	.2833152	0	1
aar08	3546	.0758601	.2648115	0	1
aar09	3546	.0820643	.2745014	0	1
aar10	3546	.0817823	.2740714	0	1
aar11	3546	.1020869	.3028052	0	1
aar12	3546	.0888325	.2845419	0	1
aar13	3546	.0893965	.2853554	0	1
aar14	3546	.0846024	.2783283	0	1

Boligtype og eierform

Ut i fra tabellen ovenfor kan vi lese at de to boligtypene enebolig og leilighet dominerer ved at de utgjør omtrent 80 prosent av de solgte boligene i perioden. Dette ser ut til å stemme relativt greit med virkeligheten, i 2011 utgjorde eneboliger 52,3 prosent av boligmassen i Askim (Statistisk sentralbyrå, 2013a). Boligtypene leilighet og rekkehus er definert ulikt hos statistisk sentralbyrå, kontra analysen. Det kan tenkes at leiligheter er hyppigere omsatt, som kan bety at de er solgt flere ganger i løpet av perioden som analyseres. Når det gjelder eierform ser det ut som om datagrunnlaget er bredt og reflekterer virkeligheten rimelig bra. Både i 2001 og i 2011 utgjorde selveier omtrent 80 prosent av markedet (Statistisk sentralbyrå, 2011a).

Postnummer

Det er blitt solgt flest boliger i postnummer 1832, etterfulgt av 1809 og 1807. Boliger som er plassert innenfor postnummer 1832 har fått en gjennomsnittslengde på 4,0 km til kommune sentrum, mens 1809 og 1807 har fått henholdsvis 1,60 km og 1,20 km.

Luftavstand til motorvei inndelt i grupper

Tabell 5 viser at 10,38 prosent av de solgte boligene i den gitte perioden ligger innenfor 500 fra motorveien meter målt i luftlinje. 20,81 prosent ligger mellom 501-1000 meter fra motorveien, mens hovedtyngden med 68,81 prosent ligger over 1000 meter unna målt i luftlinje.

År

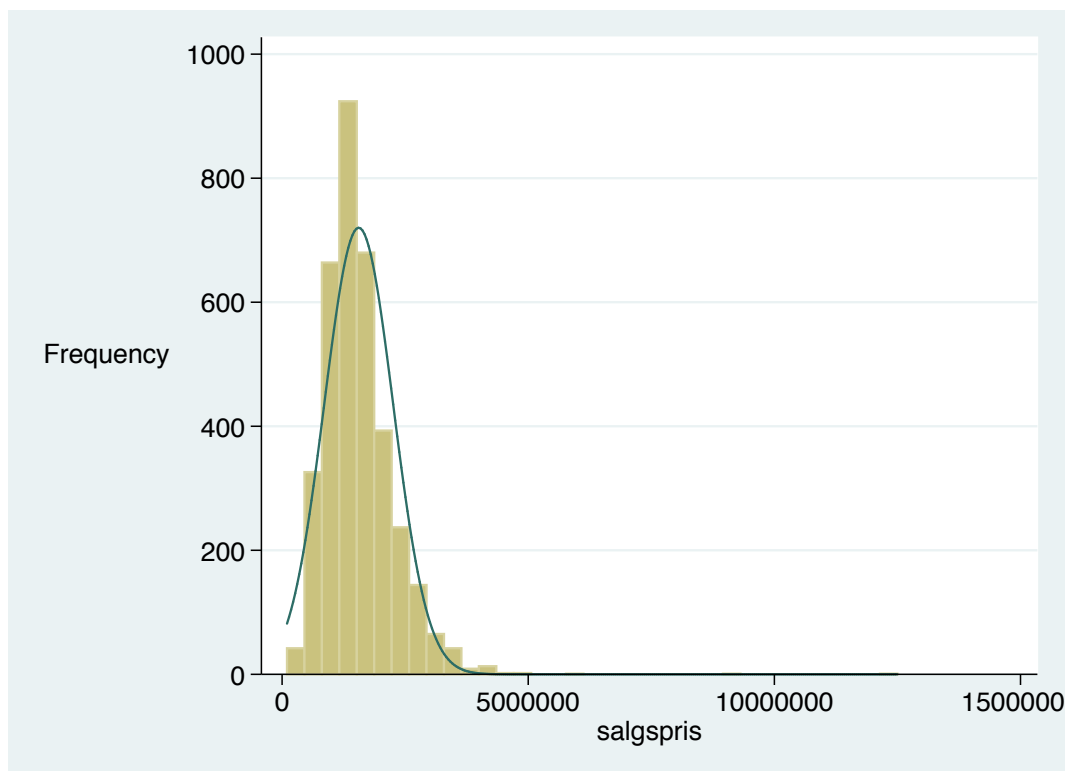
I datasettet har det blitt konkretisert hvilket årstall boligen er solgt. I år 2011 ser vi at det ble solgt flest boliger. Boligsalget vært relativt jevnt over hele perioden.

Tabell 6: kontinuerlige variabler i analysen

Variable	Obs	Mean	Std. Dev.	Min	Max
salgspris	3546	1556162	698094.5	100000	1.25e+07
fellesgjeld	3546	43233.12	142816	0	1295247
BOA	3546	112.869	55.47377	20	456
alder	3546	35.16328	23.33133	0	208
sentrum	3546	2.179851	1.276628	.1	4.5
avstpåkjm~i	3546	2549.817	859.2346	350	9300

Salgspris

Av tabell 6 kan vi lese at gjennomsnittsprisen på boligene solgt i perioden 2003-2014 er 1 556 162 kr. Den laveste salgsprisen som er registrert i perioden er 100 000 kr, mens den høyeste er omtrent 12 500 000 kr. Grunnen til at den laveste prisen er såpass lav er at fellesgjelden ikke er hensyntatt, som i dette tilfellet utgjør 1 272 000 kr. På bakgrunn av at dette spranget er såpass stort blir standardavviket relativt høyt, i dette tilfellet er det beregnet til omtrent 698 095 kr. Den burformende linjen i histogrammet nedenfor indikerer hvordan boligprisene skulle ha vært dersom vi hadde hatt perfekt normalfordeling. Ut i fra figuren ser vi at den er litt feilfordelt ved at den største andelen av de solgte boligene ligger nærmere minimumsverdien enn maksimumsverdien.

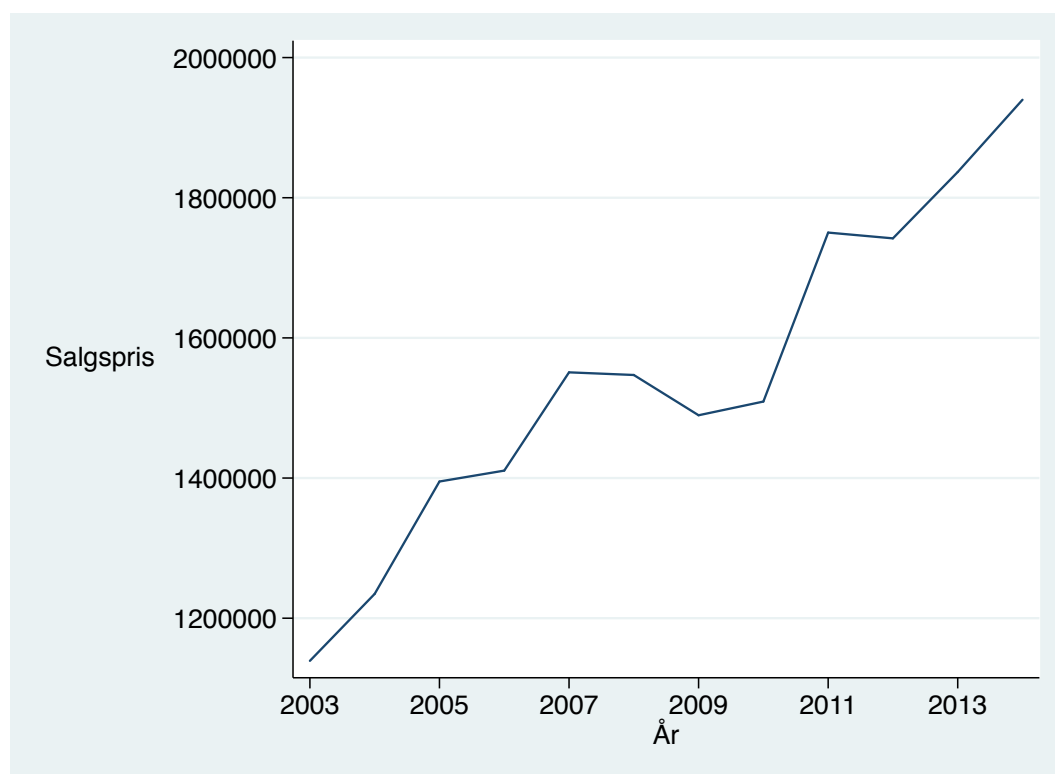


Figur 16: Solgte boliger med hensyn på salgspris, 2003-2014

Siden datasettet dekker mange år er det nedenfor utformet en tabell som viser gjennomsnitt, standardavvik, minimum- og maksimums verdier for de ulike årene. Av denne kan vi lese at gjennomsnittsprisen på solgte boliger i den valgte tidsperioden har økt fra 1.139.172 kr til 1.939.787 kr. Dette er en økning på 70,28 prosent.

Tabell 7: Gjennomsnittlig salgspris for boliger solgt i Askim, sortert på år

Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maximum
salgspris 2003	308	1139172	461744.2	400000	3200000
salgspris 2004	281	1234687	450812.3	440000	3450000
salgspris 2005	259	1395062	831609.5	450000	1.25e+07
salgspris 2006	242	1410599	482226.5	525000	3250000
salgspris 2007	312	1550828	582676.4	230000	3700000
salgspris 2008	269	1547082	572210.6	166000	4450000
salgspris 2009	291	1489605	603511.9	310000	3450000
salgspris 2010	290	1509166	753653.6	100000	9000000
salgspris 2011	362	1750246	725202.6	220000	4775000
salgspris 2012	315	1742024	749146.7	320000	4800000
salgspris 2013	317	1836703	728073.9	350000	4250000
salgspris 2014	300	1939787	784574.7	415000	6100000
salgspris 2003-2014	3546	1556162	698094.5	100000	1.25e+07



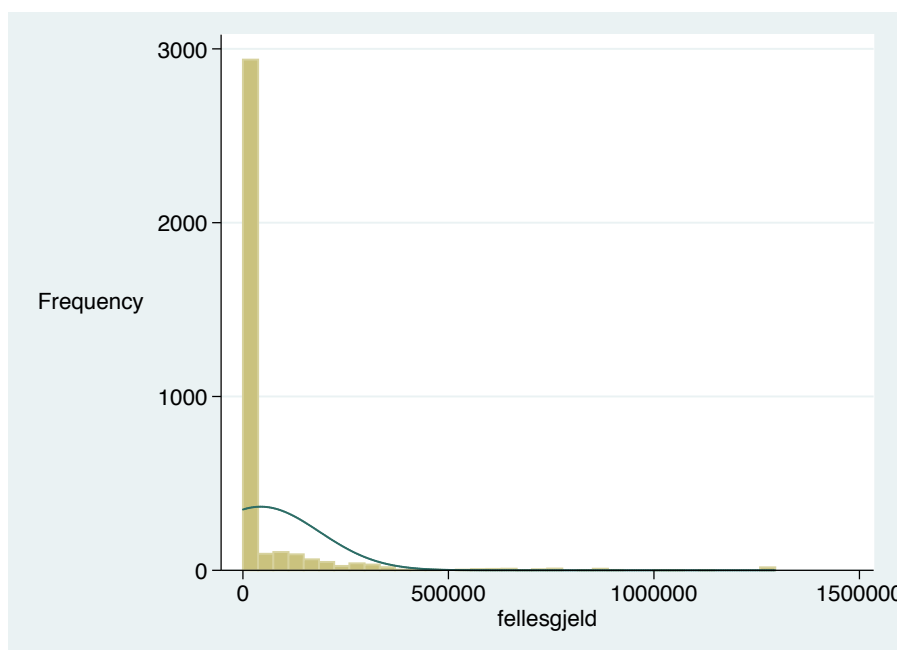
Figur 17: Gjennomsnittlig boligpriser fra 2003-2014

Figur 17 gir et mer visuelt bilde av hvordan utviklingen i de gjennomsnittlige boligprisene har vært. Det ser ut til at det har vært en jevn økning hele veien, med unntak av 2007-2009. Det kan være utallige årsaker til dette, men det var i denne perioden finanskrisen inntraff Norge.

En figur som viser gjennomsnittsprisene på de ulike postnumrene vil bli veldig uoversiktlig. I vedlegg 2 finnes det en oversikt som viser gjennomsnittlig salgspris for de ulike postnumrene i perioden. Denne viser at det ikke er så store forskjeller, men at boligene med postnummer 1815 og 1831 har den høyeste gjennomsnittlige salgsprisen.

Fellesgjeld

I den gitte perioden er gjennomsnittlig fellesgjeld i Askim omtrent 43 233 kr, hvor standardavviket utgjør 142 816 kr. Den minste observerte verdien på fellesgjeld er null, mens den høyeste er 1 295 247 kr. Grunnen til at gjennomsnittsverdien blir såpass lav er at datasettet inneholder hele 2 765 boliger som ikke har fellesgjeld. Den store spredningen mellom minimums- og maksimumsverdien gjør at standardavviket er over tre ganger så stort som gjennomsnittsverdien.



Figur 18: Solgte boliger med hensyn på fellesgjeld, 2003-2014

Tabell 8: Gjennomsnittlig fellesgjeld for boliger solgt i Askim, sortert på år

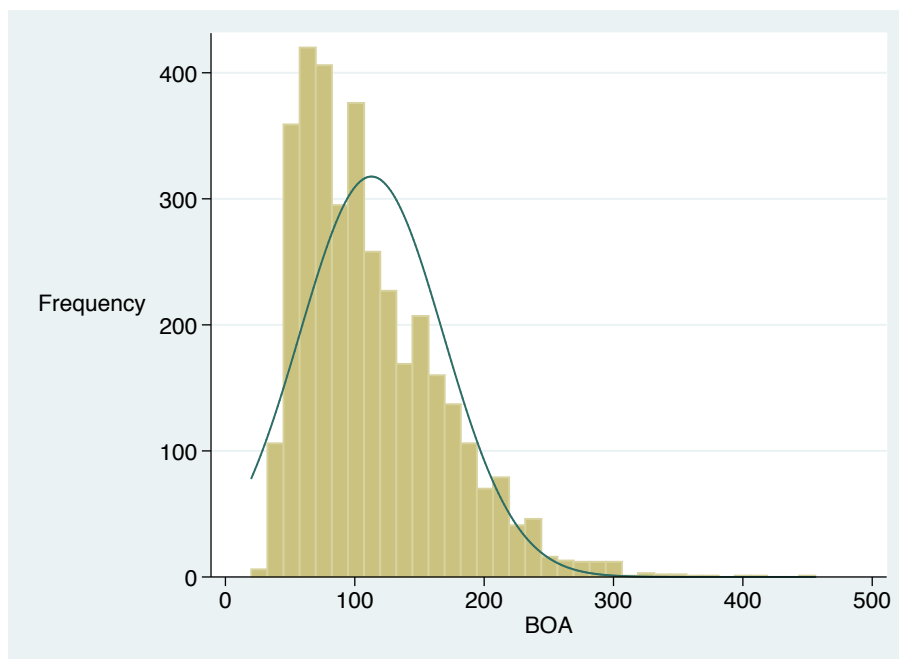
Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
Fellesgjeld 2003	308	10835.66	50305.34	0	670000
Fellesgjeld 2004	281	20548.02	109503.7	0	887538
Fellesgjeld 2005	259	17526.09	82483.88	0	745176
Fellesgjeld 2006	242	30548.7	128641.3	0	884594
Fellesgjeld 2007	312	19857.18	82142.62	0	890959
Fellesgjeld 2008	269	27080.48	94387.3	0	884594
Fellesgjeld 2009	291	40169.36	106942.6	0	884593
Fellesgjeld 2010	290	73603.41	200462.2	0	1272000
Fellesgjeld 2011	362	61336.28	186730.2	0	1272000
Fellesgjeld 2012	315	54108.74	142531.4	0	1292360
Fellesgjeld 2013	317	70394.69	156970.7	0	1295247
Fellesgjeld 2014	300	80612.56	218391.8	0	1288294

Av tabell 8 ovenfor kan vi lese at størrelsen på den gjennomsnittlige fellesgjelden har variert en del gjennom perioden vi analyserer. Den gjennomsnittlige størrelsen på fellesgjelden har vært relativt mye høyere fra år 2009-2014, dette kan eksempelvis indikere at det i nyere tider har blitt solgt flere boliger med høyere fellesgjeld. Den høyeste observerte størrelsen på fellesgjelden de siste 5 årene har vært relativt mye høyere enn de foregående årene.

Vedlegg 3 viser hvordan fellesgjelden er fordelt på de ulike postnumrene. Tabellen viser en indikasjon på at det er mindre fellesgjeld i områdene hvor gjennomsnittsprisen var høyest, i områdene med postnummer 1815 og 1831.

Boareal

Solgte boliger i den valgte perioden har et gjennomsnittlig boareal på omtrent 113 kvadratmeter, hvor standardavviket utgjør 55 kvadratmeter. Det minste boarealet var på 20 kvadratmeter, mens det største var på 456. Figur 19 nedenfor har en buet linje som viser hvordan boarealet skulle vært dersom det var perfekt normalfordelt. Figuren viser at det er flere boliger med lavt boareal enn hva normalfordelingen viser. Vi har dermed en skjev fordeling mot venstre.



Figur 19: Solgte boliger med hensyn på boareal, 2003-2014

Tabell 9: Gjennomsnittlig boareal for boliger solgt i Askim, sortert på år

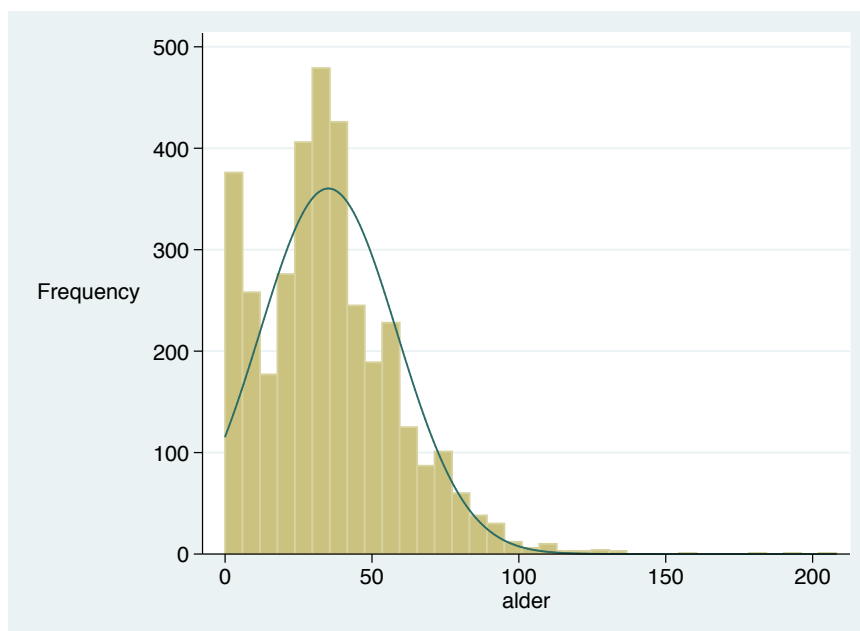
Variabel	Observasjoner	Gjennomsnitt	standardavvik	Minimum	Maksimum
BOA 2003	308	115.0181	57.86505	33	351
BOA 2004	281	121.3041	54.49954	33	323
BOA 2005	259	120.7915	58.36535	37	415
BOA 2006	242	117.6659	54.1211	39	284
BOA 2007	312	113.641	56.37871	22	301
BOA 2008	269	110.8124	54.43163	39	375
BOA, 2009	291	107.6419	51.87287	36	333
BOA 2010	290	105.9323	55.89265	20	456
BOA 2011	362	113.7735	53.04907	38	300
BOA 2012	315	109.8286	54.75396	21	294
BOA 2013	317	111.2326	56.28252	32	364.9825
BOA 2014	300	108.7	56.81058	33	351

Når det gjelder boarealet ser vi av tabell 9 at det ikke store gjennomsnittlige forandringer gjennom perioden som skal analyseres. De fire første årene er gjennomsnittet noen kvadratmeter høyere enn fra 2007-2014.

