



UNIVERSITETET I AGDER

**Empirisk analyse av boligmarkedet i Haugesund og
Stavanger**

*Hvilken betydning har utvalgte boligattributter på boligprisen
i Haugesund og Stavanger?
Perioden 2009-2013*

Ole Christian Langeland

Veileder

Anne Wenche Emblem

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

Universitetet i Agder, 2014
Handelshøyskolen ved UiA

Forord

Masteroppgaven BE-501 utgjør 30 studiepoeng og er skrevet som en avsluttende del av et femårig løp ved masterprogrammet i økonomi og administrasjon ved Universitetet i Agder. Målet med oppgaven er at studentene skal vise selvstendighet og anvende teori og kunnskap de har tilegnet seg i løpet av de 5 årene ved universitetet.

Denne masteroppgaven har opphav fra min interesse og nysgjerrighet for boligmarkedet. Høsten 2013 hadde jeg et kurs i eiendomsøkonomi ved UiA, som var med å styrke interessen. Jeg bestemte meg på et forholdsvis tidlig tidspunkt om at jeg ville ta utgangspunkt i boligmarkedet i Haugesund og Stavanger. Kombinasjonen av boligmarkedenes omtale i media og en forholdsvis nær tilknytting til byene bidro til min interesse for å undersøke hva som har betydning for boligprisene. Det var i så måte naturlig å se på hvilke attributter det er betalingsvillighet for og i hvilken grad de påvirker boligprisene i byene.

Jeg ønsker å rette en takk til veileder Førsteamanuensis Anne Wenche Emblem for god rådgivning og veiledning underveis i prosessen. Videre vil jeg takke Irene Bredal for hjelp til korrekturlesning.

Kristiansand 3.juni 2014.

Ole Christian Langeland

Sammendrag

Målet med oppgaven er å undersøke ulike boligattributters betydning for boligprisen i Haugesund og Stavanger, og om attributtene varierer i betydningsgrad byene seg i mellom. Dette er blitt sett på gjennom en femårsperiode, fra 2009 til 2013.

De siste årene har det vært mye fokus på de svært store prisforskjellene mellom Haugesund og Stavanger som har en forholdsvis geografisk nærhet. Det oppfattes som om man får mer bolig for pengene i Haugesund enn i Stavanger. Dette utgjorde i hovedsak min motivasjon for oppgaven.

Jeg tror og håper at masteroppgaven vil være interessant lesning, og at resultatene som er fremlagt kan være av interesse for eiendomsめglere, boligeiere, potensielle bolig kjøpere og andre interessenter.

Mine 14 uavhengige variabler har en forklaringskraft på hele 72,36% og 76,11% på boligprisen i henholdsvis Haugesund og Stavanger. Jeg kunne konkludere på et 5% signifikansnivå at boligareal har større betydning for boligprisen i Stavanger enn i Haugesund. Videre er funnet av at avstand til sentrum har større betydning for boligprisen i Haugesund enn i Stavanger svært interessant. Boligalder har også større betydning for boligprisen i Haugesund enn i Stavanger. Videre kunne jeg med sikkerhet si at boligtype har betydning for boligprisen i begge byer.

Innledningsvis i oppgaven blir vi bedre kjent med Rogalandsbyene. Tabeller og statistikker er presentert for å gi et innblikk i boligmarkedene og prisutviklingen. Oppgaven tar utgangspunkt i teori om prisdannelse i boligmarkedet, teori om tomtemarkedet og den hedonistiske metoden. Basert på teorien ble attributtene som antas å ha høyest forklaringskraft på boligprisen valgt. Attributtene som undersøkes er boligareal, boligalder, tomteareal, avstand til sentrum, avstand til sjø, boligtype og om man må i bil for å komme til barnehage. Videre er datamaterialet presentert før ulike regresjonsmodeller er gjennomført. Avslutningsvis drøfter jeg funnene og konkluderer.

Innholdsfortegnelse

FORORD	2
SAMMENDRAG	3
FIGUROVERSIKT	6
TABELLOVERSIKT	7
1. INNLEDNING	8
2. BAKGRUNN	9
2.1 INTRODUKSJON AV HAUGESUND OG STAVANGER	10
2.2 HAUGESUND.....	10
2.3 STAVANGER.....	11
2.4 BOLIGMARKEDET I HAUGESUND OG STAVANGER	12
2.4.1 HAUGESUND	12
2.4.2 STAVANGER.....	14
3. TEORI	18
3.1 KJENNETEGN VED BOLIGMARKEDET	18
3.2 TEORI OM PRISDANNELSEN I BOLIGMARKEDET	20
3.2.1 ETTERSPØRSEL ETTER BOLIGER.....	20
3.2.2 TILBUD AV BOLIGER.....	21
3.2.3 MARKEDSPRIS.....	21
3.3 TEORI OM TOMTEMARKEDET	22
3.3.1 ALONSO-MUTH-MILLS MODELLEN	23
3.4 DEN HEDONISTISKE METODEN	27
3.5 HYPOTESER	37
4.0 DATAINNSAMLING OG BESKRIVELSE AV DATAMATERIALET	42
4.1 DATAINNSAMLING	42
4.2 UTVALGSPROESS	43
4.3 PRESENTASJON OG OPERASJONALISERING AV VARIABLENE	44
4.3.1 DEN AVHENGIGE VARIABLEN	45
4.3.2 UAVHENGIGE VARIABLER	45
4.3.3 DUMMYVARIABLER.....	47
4.4 KODING	48
4.5 DATARENSING	49
4.6 DESKRIPTIV STATISTIKK	50
4.7 DESKRIPTIV STATISTIKK AV HVER ENKELT VARIABEL	52
4.7.1 BOLIGPRIS	52
4.7.2 BOLIGTYPE	53
4.7.3 EIEFORM	53
4.7.4 SALGSÅR.....	54
4.7.5 BOLIGAREAL.....	55
4.7.6 TOMTEAREAL TIL SELVEIDE ENEBOLIGER.....	55
4.7.7 BOLIGALDER	56
4.7.8 AVSTAND TIL SENTRUM	57
4.7.9 AVSTAND TIL SJØ.....	57
4.7.10 MÅ I BIL FOR Å KOMME TIL BARNEHAGE.....	58

4.8 KORRELASJON MELLOM VARIABLENE	58
5. ANALYSE	62
5.1 KORT OM REGRESJONSANALYSE	62
5.2 BIVARIAT REGRESJONSANALYSE	63
5.2.1 EKSEMPEL PÅ BIVARIAT REGRESJONSANALYSE	63
5.2.2 VURDERING AV FEILLEDDENE.....	65
5.3 LINEÆR REGRESJON MED RELEVANTE VARIABLER (MULTIVARIAT)	68
5.3.1 TEST PÅ MULTIKOLLINARITET.....	70
5.3.2 VURDERING AV FEILLEDDENE.....	71
5.4 LOGISTISKE MODELLER	72
5.4.1 SEMILOGARITMISK MODELL	73
5.4.2 VURDERING AV FEILLEDDENE.....	74
5.4.3 DOBBELTLOGARITMISK MODELL	75
5.4.4 VURDERING AV FEILLEDDENE.....	77
5.5 VURDERING AV MODELLENE - VALG AV MODELL	78
5.5.1 HAUGESUND	78
5.5.2 STAVANGER.....	78
5.6 SAMMENLIGNING AV TO REGRESJONER (STATISTISK REGRESJON)	79
5.7 EKSEMPEL VED BRUK AV LINEÆR MODELL	84
5.8 RESULTATER - HYPOTSETESTING	85
6. DRØFTING AV RESULTATENE	89
6.1 DRØFTING AV PROBLEMSTILLINGEN	89
6.2 KRITISKE VURDERINGER.....	92
7. KONKLUSJON	94
7.1 FORSLAG TIL VIDERE FORSKNING	94
LITTERATURLISTE	96
VEDLEGG	99