Vedlegg 4 viser hvordan gjennomsnittlig boareal er fordelt på de ulike postnumrene. Ut i fra denne tabellen kan vi lese at det er høyest gjennomsnittlig boareal i boliger med postnummer 1815 og 1831. Dette er steder som også har den høyeste gjennomsnittlige salgsprisen. Boliger som ligger i sentrum med postnummer 1830 har lavest gjennomsnittlig boareal, men ikke lavest salgspris.

Alder

Alderen på boligene i analysen er definert som salgsår minus byggeår. Den gjennomsnittlige alderen på de solgte boligene vi studerer er på 35 år, med et standardavvik på rundt 23 år. Den laveste verdien på alder som er observert er null år og den høyeste observerte er 208 år. Det at en bolig har en alder på null år forteller oss at det er at boligen er solgt det året den var ferdigstilt. Figuren nedenfor viser hvordan alderen er i forhold til hvordan den ville vært dersom den var perfekt normalfordelt.



Figur 20: Solgte boliger med hensyn på alder, 2003-2014

Tabell 10: Gjennomsnittlig alder for solgte boliger i Askim, sortert på år

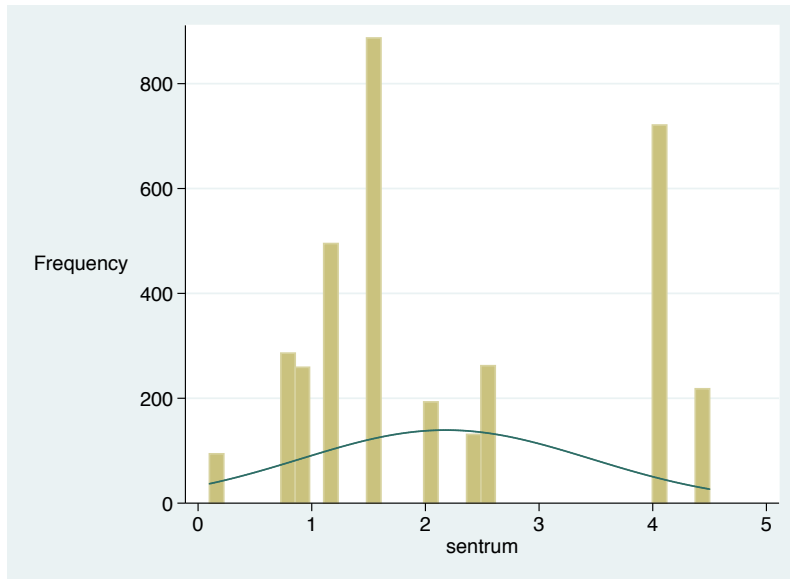
Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
alder 2003	308	31.94156	21.08143	0	125
alder 2004	281	34.91459	22.2991	0	121
alder 2005	259	31.64093	22.18925	0	125
alder 2006	242	35.69008	26.21644	0	156
alder 2007	312	33.89103	24.86868	0	128
alder 2008	269	32.43866	24.49598	0	208
alder 2009	291	36.2646	23.34188	0	109
alder 2010	290	35.29655	23.72044	0	131
alder 2011	362	35.62155	23.78581	0	182
alder 2012	315	36.85397	21.62	0	132
alder 2013	317	38.78864	22.84433	0	113
alder 2014	300	37.73	22.78481	1	194

Av tabell 10 over kan vi lese at gjennomsnittlig alder på solgte boliger har vært relativt jevn gjennom hele tidsperioden. Den gjennomsnittlige alderen var høyest i 2013, merk at her er den høyeste observerte alderen kun 113 år.

Vedlegg 5 viser gjennomsnittlig alder på de solgte boligene sortert etter postnummer. Her ser vi at postnummer 1815 har den høyeste gjennomsnittlige alderen på 57 år.

Avstand til sentrum

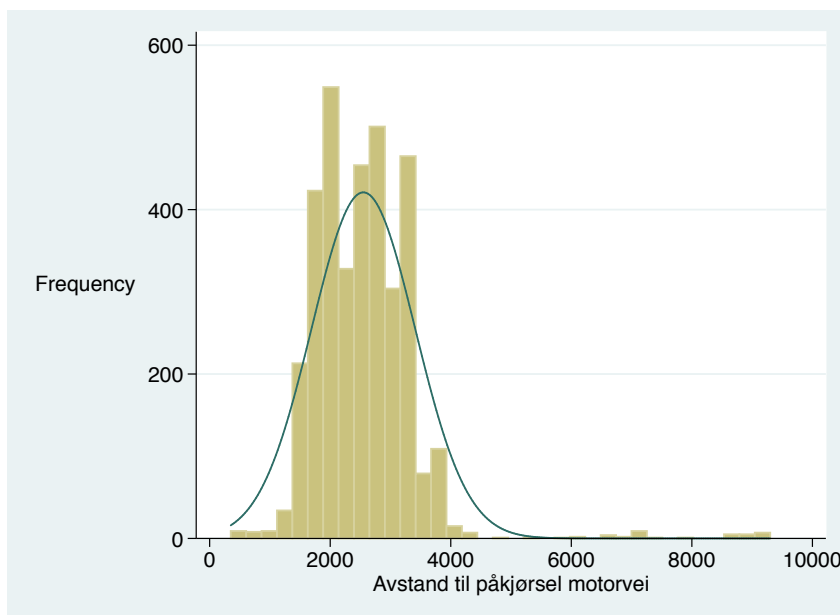
Avstand til sentrum er basert på de ulike postnumrene, slik at alle boliger med likt postnummer har fått lik avstand til sentrum. Den gjennomsnittlige avstanden er beregnet til å være omtrent 2,2 kilometer, hvor standardavviket utgjør rundt 1,3 kilometer. Den minste avstanden som er med i analysen er 0,1 kilometer og den lengste avstanden er 4,5 kilometer.



Figur 21: Solgte boliger med hensyn på avstand til sentrum, 2003-2014

Avstand til påkjørsel motorvei

Denne avstanden er som tidligere nevnt den kortest mulige kjøreavstanden målt i meter fra hver enkelt bolig til nærmeste påkjørsel på motorveien. Den gjennomsnittlige avstanden til nærmeste påkjørsel utgjør 2 550 meter, hvor standardavviket er på omtrent 859 meter. Den korteste veien som er observert er 350 meter, mens den lengste kjøreveien er observert til 9300 meter.



Figur 22: Solgte boliger med hensyn på avstand til påkjørsel motorvei, 2003-2014

5.4 Korrelasjon

En enkel forklaring på korrelasjon er at den måler graden av lineær sammenheng mellom to variabler (Brooks, 2008). Korrelasjonskoeffisienten vil alltid ha en verdi mellom -1 og 1, hvor fortegnet indikerer retning på sammenhengen mellom de to variablene. Dersom koeffisienten er lik 0 eksisterer det ikke noen lineær sammenheng. En positiv korrelasjon indikerer at dersom en variabel øker vil den andre også øke. En negativ korrelasjon indikerer at dersom en variabel øker vil den andre synke, de vil dermed gå motsatt vei av hverandre.

På grunn av mangel på plass har jeg utelatt år og postnummer i korrelasjonsmatrisen.

Fullstendig matrise kan ses i vedlegg 6. Kjøreavstand til påkjørsel motorvei målt i meter og dummy-variablene som viser luftavstand i meter til motorveien vil ikke bli brukt i samme regresjon.

Tabell 11: Korrelasjonsmatrise

	salgsp~s	felles~d	avstp~i	BOA	alder	enebolig	tomann~g
salgspris	1.0000						
fellesgjeld	-0.2571*	1.0000					
avstp~i	0.0023	-0.0888*	1.0000				
BOA	0.7235*	-0.2003*	0.0680*	1.0000			
alder	0.0270	-0.1395*	0.0512*	0.2165*	1.0000		
enebolig	0.4865*	-0.1923*	0.1218*	0.6423*	0.2666*	1.0000	
tomannsbolig	0.0167	-0.1123*	-0.0533*	0.0739*	0.1074*	-0.3052*	1.0000
rekkehus	-0.0789*	0.0414*	0.0231	-0.0709*	-0.0626*	-0.2236*	-0.1008*
leilighet	-0.4564*	0.2453*	-0.0982*	-0.6541*	-0.3054*	-0.6810*	-0.3071*
selveier	0.3900*	-0.5916*	-0.0544*	0.3675*	0.0123	0.3571*	0.1798*
borettslag	-0.3900*	0.5916*	0.0544*	-0.3675*	-0.0123	-0.3571*	-0.1798*
sentrum	-0.0338*	-0.0100	0.1987*	-0.0083	-0.1954*	0.0364*	-0.1535*
innenfor500	0.0416*	0.0485*	-0.2950*	-0.0151	-0.0995*	-0.0066	-0.0383*
innenfor1000	0.0343*	-0.0933*	-0.2757*	0.1083*	0.1217*	0.0725*	0.1336*
over1000	-0.0574*	0.0498*	0.4358*	-0.0850*	-0.0411*	-0.0591*	-0.0918*
	rekkehus	leilig~t	selveier	borett~g	sentrum	inne~500	inn~1000
rekkehus	1.0000						
leilighet	-0.2250*	1.0000					
selveier	-0.0858*	-0.4319*	1.0000				
borettslag	0.0858*	0.4319*	-1.0000	1.0000			
sentrum	0.0208	0.0548*	0.0622*	-0.0622*	1.0000		
innenfor500	0.0244	0.0195	0.0740*	-0.0740*	0.2536*	1.0000	
innenfor1000	-0.1229*	-0.0977*	0.1448*	-0.1448*	0.0358*	-0.1745*	1.0000
over1000	0.0916*	0.0728*	-0.1756*	0.1756*	-0.1983*	-0.5054*	-0.7615*
	over1000						
over1000	1.0000						

Dersom en koeffisient i matrisen er merket med stjerne betyr dette at den er signifikant på 5 prosent nivå, og sammenhengen mellom de to variablene kan fastslås med 95 prosent sannsynlighet. Det første blikkfanget i matrisen er at pris og boareal er ganske sterkt positivt korrelerte, dette indikerer at høy pris henger som oftest sammen med et stort boareal. Videre kan vi se at fellesgjeld har en signifikant svak negativ korrelasjon med salgsprisen, dette forteller at en økning i fellesgjeld vil innebære lavere salgspris. Av korrelasjonsmatrisen vises det at avstand til påkjørsel motorvei ikke er signifikant på 5 prosent nivå. Det faktum at denne korrelasjonen er svakt positiv indikerer at dersom kjøreavstand til nærmeste påkjørsel til motorveien øker vil boligprisene også øke i en viss grad. Dummy-variablene selveier og borettslag utelukker hverandre ved at de er perfekt negativt korrelerte. En av disse kan dermed utelukkes i den videre analysen.

6. Analyse

I kapittel 4 ble det presentert fire ulike regresjonsmodeller. I Analysen vil jeg utelukke den enkle regresjonen av den grunn at boligprisen ikke kan bestemmes ut i fra en enkel attributt. De tre gjenværende modellene vil i dette kapittelet bli testet og vurdert slik at den beste modellen blir valgt før vi kommer til hypotesetestingen. Som nevnt tidligere vil det for hver regresjonsmodell bli utført to regresjoner; en hvor kjøreavstand i meter til nærmeste påkjørsel på motorveien er inkludert som en kontinuerlig variabel og en hvor luftavstand i meter til motorveien er inkludert som kategoriserte dummy-variabler.

6.1 Gjennomgang av regresjonsmodellene

6.1.1 Multippel lineær regresjon

Dummyvariablene leilighet, selveier, aar03, postnummer postnr1830 og postnr1831 blir ekskludert for å unngå multikollinearitet. En slik bolig med boareal lik null vil omtales som basisbolig. Basisboligen i denne oppgaven vil dermed være en selveier leilighet med postnummer 1830 eller 1831 solgt i år 2003.

I den første regresjonsanalysen ble dummyvariablene som viser postnummer (med unntak av 1830 og 1831) samt variabelen som viser avstand til sentrum inkludert. Forklaringskraften til denne modellen var på 74,47 prosent, hvor noen av dummy-variablene ikke er signifikante på 5 prosent nivå. En VIF-test (Variance Inflation Factor – test) ble gjennomført for hver enkelt uavhengig variabel. Denne testen måler den kombinerte effekten av avhengigheter mellom forklaringsvariablene på variansen, og dersom $VIF > 10$ kan dette være en indikasjon på at vi vil få problemer med multikollinearitet. Denne testen avslørte en ganske høy verdi på blant annet «sentrum» på 206,59 i tillegg hadde mange av postnummerdummyene veldig høye verdier.

Grunnet multikollinearitet blir den uavhengige variabelen «sentrum» ekskludert.

Forklaringskraften til modellen går ned til 74,19 prosent, noe som fortsatt er veldig bra. Regresjonen ser nå mye bedre ut ved at alle postnummerdummyene er signifikante på 5 prosent nivå samt at $VIF < 5$ for alle forklaringsvariablene i regresjonen. Postnumrene vil kunne fungere og si noe om avstand til sentrum. Regresjonsanalysen er vist i tabell 12 nedenfor.

Tabell 12: Multipl linear regresjon

Source	SS	df	MS	Number of obs = 3546		
Model	1.2853e+15	28	4.5902e+13	F(28, 3517) = 364.95		
Residual	4.4235e+14	3517	1.2578e+11	Prob > F = 0.0000		
Total	1.7276e+15	3545	4.8734e+11	R-squared = 0.7440		
				Adj R-squared = 0.7419		
				Root MSE = 3.5e+05		

salgspris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	8541.069	155.4331	54.95	0.000	8236.321	8845.818
fellesgjeld	-.7039612	.0546418	-12.88	0.000	-.811094	-.5968285
avstpåkjemotorvei	-25.62216	8.223158	-3.12	0.002	-41.74481	-9.499523
aar04	56117.52	29350.68	1.91	0.056	-1428.56	113663.6
aar05	191598.4	29982.34	6.39	0.000	132813.8	250382.9
aar06	268212.7	30631.93	8.76	0.000	208154.5	328270.8
aar07	447152.7	28585.03	15.64	0.000	391107.8	503197.6
aar08	475961.5	29651.47	16.05	0.000	417825.7	534097.3
aar09	480904.5	29109.37	16.52	0.000	423831.6	537977.5
aar10	542504.5	29249.94	18.55	0.000	485156	599853.1
aar11	695369.3	27661.04	25.14	0.000	641136	749602.6
aar12	728864.5	28595.15	25.49	0.000	672799.7	784929.2
aar13	842658.4	28686.07	29.38	0.000	786415.3	898901.4
aar14	959020.5	29050.64	33.01	0.000	902062.7	1015978
alder	-6117.736	295.0993	-20.73	0.000	-6696.32	-5539.153
enebolig	107836.5	21378.72	5.04	0.000	65920.53	149752.4
tomannsbolig	-41834.48	23608.93	-1.77	0.076	-88123.05	4454.101
rekkehus	-18083.19	26186.01	-0.69	0.490	-69424.5	33258.13
borettslag	-133201.7	22604.42	-5.89	0.000	-177520.8	-88882.63
postnr1807	-183988.7	25854.37	-7.12	0.000	-234679.8	-133297.6
postnr1808	-203870.7	34307.89	-5.94	0.000	-271136.1	-136605.3
postnr1809	-196682	25536.77	-7.70	0.000	-246750.4	-146613.6
postnr1811	-199800.9	26678.28	-7.49	0.000	-252107.4	-147494.5
postnr1812	-159144	31400.96	-5.07	0.000	-220710	-97578.07
postnr1813	-148330.3	29372.69	-5.05	0.000	-205919.5	-90741.05
postnr1814	-137468.7	31879.41	-4.31	0.000	-199972.7	-74964.71
postnr1815	-122493.9	37047.56	-3.31	0.001	-195130.8	-49857.02
postnr1832	-279972.6	23386.25	-11.97	0.000	-325824.6	-234120.6
_cons	583221.9	34455.92	16.93	0.000	515666.3	650777.5

Den estimerte lineære modellen har et konstantledd på 583.221,9. Dette konstantleddet indikerer verdien på basisboligen med boareal lik null kvadratmeter. Hver enkelt koeffisient angir hvor mye prisen øker dersom den tilhørende variabelen øker med en enhet og alt annet holdes konstant. For hver kvadratmeter boarealet i boligen øker, vil salgsprisen på boligen øke med 8.541,069 kr. Denne modellen viser at boligprisen vil synke med 25,62 kr per økte meter med kjøreavstand til nærmeste påkjørsel på motorveien. Dersom en dummy-variabel inntreffer og får verdien 1, viser de tilhørende koeffisientene hvor mye boligprisen vil endre seg. Dersom vi for eksempel har med en enebolig å gjøre så vil salgsprisen på boligen være 107.836,50 kr høyere enn basisboligen som er en leilighet. Fra kapittel 2.3 har vi at traseen

Sekkelsten- Krosby, som er den delen av motorveien som går gjennom Askim, ble åpnet for trafikk i år 2005. Den multiple lineære regresjonen viser en ganske kraftig prisøkning fra år 2004-2005. Etter at strekket Krosby-Knapstad ble ferdigstilt november 2010 viser regresjonen en relativt stor økning i salgspris.