Figuroversikt

FIGUR 1.1: AVSTAND FRA HAUGESUND TIL STAVANGER (GOOGLE MAPS, 2014).....	12
FIGUR 3.1: SAMLET TILBUD OG SAMLET ETTERSPOERSEL I MARKEDET FOR BRUKTE BOLIGER PÅ KORT SIKT (NOU, 2002:2)	21
FIGUR 3.2: ENDRINGER I PRISEN PÅ BOLIGER VED ØKNING I TILBUDET AV OG ETTERSPOERSEL ETTER BOLIGER (NOU, 2002:2).....	22
FIGUR 3.3: HUSLEIEGRADIENTEN (JON P. KNUDSEN OG SIGBJØRN SØDAL, 2010).....	25
FIGUR 3.4: HUSHOLDNINGENES BUDFUNKSJONER (OSLAND, 2003)	32
FIGUR 3.5: PRODUSENTENES OFFERFUNKSJON (OSLAND, 2001)	36
FIGUR 3.6: MARKEDSLIKEVEKT (OSLAND, 2001)	37
FIGUR 4.1: FORDELING AV BOLIGER ETTER BOLIGPRIS I PERIODEN 2009-2013	53
FIGUR 4.2: HISTOGRAMMER FOR SALGSÅRENE 2009-2013.....	54
FIGUR 4.3: FORDELING AV BOLIGER ETTER BOLIGAREAL I KVADRATMETER	55
FIGUR 4.4: FORDELING AV TOMTEAREAL TIL ENEBOLIGER I HAUGESUND OG STAVANGER.....	56
FIGUR 4.5: FORDELING AV BOLIGALDER ETTER ÅR FOR HAUGESUND OG STAVANGER	56
FIGUR 4.6: ANTALL BOLIGER FORDELT ETTER AVSTAND TIL SENTRUM I HAUGESUND OG STAVANGER	57
FIGUR 4.7: ANTALL BOLIGER FORDELT ETTER AVSTAND TIL SJØEN I HAUGESUND OG STAVANGER.....	58
FIGUR 4.8: KORRELASJONSMØNSTER (ZIKMUND ET AL., 2013).....	59
FIGUR 5.1: KORRELASJONSPLOTT MELLOM BOLIGPRIS OG BOLIGAREAL, HAUGESUND & STAVANGER.....	66
FIGUR 5.2: NORMALSKRÅPLOTT FOR DEN ENKLE LINEÆRE MODELLEN, HAUGESUND & STAVANGER	66
FIGUR 5.3: RESIDUALPLOTT TIL DEN LINEÆRE MODELLEN, HAUGESUND & STAVANGER.....	71
FIGUR 5.4: NORMALSKRÅPLOTT TIL DEN LINEÆRE MODELLEN, HAUGESUND & STAVANGER	72
FIGUR 5.5: RESIDUALPLOTT TIL DEN SEMILOGARITMISKE MODELLEN, HAUGESUND & STAVANGER.....	75
FIGUR 5.6: NORMALSKRÅPLOTT TIL DEN SEMILOGARITMISKE MODELLEN, HAUGESUND & STAVANGER.....	75
FIGUR 5.7: RESIDUALPLOTT TIL DEN DOBBELLOGARITMISKE MODELLEN, HAUGESUND & STAVANGER.....	77
FIGUR 5.8: NORMALSKRÅPLOTT TIL DEN DOBBELLOGARITMISKE MODELLEN, HAUGESUND & STAVANGER.....	77

Tabelloversikt

TABELL 2.1: BOLIGER ETTER BOLIGTYPE I 2013 (STATISTISK SENTRALBYRÅ, 2014A)	12
TABELL 2.2: SELVEIERBOLIGER ETTER BOLIGTYPE 2009-2013 (STATISTISK SENTRALBYRÅ, 2014B)	14
TABELL 2.3: BOLIGER ETTER BOLIGTYPE I 2013 (STATISTISK SENTRALBYRÅ, 2014A)	15
TABELL 2.4: SELVEIERBOLIGER ETTER BOLIGTYPE 2009-2013 (STATISTISK SENTRALBYRÅ, 2013B)	15
TABELL 2.5: BRUTTOINNTÉKT BOSATTE PERSONER 17 ÅR OG ELDRE I 2012 (STATISTISK SENTRALBYRÅ, 2014C)	16
TABELL 2.6: ANTALL VIRKSOMHETER I 2014 (STATISTISK SENTRALBYRÅ, 2014D)	17
TABELL 4.1: KODING	49
TABELL 4.2: OVERSIKT OVER GJENNOMFØRINGEN AV DATARENSING.....	50
TABELL 4.3: DESKRIPTIV STATISTIKK, HAUGESUND & STAVANGER	51
TABELL 4.4: BOLIGPRIS FOR HAUGESUND OG STAVANGER (01.01.09-31.12.13).....	52
TABELL 4.5: BOLIGTYPE I HAUGESUND OG STAVANGER I ANTALL OG PROSENT.....	53
TABELL 4.5: EIEFORM I ANTALL OG PROSENT FOR HAUGESUND OG STAVANGER.....	54
TABELL 4.6: ANTALL BOLIGER SOLGT HVERT ÅR (2009-2013).....	54
TABELL 4.7: BOLIGAREAL I KVM ² FOR HAUGESUND OG STAVANGER.....	55
TABELL 4.8: TOMTEAREAL TIL ENEBOLIG I HAUGESUND OG STAVANGER I KVM ²	55
TABELL 4.9: BOLIGALDER ETTER ANTALL ÅR.....	56
TABELL 4.10: AVSTAND TIL SENTRUM I KILOMETER.....	57
TABELL 4.11: AVSTAND TIL SJØ I KILOMETER	58
TABELL 4.12: MÅ I BIL FOR Å KOMME TIL BARNEHAGE (ANTALL OG PROSENT).....	58
TABELL 4.13: KORRELASJONSMATRISSE, HAUGESUND	60
TABELL 4.14: KORRELASJONSMATRISSE, STAVANGER	60
TABELL 5.1: BIVARIAT REGRESJONSANALYSE, HAUGESUND	63
TABELL 5.2: BIVARIAT REGRESJONSANALYSE, STAVANGER	64
TABELL 5.3: LINEÆR REGRESJON MED ALLE RELEVANTE VARIABLER, HAUGESUND.....	69
TABELL 5.4: LINEÆR REGRESJON MED ALLE RELEVANTE VARIABLER, STAVANGER.....	69
TABELL 5.5: VIF-TEST, HAUGESUND & STAVANGER	71
TABELL 5.6: SEMILOGARITMISK REGRESJONSANALYSE, HAUGESUND	73
TABELL 5.7: SEMILOGARITMISK REGRESJONSANALYSE, STAVANGER	74
TABELL 5.8: DOBBELTLOGARITMISK REGRESJON, HAUGESUND	76
TABELL 5.9: DOBBELTLOGARITMISK REGRESJON, STAVANGER	76
TABELL 5.10: INTERAKSJONSMODELL MED RELEVANTE VARIABLER.	81

1. Innledning

De høye boligprisene i Stavanger er et debattert tema. Det dukker stadig opp i media at det er en økende trend at folk både velger og blir tvunget til å bosette seg litt lengre utenfor byen på grunn av de høye boligprisene i Stavanger. Flere velger også å flytte til Haugesund¹. Overskrifter som "Valgte det gode liv i Haugesund – de flyttet over Boknafjorden fra rådyre Stavanger til langt rimeligere Haugesund. Mange vurderer det samme" (nrk.no, 2012) er et eksempel.

Mens Stavanger er den dyreste byen i Norge å bosette seg i ², har Haugesund har et mer moderert boligmarked, og kan gjerne sies å ha et uforløst potensiale. Dersom man velger å bosette seg nord istedenfor sør for Boknafjorden, kan man gjerne ha råd til en hytte ved sjøen eller på fjellet i tillegg.

Av "Omdømmesbarometeret 2013" for Haugesundregionen, mener 32% av respondentene at man får mye mer for pengene i boligmarkedet i Haugesundregionen. I Stavangerregionen mener kun 7% det samme³.

Med boligprisene i Haugesund og Stavanger på dagsorden synes jeg det hadde vært av interesse å se hvilke attributter som har betydningen for boligprisen i både Stavanger og Haugesund.

Jeg har dermed utarbeidet denne problemstillingen: *Hvilken betydning har utvalgte boligattributter på boligprisen i Haugesund og Stavanger? Perioden 2009-2013.*

Områdene som skal studeres er avgrenset til kommunene Haugesund og Stavanger. Perioden jeg har valgt er de fem siste årene, fra 2009 til og med 2013.

¹ <http://www.nrk.no/rogaland/store-prisforskjeller-pa-bolig-1.8268358>

² <http://www.dinside.no/921053/stavanger-er-den-dyreste-byen-aa-etablere-seg-i>

³ Omdømmesbarometeret utarbeidet av Ordkraft og utgitt av Hovedorganisasjonen VIRKE

I denne oppgaven vil først utvalgsområdene bli presentert for å bli bedre kjent med områdene som skal undersøkes. Videre i kapittel tre vil jeg presentere generelle trekk ved boligmarkedet, teori om totemarkedet og den hedonistiske metoden. Kapittel fire redegjør for innsamling og behandling av datamaterialet. Videre vil jeg presentere datamaterialet på en oversiktlig måte. I kapittel fem er selve analysen av datamaterialet. I kapittel seks drøfter jeg resultatene før oppgaven avsluttes med en konklusjon i kapittel syv.

2. Bakgrunn

I dette kapittelet vil jeg innledningsvis kort kommentere boligmarkedet i Norge og Rogaland generelt. Videre vil jeg presentere generell fakta om Haugesund og Stavanger, før jeg avslutningsvis vil beskrive boligmarkedene til byene.

Helt siden 2009 har Norge opplevd en stigende trend i boligprisene. De siste par årene har vi likevel sett tendenser til at markedet har kjølnet litt og moderert seg. Dette kan underbygges av at vi fikk et fall på boligmarkedet høsten 2013 for første gang siden finanskrisen, som ikke kan forklares av sesongvariasjoner⁴.

Christian Vammervold Dreyer, administrerende direktør i Eiendomsmeglerforetakenes Forening uttalte i januar 2014 at Norge har et velfungerende og mobilt marked som fungerer slik det skal⁵.

Rogaland er Norges fjerde største fylke og er lokalisert på Vestlandet. Fylket grenser mot Hordaland i nord og Aust-Agder, Vest-Agder og Telemark i øst. Rogaland er det fylke i Norge som er klart mest preget av olje og gasseventyret, både i privat og offentlig sektor⁶. Ifølge fylkesordfører i Rogaland Janne Johnson er Rogaland det fylket hvor "alle" ønsker å bo og arbeide⁷. I 2013 var det en nettotilflytting på 4139 personer i fylket⁸. Dette fører til at Rogaland vil ha store utfordringer med å anskaffe nok boliger i fremtiden.

⁴ http://eiendomsverdi.no/app/Factsheets/Fakta_januar_2014.pdf

⁵ <http://www.dinside.no/926037/boligprisene-ned-i-desember>

⁶ <http://www.bnl.no/article.php?articleID=2379&categoryID=6>

⁷ http://www.bnl.no/getfile.php/Filer/Publikasjoner/BNL_Rogaland_web%282%29.pdf

⁸ Statistisk Sentralbyrå, flyttinger, Rogaland, 2013

2.1 Introduksjon av Haugesund og Stavanger

For å bli bedre kjent med områdene som skal undersøkes vil jeg presentere byene hver for seg.

2.2 Haugesund

Haugesund er en kystby lokalisert på Vestlandet "midt" mellom Stavanger og Bergen. Haugesund kommune er relativt liten med et areal på 72km².

Kommunen grenser i nord mot Sveio, i vest mot Bømlo, i øst mot Tysvær og i sør mot Karmøy. Byen ligger åpent mot Nordsjøen, men blir til dels skjermert av øyene Røvær og Karmøy som tilhører Karmøy kommune⁹. Det urbane sentrum i byen er preget av et kvadratur og strekker seg langs Smedasundet og over broer til småøyene Risøy og Hasseløy. Haugesund er i dag en av Norges fremste handelsbyer, med en handel over dobbelt så stor som folketallet skulle tilsi¹⁰. Medvirkende årsaker til det er tilgang til flere store kjøpesenter og et variert handelstilbud.¹¹

Befolkningsveksten har økt fra 31 361 innbyggere ved inngangen av 2004 til 36 107 ved utgangen av 2013, ifølge Statistisk Sentralbyrå. Dette vil tilsi en samlet vekst på 15,13%. Videre er det knyttet usikkerhet til om Haugesund klarer å tiltrekke seg den yngre generasjon, slik som andre storbyer, og om hvor stor del av de som flytter fra Haugesund vil vende tilbake igjen til byen etter for eksempel endt studier (Scenarier2029, 2012).

Byen er kjent for et differensiert næringsliv som gjør at konjunkturbølger på landsbasis har mindre effekt i Haugesund enn andre steder. Den økende aktiviteten i Nordsjøen og i oljenæringen har bidratt til god vekst i regionen. Flere store selskaper tilknyttet dette næringslivet er lokalisert i Haugesundregionen, blant annet Statoil, Aker Stord og Aibel¹².

Tradisjonelt sett har sjøveien vært den viktigste transport veien i dette området, men de siste to tiårene har riksveiene E134 og E39 henholdsvis østover mot Oslo

⁹ <http://www.bysiden.no/sted/Haugesund/places/info>

¹⁰ <http://www.bysiden.no/sted/Haugesund/places/info>

¹¹ <http://www.bysiden.no/sted/Haugesund/places/info>

¹² <http://www.estatenyheter.no/index.php/component/content/article/5-nyheter/nyheter/1006-full-rullepa-haugalandet>

og sørover mot Stavanger bidratt til forbedret kommunikasjon og utvikling av byen. I tillegg har det vært en merkbar trafikkøkning på Helganes flyplass.

2.3 Stavanger

Stavanger er lokalisert 82 km¹³ sør for Haugesund og blir i mange sammenhenger referert som oljehovedstaden i Norge. Denne betegnelsen har som opphav fra den store olje- og petroleumsaktiviteten i område. Målt i reisetid bruker man rett i underkant av to timer med bil fra bykjernen i Haugesund, inkludert en ferjeovergang på 25 minutter (Arsvågen-Mortavika) til Stavanger sentrum¹⁴.

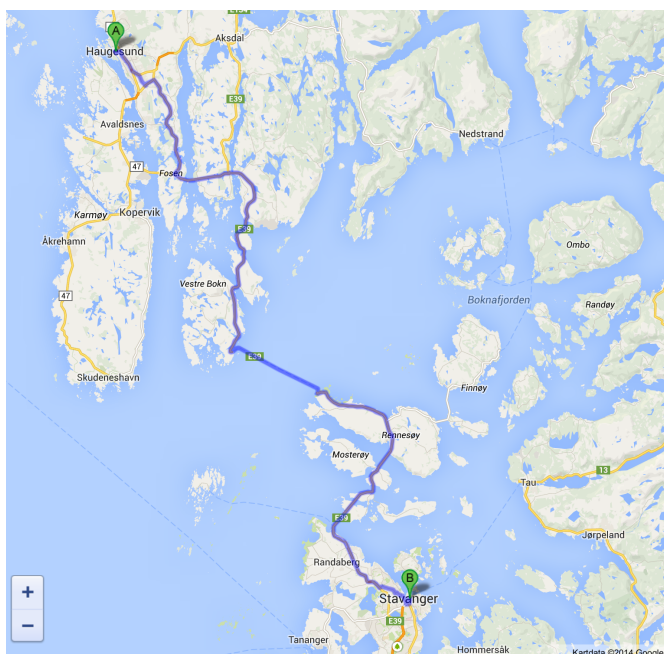
Stavanger kommune består av fastlandet og et stort antall bebodde øyer. Disse øyene er Hundvåg/Buøy, Austre Åmøy, Langøy, Bjørnøy, Roaldsøy, Ormøy, Steinsøy, Engøy, Sølyst, Grasholmen, Vassøy, Lindøy, Hellesøy og Kalvøy. Totalt utgjør dette et areal på 71 km², omtrent samme størrelse som Haugesund. Stavanger Kommune grenser til nabokommunene Rennesøy, Randaberg, Sola og Sandnes (Stavanger-kommune.no, Stavanger-statistikken).

Stavangers befolkningsvekstrate de ti siste årene er marginalt høyere enn hva den er i Haugesund med 16,35%. Det er viktig å presisere at folkeantallet er langt høyere. Ifølge Statistisk Sentralbyrå var befolkningen 112 405 ved inngangen av 2004 og 130 788 ved utgangen av 2013, noe som gjør den til Norges mest befolkningstettede kommune¹⁵.

¹³ Tall hentet fra Google Maps

¹⁴ Avstand hentet fra Google Maps

¹⁵ <http://www.husbanken.no/boligsosialt-arbeid/boligsosial-planlegging/boligsosiale-utviklingsprogram/region-vest/programkommunene/stavanger-kommune/>



Figur 1.1: Avstand fra Haugesund til Stavanger (Google Maps, 2014)

2.4 Boligmarkedet i Haugesund og Stavanger

Nå vil jeg presentere Haugesund og Stavanger og deres tilhørende boligmarkeder.

2.4.1 Haugesund

Boligmarkedet i Haugesund er variert og stabilt i den forstand at det ikke i stor grad blir påvirket av eventuelle opp- eller nedgangskonjunkturer i boligmarkedet ellers i landet. Dette gjør at man ikke får de store prisvariasjonene (Horisont - TV Haugaland, 2013). I tabell 2.1 er det gitt en oversikt over beholdningen av de ulike boligtypene i Haugesund. Eneboliger var i 2013 den boligtypen som hadde klart høyest beholdningen med en andel på 45%.

Tabell 2.1: Boliger etter boligtype i 2013 (Statistisk Sentralbyrå, 2014a)

	2013	Andel i %
Haugesund	Boliger (beboede og ubebodde)	
Enebolig	8002	45 %
Tomannsbolig	2034	11 %
Rekkehus, kjedehus og andre småhus	3317	19 %
Boligblokk	3164	18 %
Bygning for bofellesskap	144	1 %
Andre bygningstyper	1245	7 %
SUM	17906	100 %

* Andre bygningstyper inkluderer i hovedsak boliger i garasjer, næringsbygninger og andre bygningstyper som ikke er boligbygninger.

Boligarealreserven i Haugesund er forholdsvis stor, noe som gjør at det er mange arealer velegnet til både boligbygging og næringsbygging tilgjengelig. Videre er potensialet for regionforstørrelse betydelig i regionen. Vei-prosjektet T-sambandet som ble ferdigstilt i 2013 knytter fastlandet og Karmøy nærmere hverandre. Videre vil veiprojektet Rogfast, som innebærer muligheten til å krysse Boknafjorden og Kvitsøyfjorden med undersjøisk tunnel, kunne knytte Haugesund og Stavanger nærmere hverandre¹⁶. Når Rogfast er ferdigstilt i 2022 eller 2013 vil det kunne endre dynamikken i boligmarkedet i både Haugesund og Stavanger¹⁷. Et markert lavere prisnivå i Haugesund i forhold til Stavanger, kombinert med eneboligtomter kan bidra til at attraktiviteten til Haugesund økes for arbeidstakere i Stavanger (Scenarier2029, 2012).