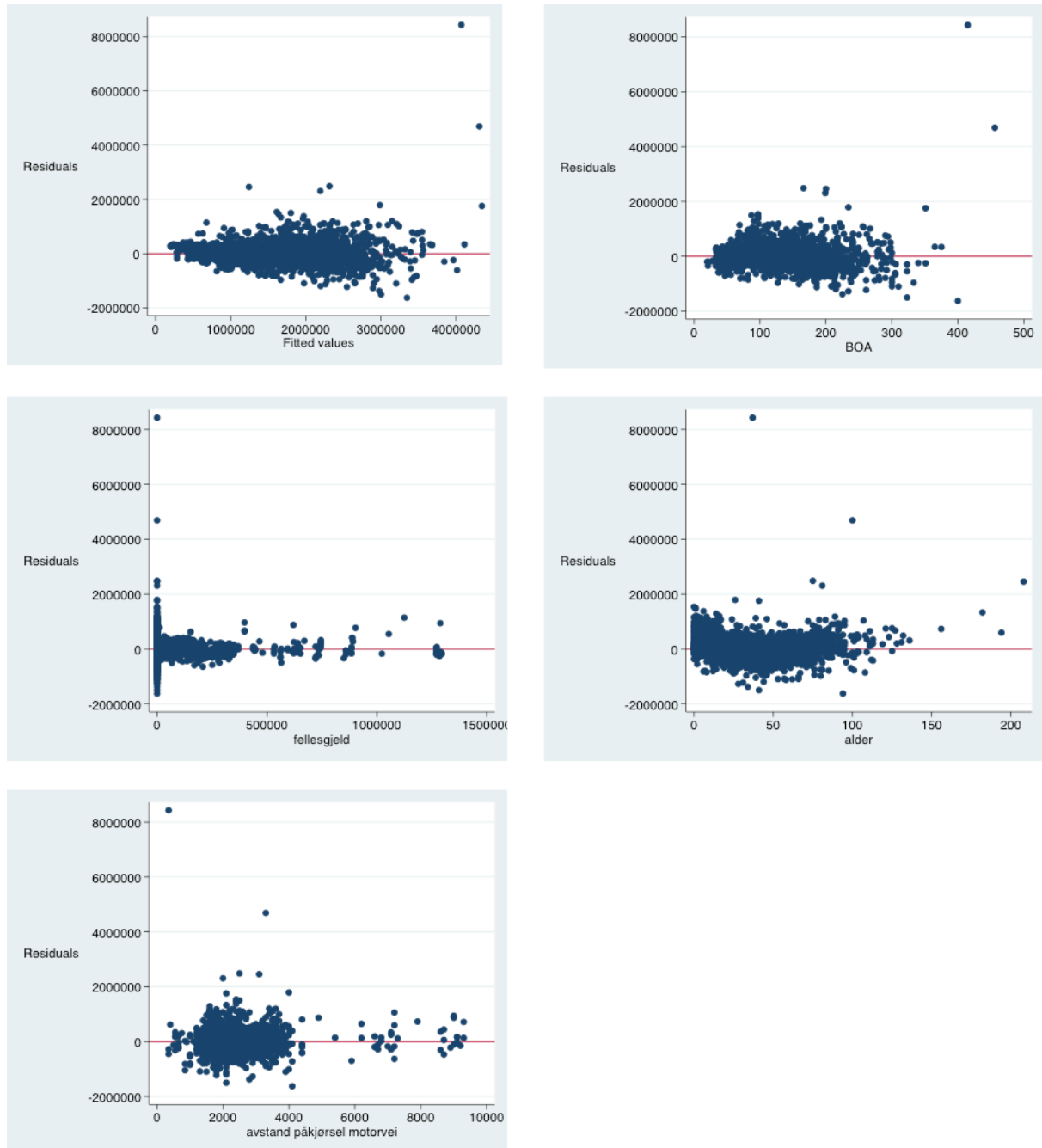
Forklaringskraften til denne modellen er på 0,7419. Dette innebærer at de uavhengige variablene som er inkludert i modellen forklarer hele 74,19 prosent av variasjonen i boligprisen. De gjenværende 25,81 prosentene blir ikke forklart og fanges dermed opp i feilleddet ε . Selv om modellen har høy forklaringskraft må vi se om den holder forutsetningene for OLS som ble presentert i kapittel 4.3. Vi må dermed kontrollere for multikollinearitet, homoskedastiskitet og om feilleddet er normalfordelt.

VIF-testen er summert i tabellen nedenfor og viser at vi ikke har noen problemer med multikollinearitet. Dette innebærer at ingen av de forklarende variablene er en perfekt lineær funksjon av en annen forklarende variabel. Gjennomsnittlig VIF for modellen er 1,86 hvor den høyeste verdien er ligger på 3,10.

Tabell 13: VIF-test, multippel lineær regresjon

Variable	VIF	1/VIF
enebolig	3.10	0.322419
postnr1832	2.50	0.400368
borettslag	2.32	0.431674
postnr1807	2.26	0.441794
postnr1809	2.22	0.451324
B0A	2.10	0.477218
aar11	1.98	0.505727
postnr1811	1.96	0.509117
postnr1812	1.90	0.525708
aar13	1.89	0.529499
aar12	1.87	0.535923
aar07	1.85	0.540957
aar14	1.84	0.542692
aar10	1.81	0.552080
aar09	1.80	0.555680
aar04	1.77	0.564298
aar08	1.74	0.575458
fellesgjeld	1.72	0.582608
aar05	1.72	0.582779
postnr1808	1.71	0.585539
aar06	1.68	0.594471
tomannsbolig	1.67	0.598395
postnr1814	1.65	0.604887
postnr1813	1.65	0.607222
avstpåkjm~i	1.41	0.710688
postnr1815	1.38	0.726363
alder	1.34	0.748452
rekkehus	1.24	0.807289
Mean VIF	1.86	

Videre må vi kontrollere at feilleddet har en konstant varians. Dette er det vi kaller «homoskedastiskitet» og innebærer at vi har konstant varians for alle observasjoner av feilleddet. Dersom variansen til restleddet avhenger av verdiene til forklaringsvariablene har vi at feilleddet er heteroskedastisk og en OLS forutsetning blir brutt (Midtbø, 2012).



Figur 23: Residualplot, multipel lineær regresjon

Figur 23 ovenfor viser residualplot for de kontinuerlige variablene som er inkludert i denne regresjonsmodellen. Dersom plottene har et mønster i form av at en sky av prikker som enten

blir smalere eller bredere når de predikerte verdiene blir større er dette en indikasjon på at vi har heteroskedastisitet (Midtbø, 2012). Plottene som viser boareal og alder ser ut til å ha en bredere sky av prikker ettersom de predikerte verdiene blir høyere. Fellesgjeld og avstand til påkjørsel motorvei ser ut til å ha en smalere sky av prikker. Plottene inneholder mange observasjoner, slik at det kan være vanskelig å konkludere kun ut i fra disse. Det blir dermed utført en Breusch-Pagan test:

Tabell 14: Breusch – Pagan test, multippel lineær regresjon

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of salgspris	
chi2(1)	= 3735.19
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: B0A	
chi2(1)	= 5443.44
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fellesgjeld	
chi2(1)	= 37.30
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: alder	
chi2(1)	= 273.93
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: avstpåkjemotorvei	
chi2(1)	= 240.21
Prob > chi2	= 0.0000

Av tabellen ovenfor kan vi lese at nullhypotesen kan forkastes med et signifikansnivå på 5 prosent for hele modellen og de kontinuerlige variablene. Nullhypotesen om homoskedastisitet kan forkastes og konklusjonen forteller at heteroskedastisitet er et problem. Det faktum at vi har heteroskedastisitet kan ha flere årsaker, en av de kan være at modellen er feilspesifisert. Siden det kan være vanskelig å identifisere formen for heteroskedastisitet er det blitt vanlig å korrigere standardfeilene uten å gå bort i fra OLS. Denne korleksjonen gir robuste standardfeil som er gyldige selv når formen for heteroskedastisitet er ukjent (Midtbø, 2012). Tabellen nedenfor viser de robuste standardfeilene. Signifikansnivået til kjøreavstand i

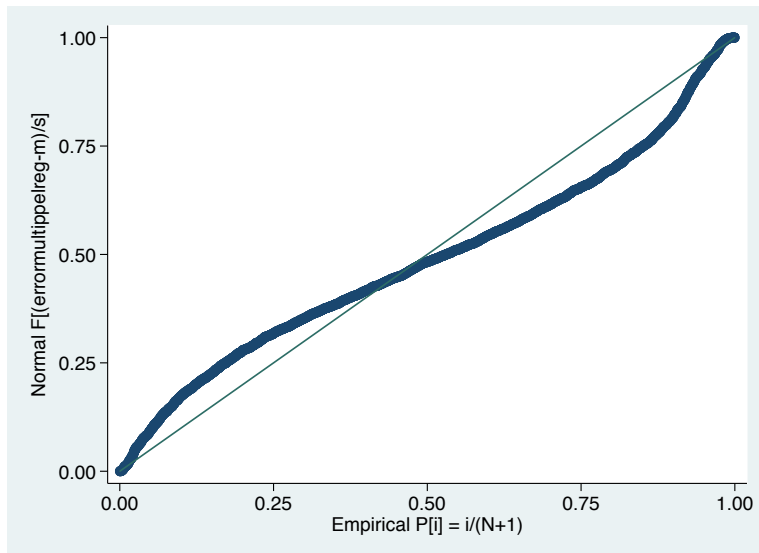
meter til nærmeste påkjørsel blir svekket, men er fortsatt signifikant på 5 prosent nivå.

Tomannsbolig, rekkehus og postnummer 1815 er ikke signifikant på ønsket nivå.

Tabell 15: Robuste standardfeil, multipl linear regresjon

Linear regression		Number of obs = 3546				
		F(28, 3517) = 314.61				
		Prob > F = 0.0000				
		R-squared = 0.7440				
		Root MSE = 3.5e+05				
salgspris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	8541.069	542.5529	15.74	0.000	7477.319	9604.82
fellesgjeld	-.7039612	.049081	-14.34	0.000	-.8001913	-.6077312
avstpåkjmotorvei	-25.62216	12.44023	-2.06	0.040	-50.01295	-1.231378
aar04	56117.52	23293.94	2.41	0.016	10446.51	101788.5
aar05	191598.4	40099.63	4.78	0.000	112977.5	270219.2
aar06	268212.7	24768.05	10.83	0.000	219651.5	316773.9
aar07	447152.7	23764.03	18.82	0.000	400560	493745.4
aar08	475961.5	25767.69	18.47	0.000	425440.4	526482.6
aar09	480904.5	25504.15	18.86	0.000	430900.1	530908.9
aar10	542504.5	28437.82	19.08	0.000	486748.3	598260.8
aar11	695369.3	25231.77	27.56	0.000	645898.9	744839.7
aar12	728864.5	27242.39	26.75	0.000	675452	782276.9
aar13	842658.4	26471.96	31.83	0.000	790756.4	894560.3
aar14	959020.5	27403.25	35.00	0.000	905292.6	1012748
alder	-6117.736	492.7377	-12.42	0.000	-7083.817	-5151.656
enebolig	107836.5	46763.74	2.31	0.021	16149.68	199523.3
tomannsbolig	-41834.48	38868.7	-1.08	0.282	-118042	34373
rekkehus	-18083.19	22137.58	-0.82	0.414	-61486.99	25320.61
borettslag	-133201.7	18293.6	-7.28	0.000	-169068.9	-97334.6
postnr1807	-183988.7	31586.58	-5.82	0.000	-245918.6	-122058.9
postnr1808	-203870.7	37049.24	-5.50	0.000	-276510.9	-131230.6
postnr1809	-196682	28836.54	-6.82	0.000	-253220	-140144
postnr1811	-199800.9	26763.48	-7.47	0.000	-252274.4	-147327.4
postnr1812	-159144	31081.4	-5.12	0.000	-220083.4	-98204.62
postnr1813	-148330.3	28682.95	-5.17	0.000	-204567.2	-92093.39
postnr1814	-137468.7	34785.08	-3.95	0.000	-205669.7	-69267.73
postnr1815	-122493.9	90016.66	-1.36	0.174	-298984	53996.26
postnr1832	-279972.6	25933.65	-10.80	0.000	-330819.1	-229126.1
_cons	583221.9	52911.59	11.02	0.000	479481.4	686962.4

Til slutt må vi kontrollere om feilledet er normalfordelt. For at feilledet skal være perfekt normalfordelt må den tykke blå linjen i figuren nedenfor samsvare med den tynne 45 graders linjen. Figur 24 viser at restleddet ikke er perfekt normalfordelt.



Figur 24: Normalfordeling i feilledet, multippel lineær regresjon

6.1.2 Multippel lineær regresjon definert med avstands-dummyer

For å kunne undersøke om boliger som ligger nærme motorveien målt i luftlinje har lavere pris enn boliger som ligger lenger unna, har det i datasettet blitt utarbeidet informasjon om avstand i meter fra hver enkelt bolig til motorvei i luftlinje. Denne informasjonen har videre blitt delt inn i ulike dummy-variabler slik at det er mulig å kjøre en egen regresjon basert på denne informasjonen. Basisboligen i denne analysen vil dermed være en selveier leilighet som ligger i postnummer 1830 eller 1831, som ligger over 1000 meter fra motorveien i luftlinje og er solgt i år 2003. Både årstall og postnummer er inkludert i analysen men blir utelatt i beskrivelsen nedenfor på grunn av at de tar opp mye plass. Den fullstendige regresjonen finnes i vedlegg 7.

Tabell 16: Multipl linear regresjon, avstandsdummyer

Source	SS	df	MS	Number of obs = 3546		
Model	1.2871e+15	29	4.4383e+13	F(29, 3516) = 354.25		
Residual	4.4051e+14	3516	1.2529e+11	Prob > F = 0.0000		
Total	1.7276e+15	3545	4.8734e+11	R-squared = 0.7450		
				Adj R-squared = 0.7429		
				Root MSE = 3.5e+05		
salgspris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	8530.431	155.2939	54.93	0.000	8225.955	8834.906
fellesgjeld	-.6975184	.0545292	-12.79	0.000	-.8044305	-.5906064
alder	-6041.464	296.3425	-20.39	0.000	-6622.485	-5460.443
enebolig	114148.7	21457.09	5.32	0.000	72079.1	156218.3
tomannsbolig	-30316.3	23758.09	-1.28	0.202	-76897.33	16264.72
rekkehus	-23557.48	26283.11	-0.90	0.370	-75089.18	27974.21
borettslag	-140097.4	22418.32	-6.25	0.000	-184051.7	-96143.18
innenfor500	-824.4655	24778.8	-0.03	0.973	-49406.74	47757.81
innenfor1000	-87960.56	18969.87	-4.64	0.000	-125153.6	-50767.5
_cons	555650.9	31652.64	17.55	0.000	493591.5	617710.3

Forklaringskraften til modellen blir omtrent den samme når kjøreavstand blir erstattet med luftavstands-dummyer. Denne regresjonsmodellen viser at avstandsvariabelen, «innenfor500», som inkluderer solgte boliger som ligger innenfor en radius på 500 meter fra motorveien i luftavstand ikke er signifikant på 5 prosent nivå. Modellen indikerer at prisen på basisboligen reduseres med 87.960 kr dersom boligen ligger mellom 501-1000 meter i luftlinje fra motorveien i forhold til om den ligger over 1000 meter unna. På bakgrunn av analysene ovenfor vil andre regresjonsmodeller bli vurdert.

6.1.3 Dobbelt-logaritmisk regresjon

I den dobbelt-logaritmiske regresjonen får vi en funksjonsform hvor vi tar den naturlige logaritmen (ln) både av den avhengige variabelen salgspris og de uavhengige variablene som er kontinuerlige. I denne modellen har vi at dersom en av de kontinuerlige variablene øker med 1 prosent, og alle andre variabler holdes konstant, vil prisen endres med en prosent tilsvarende koeffisienten. Den naturlige logaritmen av negative tall er ikke definert. Dette innebærer at variabler som fellesgjeld og alder får pluss på verdien en, ved at det er noen observasjoner av disse variablene har verdien null. Tabell 17 nedenfor viser den dobbelt-logaritmiske regresjonen:

Tabell 17: Dobbel-logaritmsk regresjon

Source	SS	df	MS	Number of obs = 3546		
Model	553.17373	28	19.7562046	F(28, 3517) = 497.16		
Residual	139.758856	3517	.039738088	Prob > F = 0.0000		
Total	692.932585	3545	.195467584	R-squared = 0.7983		
				Adj R-squared = 0.7967		
				Root MSE = .19934		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6489544	.0112182	57.85	0.000	.6269596	.6709491
lnfellesgjeld	-.0110736	.0015428	-7.18	0.000	-.0140986	-.0080487
lnalder	-.0911465	.0039711	-22.95	0.000	-.0989323	-.0833606
lnavstpåkjemotorvei	-.0039077	.0133253	-0.29	0.769	-.0300338	.0222184
aar04	.0531806	.0164978	3.22	0.001	.0208345	.0855268
aar05	.1374561	.0168543	8.16	0.000	.104411	.1705013
aar06	.1974793	.0172007	11.48	0.000	.1637548	.2312037
aar07	.3133616	.0160597	19.51	0.000	.2818744	.3448489
aar08	.3427	.0166711	20.56	0.000	.3100141	.3753859
aar09	.3288556	.0163717	20.09	0.000	.2967566	.3609547
aar10	.3542468	.0164789	21.50	0.000	.3219377	.3865559
aar11	.4484912	.0155907	28.77	0.000	.4179235	.4790589
aar12	.4880177	.016143	30.23	0.000	.4563671	.5196682
aar13	.5589208	.016158	34.59	0.000	.5272409	.5906007
aar14	.6271653	.0163635	38.33	0.000	.5950825	.6592482
enebolig	.020881	.0131768	1.58	0.113	-.0049539	.046716
tomannsbolig	-.0598866	.013914	-4.30	0.000	-.0871669	-.0326062
rekkehus	-.0195181	.0151904	-1.28	0.199	-.049301	.0102648
borettslag	-.1220849	.0183895	-6.64	0.000	-.1581401	-.0860298
postnr1807	-.0927937	.0147986	-6.27	0.000	-.1218083	-.063779
postnr1808	-.1236388	.0195466	-6.33	0.000	-.1619626	-.0853149
postnr1809	-.1433375	.014647	-9.79	0.000	-.172055	-.1146201
postnr1811	-.1091116	.0149856	-7.28	0.000	-.1384929	-.0797302
postnr1812	-.0401242	.0178085	-2.25	0.024	-.0750402	-.0052082
postnr1813	-.0590757	.0165131	-3.58	0.000	-.091452	-.0266994
postnr1814	-.0981198	.0175238	-5.60	0.000	-.1324777	-.0637619
postnr1815	-.1579137	.0209508	-7.54	0.000	-.1989907	-.1168368
postnr1832	-.1739743	.0131372	-13.24	0.000	-.1997317	-.1482169
_cons	11.32893	.1127045	100.52	0.000	11.10796	11.5499

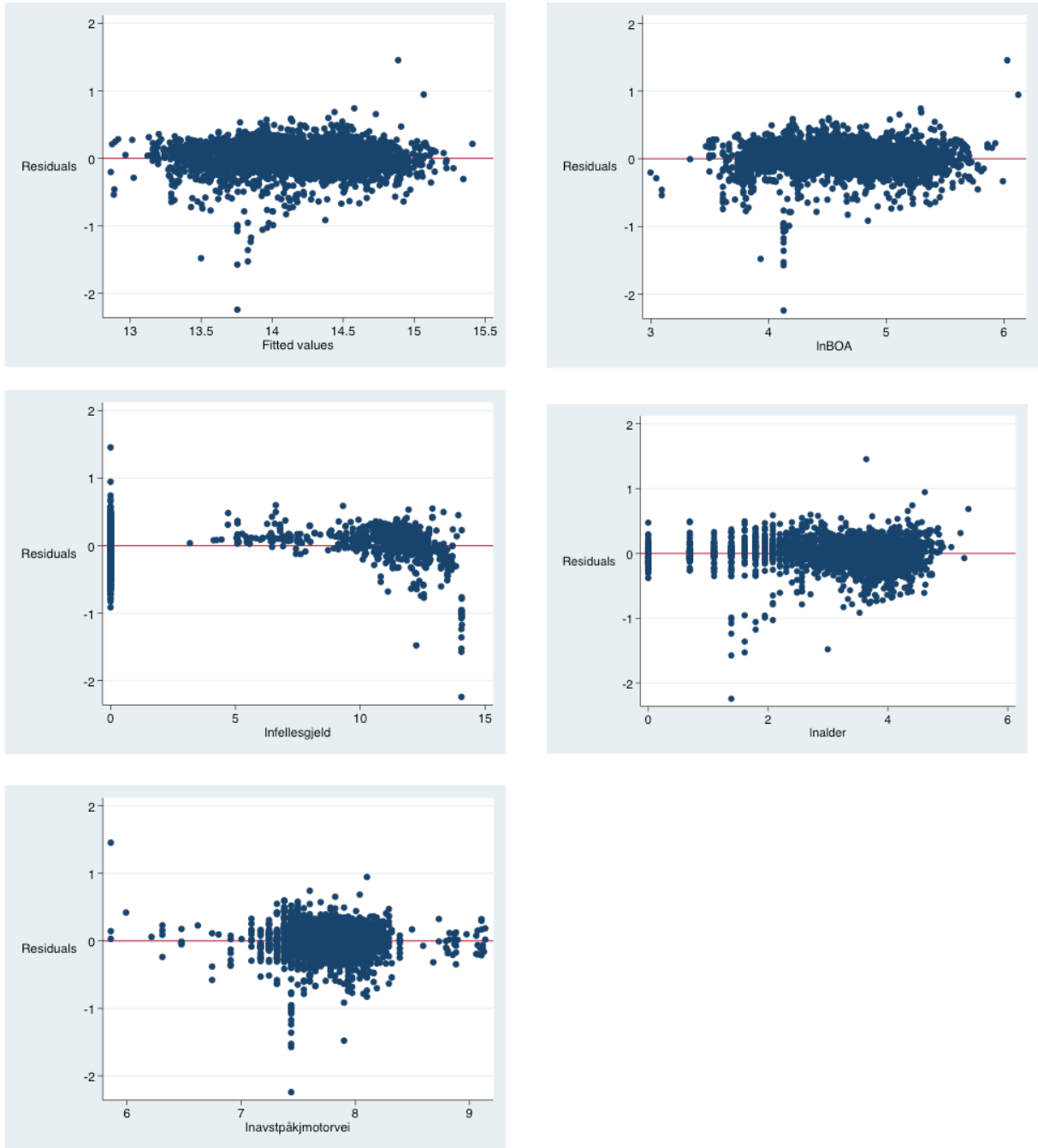
Denne regresjonsmodellen forteller at dersom BOA øker med 1 prosent vil prisen på boligen øke med rundt 64,90 prosent. Koeffisienten til kjøreavstanden målt i meter til nærmeste påkjørsel til motorveien forteller at dersom avstanden øker med 1 prosent vil prisen bli redusert med 3,91 prosent. Forklaringskraften er på hele 79,67 prosent, men allerede her ser det ut som om modellen er dårlig egnet ved at kjøreavstanden til nærmeste påkjørsel ikke er signifikant på et ønsket nivå. Også her må vi kontrollere for multikollinearitet, homoskedastiskitet og om feilledet er normalfordelt.