Hva som er å forvente som attraktive boligområder forandres kontinuerlig. Folk har ulike behov og syn på hva de mener er attraktivt. Det er blitt gjort få undersøkelser av hvilke boligpreferanser, flyttemønstre og flyttemotiver som finnes i Haugesund. Man kan likevel peke på noen trender. Tendensene i markedet er at eneboliger er attraktivt og disse skifter raskt eier når de blir lagt ut for salg (Horisont – TV Haugaland, 2013). Hovedårsaken til dette er fordi det er en forholdsvis liten tilgang på eneboliger i regionen (Scenarier2029, 2012).

I tabell 2.2 er boligprisutviklingen i Haugesund og hele landet presentert. Den viser at Haugesund har opplevd en god positiv prisvekst i boligmarkedet de siste årene. Boligprisene har likevel hatt en lavere prisvekst i forhold til hva som er tilfellet på landsbasis. Prisnivået er også lavere i Haugesund enn hva det er på landsbasis.

¹⁶ <http://www.vegvesen.no/Europaveg/e39rogfast/Fakta>

¹⁷ <http://www.h-avis.no/nyheter/billigere-a-kjore-rogfast-1.8093560>

Tabell 2.2: Selveierboliger etter boligtype 2009-2013 (Statistisk Sentralbyrå, 2014b)

		2009	2013	2009-2013
Område	Boligtype	Gj.snitt kvadratmeterpris (kr)	Gj.snitt kvadratmeterpris (kr)	%
Haugesund	Eneboliger	15 466	18 446	19,4 %
	Småhus	15 162	18 385	21,3 %
	Blokkleiligheter	20 956	24 963	19,1 %
Hele landet	Eneboliger	16 847	20 780	23,3 %
	Småhus	21 143	27 648	30,8 %
	Blokkleiligheter	31 326	41 506	32,5 %

**Prisen er pr. kvadratmeter primære rom (P-ROM)*

2.4.2 Stavanger

Stavanger er preget av pressede boligpriser, som følge av høy etterspørsel og en liten tilgang på boliger. Den høye oljeaktiviteten i området og en liten geografiske utstrekning er underbyggende årsaker til det. Man har begynt å se tegn på et flyttemønster av at folk velger å bosette seg i periferien på grunn av de høye boligprisene. Spesielt studenter og nyetablerere har vanskeligheter med å komme seg inn på markedet. Videre er folk i Stavanger mer indifferent i forhold til andre byer om man bosetter seg i sentrum eller litt utenfor (Scenarier2029, 2012).

Stavanger har ikke like store arealområder tilgjengelig til boligbygging som Haugesund. Byveksten de siste årene har i hovedsak foregått gjennom gjenbruk av eksisterende bygninger, som følge av fortetting og byfornyelse (Stavanger kommuneplan, 2010-2025).

Som i Haugesund er eneboliger den dominerende boligtypen i Stavanger, ref. tabell 2.3. Andelen eneboliger utgjør likevel ikke en like stor del som i Haugesund. I motsetning til Haugesund er det viktig å merke seg at Stavanger har en høyere andel boligblokker. Hovedgrunnene til dette er mangel på arealer og nye boligtrender som har bidratt til økt storstilt utbygging av leiligheter. Felles for disse leilighetene er at de er svært kostbare, noe som er med å forsterke prispresset (Scenarier2029, 2012). Totalt sett er det over tre ganger så mange boliger i Stavanger i forhold til Haugesund.

Tabell 2.3: Boliger etter boligtype i 2013 (Statistisk Sentralbyrå, 2014a)

	2013	Andel i %
Stavanger	Boliger (beboede og ubebodde)	
Enebolig	21885	36 %
Tomannsbolig	11433	19 %
Rekkehus, kjedehus og andre småhus	9620	16 %
Boligblokk	14995	25 %
Bygning for bofellesskap	1695	3 %
Andre bygningstyper	1111	2 %
SUM	60739	100 %

* Andre bygningstyper inkluderer i hovedsak boliger i garasjer, næringsbygninger og andre bygningstyper som ikke er boligbygninger.

Videre har gjennomsnittstørrelsen på boliger i Stavanger blitt redusert de 10 siste årene. For å takle den økende befolkningsveksten har Stavanger Kommune lagt frem i sin kommuneplan at ca. to av tre nye boliger som bygges skal være minimum 55m². (Stavanger kommuneplan, 2010-2025).

Prispresset som oppleves i Stavanger gjenspeiles av kvadratmeterprisen. I tabell 2.4 er kvadratmeterprisen og dens utvikling illustrert.

Tabell 2.4: Selveierboliger etter boligtype 2009-2013 (Statistisk Sentralbyrå, 2013b)

		2009	2013	2009-2013
Område	Boligtype	Gj.snitt kvadratmeterpris (kr)	Gj.snitt kvadratmeterpris (kr)	%
Stavanger	Eneboliger	25 729	34 634	34,6 %
	Småhus	26 329	34 451	31,8 %
	Blokkleiligheter	34 258	47 169	37,7 %
Hele landet	Eneboliger	16 847	20 780	23,3 %
	Småhus	21 143	27 648	30,8 %
	Blokkleiligheter	31 326	41 506	32,5 %

*Prisen er pr. kvadratmeter primære rom (P-ROM)

Kvadratmeterprisen for eneboliger i 2013 er nesten dobbelt så høy per kvadratmeter i Stavanger i forhold til Haugesund representert med en prisdifferanse på litt over 16 000 kr. Prisdifferansen mellom byene er størst ved blokkleiligheter på omtrent 22 000 kr. Prosentvis økning de 5 siste årene er også dobbelt så stor for denne type boliger i Stavanger i forhold til Haugesund. Sett i forhold til landsbasis er kvadratmeterprisen dyrere i Stavanger og har samtidig hatt en høyere prosentvis vekst.

En regionforstørrelse og utflytting til andre plasser i regionen kan være med å lette på prispresset i Stavanger. De store vei- og tunnelprosjekter som Ryfast og Rogfast vil være med å knytte Stavanger tettere sammen med Haugesund og kan bidra til dette, men disse vil være ferdigstilt lengre frem i tid (Scenarier2029, 2012)

Ifølge Olav Askeland, eiendomsutvikler i Haugesund og Stavanger, er mange bevisste på hvor nærmeste barnehage er i forhold til hvor de velger å bosette seg. Dette stemmer overens med en fersk markedsundersøkelse av Haugesundsregionen gjort av Respons analyse på vegne av Eiendomsmegler Vest¹⁸. Videre viser det seg at det kan være en utfordring å få barnehageplass i samme bydel hvor man er bosatt. Disse utfordringene er spesielt knyttet til Stavanger¹⁹.

I tabellen nedenfor vises bruttoinntekten for bosatte i Haugesund og Stavanger. I 2012 hadde bosatte i Stavanger ca. 100 000 kr høyere bruttoinntekt enn bosatte i Haugesund. Dette indikerer en høyere betalingsevne i Stavanger i forhold til Haugesund.

Tabell 2.5: Bruttoinntekt bosatte personer 17 år og eldre i 2012 (Statistisk Sentralbyrå, 2014c)

	2012	
By:	Bruttoinntekt	
Stavanger	kr	497 200
Haugesund	kr	391 600

Stavanger kan sies å være en attraktiv by å bo for de som har råd til å bosette seg der. En av grunnene til det er antall arbeidsplasser. Mulighetene til å få arbeid kan være større i Stavanger enn hva den er i Haugesund. Årsaken til dette er antall bedrifter som er lokalisert i Stavanger i forhold til Haugesund. Tabellen under illustrerer nettopp dette.

¹⁸ <http://touch.h-avis.no/nyheter/sann-vil-vi-bo-1.8389328>

¹⁹ <http://www.nrk.no/rogaland/barnehageplass-utenfor-egen-bydel-1.8146537>

Tabell 2.6: Antall virksomheter i 2014 (Statistisk Sentralbyrå, 2014d)

By	2014
	Bedrifter - alle størrelser
Stavanger	11082
Haugesund	3315

Kort oppsummert har befolkningen økt i begge byene de ti siste årene. Eneboliger utgjør den største bolig andelen både i Haugesund og Stavanger på henholdsvis 45% og 36%. Sett opp mot landsbasis har boligprisutviklingen vært høyere i Stavanger, mens den har vært lavere i Haugesund. Det er over tre ganger så mange boliger i Stavanger enn i Haugesund. Videre har bosatte i Stavanger enn høyere inntekt enn bosatte i Haugesund. Det er også flere virksomheter i Stavanger enn i Haugesund.