Av VIF-testen i tabell 18 nedenfor kan vi også for denne regresjonsmodellen konkludere med at vi ikke har noen problemer med multikollinearitet. Den høyeste observerte verdien er 4,85 og gjennomsnittlig VIF er 2,12.

Tabell 18: VIF-test, dobbelt-logaritmisk regresjon

Variable	VIF	1/VIF
borettslag	4.85	0.206070
lnfellesgj~d	4.76	0.210267
enebolig	3.73	0.268148
lnBOA	2.59	0.386112
postnr1832	2.49	0.400851
postnr1807	2.35	0.426049
postnr1809	2.31	0.433446
aar11	1.99	0.502960
postnr1811	1.96	0.509794
postnr1812	1.94	0.516401
aar13	1.90	0.527286
aar12	1.88	0.531290
aar14	1.85	0.540412
aar07	1.85	0.541473
tomannsbolig	1.84	0.544310
aar10	1.82	0.549552
aar09	1.80	0.555025
aar04	1.77	0.564294
postnr1808	1.75	0.569916
aar08	1.74	0.575162
aar05	1.72	0.582675
aar06	1.68	0.595658
postnr1813	1.65	0.606998
postnr1814	1.58	0.632480
lnavstpåkji	1.51	0.661275
postnr1815	1.39	0.717602
lnalder	1.33	0.752409
rekkehus	1.32	0.757948
Mean VIF	2.12	

Figur 25 nedenfor viser residualplottene for hele modellen og de kontinuerlige variablene i analysen. Et tydelig mønster fremkommer for variabelen fellesgjeld, hvor spredningen ser ut til å øke ettersom de predikerte verdiene blir større. Ved at det er vanskelig å finne en konklusjon om nullhypotesen om homoskedastiskitet kan forkastes, gjennomføres det også her en Breusch- Pagan test. Resultatet av denne testen ser vi i tabell 19 nedenfor, $\chi^2 > \text{Prob}$ slik at vi også i denne modellen har problemer med heteroskedastiskitet.



Figur 25: Residualplot, dobbelt-logaritmisk regresjon

Tabell 19: Breusch – Pagan test, dobbelt-logaritmisk regresjon

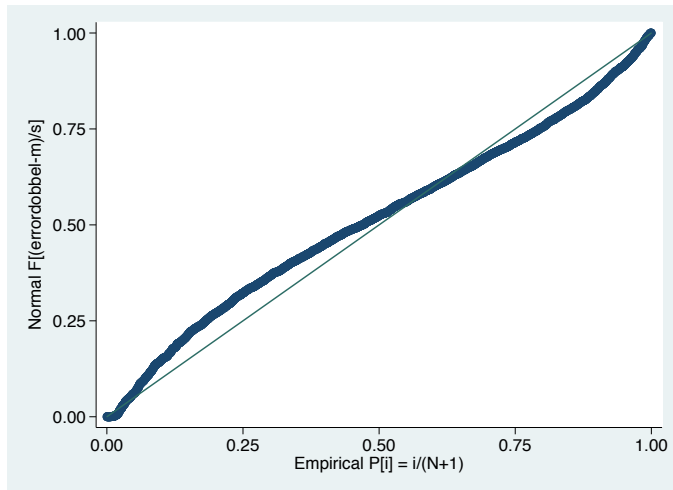
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of lnpris	
chi2(1)	= 61.94
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: lnBOA	
chi2(1)	= 41.54
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: lnfellesgjeld	
chi2(1)	= 537.85
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: lnalder	
chi2(1)	= 71.34
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: lnnavstpåkjemotorvei	
chi2(1)	= 216.79
Prob > chi2	= 0.0000

Av tabell 20 nedenfor kan vi se at de robuste standardfeilene for «lnnavstpåkjemotorvei», «enebolig» er høyere enn den opprinnelige regresjonen, mens for «rekkehus» er standardfeilen litt lavere. De resterende forklaringsvariablene er signifikante på 5 prosent nivå.

Tabell 20: robuste standardfeil, dobbelt-logaritmisk regresjon

Linear regression		Number of obs = 3546				
		F(28, 3517) = 529.81				
		Prob > F = 0.0000				
		R-squared = 0.7983				
		Root MSE = .19934				
lnpris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6489544	.0126106	51.46	0.000	.6242296	.6736791
lnfellesgjeld	-.0110736	.0014656	-7.56	0.000	-.0139472	-.0082001
lnalder	-.0911465	.0045803	-19.90	0.000	-.1001268	-.0821661
lnavstpåkjemotorvei	-.0039077	.0176586	-0.22	0.825	-.0385298	.0307144
aar04	.0531806	.0146115	3.64	0.000	.0245328	.0818285
aar05	.1374561	.016339	8.41	0.000	.1054212	.1694911
aar06	.1974793	.0158263	12.48	0.000	.1664496	.2285089
aar07	.3133616	.0144782	21.64	0.000	.2849751	.3417482
aar08	.3427	.0161434	21.23	0.000	.3110486	.3743514
aar09	.3288556	.0157628	20.86	0.000	.2979504	.3597608
aar10	.3542468	.0190511	18.59	0.000	.3168946	.3915991
aar11	.4484912	.0158749	28.25	0.000	.4173664	.4796161
aar12	.4880177	.0157722	30.94	0.000	.4570941	.5189412
aar13	.5589208	.0155116	36.03	0.000	.5285082	.5893335
aar14	.6271653	.0159776	39.25	0.000	.5958391	.6584916
enebolig	.020881	.0140278	1.49	0.137	-.0066224	.0483845
tomannsbolig	-.0598866	.0139129	-4.30	0.000	-.0871648	-.0326083
rekkehus	-.0195181	.0147417	-1.32	0.186	-.0484213	.0093851
borettslag	-.1220849	.0155995	-7.83	0.000	-.1526698	-.0915
postnr1807	-.0927937	.0146577	-6.33	0.000	-.1215321	-.0640552
postnr1808	-.1236388	.0191108	-6.47	0.000	-.1611081	-.0861695
postnr1809	-.1433375	.0149756	-9.57	0.000	-.1726992	-.1139758
postnr1811	-.1091116	.0145018	-7.52	0.000	-.1375443	-.0806788
postnr1812	-.0401242	.0167968	-2.39	0.017	-.0730566	-.0071918
postnr1813	-.0590757	.0150269	-3.93	0.000	-.0885381	-.0296133
postnr1814	-.0981198	.01598	-6.14	0.000	-.1294508	-.0667888
postnr1815	-.1579137	.0249911	-6.32	0.000	-.2069122	-.1089153
postnr1832	-.1739743	.0138074	-12.60	0.000	-.2010456	-.1469031
_cons	11.32893	.1378521	82.18	0.000	11.05865	11.59921

Til slutt må det kontrolleres om feilleddene er lineære. Figur 26 nedenfor viser at modellen ikke er perfekt normalfordelt ved at den tykke blå linjen ikke ligger jevnt med 45 graders linjen. Når dette er sagt ser vi at denne modellen er nærmere normalfordelt enn den foregående multipel lineære regresjonsmodellen.



Figur 26: Normalfordeling i feilledet, dobbelt-logaritmisk regresjon

6.1.4 Dobbelt-logaritmisk regresjon definert med avstands-dummyer

Det må også utarbeides en analyse som inkluderer avstands-dummyene for den dobbelt-logaritmiske regresjons modellen. Hele regresjonen er presentert i vedlegg 8, nedenfor er årstall og postnummer ikke vist for å spare plass.

Tabell 21: Dobbelt-logaritmisk regresjon, avstands-dummyer

Source	SS	df	MS	Number of obs = 3546		
Model	554.949961	29	19.1362056	F(29, 3516) = 487.62		
Residual	137.982624	3516	.039244205	Prob > F = 0.0000		
Total	692.932585	3545	.195467584	R-squared = 0.8009		
				Adj R-squared = 0.7992		
				Root MSE = .1981		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6450419	.0111651	57.77	0.000	.6231512	.6669326
lnfellesgjeld	-.0115971	.0015358	-7.55	0.000	-.0146083	-.0085859
lnalder	-.0915989	.0039785	-23.02	0.000	-.0993994	-.0837985
enebolig	.031275	.0131833	2.37	0.018	.0054274	.0571226
tomannsbolig	-.0489145	.0139398	-3.51	0.000	-.0762454	-.0215837
rekkehus	-.0119836	.0152074	-0.79	0.431	-.0417998	.0178326
borettslag	-.1165394	.0182727	-6.38	0.000	-.1523656	-.0807133
innenfor500	-.0728611	.0138125	-5.28	0.000	-.0999423	-.0457798
innenfor1000	-.0616135	.0106625	-5.78	0.000	-.0825188	-.0407082
_cons	11.33973	.0517579	219.09	0.000	11.23825	11.4412

Forklaringskraften forblir omtrent på det samme nivået når kjøreavstand blir erstattet med luftavstands-dummyer. Denne regresjonsmodellen forteller at dersom en bolig ligger innenfor 500 meter fra motorveien målt i luftavstand er prisen omtrent 7,29 prosent lavere enn en bolig som ligger over 1000 meter unna. For en bolig som ligger mellom 501-1000 meter unna er

prisen 6,16 prosent lavere enn en bolig som ligger over 1000 meter unna. Dette er tall som er signifikante på 5 prosent nivå.

6.1.5 Semi-logaritmisk regresjon

Den semi-logaritmiske funksjonsformen er en form hvor man står mer fritt til å velge hvilke variabler som skal uttrykkes ved hjelp av naturlige logaritmer (ln). Jeg velger å kun uttrykke den avhengige variabelen, prisen, ved hjelp av ln. I denne modellen har vi at dersom en av de uavhengige variablene øker med en *enhet*, så vil den tilhørende koeffisienten gi den *prosentvise* endringen i prisen, når alt annet holdes konstant. Tabellen nedenfor viser den semi-logaritmiske regresjonen:

Tabell 22: Semi-logaritmisk regresjon

Source	SS	df	MS	Number of obs = 3546		
Model	538.265921	28	19.2237829	F(28, 3517) = 437.13		
Residual	154.666664	3517	.043976873	Prob > F = 0.0000		
				R-squared = 0.7768		
				Adj R-squared = 0.7750		
Total	692.932585	3545	.195467584	Root MSE = .20971		

lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0045119	.0000919	49.09	0.000	.0043317	.0046921
fellesgjeld	-7.22e-07	3.23e-08	-22.34	0.000	-7.85e-07	-6.58e-07
avstpåkjemotorvei	-.0000203	4.86e-06	-4.17	0.000	-.0000298	-.0000108
aar04	.0684301	.0173553	3.94	0.000	.0344026	.1024576
aar05	.1554425	.0177288	8.77	0.000	.1206826	.1902023
aar06	.2343032	.0181129	12.94	0.000	.1987903	.2698161
aar07	.3363974	.0169026	19.90	0.000	.3032575	.3695372
aar08	.3624752	.0175332	20.67	0.000	.3280989	.3968514
aar09	.347608	.0172126	20.19	0.000	.3138603	.3813557
aar10	.3771782	.0172957	21.81	0.000	.3432675	.4110889
aar11	.4840447	.0163562	29.59	0.000	.4519761	.5161133
aar12	.5010002	.0169086	29.63	0.000	.4678486	.5341518
aar13	.5897613	.0169623	34.77	0.000	.5565043	.6230182
aar14	.6521403	.0171779	37.96	0.000	.6184606	.6858199
alder	-.0041845	.0001745	-23.98	0.000	-.0045267	-.0038424
enebolig	.1805236	.0126414	14.28	0.000	.1557383	.2053088
tomannsbolig	.0735009	.0139602	5.27	0.000	.04613	.1008717
rekkehus	.0800435	.015484	5.17	0.000	.0496849	.1104021
borettslag	-.0908375	.0133662	-6.80	0.000	-.1170438	-.0646312
postnr1807	-.1485568	.0152879	-9.72	0.000	-.1785309	-.1185827
postnr1808	-.2059707	.0202866	-10.15	0.000	-.2457453	-.1661961
postnr1809	-.1568718	.0151001	-10.39	0.000	-.1864776	-.1272659
postnr1811	-.1030373	.0157751	-6.53	0.000	-.1339665	-.072108
postnr1812	-.1178354	.0185677	-6.35	0.000	-.1542398	-.0814309
postnr1813	-.0936922	.0173683	-5.39	0.000	-.1277452	-.0596392
postnr1814	-.0823332	.0188506	-4.37	0.000	-.1192924	-.045374
postnr1815	-.1617919	.0219065	-7.39	0.000	-.2047427	-.1188411
postnr1832	-.2055755	.0138285	-14.87	0.000	-.2326882	-.1784628
_cons	13.59871	.0203741	667.45	0.000	13.55877	13.63866

Denne modellen har en forklaringskraft på 77,50 prosent. Av denne regresjonen kan vi lese at dersom BOA øker med en kvadratmeter, vil boligprisen øke med omtrent 0,45 prosent. Som vi har sett av de tidligere gjennomgåtte modellene har alle hatt en relativ høy forklaringskraft, for å kunne skille de ulike modellene må vi også her kontrollere for multikollinearitet, homoskedastisitet og at feilleddet er normalfordelt.

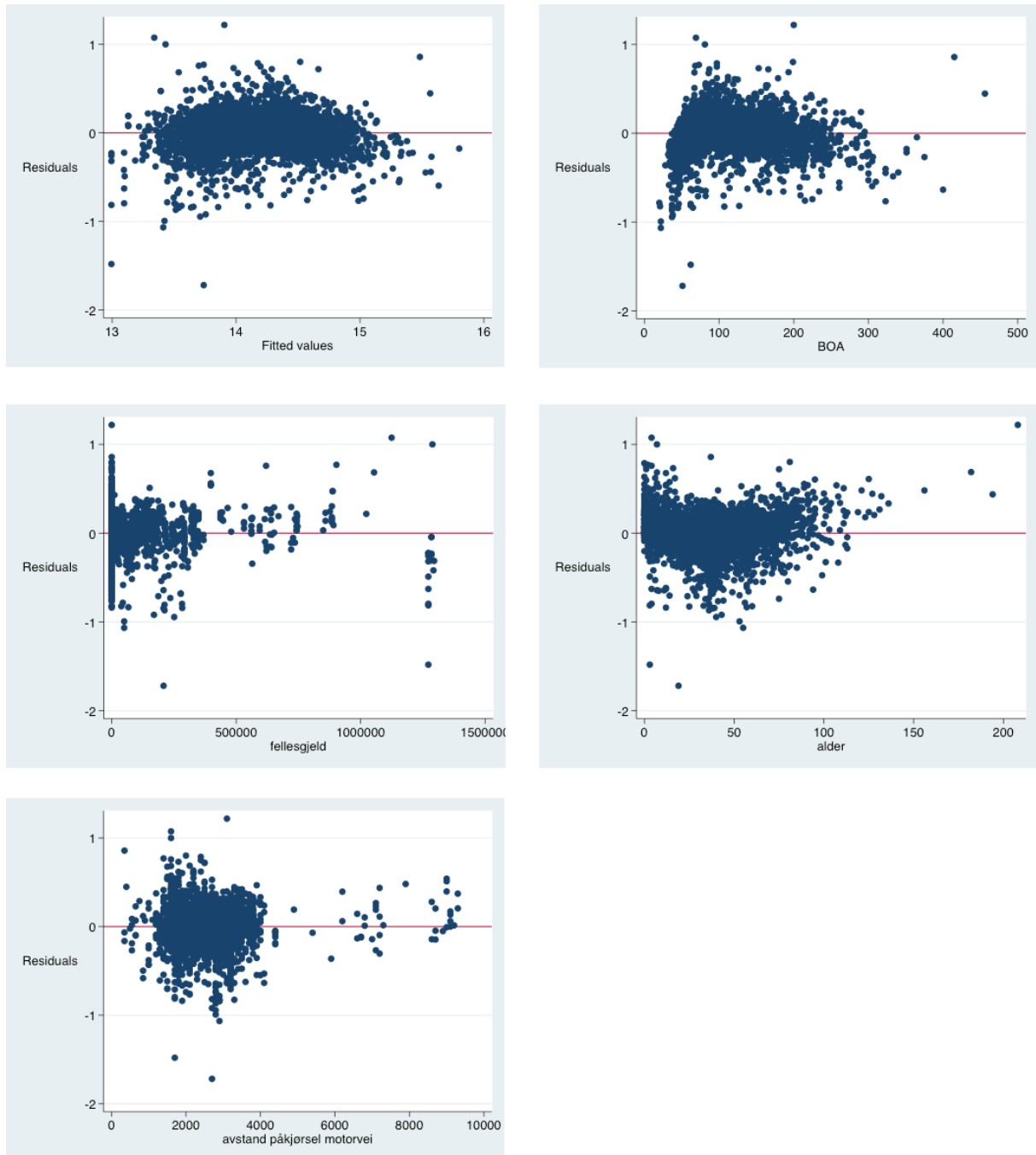
VIF-testen som er summert i tabell 23 nedenfor forteller at vi ikke vil få problemer med multikollinearitet ved at $VIF < 10$ for alle forklaringsvariablene. Gjennomsnittlig VIF for denne modellen er på 1,86.

Tabell 23: VIF-test, semi-logaritmisk regresjon

Variable	VIF	1/VIF
enebolig	3.10	0.322419
postnr1832	2.50	0.400368
borettslag	2.32	0.431674
postnr1807	2.26	0.441794
postnr1809	2.22	0.451324
BOA	2.10	0.477218
aar11	1.98	0.505727
postnr1811	1.96	0.509117
postnr1812	1.90	0.525708
aar13	1.89	0.529499
aar12	1.87	0.535923
aar07	1.85	0.540957
aar14	1.84	0.542692
aar10	1.81	0.552080
aar09	1.80	0.555680
aar04	1.77	0.564298
aar08	1.74	0.575458
fellesgjeld	1.72	0.582608
aar05	1.72	0.582779
postnr1808	1.71	0.585539
aar06	1.68	0.594471
tomannsbolig	1.67	0.598395
postnr1814	1.65	0.604887
postnr1813	1.65	0.607222
avstpåkjo~i	1.41	0.710688
postnr1815	1.38	0.726363
alder	1.34	0.748452
rekkehus	1.24	0.807289
Mean VIF	1.86	

Figur 27 nedenfor viser residualplottene til modellen og de kontinuerlige variablene i analysen. Også for denne modellen ser det ut til at heteroskedastisitet er til stede ved at de

predikerte verdiene blir trangere ettersom de øker. Dette gjelder for boareal, alder og avstand til påkjørsel motorvei. Av plottet for fellesgjeld kan det se ut til at de predikerte verdiene får mer variasjon ettersom verdiene blir større.



Figur 27: Residualplot, semi-logaritmisk regresjon

For å kunne si sikkert at heteroskedastisitet faktisk er til stede gjennomføres det også her en Breusch – Pagan test som er å finne i tabell 24 nedenfor. Denne testen viser at vi ikke kan forkaste nullhypotesen for de kontinuerlige variablene boareal og avstand til påkjørsel

motorvei. For modellen og variabelen fellesgjeld og alder kan nullhypotesen om homoskedastisitet forkastes med et signifikansnivå på 5 prosent.

Tabell 24: Breusch – Pagan test, semi-logaritmisk regresjon

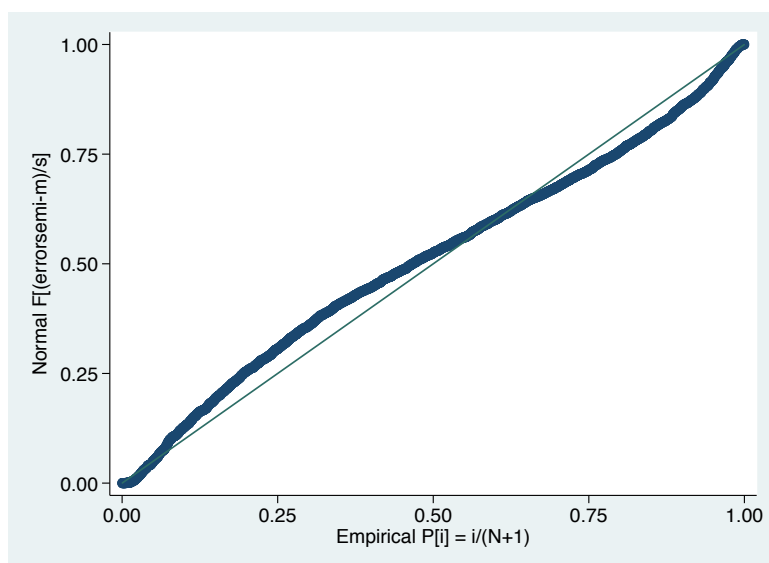
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of lnpris	
chi2(1)	= 84.69
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: BOA	
chi2(1)	= 1.23
Prob > chi2	= 0.2681
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fellesgjeld	
chi2(1)	= 426.98
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: alder	
chi2(1)	= 76.22
Prob > chi2	= 0.0000
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: avstpåkjemotorvei	
chi2(1)	= 3.10
Prob > chi2	= 0.0782

Ved at heteroskedastisitet er til stede i modellen må standardfeilene korrigeres, slik at vi fortsatt kan benytte OLS. Av tabell 25 nedenfor kan det vises at alle variablene som er inkludert i analysen er signifikante når vi har korrigert for heteroskedastisitet.

Tabell 25: robuste standardfeil, semi-logaritmisk regresjon

Linear regression		Number of obs = 3546				
		F(28, 3517) = 390.55				
		Prob > F = 0.0000				
		R-squared = 0.7768				
		Root MSE = .20971				
lnpris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0045119	.0001184	38.11	0.000	.0042797	.004744
fellesgjeld	-7.22e-07	6.78e-08	-10.64	0.000	-8.55e-07	-5.89e-07
avstpåkjemotorvei	-.0000203	5.45e-06	-3.73	0.000	-.000031	-9.61e-06
aar04	.0684301	.0159781	4.28	0.000	.0371028	.0997575
aar05	.1554425	.0169515	9.17	0.000	.1222067	.1886782
aar06	.2343032	.0169784	13.80	0.000	.2010148	.2675916
aar07	.3363974	.0165126	20.37	0.000	.3040222	.3687725
aar08	.3624752	.0180923	20.03	0.000	.3270027	.3979476
aar09	.347608	.0180407	19.27	0.000	.3122366	.3829793
aar10	.3771782	.0182347	20.68	0.000	.3414264	.4129299
aar11	.4840447	.0160809	30.10	0.000	.4525159	.5155735
aar12	.5010002	.0171082	29.28	0.000	.4674571	.5345433
aar13	.5897613	.0170146	34.66	0.000	.5564017	.6231208
aar14	.6521403	.0165092	39.50	0.000	.6197716	.6845089
alder	-.0041845	.0002548	-16.42	0.000	-.0046841	-.003685
enebolig	.1805236	.0137198	13.16	0.000	.1536241	.207423
tomannsbolig	.0735009	.0131958	5.57	0.000	.0476287	.099373
rekkehus	.0800435	.015706	5.10	0.000	.0492497	.1108373
borettslag	-.0908375	.0166619	-5.45	0.000	-.1235056	-.0581695
postnr1807	-.1485568	.0161195	-9.22	0.000	-.1801612	-.1169524
postnr1808	-.2059707	.0230286	-8.94	0.000	-.2511215	-.1608199
postnr1809	-.1568718	.0162706	-9.64	0.000	-.1887725	-.124971
postnr1811	-.1030373	.0163574	-6.30	0.000	-.1351082	-.0709663
postnr1812	-.1178354	.0178051	-6.62	0.000	-.1527447	-.082926
postnr1813	-.0936922	.016149	-5.80	0.000	-.1253547	-.0620298
postnr1814	-.0823332	.0175553	-4.69	0.000	-.1167528	-.0479137
postnr1815	-.1617919	.0250688	-6.45	0.000	-.2109427	-.1126411
postnr1832	-.2055755	.0142334	-14.44	0.000	-.233482	-.177669
_cons	13.59871	.0222746	610.50	0.000	13.55504	13.64238

Til slutt må det kontrolleres om vi har normalfordeling i feilledet. Figur 28 nedenfor viser at denne modellen heller ikke er perfekt normalfordelt ved at den tykke blå linjen ikke ligger jevnt med 45 graders linjen. Til tross for dette ligger de predikerte verdiene i denne modellen nærmere 45 graders linjen enn tidligere.



Figur 28: Normalfordeling i feilledet, semi-logaritmisk regresjon

6.1.6 Semi-logaritmisk regresjon definert med avstands-dummyer

Det må også her foretas en regresjonsanalyse hvor avstands-dummyene inkluderes. For å spare plass er årstall og postnummer ikke vist i tabellen nedenfor. Hele regresjonen finnes i vedlegg 9.

Tabell 26: Semi-logaritmisk regresjon, avstands-dummyer

Source	SS	df	MS	Number of obs = 3546		
Model	539.557117	29	18.6054178	F(29, 3516) =	426.51	
Residual	153.375468	3516	.043622147	Prob > F =	0.0000	
Total	692.932585	3545	.195467584	R-squared =	0.7787	
				Adj R-squared =	0.7768	
				Root MSE =	.20886	
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0044915	.0000916	49.02	0.000	.0043118	.0046711
fellesgjeld	-7.08e-07	3.22e-08	-22.02	0.000	-7.72e-07	-6.45e-07
alder	-.0041647	.0001749	-23.82	0.000	-.0045075	-.0038218
enebolig	.1874957	.0126611	14.81	0.000	.1626719	.2123196
tomannsbolig	.0838402	.0140188	5.98	0.000	.0563543	.111326
rekkehus	.0797876	.0155088	5.14	0.000	.0493805	.1101947
borettslag	-.097507	.0132283	-7.37	0.000	-.1234428	-.0715711
innenfor500	-.0361341	.0146211	-2.47	0.014	-.0648008	-.0074674
innenfor1000	-.0768569	.0111935	-6.87	0.000	-.0988032	-.0549105
_cons	13.58397	.0186771	727.31	0.000	13.54735	13.62059

Forklaringskraften blir omtrent den samme når vi benytter avstand til motorvei i luftavstand definert som dummy-variabler. Denne regresjonen forteller at dersom en bolig ligger innenfor

500 meter i luftavstand fra motorveien vil prisen på denne boligen være 3,61 prosent lavere enn en bolig som ligger over 1000 meter unna. Dersom boligen ligger mellom 501-1000 meter i luftlinje fra motorveien viser regresjonen at denne boligen vil være 7,68 prosent lavere i pris enn en bolig som ligger over 1000 meter fra motorveien. Alle variablene i denne analysen er signifikante på 5 prosent nivå.

6.2 Valg av modell

Før hypotesetestingen kan gjennomføres må den beste tilpassede modellen til datasettet velges. Den modellen som beskriver datasettet på en god måte og ivaretar restleddets forsetninger vil bli benyttet.

Av de tre gjennomgatte regresjonsmodellene har alle en god forklaringskraft ved at \bar{R}^2 er relativt høy. Ingen av modellene har multikollinearitet og alle må forkaste nullhypotesen om homoskedastisitet. Restleddet er best tilpasset normalfordelingen i den semi-logaritmiske regresjonsmodellen. Alle koeffisientene i denne modellen er signifikante, selv når problemet med heteroskedastisitet er løst ved hjelp av robuste standardfeil. Den dobbelt-logaritmiske regresjonsmodellen skiller seg ut ved at koeffisienten til kjøreavstand målt i meter til motorvei ikke er signifikant og kan dermed ikke si noe om det er positivt eller negativt å bo nærme en påkjørsel til motorveien målt i kjøreavstand. På bakgrunn av tilnærmet normalfordelte feilledd og at signifikansen til forklaringsvariablene ikke ble svekket på grunn av heteroskedastisitet, vil den semi-logaritmiske modellen bli brukt til hypotesetestingen.

6.3 Hypotesetesting

Når hypotesene skal testes vil P-verdiene bli benyttet som en alternativ tilnærming til t-testen som ble gjennomgått i kapittel 4. P- verdien angir det laveste nivået av signifikans som nullhypotesen kan forkastes på. Hypotesene vil bli testet ved bruk av den semi-logaritmiske modellen med avstand til nærmeste påkjørsel til motorveien som en kontinuerlig variabel, bortsett fra hypotese 2 hvor den semi-logaritmiske modellen med avstands-dummys målt i luftavstand vil bli benyttet.

Hypotese 1 - kjøreavstand til nærmeste påkjørsel på motorveien

H_0^1 : Kjøreavstand til påkjørsel motorvei målt i meter har ingen betydning for boligprisen

H_A^1 : Kjøreavstand til påkjørsel motorvei målt i meter har betydning for boligprisen

Koeffisienten som tilhører variabelen «avstpåkjmotorvei» indikerer at boligprisen synker med 0,00203 prosent for hver meter kjøreavstanden til motorveien øker. Ved at den tilhørende P-verdien er 0,000 er variabelen signifikant på 5 prosent nivå og nullhypotesen kan forkastes med 95 prosent sannsynlighet. Konklusjonen blir at kjøreavstand til motorvei har betydning for boligprisen. Analysen viser en negativ sammenheng mellom kjøreavstand til motorveien og boligpriser i Askim.

Hypotese 2: Negativ effekt å bo nærme motorveien målt i luftavstand

H_0^2 : Prisen på en bolig som er lokalisert nærme motorveien målt i luftavstand vil være høyere enn en bolig som ligger lenger unna.

H_A^2 : Prisen på en bolig som er lokalisert nærme motorveien målt i luftavstand vil være lavere enn en bolig som ligger lenger unna.

Koeffisienten til boligene som ligger innenfor 500 meter målt i luftavstand fra motorveien har en P-verdi på 0,014. Dette betyr at variabelen vil være signifikant på 5 prosent nivå, og med 95 prosent sannsynlighet kan vi påstå at en bolig innenfor 500 meter målt i luftavstand fra motorveien vil ha en pris som er 3,613 prosent lavere enn basisboligen som er lokalisert over 1000 meter unna motorveien.

Koeffisienten til boligene som ligger mellom 501-1000 meter målt i luftavstand fra motorveien har en P-verdi på 0,000. Denne variabelen er derfor også signifikant på 5 prosent nivå, og med 95 prosent sannsynlighet kan vi påstå at en bolig med denne lokaliseringen vil ha en pris som er 7,686 prosent lavere enn basisboligen som ligger over 1000 meter unna motorveien.

I kapittel 5 ble det antatt at boliger som ligger innenfor 500 meter i luftavstand fra motorveien vil være ganske strekt påvirket av eksternalitetene som følger av trafikken fra motorveien, boliger mellom 501-1000 meter vil heller ikke være upåvirket ved at lyden reiser langt blant

annet på grunn av flate områder. Boliger som ligger over 1000 meter unna motorveien i luftlinje antas å være upåvirket. Koeffisientene til innenfor 500 meter og 501-1000 meter er henholdsvis -3,613 prosent og -7,686 prosent. Det at begge koeffisientene er negative indikerer at boliger som er lokalisert forholdsvis nærme motorveien målt i luftavstand vil ha en lavere pris enn boliger som er lokalisert lenger unna. Boliger som ligger innenfor 1000 meter fra motorveien målt i luftlinje vil altså ha en lavere pris enn boliger som er lokalisert over 1000 meter unna. Det at boliger som ligger innenfor 500 meter i luftlinje fra motorveien har en høyere pris enn de som ligger mellom 501-1000 meter unna var et uventet resultat. En skulle tro at eksternalitetene som følger den nye motorveien ville gjøre at boligprisene innenfor en radius av 500 meter var lavere enn boligene som ligger litt lenger unna. Det faktum at boligprisene er lavere når boligen er lokalisert innenfor 1000 meter fra motorveien i forhold basisboligen som er lokalisert over 1000 meter unna gjør at nullhypotesen kan forkastes. Konklusjonen blir at en bolig som er lokalisert nærme motorveien målt i luftlinje vil ha en lavere pris enn en bolig som ligger lenger unna.

Hypotese 3: Sammenheng mellom boareal og boligpris

H_0^3 : Det er ingen sammenheng mellom boareal og boligpris

H_A^3 : Det er sammenheng mellom boareal og boligpris

Koeffisienten som tilhører variabelen «BOA» indikerer at boligprisen øker med 0,45119 prosent per kvadratmeter. Den tilhørende P-verdien er 0,000 og med 95 prosent sikkerhet kan nullhypotesen forkastes. Det kan konkluderes med at det er en positiv sammenheng mellom boareal og pris.

7. Nærmere drøfting av hovedproblemstilling

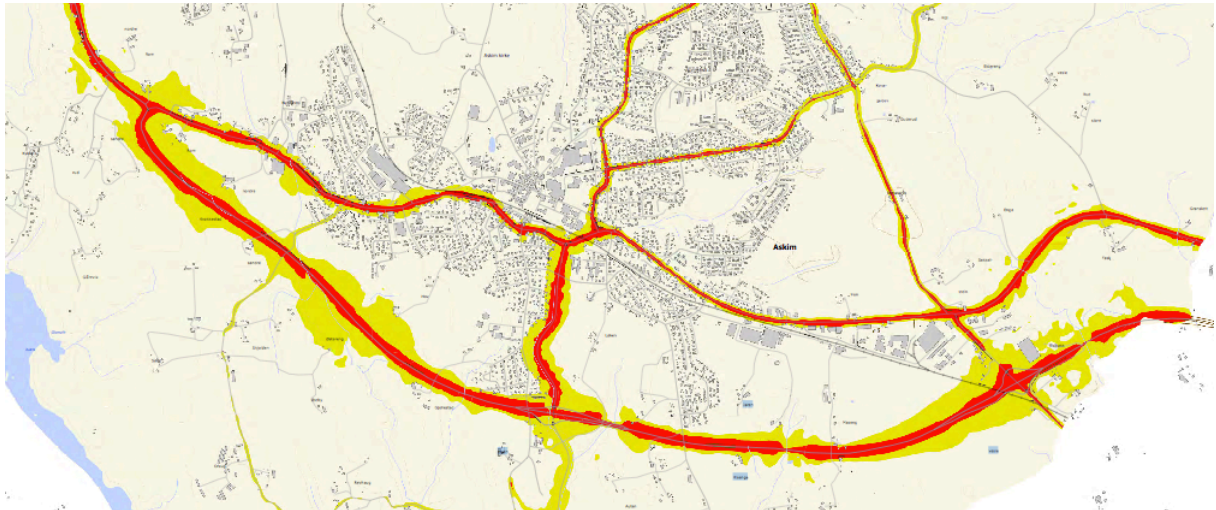
I forrige kapittel kunne vi konkludere med at lengre kjøreavstand til motorveien hadde en negativ effekt på boligprisene. Samtidig fant vi at boliger som er lokalisert innenfor 500 meter og mellom 501-1000 meter fra motorveien målt i luftavstand har en lavere pris enn boliger som er lokalisert over 1000 meter unna.

7.1 Tolkning av resultatene

Under hypotese 1 ble det undersøkt om kjøreavstand i meter til nærmeste påkjørsel på motorveien har hatt innvirkning på boligprisene i Askim. Resultatet ble som forventet ved at boligprisene ble redusert med lenger kjøreavstand til motorveien. Dette er i tråd med teorien; redusert reisetid og reisekostnader resulterer i økte boligpriser.

Under hypotese 2 ble det undersøkt om boliger som er lokalisert nærme motorveien målt i luftlinje vil ha en lavere pris enn boliger som er lokalisert lenger unna. Analysen viste at boliger som ligger innenfor 1000 meter fra motorveien har en lavere pris enn de som ligger over 1000 meter unna. Resultatene ble ikke helt som forventet ved at boliger som var lokalisert mellom 501-1000 meter i luftlinje fra motorveien hadde den laveste boligprisen. Det som kanskje er mer intuitivt å forvente er at de boligene som er plassert nærmest motorveien har lavest pris på grunn av mest støy og forurensing.