3. Teori

I dette kapitlet vil jeg presentere teori som jeg vil anvende senere i oppgaven og som vil hjelpe å belyse problemstillingen min. Jeg vil starte med å presentere hovedtrekkene i eiendomsmarkedet og si hva som kjennetegner et boligmarked. Dette er viktige momenter som vil bidra til en bedre forståelse av mekanismene i boligmarkedet. Videre vil jeg presentere teori om totemarkedet ved Alonso-Muth-Mills sin modell som gir en enkel forklaring på hvordan hus- og tomteleien bestemmes av ulik lokalisering i forhold til sentrum. Avslutningsvis vil den hedonistiske metoden bli utledet for å vise hvordan boligpriser bestemmes ut ifra forskjellige attributter. Teorien gjør at vi til slutt kan presentere hypoteser basert på de forskjellige attributtenes påvirkning på boligprisen.

3.1 Kjennetegn ved boligmarkedet

En bolig er et økonomisk gode som skiller seg fra andre økonomiske goder man ellers finner i hverdagen. Det kan derfor oppsummeres med flere ulike kjennetegn. Kjennetegnene bygger på forelesningsnotatene til dosent Karl Robertsen fra kurset Eiendomsøkonomi (BE-409), høsten 2013.

Boligmarkedets kjennetegn er følgende:

Varig konsumkapital.

En bolig har lang levetid og det er meningen at den skal stå på fast grunn i lang tid fremover. En boliginvestering medfører store transaksjonskostnader og flyttekostnader, noe som gjør at folk ikke foretar denne investeringen ofte.

Immobilitet

At godet er immobilt vil si at det svært vanskelig og krevende å flytte boligen. I tillegg vil flytting av bolig være knyttet til store kostnader. Man velger å kjøpe en bolig der man ønsker den skal være.

Heterogent

Boligmarkedet består av heterogene goder. Ingen boliger er identiske fordi alle boligene har ulike attributter som gjør dem unike. Hovedfaktoren til dette er lokaliseringen.

Boligkapitalen produserer bolig tjenester.

Boligen vil produsere tjenester over tid etter våre behov. Behovene kan gjerne endres over tid.

Uelastisk kortsiktig tilbud.

På kort sikt er tilbudet av boliger gitt. Dette fordi det er en lenger prosess å bygge boliger.

Tilbudet kan økes på lang sikt.

På et lengre perspektiv kan tilbudet av boliger økes ved mer nybygging. Dette vil føre til økt boligbeholdning.

Markedstransaksjoner er forbundet med store søkekostnader og transaksjonskostnader.

Det kan ta tid å finne den rette boligen som tilfredsstillende preferanser og behov. Dette medfører søkekostnader. Ved et boligkjøp vil det påløpe transaksjonskostnader som flyttekostnader, meglerhonorar og dokumentavgifter. Disse kostnadene kan utgjøre omlag 8-10% av omsetningsverdien av boligen.

Fysisk beliggenhet er en del av eiendomsverdien.

Beliggenheten har ofte stor betydning for eiendomsverdien. Utsikt, solforhold, hage, nærhet til fasiliteter, turområder, avstand til skole, barnehage, jobb, bysentrum, er alle ulike faktorer/attributter som har betydning for eiendomsverdien.

På grunn av låneopptak, påvirker boligmarkedet folk sparinger.

Boligkjøp er ofte den største finansielle investeringen folk gjør i løpet av livet. De fleste må dermed ta opp lån ved et boligkjøp. Dette vil påvirke folks evne til å spare. Store deler av den disponible inntekten vil gå med til å betjene boliglånet, som igjen fører til at man må tilpasse seg situasjonen som boligeier.

Videre vil jeg presentere boligmarkedet, basert på Norges offentlige utredninger (NOU). Dette gjøres for å få en bedre forståelse om hvordan mekanismene i boligmarkedet fungerer.

3.2 Teori om prisdannelsen i boligmarkedet

Boligmarkedet er et komplisert og uoversiktlig marked som består av flere delmarkeder. Aktørene i markedet er ulike og beslutningene de tar er ofte med et langsiktig perspektiv. Jeg vil her gi en kort innføring i teori for å belyse drivkreftene i boligmarkedet. For å få frem sentrale sammenhenger og anvende teorien om prisdannelsen må man gjøre forenklerende forutsetninger. Videre vil det forutsettes at alle boliger er like, selv om dette bryter med et av boligenes kjennetegn. Samtidig vil det forutsettes at boligene er eierboliger. Leiemarkedet vil derfor ikke bli tatt hensyn til (NOU, 2002:2).

3.2.1 Etterspørsel etter boliger

Etterspørerne i boligmarkedet er alle som ønsker seg egen bolig hvis bare prisen er lav nok. Etterspørselsfunksjonen blir til ved å rangere alle markedsdeltakerne etter betalingsvilje. Den markedsdeltakeren som er villig til å betale mest, står først i rekken. Etterspørselsfunksjonen blir en fallende kurve som viser hvor mange etterspørere som er villig til å betale for boliger i ulike prisklasser (NOU, 2002:2).

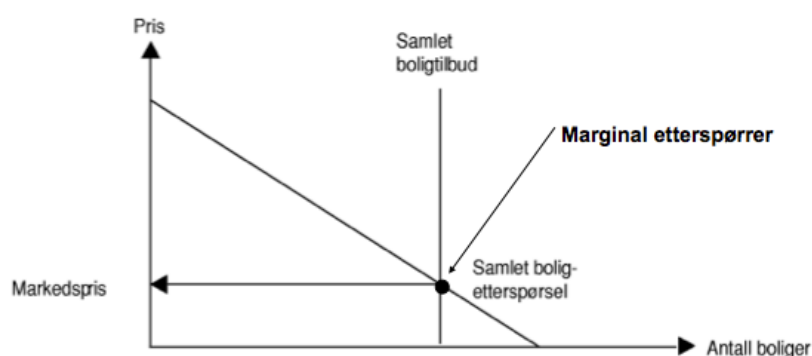
Husholdningenes betalingsvilje er i hovedsak bestemt av deres betalingsevne. Blant de med høyest betalingsvilje vil vi finne de med relativt høye inntekter eller formuer. Den prisen man evner å betale avhenger også av rentenivået og andre forhold som påvirker de årlige bokostnadene (det man gir avkall på av goder for å bruke på en bolig i en periode). Videre avhenger betalingsviljen av hvordan en vektlegger en bolig i forhold til andre konsumgoder og hvordan en vurderer investering i bolig i forhold til andre investeringer. Fremtidige forventninger og risiko kan også spille inn. Betalingsvilligheten vil også avhenge av preferanser og boligbehov, noe som betyr at det ikke alltid er de med høyest inntekt og formue som har størst betalingsvilje. Folk med samme betalingsevne kan derfor ha noe forskjellig betalingsvilje (NOU, 2002:2).

3.2.2 Tilbud av boliger

Tilbyderne i markedet vil være alle som er eiere av boliger. Disse vil selge boligene de ikke skal bruke. Tilbudet av boliger består av den eksisterende boligbeholdningen. Tilbudet av boliger endres derimot ved nybygging og avgang (fraflytting, riving og ombygging). Nybygging av boliger utgjør kun en liten del av den totale bolig andelen. I Norge utgjør nybyggingen ca. 1% av boligbeholdningen. Dette medfører at tilbudet ofte antas på kort sikt å være gitt, uavhengig av prisen (NOU, 2002:2)

3.2.3 Markedspris

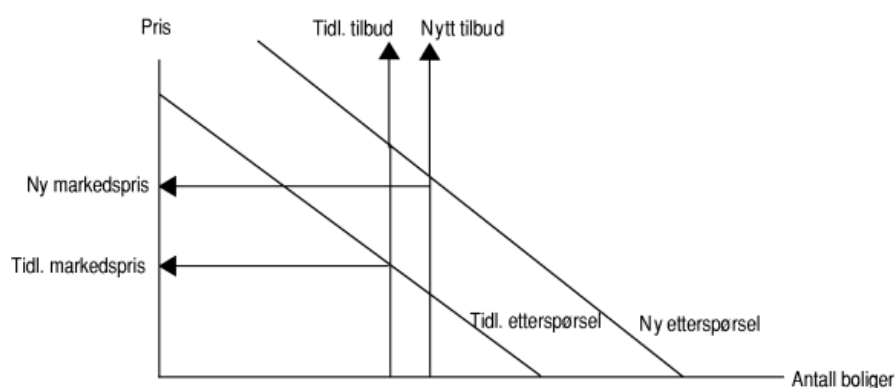
Ved å se på samlet etterspørsel og samlet tilbud av boliger kan man finne markedsprisen på bolig. Dersom prisen på bolig er lav nok, vil det være flere som etterspør boliger enn antallet boliger som finnes. Prisen vil da bli presset opp slik at mange etterspørrere trekker seg og det blir en bolig til de gjenværende. Dette medfører at man får et skjæringspunkt mellom etterspørselskurven og tilbudskurven. Dette skjæringspunktet bestemmer markedsprisen. De med høyere betalingsvilje enn dette punktet får egen bolig. De med lavere betalingsvillighet enn dette punktet får ikke bolig. Den "siste" som får bolig kalles gjerne den marginale etterspørreren. Denne personens betalingsvillighet spiller en sentral rolle fordi det er han eller hun som til syvende og sist bestemmer markedsprisen (NOU, 2002:2).



Figur 3.1: Samlet tilbud og samlet etterspørsel i markedet for brukte boliger på kort sikt (NOU, 2002:2)

På lenger sikt vil situasjonen endre seg da tilbudet av boliger vil øke som følge av nybygging. Man kan dermed ikke si at tilbudet av boliger lenger er gitt. Dette

medfører at vi får en ny tilbudskurve og økt boligbeholdning. Etterspørselen vil også øke som følge av befolknings- og inntektsvekst. Dersom etterspørselen har en sterkere vekst enn tilbudet får vi en etterspørselskurve som skifter utover i større grad enn tilbudskurven. Dette fører til at boligprisene presses opp og vi får en ny markedspris. Nedenfor er dette illustrert grafisk. I et motsatt tilfelle, hvor etterspørselsveksten er lavere enn nybyggingen, ville vi fått etterspørselskurve som skiftet innover og dermed fått en lavere markedspris (NOU, 2002:2).



Figur 3.2: Endringer i prisen på boliger ved økning i tilbudet av og etterspørsel etter boliger (NOU, 2002:2)

Tilbudet av boliger endres langsomt, det vil si at store prisbevegelser på kort tid vil komme fra forhold på etterspørselssiden. Antall etterspørrere i Norge endres ikke så veldig mye fra år til år. Det antas derfor at endringene i etterspørselen vil komme fra etterspørrernes betalingsvilje. Dette påvirkes igjen av betalingsevnen, vurderingen av bolig som et investeringsobjekt og preferanser (NOU, 2002:2).

3.3 Teori om tomtemarkedet

Jeg vil nå presentere en teori som baserer seg på det urbane tomtemarkedet som vil gi en forklaring på hvordan tomtepris varierer med beliggenhet.

For min oppgave er det av interesse å undersøke hvordan prisen på boliger og tomter varierer med ulik lokalisering fra sentrum. Da jeg opererer med to byer i denne oppgaven vil jeg derfor ha to ulike bysentrum. Haugesund og Stavanger har ulik bystørrelse, og ulik avstand til sentrum. Det er derfor naturlig at byene vil bli sett på hver for seg. Modellen som jeg skal se nærmere på har som formål å gi en forklaring av nettopp det jeg ønsker å undersøke.

3.3.1 Alonso-Muth-Mills modellen

Alonso-Muth-Mills modellen forklarer hvordan boligpriser i en by varierer med avstand fra sentrum. Utledningen er hentet fra en artikkel i "Økonomi og Tid" av Jon P. Knudsen og Sigbjørn Sødal (2010), der Karl Robertsen og Theis Theisen for seg DiPasquale & Wheaton (1996) sin forenklende fremstilling av Alonso-Muth-Mills modellen.

For mange vil beliggenhet ved boligen være av stor betydning, og det er nettopp dette denne modellen ønsker å forklare på en forenklet måte. På bakgrunn av dette kan vi fastslå at tomtearealer er fullstendig differensiert goder, som vil si at ingen tomter er identiske. Årsaken til dette er fordi hver enkelt tomt har ulik beliggenhet som er unik. Tilbudet av tomteareal på et avgrenset sted er uelastisk, mens etterspørselen er elastisk. Dette betyr at det dermed er etterspørselen som bestemmer tomteprisen og denne vil følgende variere med beliggenheten (Robertsen og Theisen, 2010).

Modellen tar utgangspunkt i en rekke forenklende forutsetninger i forhold til virkeligheten og de er presentert følgende (Robertsen og Theisen, 2010):

- Det er en monosentrisk by med ett bysenter hvor alle jobbene er. I mitt tilfelle har jeg to slike sentrum, representert ved Haugesund og Stavanger.
- Byen har en gitt bygningsstruktur som er bestemt av historisk bygging, det er ikke mulig å substituere land med høyere bygninger.
- Folk pendler til sentrum langs en rett linje. Det koster k kr pr km, pr år i transportkostnader. Pendlingsvariabelen er d .
- Alle husholdningene er identiske, og den eksogent gitte inntekten y brukes til pendling, husleie og annet konsum.
- Alle boliger er identiske og vi får med det en husleiefunksjon uttrykt ved $R(d)$.
- Husleietjenester produseres av tomtearealet q pr hus og annen innsats c som er et uttrykk for bygningskostnader, som vil si leie av selve bygningen isolert fra tomten.
- Boligene leies ut til den leietaker som har størst betalingsvillighet.

Når alle husholdninger har samme nyttefunksjon, inntekt, størrelse mv., må forskjeller i husleie $R(d)$ tilsvare eksakt forskjeller i transportkostnader. Videre betyr dette at alle husholdninger har likt annet konsum, det vil si $x = x^0$ (annet konsum) for alle. Husleien blir dermed for en gitt avstand d fra sentrum ((Robertsen og Theisen, 2010):

$$(1) R(d) = y - kd - x^0.$$

I sentrum er da husleien:

$$(2) R(0) = y - x^0.$$

Grunnen til dette er fordi man antar at arbeidsplassene er i sentrum og dermed faller pendlingskostnaden kd bort. De som bor i sentrum vil ikke ha noen form for transportkostnad til og fra jobb. Ved bevegelse ut fra sentrum derimot, vil pendlingskostnaden spille en viktig rolle. Husleien vil falle med økt avstand fra sentrum og arbeidsplassene, dette vises med et negativt fortegn i husleiefunksjonen (1). Jo lenger man beveger seg ut ifra sentrum, jo høyere blir pendlingskostnaden.

Videre antar vi at $d = b$, hvor b representerer avstand til bygrensa. Utenfor bygrensa b er jordbruk den eneste alternative arealbruken. Ved å bruke arealet til jordbruk får man jordleie r^a per mål. Dersom man befinner seg på bygrensa vil tomteleien være $r^a q$, hvor q uttrykker tomtearealet. I tillegg har man byggeleie uttrykket ved c . Byggeleien er annuiteten av bygningskostnadene. Husleia på bygrensa vil dermed være summen av tomteleie og byggeleien (Robertsen og Theisen, 2010):

$$(3) R(d) = r^a q + c$$

Annent konsum for alle husholdninger er lik og er følgende:

$$(4) x^0 = y - kb - (r^a q + c)$$

En husholdning som bor i en avstand d fra sentrum vil ha husleien som følger (Jon P. Knudsen og Sigbjørn Sødal, 2010):

$$(5) R(d) = y - kd - x^0$$

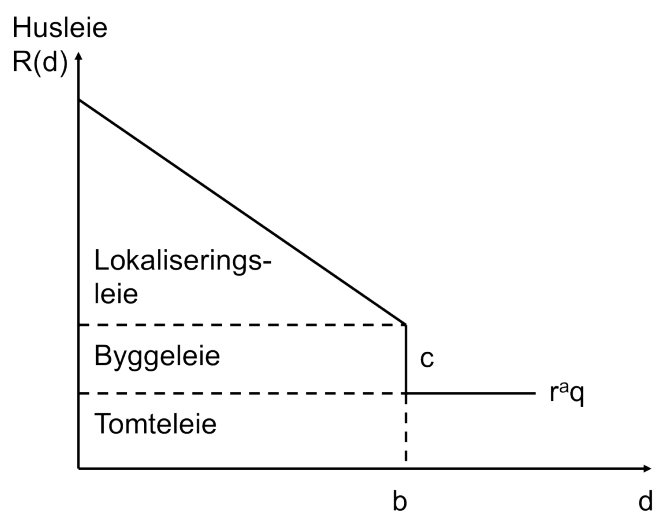
Man kan nå sette (4) inn for x^0 i (5) og får dermed:

$$(6) R(d) = y - kd - y + kb + (r^a q + c)$$

Ved å rydde opp i (6) kan husleiegradienten for en husholdning i avstand d fra sentrum formuleres slik:

$$(7) R(d) = (r^a q + c) + k(b - d)$$

Likning (7) illustrert i en modell/figur:



Figur 3.3: Husleiegradienten (Jon P. Knudsen og Sigbjørn Sødal, 2010)

Sentrum er illustrert nede til venstre i figuren. Videre kan man se at tomteleien og byggeleien holder seg konstant for alle avstander fra sentrum. Det som ikke holder seg konstant er lokalisierungsleien. Jo lenger man beveger seg ut ifra sentrum, jo mer synker lokalisierungsleien. Lokalisierungsleien er høyest der avstanden til sentrum er lik null.

Ved bevegelse utover fra sentrum faller husleien med nøyaktig forskjellen i transportkostnader. Dette bekreftes av husleiegradienten som viser helningen til gradienten. Denne får vi ved å derivere $R(d)$ med hensyn på avstand d :

$$(8) \quad \frac{\partial R(d)}{\partial d} = -k$$

I likning (8) ser vi at husleien faller i den størrelsesorden man beveger seg ut ifra sentrum.