I figur 29 nedenfor ser vi et oversiktlig støykart over Askim som viser hvor bebyggelse og hovedveier ligger i forhold til hverandre. Den nederste markerte veien viser ny motorvei, mens veien ovenfor er gamle E18 som nå er blitt fylkesvei 128. Ved at motorveien er lagt utenfor de største tettbebyggelsene kan mangel på observasjoner av nærliggende boliger målt i luftlinje være en grunn til at resultatene i analysen ikke ble som forventet. Av kartet kan det være rimelig å anta at mange av boligene som ligger mellom 501-1000 meter i luftlinje fra motorveien ligger ganske nærme fylkesveien. Ved at fylkesveien må benyttes for å komme på motorveien kan det være at denne veien har blitt mer trafikkert. Dersom denne økte trafikken skaper mer støy og forurensing, kan dette være grunnen til at boligene som ligger mellom 501-1000 meter unna motorveien i luftlinje har lavest boligpris. Dette er ikke en problemstilling i denne oppgaven, men det kan anses som en mulig forklaring til funnene i analysen.



Figur 29: Støykart over Askim (Statens vegvesen, 2013)

7.2 Undersøkelse av verdiendring

Her vil det først og fremst bli sett på ulike scenarier hvor relevante variabler i analysen blir forandret. Basisboligen er som nevnt tidligere en selveier leilighet lokalisert med postnummer 1830 eller 1831 som er solgt i år 2003. Dersom vi setter et boareal på basisboligen lik 90 kvadratmeter, vil prisen være:

$$P = e^{13,59871+(0,0045119*90)} = 1.208.368$$

Når vi endrer basisboligen til bli solgt i år 2013, vil prisen på denne boligen være:

$$P = e^{13,59871+(0,0045119*90)+(0,5897613)} = 2.179.361$$

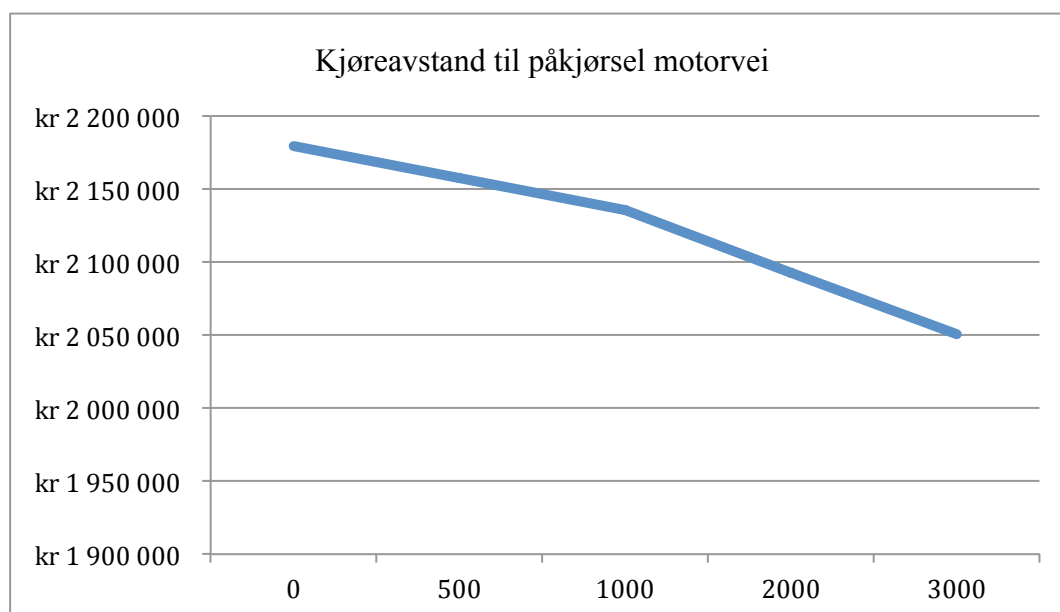
Når kjøreavstanden til nærmeste påkjørsel settes til 100 meter, vil boligprisen være:

$$P = e^{13,59871+(0,0045119*90)+(0,5897613)+(-0,0000203*100)} = 2.174.941$$

Prisen har blitt redusert med 4.420 kr når kjøreavstanden til motorveien økte fra 0 til 100 meter. I tabellen nedenfor er det blitt estimert pris på denne type bolig for ulike kjøreavstander til motorveien.

Tabell 27: Verdiendring med avstand som kontinuerlig variabel

Kjøreavstand (meter)	Boligpris	Endring i pris
0	2 179 361	
500	2 157 352	- 22 009
1000	2 135 566	- 21 786
2000	2 092 651	- 42 915
3000	2 050 599	- 42 052



Figur 30: Boligpris basert på kjøreavstand til motorvei

Tabellen og figuren ovenfor viser at boligprisen synker med økende kjøreavstand til påkjørsel på motorveien.

Videre skal vi se på hvordan verdiendringen er som følge av at avstand til motorveien blir målt i luftavstand. Disse målingene er som vi vet definert ved hjelp av dummy-variabler i analysen. Basisboligen ligger over 1000 meter fra motorveien målt i luftavstand. Dersom boligen ligger innenfor 500 meter fra motorveien målt i luftavstand vil en basisbolig på 90 kvm solgt i år 2013 vil ha prisen:

$$P = e^{13,58397+(0,0044915*90)+0,5891276+(-0,0361341)} = 2.066.152$$

Hvis boligen i stedet ligger mellom 501-1000 meter i luftavstand fra motorveien, vil prisen være:

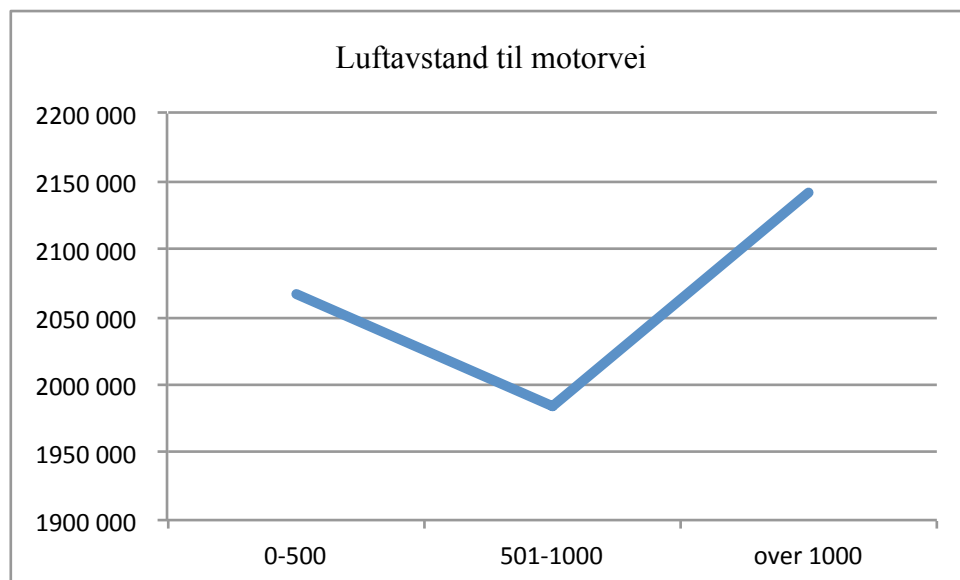
$$P = e^{13,58397+(0,0044915*90)+0,5891276+(-0,0768569)} = 1.983.703$$

Hvis boligen ligger over 1000 meter fra motorveien, vil prisen være:

$$P = e^{13,58397+(0,0044915*90)+0,5891276} = 2.142.176$$

Tabell 28: Verdiendring med avstands-dummyer

Luftavstand (meter)	Boligpris	Ending i pris
0-500	2 066 152	
501-1000	1 983 703	- 82 449
over 1000	2 142 176	158 473



Figur 31: Boligpris basert på luftavstand til motorvei

Dersom denne basisboligen med 90 kvadratmeter solgt i år 2013 ligger innenfor 500 meter i luftlinje fra motorveien vil verdien på boligen være 82.449 kr høyere enn hvis den hadde

ligget mellom 501-1000 meter i luftlinje fra motorveien. Dersom boligen ligger over 1000 meter i luftlinje fra motorveien vil verdien være høyest. Prisen vil være 76.024 kr høyere enn boligen som ligger innenfor 500 meter fra motorveien, mens den vil være 158.473 kr høyere enn boligen som ligger mellom 501-1000 meter unna motorveien i luftlinje.

7.3 Kjente begrensninger i analysen

Virkeligheten er veldig kompleks noe som gjør den vanskelig å analysere. Det finnes variabler som ikke er inkludert i analysen, men som i virkeligheten er med på å forklare utviklingen i boligprisene. Tidligere i dette kapittelet så vi at dersom man skal på en av påkjørslene til motorveien i Askim så må fylkesvei 128 benyttes. En svakhet i analysen er dermed at jeg ikke har tatt hensyn til at mange av de solgte boligene som ligger langs denne veien også opplever økt støy og forurensning på grunn av økt trafikk. Innledningsvis i oppgaven så vi at mange som bor i Askim jobber i en større kommune. Det kan dermed tenkes at avstand og tilgjengelighet til kollektivtrafikk spiller inn når en bolig skal kjøpes. Dette er heller ikke hensyntatt i analysen. Det samme gjelder nærhet til skoler og barnehager. Alle attributtene som er nevnt ovenfor er viktige kriterier når en bolig skal kjøpes.

I denne oppgaven har fellesgjeld vært en kontinuerlig variabel og blitt behandlet som «vanlig gjeld». Det er blitt gjort en del studier på dette området og det er funnet empirisk støtte for at fellesgjelden har en rentediskonterings effekt, som betyr at denne gjelden ofte er billigere enn vanlig boliglån. Dette bidrar til høyere egenkapital pris for boliger i borettslag enn for selveide boliger (Robertsen & Theisen, 2011). Dette kan bli sett på som en begrensning ved at fellesgjelden i denne oppgaven er overpriset.

8. Konklusjon

Målet med denne masteroppgaven var å undersøke hvilken betydning avstand til ny E18 har hatt for boligprisene i Askim. For å belyse dette ble flere variabler som kunne ha innvirkning på boligprisene inkludert, hvor alle variablene ble analysert i den semi-logaritmiske modellen som var best tilpasset datasettet.

For hypotese 1 som omhandlet hvorvidt kjøreavstand til nærmeste påkjørsel på motorveien har noen innvirkning på boligprisene, viste analysen at for hver meter økning i kjøreavstand synker boligprisene med 0,00203 prosent. Dette er signifikant innenfor 5 prosent og konklusjonen blir at kjøreavstanden til motorveien har en innvirkning på boligprisene i Askim. Boligprisene synker ettersom kjøreavstanden til motorveien øker.

Hypotese 2 sa noe om det er negativt å bo nærme motorveien målt i luftavstand i forhold til å bo lenger unna. Ved at disse målingene ble delt inn i dummy-variabler viste analysen at dersom boligen ligger innenfor 500 meter fra motorveien vil prisen være 3,613 prosent lavere enn en bolig som er lokalisert over 1000 meter fra motorveien. Videre så vi at boliger som lå mellom 501-1000 meter i luftlinje fra motorveien ville ha en pris som var 7,686 prosent lavere enn en bolig som var lokalisert over 1000 meter fra motorveien. Disse koeffisientene bekrefter at boliger som ligger nærme motorveien målt i luftlinje vil ha en lavere pris enn boliger som ligger over 1000 meter unna. Med et signifikansnivå på 5 prosent kan vi konkludere med at boliger som ligger mellom 0-1000 meter målt i luftlinje fra motorveien har en lavere pris enn boliger som ligger over 1000 meter unna. Det som er interessant er at boliger som er lokalisert innenfor 500 meter har i følge analysen en høyere pris enn boliger som er lokalisert mellom 501-1000 meter unna.

Kontrollhypotesen som omhandlet at det var sammenheng mellom boareal og boligpris viste seg som forventet å stemme. Analysen viste at boligprisen økte med 0,45119 prosent for hver kvadratmeter boligen økte med.

Et forslag til videreføring av dette arbeidet er å se om det faktisk er fylkesvei 128 som gjør at boligprisene er lavest for boligene som er lokalisert mellom 501-1000 meter i luftlinje fra motorveien. Et annet forslag er å se om det er sentraliseringen av Askim i forhold til større kommuner som gjør at boligprisene ser ut til å vokse relativt mye. I denne sammenheng kan

for eksempel kollektivtrafikk komme inn i bildet. Ved at kollektivtrafikken skal utbedres vil det være interessant å gjennomføre en slik analyse om noen år.

Litteraturliste

- Anundsen, A., & Jansen, E. (2013). Self-reinforcing effects between housing prices and credit. *J. Hous. Econ.*, 22(3), 192-212. doi: 10.1016/j.jhe.2013.07.001
- Askim kommune. (2009). Askim på 9. Plass av Norges beste næringsbyer 2009 Hentet 23.02.2015, fra <http://www.askim.kommune.no/askim-paa-9-plass-av-norges-beste-naeringsbyer-2009.4984213-195261.html>
- Boligareal. (2010). Store norske leksikon. Hentet fra <https://snl.no/boligareal>
- Bolstad, E. (2014). Postnummer med koordinater. Hentet 16.03.2015, fra <http://www.erikbolstad.no/geo/noreg/postnummer/>
- Brooks, C. (2008). *Introductory econometrics for finance* (2nd ed. utg.). Cambridge: Cambridge University Press.
- DiPasquale, D., & Wheaton, W. C. (1996). *Urban economics and real estate markets*. Englewood Cliffs, N.J: Prentice Hall.
- Eiendomsverdi. (2015). Eiendomsinformasjon satt i system Hentet 03.03.2015, fra <https://eiendomsverdi.no>
- Emblem, A. W. (2014). [Forelesningsnotater be-409, eiendomsøkonomi].
- Gersh, S. (2013). Nabostøy: Den skjulte plage (13005/2013) Hentet fra [http://www.stoyforeningen.no/content/download/1635/14000/file/Nabost%C3%B8yrapport 2013 Hdir - endelig.pdf](http://www.stoyforeningen.no/content/download/1635/14000/file/Nabost%C3%B8yrapport%202013%20Hdir%20-%20endelig.pdf).
- Gripsrud, G. (2010). *Metode og dataanalyse : Beslutningsstøtte for bedrifter ved bruk av jmp : Oppgavesamling* (2. utg. utg.). Kristiansand: Høyskoleforl.
- Høydahl, E. (2010). Befolkningsvekst rundt oslo. *Statistisk Sentralbyrå*. Hentet 23.02.2015, fra <https://http://www.ssb.no/befolkning/artikler-og-publikasjoner/befolkningsvekst-rundt-oslo>
- Midtbø, T. (2012). *Stata : En entusiastisk innføring*. Oslo: Universitetsforl.
- Mieszkowski, P., & Saper, A. M. (1978). An estimate of the effects of airport noise on property values. *Journal of Urban Economics*, 5(4), 425-440.
- Nelson, J. P. (1978). Residential choice, hedonic prices, and the demand for urban air quality. *Journal of Urban Economics*, 5(3), 357-369. doi: 10.1016/0094-1190(78)90016-5
- Osland, L. (2001). Den hedonistiske metoden og estimering av attributtpriser *Norsk Økonomisk tidsskrift*(115), 1-22.

- Robertson, K., & Theisen, T. (2010). Boligmarkedet i kristiansand. Bergen: Fagbokforl., cop. 2010.
- Robertson, K., & Theisen, T. (2011). The impact of financial arrangements and institutional form on housing prices. *Journal of Real Estate Finance and Economics*, 42(3), 371-392. doi: 10.1007/s11146-009-9213-z
- Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, 82(1), 34-55.
- Statens vegvesen. (2006). Vedlegg 12 støymålinger ved boliger langs ny e18. Hentet 08.05.2015, fra http://www.vegvesen.no/attachment/60155/binary/10407?fast_title=Vedlegg+12+St%C3%B8ym%C3%A5linger+ved+boliger+langs+ny+E18
- Statens vegvesen. (2013). Støyvarselkart askim Hentet 08.05.2015, fra http://www.vegvesen.no/attachment/212571/binary/409126?fast_title=St%C3%B8yvarselkart+Askim+A0+%281%2C9+MB%29.pdf
- Statens vegvesen. (2014a). Bakgrunn. Hentet 24.02.2015, fra <http://www.vegvesen.no/Vegprosjekter/e18ostfold/Bakgrunn>
- Statens vegvesen. (2014b). Delstrekninger. Hentet 24.02.2015, fra <http://www.vegvesen.no/Vegprosjekter/e18ostfold/Delprosjekter>
- Statistisk sentralbyrå. (2007). Bolig, omgivelser og miljø. Hentet 26.02.2015, fra <https://http://www.ssb.no/a/samfunnsspeilet/utg/200705/11/tab-2007-12-12-01.html>
- Statistisk sentralbyrå. (2011a). Folke- og boligtellingsen, boliger. Hentet 16.04.2015, fra <https://http://www.ssb.no/statistikkbanken/SelectVarVal/saveselections.asp>
- Statistisk Sentralbyrå. (2011b). Sysselsatte 15-74 år, etter arbeidssted og innpendling (k). Hentet 23.02.2015, fra <https://http://www.ssb.no/statistikkbanken/SelectVarVal/Define.asp?MainTable=FOBsysInnpending&KortNavnWeb=fobsysut&PLanguage=0&checked=true>
- Statistisk sentralbyrå. (2011c). Sysselsatte 15-74 år, etter bosted og utpendling (k). Hentet 23.02.2015, fra <https://http://www.ssb.no/statistikkbanken/SelectVarVal/Define.asp?MainTable=FOBsysUtpending&KortNavnWeb=fobsysut&PLanguage=0&checked=true>
- Statistisk sentralbyrå. (2013a). Folke- og boligtellingsen, boliger, 19. November 2011. Hentet 16.04.2015, fra

<https://http://www.ssb.no/befolkning/statistikker/fobbolig/hvert-10-aar/2013-02-26?fane=tabell&sort=nummer&tabell=101959>

Statistisk sentralbyrå. (2013b). Sysselsetting, registerbasert. Hentet 20.04.2015, fra

<https://http://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=SysPendling&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=arbeid-og-lonn&KortNavnWeb=regsys&StatVariant=&checked=true>

Statistisk sentralbyrå. (2014a). Beregnet folkemengde 31. Desember og beregnet folketilvekst i året (k). Hentet 23.02.2015, fra

<https://http://www.ssb.no/statistikkbanken/SelectVarVal/Define.asp?MainTable=BeregnFolkem&KortNavnWeb=folkemengde&PLanguage=0&checked=true>

Statistisk sentralbyrå. (2014b). Boligprisindeksen. Hentet 20.04.2015, fra

<https://http://www.ssb.no/statistikkbanken/SelectVarVal/saveselections.asp>

Statistisk sentralbyrå. (2014c). Folkemengde. Hentet 20.04.2015, fra

<https://http://www.ssb.no/statistikkbanken/SelectVarVal/saveselections.asp>

Statistisk sentralbyrå. (2014d). Kvadratmeterpriser for eneboliger, 2013. Hentet 23.02.2015,

fra <https://http://www.ssb.no/priser-og-prisindekser/statistikker/kvadenebol>

Statistisk sentralbyrå. (2014e). Nye foretak untatt offentlig forvaltning og primærmæringene.

Hentet 23.02.2015, fra

<https://http://www.ssb.no/statistikkbanken/SelectVarVal/Define.asp?MainTable=ForetakNyeNye&KortNavnWeb=foretak&PLanguage=0&checked=true>

Statistisk sentralbyrå. (2015). Folkemengde 1. Januar og endringer i kalenderåret (k). Hentet 24.02.2015, fra

<https://http://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=Folkemengd1951&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=befolkning&KortNavnWeb=folkendrkv&StatVariant=&checked=true>

Stock, J. H., & Watson, M. W. (2012). *Introduction to econometrics* (3rd ed., global ed. utg.). Boston, Mass: Pearson.

Studenmund, A. H. (2011). *Using econometrics a practical guide* (6th ed. utg.). Boston: Pearson.

Thorsnæs, G. (2014). Askim. *Store Norske Leksikon* Hentet 23.02.2014, fra

<https://snl.no/Askim>

Wilhelmsson, M. (2000). The impact of traffic noise on the values of single-family houses.

Journal of Environmental Planning and Management, 43(6), 799-815.

Østfold fylkeskommune. (2012). Regional utvikling Hentet 23.02.2015, fra

<http://www.ostfold->

[f.kommune.no/modules/module_123/proxy.asp?D=2&C=360&I=16890&mnusel](http://www.ostfold-f.kommune.no/modules/module_123/proxy.asp?D=2&C=360&I=16890&mnusel)

[=1034a](http://www.ostfold-f.kommune.no/modules/module_123/proxy.asp?D=2&C=360&I=16890&mnusel)

Vedlegg 1: Postnummer med avstand til sentrum

Postnummer	Avstand til sentrum
1807	1.20 km
1808	2.00 km
1809	1.60 km
1811	1.60 km
1812	2.60 km
1813	0.95 km
1814	4.50 km
1815	2.40 km
1830	0.10 km
1831	0.75 km
1832	4.00 km

Vedlegg 2: Gjennomsnittlig salgspris etter postnummer, 2003-2014

Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
salgspris, 1807	495	1524413	634697.6	310000	4250000
Salgspris, 1808	193	1601373	777042.4	230000	4350000
Salgspris, 1809	497	1514073	684183.3	166000	4500000
Salgspris, 1811	390	1404628	512094.2	450000	3550000
Salgspris, 1812	262	1534696	696823.9	550000	4200000
Salgspris, 1813	259	1534853	602432.6	600000	3550000
Salgspris, 1814	218	1781845	673222.6	750000	4775000
Salgspris, 1815	131	1875115	1330033	700000	1.25e+07
Salgspris, 1830	94	1578000	564903.8	550000	3250000
Salgspris, 1831	286	1878741	704377.8	750000	6100000
Salgspris, 1832	721	1435300	626919.6	100000	4300000

Vedlegg 3: Gjennomsnittlig fellesgjeld etter postnummer, 2003-2014

Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
Fellesgjeld, 1807	495	31923.43	71615.62	0	305801
Fellesgjeld, 1808	193	25961.25	40362.86	0	123583
Fellesgjeld, 1809	497	55687.81	186268.4	0	1022674
Fellesgjeld 1811	390	121076.2	206068.8	0	1288294
Fellesgjeld,1812	262	49108.9	77750.43	0	294844
Fellesgjeld, 1813	259	26413.47	59013.19	0	250000
Fellesgjeld, 1814	218	341	2512.97	0	33600
Fellesgjeld, 1815	131	0	0	0	0
Fellesgjeld, 1830	94	34198.18	144882.1	0	902734
Fellesgjeld, 1831	286	90.81469	996.0169	0	16330
Fellesgjeld, 1832	721	47951.35	195275.7	0	1295247

Vedlegg 4: Gjennomsnittlig boareal etter postnummer, 2003-2014

Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
BOA, 1807	495	108.0753	51.22604	36	270
BOA, 1808	193	121.2021	58.57337	20	295
BOA, 1809	497	117.9543	54.76105	38	292
BOA, 1811	390	101.4128	48.16607	42	323
BOA, 1812	262	111.2061	56.78429	43	301
BOA, 1813	259	113.8456	55.99162	43	300
BOA, 1814	218	125.8335	60.40953	70	310
BOA, 1815	131	148.435	68.21418	64	456
BOA, 1830	94	75.20213	15.88844	28	165
BOA, 1831	286	131.9801	59.46111	53	351
BOA, 1832	721	103.8225	51.32876	32	333

Vedlegg 5: Gjennomsnittlig alder etter postnummer, 2003-2014

Variabel	Observasjoner	Gjennomsnitt	Standardavvik	Minimum	Maksimum
alder, 1807	495	36.2202	18.67336	0	99
alder, 1808	193	40.44041	17.19096	3	94
alder, 1809	497	41.07847	20.09864	0	108
alder, 1811	390	28.61282	22.00681	0	182
alder, 1812	262	37.17939	11.00689	2	83
alder, 1813	259	49.28185	21.84498	0	107
alder, 1814	218	30.3578	33.04724	0	194
alder, 1815	131	57.26718	38.24814	0	136
alder, 1830	94	29.46809	28.86703	0	80
alder, 1831	286	36.4965	25.18434	0	103
alder, 1832	721	24.33703	17.55781	0	208

Vedlegg 6: Fullstendig korrelasjonsmatrise

	salgspr>s	BOA	felles~d	avstpå~i	avstlu~i	alder	aar03	aar04	aar05	
salgspris	1.0000									
BOA	0.7235	1.0000								
fellesgjeld	-0.2571	-0.2003	1.0000							
avstpåkjm~i	0.0023	0.0680	-0.0888	1.0000						
avstluftlm~i	0.0186	0.0728	-0.0225	0.7572	1.0000					
alder	0.0270	0.2165	-0.1395	0.0512	0.1435	1.0000				
aar03	-0.1843	0.0119	-0.0700	0.0081	0.0164	-0.0426	1.0000			
aar04	-0.1351	0.0446	-0.0466	0.0016	-0.0126	-0.0031	-0.0905	1.0000		
aar05	-0.0648	0.0401	-0.0505	-0.0128	-0.0397	-0.0424	-0.0866	-0.0823	1.0000	
aar06	-0.0564	0.0234	-0.0240	-0.0060	-0.0055	0.0061	-0.0835	-0.0794	-0.0760	
aar07	-0.0024	0.0043	-0.0508	-0.0009	-0.0176	-0.0169	-0.0958	-0.0911	-0.0872	
aar08	-0.0037	-0.0106	-0.0324	0.0229	0.0141	-0.0335	-0.0884	-0.0841	-0.0804	
aar09	-0.0285	-0.0282	-0.0064	0.0015	0.0005	0.0141	-0.0922	-0.0877	-0.0839	
aar10	-0.0201	-0.0373	0.0635	-0.0039	0.0009	0.0017	-0.0920	-0.0876	-0.0838	
aar11	0.0938	0.0055	0.0427	-0.0102	-0.0126	0.0066	-0.1040	-0.0989	-0.0946	
aar12	0.0831	-0.0171	0.0238	0.0020	0.0134	0.0226	-0.0963	-0.0916	-0.0876	
aar13	0.1259	-0.0092	0.0596	0.0096	0.0297	0.0487	-0.0966	-0.0919	-0.0880	
aar14	0.1671	-0.0229	0.0796	-0.0117	0.0102	0.0334	-0.0938	-0.0892	-0.0853	
enebolig	0.4865	0.6423	-0.1923	0.1218	0.1394	0.2666	0.0423	0.0545	0.0320	
tomannsbolig	0.0167	0.0739	-0.1123	-0.0533	-0.0440	0.1074	-0.0223	0.0160	0.0122	
rekkehus	-0.0789	-0.0709	0.0414	0.0231	0.0435	-0.0626	0.0151	0.0027	-0.0121	
leilighet	-0.4564	-0.6541	0.2453	-0.0982	-0.1324	-0.3054	-0.0352	-0.0665	-0.0338	
selveier	0.3900	0.3675	-0.5916	-0.0544	-0.1734	0.0123	0.0400	0.0562	0.0653	
borettslag	-0.3900	-0.3675	0.5916	0.0544	0.1734	-0.0123	-0.0400	-0.0562	-0.0653	
postnr1807	-0.0183	-0.0348	-0.0319	0.1591	0.3181	0.0182	0.0000	-0.0248	-0.0192	
postnr1808	0.0155	0.0360	-0.0290	0.2377	0.3663	0.0543	0.0187	-0.0290	-0.0100	
postnr1809	-0.0243	0.0370	0.0352	0.0509	0.0990	0.1024	0.0024	0.0500	-0.0290	
postnr1811	-0.0763	-0.0726	0.1916	-0.1747	-0.1555	-0.0987	0.0004	-0.0297	0.0087	
postnr1812	-0.0087	-0.0085	0.0116	0.1296	0.3668	0.0244	0.0086	-0.0190	-0.0213	
postnr1813	-0.0086	0.0049	-0.0331	-0.2161	-0.2689	0.1699	-0.0096	0.0099	0.0087	
postnr1814	0.0828	0.0598	-0.0769	0.1995	0.0348	-0.0527	-0.0289	0.0075	0.0184	
postnr1815	0.0895	0.1256	-0.0593	-0.1528	-0.1125	0.1856	0.0086	-0.0076	-0.0090	
postnr1830	0.0052	-0.1121	-0.0104	-0.1538	-0.0395	-0.0403	0.0177	0.0296	-0.0058	
postnr1831	0.1369	0.1021	-0.0895	-0.1544	-0.2045	0.0169	-0.0215	0.0013	0.0562	
postnr1832	-0.0875	-0.0824	0.0167	-0.0052	-0.3233	-0.2345	0.0084	0.0100	0.0063	
sentrum	-0.0338	-0.0083	-0.0100	0.1987	-0.0921	-0.1954	-0.0035	0.0004	0.0005	
innenfor500	0.0416	-0.0151	0.0485	-0.2950	-0.4682	-0.0995	-0.0229	-0.0211	0.0289	
innenfor1000	0.0343	0.1083	-0.0933	-0.2757	-0.4088	0.1217	0.0071	0.0348	0.0216	
over1000	-0.0574	-0.0850	0.0498	0.4358	0.6665	-0.0411	0.0088	-0.0166	-0.0379	
		aar06	aar07	aar08	aar09	aar10	aar11	aar12	aar13	aar14
aar06	1.0000									
aar07	-0.0841	1.0000								
aar08	-0.0775	-0.0890	1.0000							
aar09	-0.0809	-0.0929	-0.0857	1.0000						
aar10	-0.0808	-0.0927	-0.0855	-0.0892	1.0000					
aar11	-0.0913	-0.1047	-0.0966	-0.1008	-0.1006	1.0000				
aar12	-0.0845	-0.0970	-0.0895	-0.0934	-0.0932	-0.1053	1.0000			
aar13	-0.0848	-0.0973	-0.0898	-0.0937	-0.0935	-0.1056	-0.0978	1.0000		
aar14	-0.0823	-0.0944	-0.0871	-0.0909	-0.0907	-0.1025	-0.0949	-0.0953	1.0000	
enebolig	0.0213	-0.0120	-0.0099	-0.0239	-0.0378	0.0017	-0.0245	-0.0260	-0.0125	
tomannsbolig	0.0299	0.0374	-0.0018	0.0088	-0.0160	-0.0023	-0.0216	-0.0223	-0.0133	
rekkehus	-0.0250	-0.0136	0.0189	0.0121	0.0042	-0.0217	0.0404	-0.0149	-0.0066	
leilighet	-0.0283	-0.0058	0.0013	0.0118	0.0462	0.0110	0.0180	0.0485	0.0248	
selveier	0.0522	0.0567	0.0031	-0.0190	-0.0631	-0.0049	-0.0360	-0.0867	-0.0520	
borettslag	-0.0522	-0.0567	-0.0031	0.0190	0.0631	0.0049	0.0360	0.0867	0.0520	
postnr1807	-0.0187	0.0128	0.0106	-0.0048	0.0164	-0.0041	0.0115	0.0050	0.0121	
postnr1808	0.0041	-0.0087	-0.0171	0.0053	-0.0036	-0.0111	0.0081	0.0120	0.0298	
postnr1809	-0.0030	-0.0250	0.0132	-0.0023	0.0011	0.0061	0.0110	0.0159	-0.0410	
postnr1811	-0.0165	0.0213	-0.0122	0.0164	-0.0062	0.0184	-0.0115	-0.0090	0.0162	
postnr1812	-0.0209	-0.0306	0.0127	0.0098	0.0023	-0.0062	0.0141	0.0324	0.0149	
postnr1813	-0.0373	0.0391	-0.0149	-0.0168	-0.0047	0.0020	-0.0038	0.0108	0.0120	
postnr1814	0.0425	-0.0049	0.0109	-0.0209	0.0093	-0.0165	0.0026	0.0062	-0.0187	
postnr1815	0.0300	0.0078	0.0060	-0.0204	0.0016	-0.0018	-0.0139	-0.0090	0.0103	
postnr1830	0.0180	0.0231	-0.0075	-0.0110	-0.0044	-0.0266	0.0102	-0.0209	-0.0186	
postnr1831	0.0472	-0.0262	-0.0184	0.0171	-0.0166	-0.0041	-0.0160	-0.0129	0.0030	
postnr1832	-0.0117	0.0014	0.0061	0.0098	0.0001	0.0171	-0.0100	-0.0281	-0.0101	
sentrum	0.0089	-0.0138	0.0182	-0.0023	0.0082	0.0076	-0.0040	-0.0071	-0.0114	
innenfor500	-0.0041	0.0020	-0.0172	-0.0175	0.0199	-0.0048	0.0140	-0.0029	0.0261	
innenfor1000	-0.0010	-0.0023	-0.0052	0.0188	-0.0161	-0.0077	-0.0185	-0.0267	-0.0011	
over1000	0.0036	0.0007	0.0159	-0.0050	0.0010	0.0099	0.0070	0.0253	-0.0163	

	enebolig tomann~g rekkehus leilig~t selveier borett~g pos~1807 pos~1808 pos~1809									
enebolig	1.0000									
tomannsbolig	-0.3052	1.0000								
rekkehus	-0.2236	-0.1008	1.0000							
leilighet	-0.6810	-0.3071	-0.2250	1.0000						
selveier	0.3571	0.1798	-0.0858	-0.4319	1.0000					
borettslag	-0.3571	-0.1798	0.0858	0.4319	-1.0000	1.0000				
postnr1807	0.0319	0.0652	-0.0806	-0.0336	-0.0108	0.0108	1.0000			
postnr1808	0.1168	-0.0471	-0.0210	-0.0746	-0.0655	0.0655	-0.0966	1.0000		
postnr1809	0.0256	0.0271	0.0924	-0.0911	-0.0464	0.0464	-0.1626	-0.0969	1.0000	
postnr1811	-0.1238	0.0188	0.0860	0.0668	-0.1472	0.1472	-0.1416	-0.0843	-0.1419	
postnr1812	-0.0060	-0.0916	0.2001	-0.0363	-0.1888	0.1888	-0.1138	-0.0678	-0.1140	
postnr1813	-0.0453	0.1185	-0.0763	0.0059	-0.0131	0.0131	-0.1131	-0.0673	-0.1133	
postnr1814	0.0599	-0.0589	-0.0649	0.0128	0.1285	-0.1285	-0.1031	-0.0614	-0.1033	
postnr1815	0.1284	0.0924	-0.0532	-0.1621	0.0983	-0.0983	-0.0789	-0.0470	-0.0791	
postnr1830	-0.1357	-0.0612	-0.0449	0.1993	0.0697	-0.0697	-0.0665	-0.0396	-0.0666	
postnr1831	0.0371	0.0680	-0.0642	-0.0491	0.1487	-0.1487	-0.1193	-0.0711	-0.1196	
postnr1832	-0.0528	-0.1144	-0.0211	0.1395	0.0807	-0.0807	-0.2035	-0.1212	-0.2040	
sentrum	0.0364	-0.1535	0.0208	0.0548	0.0622	-0.0622	-0.3092	-0.0338	-0.1834	
innenfor500	-0.0066	-0.0383	0.0244	0.0195	0.0740	-0.0740	-0.1371	-0.0816	-0.1374	
innenfor1000	0.0725	0.1336	-0.1229	-0.0977	0.1448	-0.1448	-0.2065	-0.1230	-0.2070	
over1000	-0.0591	-0.0918	0.0916	0.0728	-0.1756	0.1756	0.2712	0.1615	0.2718	
	pos~1811	pos~1812	pos~1813	pos~1814	pos~1815	pos~1830	pos~1831	pos~1832	sentrum	
postnr1811	1.0000									
postnr1812	-0.0993	1.0000								
postnr1813	-0.0987	-0.0793	1.0000							
postnr1814	-0.0900	-0.0723	-0.0718	1.0000						
postnr1815	-0.0689	-0.0553	-0.0550	-0.0501	1.0000					
postnr1830	-0.0580	-0.0466	-0.0463	-0.0422	-0.0323	1.0000				
postnr1831	-0.1041	-0.0837	-0.0831	-0.0758	-0.0580	-0.0489	1.0000			
postnr1832	-0.1776	-0.1427	-0.1418	-0.1293	-0.0989	-0.0834	-0.1496	1.0000		
sentrum	-0.1597	0.0930	-0.2705	0.4652	0.0338	-0.2689	-0.3318	0.7204	1.0000	
innenfor500	-0.1167	-0.0961	0.1319	0.4018	0.1638	-0.0562	0.0147	0.0464	0.2536	
innenfor1000	0.0440	-0.1448	0.3500	-0.0763	0.0837	-0.0803	0.1594	0.1898	0.0358	
over1000	0.0382	0.1902	-0.3935	-0.1977	-0.1812	0.1073	-0.1493	-0.1968	-0.1983	
	inne~500	inn~1000	over1000							
innenfor500	1.0000									
innenfor1000	-0.1745	1.0000								
over1000	-0.5054	-0.7615	1.0000							

Vedlegg 7: Fullstendig multippel lineær regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 3546		
Model	1.2871e+15	29	4.4383e+13	F(29, 3516) =	354.25	
Residual	4.4051e+14	3516	1.2529e+11	Prob > F =	0.0000	
Total	1.7276e+15	3545	4.8734e+11	R-squared =	0.7450	
				Adj R-squared =	0.7429	
				Root MSE =	3.5e+05	

salgspris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
B0A	8530.431	155.2939	54.93	0.000	8225.955	8834.906
fellesgjeld	-.6975184	.0545292	-12.79	0.000	-.8044305	-.5906064
aar04	57216.22	29296.55	1.95	0.051	-223.7342	114656.2
aar05	190262.8	29930.93	6.36	0.000	131579.1	248946.5
aar06	267343.4	30577	8.74	0.000	207393	327293.9
aar07	442059.4	28539.97	15.49	0.000	386102.9	498016
aar08	475198.1	29594.33	16.06	0.000	417174.3	533221.9
aar09	482631.5	29056.74	16.61	0.000	425661.7	539601.3
aar10	542417.9	29196.94	18.58	0.000	485173.2	599662.5
aar11	693165.2	27611.81	25.10	0.000	639028.5	747302
aar12	728794	28544.8	25.53	0.000	672827.9	784760
aar13	841365.9	28633.27	29.38	0.000	785226.4	897505.3
aar14	958659.1	29007.65	33.05	0.000	901785.5	1015533
alder	-6041.464	296.3425	-20.39	0.000	-6622.485	-5460.443
enebolig	114148.7	21457.09	5.32	0.000	72079.1	156218.3
tomannsbolig	-30316.3	23758.09	-1.28	0.202	-76897.33	16264.72
rekkehus	-23557.48	26283.11	-0.90	0.370	-75089.18	27974.21
borettslag	-140097.4	22418.32	-6.25	0.000	-184051.7	-96143.18
postnr1807	-234623.4	26064.9	-9.00	0.000	-285727.3	-183519.5
postnr1808	-266950	33557.62	-7.95	0.000	-332744.4	-201155.6
postnr1809	-240753.8	26186.45	-9.19	0.000	-292096	-189411.6
postnr1811	-204583.4	26829.74	-7.63	0.000	-257186.8	-151979.9
postnr1812	-206366.5	31485.57	-6.55	0.000	-268098.4	-144634.7
postnr1813	-111308.2	30702.03	-3.63	0.000	-171503.8	-51112.61
postnr1814	-188145	32304.53	-5.82	0.000	-251482.5	-124807.5
postnr1815	-118798.4	37620	-3.16	0.002	-192557.6	-45039.12
postnr1832	-287433.7	22952.62	-12.52	0.000	-332435.5	-242431.9
innenfor500	-824.4655	24778.8	-0.03	0.973	-49406.74	47757.81
innenfor1000	-87960.56	18969.87	-4.64	0.000	-125153.6	-50767.5
_cons	555650.9	31652.64	17.55	0.000	493591.5	617710.3