Den urbane tomteleien består av tomteleie og lokaliseringsleie per mål:

$$(9) \quad r(d) = \frac{R(d)-c}{q}$$

I uttrykket (9) er byggeleien trukket fra, for å så delt på antall mål q . Her er q gitt og konstant. Vi setter inn uttrykket (7) inn i (9) og får:

$$(10) \quad r(d) = \frac{(r^a q + c) + k(b-d) - c}{q}$$

Dette uttrykket kan vi rydde opp og vi får følgende uttrykk:

$$(11) \quad r(d) = r^a + \frac{k(b-d)}{q}$$

Det første leddet i uttrykket er tomteleie, mens det siste leddet er sparte transportkostnader per mål. Tomteleie avtar med økningen i transportkostnader pr mål:

$$(12) \quad \frac{\partial r(d)}{\partial d} = -\frac{k}{q}$$

Dette betyr husleiegradienten er avtagende og har en helning lik: $-\frac{k}{q}$

Basert på denne forenklede modellen kan vi trekke noen konklusjoner:

- Hus- og lokaliseringsleie ville vært høyere for alle lokaliseringer innenfor bygrensa (b) dersom bygrensa hadde vært lenger ute.
- Dersom pendlingskostnadene hadde vært på et høyere nivå, ville dette ført til at hus- og lokaliseringsleie for alle lokaliseringer innenfor bygrensa ville vært høyere.
- Ved høyere byggekostnader (c) eller en bedre avkastning i landbruket ($r^a q$) ville man fått en høyere husleie.
- Til sist kan vi konkludere med at dersom bygningstettheten var større (mindre (q)), ville gradienten for tomteleie blitt brattere og vi ville fått relativt høyere tomteleie i sentrum.

3.4 Den hedonistiske metoden

Presentasjonen av den hedonistiske metoden er basert på grunnlag av artikkelen av Liv Osland "Den hedonistiske metoden og estimering av attributtpriser".

Denne metoden er ofte brukt til å analysere boligmarked og variasjonene i boligprisene. Betegnelsen *hedonisme* er gresk og stammer fra ordet *hedone*, som betyr lyst eller glede (Osland, 2001). Ingen boliger er identiske og er følgende heterogene goder bestående av ulike enkeltattributter som gir glede eller nytte for konsumenten. Attributter er ulike egenskaper ved en bolig som for eksempel, avstand til sentrum, utsikt, alder, størrelse, etc. Disse attributtene vil samlet forklare og påvirke prisen på en bolig. Hedonistiske prisfunksjoner viser av den grunn hvordan prisen på en bolig avhenger av attributtprisene. Ved å benytte hedonistiske prisfunksjoner vil jeg forsøke å forklare pris på ulike attributter i to ulike byer.

Konsumenter har ulike preferanser når det kommer til hvilke attributter som gir nytte eller glede. Dette påvirker dermed konsumentenes betalingsvillighet for en bolig. Attributter gir nytte for en konsument, mens de genererer kostnad for produsentene da det antas at de vil tilpasse seg disse preferansene. Siden det eksisterer tilbud og etterspørsel etter enkeltattributter, som gir ulike nytte, blir det også sentralt å se på prisstrukturen til attributtene.

Indirekte priser, hedonistiske priser, marginale priser eller implisitte priser er alle fellesbetegnelser for attributtpriser. Attributtprisene kan observeres

indirekte via totalprisen på godet og er definert som økning i samlet pris på godet ved en marginal partiell økning i mengden av et attributt. (Osland, 2011). Formulert på en enklere måte, dersom man øker mengden av attributter til en bolig og med antagelsen at konsumentene oppnår nytte eller glede av attributtene, vil dette føre til økt boligpris. Dermed kan vi skrive at totalprisen blir en funksjon av mengden attributter og deres implisitte pris (Osland, 2011):

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

Z er vektoren og n er attributtet ved boligen. Den hedonistiske prisfunksjonen kan uttrykkes ved $P(Z)$, som viser at totalprisen, P , for en bolig er avhengig av sammensetningen av de ulike attributtene (Osland, 2011);

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

Den hedonistiske prisfunksjonen har som hovedformål å forklare samspillet mellom tilbud og etterspørsel i markedet for det heterogene godet.

Dette samspillet vil jeg forklare nærmere i de neste avsnittene. Innledningsvis vil jeg beskrive teorigrunnet for den hedonistiske metoden, deretter etterspørselssiden av markedet hvor budfunksjonen til husholdningene spiller en sentral rolle. Videre vil jeg forklare den andre siden av markedet, nemlig tilbudssiden, hvor offerfunksjonen til produsentene er sentral. Avslutningsvis vil jeg beskrive markedslikevekt mellom de to aktørene tidligere nevnt.

Teorigrunnet for den hedonistiske metoden

Den hedonistiske metoden forbindes med heterogene goder. Disse godene er sammensatt av flere attributter som gir nytte for konsumentene og som hver for seg har en implisitt pris. De første hedonistiske analysene bygger implisitt på Lancaster (1966) sin teori. Den hedonistiske metoden ble videreutviklet av Rosen i 1974, hvor han presenterte en teoretisk forklaring på sammenhengen mellom den hedonistiske prisfunksjonen og tilpasningen til enkeltaktørene på begge sider av markedet. Modellen til Rosen (1974) er statisk hvor den tar utgangspunkt i at et gode kan bli sett på som en vektor bestående av n objektive målte attributter: $Z = (Z_1, Z_2, \dots, Z_n)$

Dersom man ser på attributtene for godet, boligeiendom, så kan disse deles inn i to hovedgrupper: Attributter ved selve boligen og attributter ved lokaliseringen

av boligen. Eksempelvis er attributter tilknyttet selve boligen boligareal og innredning, mens avstand til sentrum, barnehage, og nabolag er eksempler på attributter ved lokaliseringen (Osland, 2011).

Rosens modell baserer seg på noen forutsetninger (Osland, 2011):

- Det er tilgang på et stort antall boliger på markedet, dette gjør at valgene mellom ulike attributter er kontinuerlige.
- Det finnes mange små aktører som hver for seg ikke har innflytelse på markedet og prisene.
- Tilpasningen skjer friksjonsfritt, med dette menes det at søke-, transaksjons- og flyttekostnader er ubetydelige.
- Alle aktørene i markedet har full informasjon om priser og attributtene til boligene.

Likevekt på etterspørselssiden av markedet (Nyttmaksimering)

På etterspørselssiden av boligmarkedet finner vi husholdningene. Disse ønsker å maksimere nytten gitt en ikke-lineær-budsjettrestriksjon. Nyttefunksjonen til husholdningen kan uttrykkes slik (Osland, 2011):

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

Den ikke-lineære budsjettrestriksjonen kan skrives slik:

$$Y_j = X + P(Z)$$

X er en vektor som representerer alle andre konsumvarer enn boligen. Denne vektoren settes lik 1. Y_j viser inntekten målt i antall enheter av X for husholdning j og α_j beskriver en vektor av parameterne som karakteriserer preferansene til husholdningen. Det antas at hver husholdning kun kjøper en bolig og den er et konsumgode. Videre antas det at nyttefunksjonen er strengt konkav. Den første-

og andreordensderiverte av prisfunksjonen $P(Z)$ eksisterer, men har ubestemt fortegn. Den marginale substitusjonsrate er i optimum mellom Z_i og X , den vil da være lik den partiellderiverte av prisfunksjonen med hensyn til de respektive boligattributter (Osland, 2001):

$$(1) \quad \frac{\frac{\partial U_j}{\partial Z_i}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_i}$$

Til høyre side for likhetstegnet i (1) finner vi de hedonistiske priser for attributt i , eller det som kalles marginale implisitte priser. Den angir helningen til prisfunksjonen i optimale mengder av Z_i .

Husholdningenes budfunksjon

Budfunksjonen til husholdningene er sentral på etterspørselssiden når man skal forklare markedslikevekten for heterogene goder. Budfunksjonen defineres av Osland i artikkelen som maksimal betalingsvillighet for ulike hustyper eller sammensetninger av attributtvektorer, under forutsetning av at nyttenivå og inntekt holdes konstant. Budfunksjonen uttrykkes slik: $\Theta_j = \Theta(Z, Y_j, U_j, \alpha_j)$.

Ut av denne budfunksjonen kommer det grafisk frem en indifferenskurve. Her vil det være mulig å studere alternative kombinasjoner av boligattributter i relasjon til subjektive priser og markedspriser, istedenfor opp mot et annet gode.

For å utlede budfunksjonen må man ta utgangspunkt i de optimale verdiene for boligvektoren og andre konsumvarer. Disse påføres en stjerne for å indikere optimale verdier, slik at symbolene blir henholdsvis Z^* og X^* . Tar man med inntekten Y_j , får vi funksjonen: $X^* = Y_j - P(Z^*)$. Settes dette videre inn i nyttefunksjonen får vi følgende: (Osland, 2001):

$$(2) \quad U_j = U(Z^*, Y_j - P(Z^*), \alpha_j) = U_j^*$$

Det antas at nyttenivået U^* er konstant og at inntekten er gitt. Videre vil det være naturlig å forutsette at den maksimale betalingsvilligheten Θ er lik den prisen

man faktisk betaler $P(Z^*)$. Dette fører til følgende uttrykk for nyttefunksjonen (Osland, 2001):

$$(3) U_j = U(Z^*, Y_j - P(Z^*), \alpha_j) = U^*_j = U(Z, Y_j - \Theta_j, \alpha_j)$$

Utrykket ovenfor viser en sammenheng mellom maksimal betalingsvillighet ved andre sammensetninger av boligattributter enn hva som er optimalt.