Vedlegg 8: Fullstendig dobbelt-logaritmisk regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 3546		
Model	554.949961	29	19.1362056	F(29, 3516) = 487.62		
Residual	137.982624	3516	.039244205	Prob > F = 0.0000		
Total	692.932585	3545	.195467584	R-squared = 0.8009		
				Adj R-squared = 0.7992		
				Root MSE = .1981		
lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnBOA	.6450419	.0111651	57.77	0.000	.6231512	.6669326
lnfellesgjeld	-.0115971	.0015358	-7.55	0.000	-.0146083	-.0085859
lnalder	-.0915989	.0039785	-23.02	0.000	-.0993994	-.0837985
aar04	.0531054	.0163955	3.24	0.001	.0209597	.0852511
aar05	.138778	.0167532	8.28	0.000	.105931	.171625
aar06	.1958356	.0170948	11.46	0.000	.1623189	.2293523
aar07	.311386	.0159652	19.50	0.000	.2800841	.3426879
aar08	.342081	.0165675	20.65	0.000	.309598	.3745639
aar09	.331201	.0162731	20.35	0.000	.2992954	.3631067
aar10	.3563165	.0163791	21.75	0.000	.324203	.38843
aar11	.4487068	.0154954	28.96	0.000	.4183259	.4790876
aar12	.4902235	.0160461	30.55	0.000	.458763	.5216841
aar13	.5596144	.0160592	34.85	0.000	.5281282	.5911007
aar14	.6304669	.0162716	38.75	0.000	.5985642	.6623696
enebolig	.031275	.0131833	2.37	0.018	.0054274	.0571226
tomannsbolig	-.0489145	.0139398	-3.51	0.000	-.0762454	-.0215837
rekkehus	-.0119836	.0152074	-0.79	0.431	-.0417998	.0178326
borettslag	-.1165394	.0182727	-6.38	0.000	-.1523656	-.0807133
postnr1807	-.1225883	.014668	-8.36	0.000	-.1513469	-.0938297
postnr1808	-.1547003	.0187996	-8.23	0.000	-.1915595	-.1178412
postnr1809	-.1731127	.0148816	-11.63	0.000	-.20229	-.1439353
postnr1811	-.1198412	.0149853	-8.00	0.000	-.149222	-.0904604
postnr1812	-.0694771	.0175882	-3.95	0.000	-.1039612	-.034993
postnr1813	-.0242706	.0171542	-1.41	0.157	-.0579038	.0093626
postnr1814	-.0783002	.0180453	-4.34	0.000	-.1136806	-.0429199
postnr1815	-.1380677	.0209301	-6.60	0.000	-.1791041	-.0970313
postnr1832	-.1691413	.0127881	-13.23	0.000	-.1942142	-.1440684
innenfor500	-.0728611	.0138125	-5.28	0.000	-.0999423	-.0457798
innenfor1000	-.0616135	.0106625	-5.78	0.000	-.0825188	-.0407082
_cons	11.33973	.0517579	219.09	0.000	11.23825	11.4412

Vedlegg 9: Fullstendig semi-logaritmisk regresjon med avstandsdummyer

Source	SS	df	MS	Number of obs = 3546		
Model	539.557117	29	18.6054178	F(29, 3516) = 426.51		
Residual	153.375468	3516	.043622147	Prob > F = 0.0000		
Total	692.932585	3545	.195467584	R-squared = 0.7787		
				Adj R-squared = 0.7768		
				Root MSE = .20886		

lnpris	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
BOA	.0044915	.0000916	49.02	0.000	.0043118	.0046711
fellesgjeld	-7.08e-07	3.22e-08	-22.02	0.000	-7.72e-07	-6.45e-07
aar04	.0687933	.0172869	3.98	0.000	.0349	.1026867
aar05	.1552433	.0176612	8.79	0.000	.120616	.1898705
aar06	.2329687	.0180424	12.91	0.000	.197594	.2683434
aar07	.3323368	.0168405	19.73	0.000	.2993187	.3653548
aar08	.3614509	.0174626	20.70	0.000	.327213	.3956887
aar09	.3495073	.0171454	20.38	0.000	.3158914	.3831232
aar10	.3778689	.0172281	21.93	0.000	.3440908	.411647
aar11	.4825104	.0162928	29.61	0.000	.4505661	.5144546
aar12	.5018163	.0168433	29.79	0.000	.4687926	.5348399
aar13	.5891276	.0168955	34.87	0.000	.5560016	.6222536
aar14	.6531544	.0171164	38.16	0.000	.6195953	.6867135
alder	-.0041647	.0001749	-23.82	0.000	-.0045075	-.0038218
enebolig	.1874957	.0126611	14.81	0.000	.1626719	.2123196
tomannsbolig	.0838402	.0140188	5.98	0.000	.0563543	.111326
rekkehus	.0797876	.0155088	5.14	0.000	.0493805	.1101947
borettslag	-.097507	.0132283	-7.37	0.000	-.1234428	-.0715711
postnr1807	-.1944549	.01538	-12.64	0.000	-.2246095	-.1643003
postnr1808	-.2616356	.0198012	-13.21	0.000	-.3004586	-.2228126
postnr1809	-.1977102	.0154517	-12.80	0.000	-.2280054	-.1674149
postnr1811	-.1116027	.0158313	-7.05	0.000	-.1426421	-.0805632
postnr1812	-.1614525	.0185785	-8.69	0.000	-.1978783	-.1250267
postnr1813	-.0554169	.0181162	-3.06	0.002	-.0909362	-.0198976
postnr1814	-.1068529	.0190618	-5.61	0.000	-.1442262	-.0694797
postnr1815	-.1483812	.0221983	-6.68	0.000	-.1919039	-.1048584
postnr1832	-.2105598	.0135435	-15.55	0.000	-.2371138	-.1840058
innenfor500	-.0361341	.0146211	-2.47	0.014	-.0648008	-.0074674
innenfor1000	-.0768569	.0111935	-6.87	0.000	-.0988032	-.0549105
_cons	13.58397	.0186771	727.31	0.000	13.54735	13.62059

Vedlegg 10: kommandoer i STATA

Importerer Excel fil:

```
insheet using "/Users/stinekarerud/Documents/Masteroppgave/STATA/2003-2015.txt
```

```
rename avstandpkjrselmotorvei avstpåkjemotorvei
```

```
rename avstandluftlinjemotorvei avstluftlmotorvei
```

Datarensing:

```
drop regdato
```

```
drop omshast
```

```
drop prisant
```

```
drop m2prom
```

```
drop tomt
```

```
drop megler
```

```
replace fellesgjeld=0 if fellesgjeld==.
```

```
drop if missing(salgspris)
```

```
drop if salgspris==0
```

```
drop if missing(bygger)
```

```
drop if bygger==0
```

```
drop if missing(boa) & missing(bta)
```

```
drop if boa==0 & bta==0
```

```
drop if missing(avstpåkjemotorvei)
```

```
drop if missing(avstluftlmotorvei)
```

Estimerer BOA:

```
reg boa bta
```

```
gen BOA= boa
```

```
replace BOA=(38.77722 +(.5500932*bta)) if boa==0
```

```
gen id1=_n
```

Oppretter år:

gen aar03=0

gen aar04=0

gen aar05=0

gen aar06=0

gen aar07=0

gen aar08=0

gen aar09=0

gen aar10=0

gen aar11=0

gen aar12=0

gen aar13=0

gen aar14=0

replace aar03=1 if id1<=308

replace aar04=1 if id1>308 & id1<=589

replace aar05=1 if id1>589 & id1<=848

replace aar06=1 if id1>848 & id1<=1090

replace aar07=1 if id1>1090 & id1<=1402

replace aar08=1 if id1>1402 & id1<=1671

replace aar09=1 if id1>1671 & id1<=1963

replace aar10=1 if id1>1963 & id1<= 2253

replace aar11=1 if id1>2253 & id1<=2615

replace aar12=1 if id1>2615 & id1<=2930

replace aar13=1 if id1>2930 & id1<=3247

replace aar14=1 if id1>3247 & id1<=3547

Oppretter alder:

gen alder=0

replace alder=(2003-bygger) if aar03==1

replace alder=(2004-bygger) if aar04==1

replace alder=(2005-bygger) if aar05==1

replace alder=(2006-bygger) if aar06==1

```
replace alder=(2007- bygger) if aar07==1
replace alder=(2008- bygger) if aar08==1
replace alder=(2009- bygger) if aar09==1
replace alder=(2010- bygger) if aar10==1
replace alder=(2011- bygger) if aar11==1
replace alder=(2012- bygger) if aar12==1
replace alder=(2013- bygger) if aar13==1
replace alder=(2014-bygger) if aar14==1
```

Oppretter boligtype:

```
gen 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:

```
gen enebolig=0
gen tomannsbolig=0
gen rekkehus=0
gen 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"
```

Oppretter Eierform:

```
gen Eierform=0
```

```
replace Eierform= 1 if eierform=="Selveier"
replace Eierform=2 if eierform=="Borettslag"
```

Oppretter dummyer:

```
gen selveier=0
gen borettslag=0
```

```
replace selveier=1 if eierform=="Selveier"
replace borettslag=1 if eierform=="Borettslag"
```

Oppretter postnummerdummyer:

```
gen postnr1807=0
gen postnr1808=0
gen postnr1809=0
gen postnr1811=0
gen postnr1812=0
gen postnr1813=0
gen postnr1814=0
gen postnr1815=0
gen postnr1830=0
gen postnr1831=0
gen postnr1832=0
```

```
replace postnr1807=1 if postnummer==1807
replace postnr1808=1 if postnummer==1808
replace postnr1809=1 if postnummer==1809
replace postnr1811=1 if postnummer==1811
replace postnr1812=1 if postnummer==1812
replace postnr1813=1 if postnummer==1813
replace postnr1814=1 if postnummer==1814
replace postnr1815=1 if postnummer==1815
replace postnr1830=1 if postnummer==1830
replace postnr1831=1 if postnummer==1831
replace postnr1832=1 if postnummer==1832
```

Oppretter avstand til sentrum:

```
gen sentrum=0
```

```
replace sentrum=1.2 if postnummer==1807
replace sentrum=2 if postnummer==1808
replace sentrum=1.6 if postnummer==1809
replace sentrum=1.6 if postnummer==1811
replace sentrum=2.6 if postnummer==1812
replace sentrum=0.95 if postnummer==1813
replace sentrum=4.5 if postnummer==1814
replace sentrum=2.4 if postnummer==1815
replace sentrum=0.1 if postnummer==1830
replace sentrum=0.75 if postnummer==1831
replace sentrum=4 if postnummer==1832
```

Oppretter dummy variabler for «grenser»:

```
gen innenfor500=0
gen innenfor1000=0
```


gen over1000=0

replace innenfor500=1 if avstluftmotorvei>0 & avstluftmotorvei<=500
replace innenfor1000=1 if avstluftmotorvei>500 & avstluftmotorvei<=1000
replace over1000=1 if avstluftmotorvei>1000

Oppretter koder på grensene:

gen grenser=0

replace grenser=1 if innenfor100==1
replace grenser=2 if innenfor200==1
replace grenser=3 if innenfor500==1
replace grenser=4 if innenfor1000==1
replace grenser=5 if over1000==1

***Oversikter, Tabeller og grafer**

sum enebolig tomannsbolig rekkehus leilighet selveier borettslag postnr1807 postnr1808
postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1830 postnr1831
postnr1832 innenfor500 innenfor1000 over1000 aar03 aar04 aar05 aar06 aar07 aar08 aar09
aar10 aar11 aar12 aar13 aar14

sum salgpris fellesgjeld BOA alder sentrum avstpåkjemotorvei

Salgspris:

histogram salgpris, normal freq

sum salgpris if aar03==1
sum salgpris if aar04==1
sum salgpris if aar05==1
sum salgpris if aar06==1
sum salgpris if aar07==1
sum salgpris if aar08==1
sum salgpris if aar09==1
sum salgpris if aar10==1
sum salgpris if aar11==1
sum salgpris if aar12==1
sum salgpris if aar13==1
sum salgpris if aar14==1
sum salgpris

sum salgpris if postnummer==1807
sum salgpris if postnummer==1808
...
sum salgpris if postnummer==1832

Fellesgjeld:

histogram fellesgjeld, normal freq

sum fellesgjeld if aar03==1

...

sum fellesgjeld if aar14==1

sum fellesgjeld if postnummer==1807

...

sum fellesgjeld if postnummer==1832

Boareal:

histogram BOA, norm freq

sum BOA if aar03==1

...

sum BOA if aar14==1

sum BOA if postnummer==1807

....

sum BOA if postnummer==1832

Alder:

histogram alder, norm freq

sum alder if aar03==1

...

sum alder if aar14==1

sum alder if postnummer==1807

...

sum alder if postnummer==1832

Avstand til sentrum:

histogram sentrum, norm freq

Avstand til motorvei:

histogram avstpåkjemotorvei, norm freq

sum innenfor100 if innenfor100==1

sum innenfor200 if innenfor200 ==1

sum innenfor500 if innenfor500 ==1

sum innenfor1000 if innenfor1000 ==1

sum over1000 if over1000 ==1

Boligtype:

```
sum Boligtype if enebolig==1
sum Boligtype if tomannsbolig ==1
sum Boligtype if rekkehus ==1
sum Boligtype if leilighet ==1
```

Korrelasjonsmatriser:

```
pwcorr salgspris fellesgjeld avstpåkjemotorvei BOA alder enebolig tomannsbolig rekkehus
leilighet selveier borettslag sentrum innenfor500 innenfor1000 over1000, star(.05)
```

```
corr salgspris BOA fellesgjeld avstpåkjemotorvei avstluftmotorvei alder aar03 aar04 aar05
aar06 aar07 aar08 aar09 aar10 aar11 aar12 aar13 aar14 enebolig tomannsbolig rekkehus
leilighet selveier borettslag postnr1807 postnr1808 postnr1809 postnr1811 postnr1812
postnr1813 postnr1814 postnr1815 postnr1830 postnr1831 postnr1832 sentrum innenfor500
innenfor1000 over1000
```

Multipel regresjon

```
reg salgspris BOA fellesgjeld avstpåkjemotorvei sentrum aar04 aar05 aar06 aar07 aar08 aar09
aar10 aar11 aar12 aar13 aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832
```

```
reg salgspris BOA fellesgjeld avstpåkjemotorvei aar04 aar05 aar06 aar07 aar08 aar09 aar10
aar11 aar12 aar13 aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832
```

Test for multikollinearitet:

```
vif
```

Test for heteroskedastisitet:

```
rvfplot, yline(0)
rvpplot BOA, yline (0)
rvpplot fellesgjeld, yline (0)
rvpplot alder, yline (0)
rvpplot avstpåkjemotorvei, yline (0)
```

```
estat hettest
estat hettest BOA
estat hettest fellesgjeld
estat hettest alder
estat hettest avstpåkjemotorvei
```

Robuste standardfeil i multipel regresjon:

```
reg salgspris BOA fellesgjeld avstpåkjemotorvei aar04 aar05 aar06 aar07 aar08 aar09 aar10
aar11 aar12 aar13 aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815
postnr1832, robust
```

kontrollerer for normalfordeling i residualene:

predict errormultippelreg, resid

pnorm errormultippelreg

Multippel lineær regresjon med avstandsdummyer:

reg salgspri BOA fellesgjeld aar04 aar05 aar06 aar07 aar08 aar09 aar10 aar11 aar12 aar13
aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807 postnr1808 postnr1809
postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832 innenfor500
innenfor1000

Dobbelt-logaritmisk regresjon:

gen lnpris=ln(salgspri)
gen lnalder=ln(alder+1)
gen lnfellesgjeld=ln(fellesgjeld+1)
gen lnBOA=ln(BOA)
gen lnavstpåkjemotorvei=ln(avstpåkjemotorvei)

reg lnpris lnBOA lnfellesgjeld lnalder lnavstpåkjemotorvei aar04 aar05 aar06 aar07 aar08
aar09 aar10 aar11 aar12 aar13 aar14 enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815
postnr1832

Test for multikollinearitet:

vif

Test for heteroskedastisitet:

rvfplot, yline(0)
rvpplot lnBOA, yline (0)
rvpplot lnfellesgjeld, yline (0)
rvpplot lnalder, yline (0)
rvpplot lnavstpåkjemotorvei, yline (0)

estat hettest

estat hettest lnBOA

estat hettest lnfellesgjeld

estat hettest lnalder

estat hettest lnavstpåkjemotorvei

Robuste standardfeil i dobbelt-logaritmisk regresjon:

reg lnpris lnBOA lnfellesgjeld lnalder lnavstpåkjemotorvei aar04 aar05 aar06 aar07 aar08
aar09 aar10 aar11 aar12 aar13 aar14 enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815
postnr1832, robust

kontrollerer for normalfordeling i residualene:

predict errordobbel, resi

pnorm errordobbel

Dobbelt-logaritmisk regresjon med avstandsdummyer:

reg lnpris lnBOA lnfellesgjeld lnalder aar04 aar05 aar06 aar07 aar08 aar09 aar10 aar11 aar12
aar13 aar14 enebolig tomannsbolig rekkehus borettslag postnr1807 postnr1808 postnr1809
postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832 innenfor500
innenfor1000

Semi-logaritmisk regresjon:

reg lnpris BOA fellesgjeld avstpåkjemotorvei aar04 aar05 aar06 aar07 aar08 aar09 aar10
aar11 aar12 aar13 aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832

Test for multikollinearitet:

vif

Test for heteroskedastisitet:

rvfplot, yline(0)
rvpplot BOA, yline(0)
rvpplot fellesgjeld, yline(0)
rvpplot alder, yline(0)
rvpplot avstpåkjemotorvei, yline(0)

estat hettest
estat hettest BOA
estat hettest fellesgjeld
estat hettest alder
estat hettest avstpåkjemotorvei

Robuste standardfeil i semi-logaritmisk regresjon:

reg lnpris BOA fellesgjeld avstpåkjemotorvei aar04 aar05 aar06 aar07 aar08 aar09 aar10
aar11 aar12 aar13 aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807
postnr1808 postnr1809 postnr1811 postnr1812 postnr1813 postnr1814 postnr1815
postnr1832, robust

kontrollerer for normalfordeling i residualene:

predict errorsemi, resi

pnorm errorsemi

Semi-logaritmisk regresjon med avstandsdummyer:

reg lnpris BOA fellesgjeld aar04 aar05 aar06 aar07 aar08 aar09 aar10 aar11 aar12 aar13
aar14 alder enebolig tomannsbolig rekkehus borettslag postnr1807 postnr1808 postnr1809
postnr1811 postnr1812 postnr1813 postnr1814 postnr1815 postnr1832 innenfor500
innenfor1000