Husholdninger vil likevel oppfatte disse kombinasjonene som likeverdige. Med dette menes at for alle de sammensetningene av boligattributt som ikke er optimale, vil det beregnes en subjektiv pris slik at hele inntekten brukes opp og husholdningene forblir på det optimale nyttenivået. Budfunksjonen varierer med valgt inntekt- og nyttenivå kan dermed uttrykkes ved (Osland, 2001):

$$(4) \Theta = \Theta(Z, Y_j, U_j, \alpha_j)$$

Ved og implisitt derivere (3) med hensyn på boligvektoren Z får vi følgende:

$$(5) \quad \frac{\partial \Theta_j}{\partial Z_i} = \frac{\frac{\partial U_j}{\partial Z_i}}{\frac{\partial U_j}{\partial X}} > 0 \quad i = 1 \dots n$$

Maksimal betalingsvillighet for en partiell økning i et boligattributt er uttrykt ved (Osland, 2001):

$$\frac{\partial \Theta_j}{\partial Z_i}$$

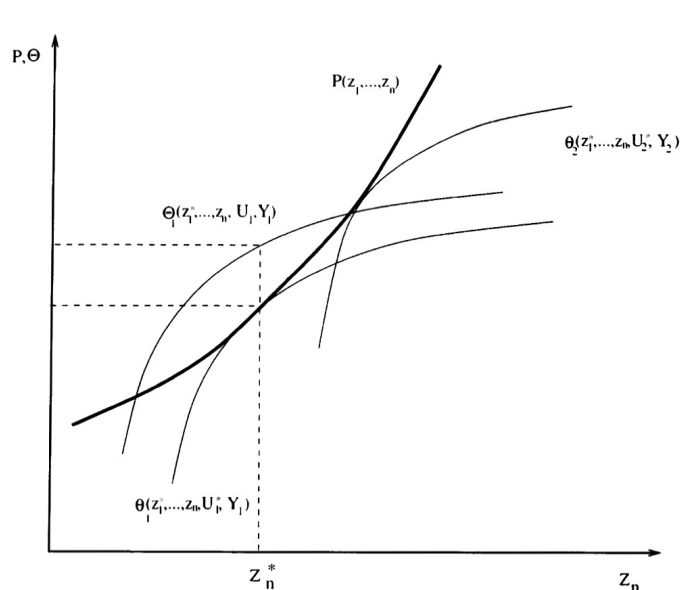
Nyttefunksjonen vil være strengt konkav dersom den andrederiverte er mindre enn null. Med dette menes at betalingsvilligheten er positiv, men avtakende for partielle økninger i boligattributter (Osland, 2001):

$$\frac{\partial^2 \Theta_j}{\partial Z^2 i} < 0$$

Nærmere forklart så viser denne at betalingsvilligheten er positiv, men avtar for partielle økninger i boligattributter (Osland, 2001).

Budfunksjonen vil gi en grafisk fremstilling av et sett indifferenskurver til hvert nyttenivå. Langs den vertikale akse måles kroner og det antas at konsumenten er optimalt tilpasset for alle attributter, med et unntak, Z_n . Dette kan eksempelvis være bruksareal. For en gitt husholdning har vi indifferenskurve Θ_1 . Man får et stigende nyttenivå dersom vi beveger oss nedover i diagrammet slik at vi får:

$$\frac{\partial^2 \Theta_j}{\partial^2 Z_i} < 0$$



Figur 3.4: Husholdningenes budfunksjoner (Osland, 2003)

Ovenfor er budfunksjonen vist grafisk med et sett indifferenskurver for hvert nyttenivå. Langs den vertikale akse måles kroner, mens ved den horisontale akse måles mengden attributter. Det antas at konsumenter er optimalt tilpasset i alle attributt utenom Z_n , som for eksempel kan være bruksareal. For en gitt husholdning har vi en indifferenskurve Θ_1 . Dersom vi beveger oss nedover i diagrammet vil nyttenivået stige slik at (Osland, 2001, Rosen 1974):

$$\frac{\partial \Theta_j}{\partial U_j} < 0$$

For å nytte maksimere vil husholdningen forsøke å finne den sammensetningen av boligattributter hvor man kommer på den lavest oppnåelige budkurven. En og enhver husholdning har ulike nyttefunksjoner og budfunksjoner, dette fremkommer av preferanseparameteren α . Lenger oppe langs prisfunksjonen i figuren finner vi husholdning 2. Denne husholdningen har budfunksjonen Θ_2 .

Husholdningen er representert ved å ha preferanser for eksempel for et større boligareal, dette kan være på grunn av stor familie eller høy inntekt.

Ved å trekke inn den eksogent gitte hedonistiske prisfunksjonen $P(Z)$ får vi likevekt for konsumentene. Den konvekse kurven i figuren ovenfor viser hvordan den hedonistiske prisfunksjonen stiger ved en partiell økning i boligareal. For å maksimere nytten vil husholdningene bevege seg langs den hedonistiske prisfunksjonen $P(Z)$, helt til den tangerer den lavest oppnåelige budfunksjonen (Osland, 2001). Betingelsen for likevekt på etterspørselssiden får man ved å sette (1) sammen med (5):

$$(6) \quad \frac{\partial \Theta_j}{\partial Z_n} = \frac{\frac{\partial U_j}{\partial Z_n}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_n} \quad j = 1, \dots, m$$

Ut ifra dette uttrykket blir nytten maksimert ved at implisitte bud eller marginal betalingsvillighet for den siste kvadratmeteren er lik den implisitte prisen på attributtet. Dette medfører identiske helninger på disse kurvene i optimum. I tillegg til tangeringsbetingelsen ovenfor, krever likevekt at $\Theta_j(Z^*, Y_j, U_j^*, \alpha_j) = P(Z)$ (Osland, 2001). Dette kan forklares, i følge Osland, med at husholdningene må betale det minste beløpet $P(Z)$ på markedet for en bolig med attributtvektoren Z , mens det maksimale beløpet husholdningen er villig til å betale er $\Theta_j(Z, Y_j, U_j, \alpha_j)$. For å underbygge dette, forutsettes det at en husholdning vil maksimere nytten ved å betale det laveste beløpet man må betale for en bolig med den optimale sammensetningen av attributter. Videre skal det tillegges at en annen husholdning kan ha andre preferanser eller inntekt som vil føre til at den er villig til å betale mer for denne boligtypen. Alt dette tatt i betraktning kan man konkludere med at den hedonistiske prisfunksjonen $P(Z)$ er et resultat av alle husholdningers budfunksjoner (Osland, 2001).

Likevekt på tilbudssiden av markedet (Produsentenes profittmaksimering)

Tilbudssiden består av tilbydere som består av mange små bedrifter, som vil tilpasse seg slik at profitten maksimeres. I det kortsiktige perspektiv innebærer

dette enten at bedriftene kan endre antall produserte enheter av en boligtype, eller endre sammensetningen av attributter. Videre skal vi se på situasjonen hvor sammensetningen av attributter kan endres. På lang sikt vil nye bedrifter komme opp og andre legges ned. Vi antar at hver bedrift spesialiserer seg og produserer en boligtype, gitt en sammensetning av attributter. Vi kan med det uttrykke profittfunksjonen til hver enkelt bedrift (Osland, 2001):

$$\pi = M * P(Z) - C(M, Z, \beta)$$

Hvor π er profitten, M er bedriftstilbud av boliger med en bestemt attributtvektor Z . Ved å multiplisere $P(Z)$ og M får vi en inntektsfunksjon som er ikke-lineær. Kostnadsfunksjonen C er en konveks stigende funksjon av antall boliger M . Vektor β er en skiftparametre som for eksempel representerer faktorpriser eller produksjonsteknologi for hver enkelt bedrift. Det skal tillegges at bedriftene oppfatter prisfunksjonen som gitt og uavhengig av antall boliger som bedriften produserer. Mange små bedrifter fører til stor variasjon i attributter, hvor de kan ha komparativ fortrinn på ulike boligtyper. Bedriftene ønsker å maksimere sin fortjeneste ved å produsere og tilby boliger M , samtidig som de ønsker å finne den sammensetningen av attributter som gir maksimal fortjeneste Z_i . Vi får dermed førsteordens betingelser for maksimal fortjeneste (Osland, 2001):

$$(7) \quad \frac{\partial P}{\partial Z_i} = \frac{\partial C}{\partial Z_i} \quad i = 1, \dots, n$$

$$(8) \quad P(Z) = \frac{\partial C}{\partial M}$$

Ligning (7) angir hvilke sammensetning av boligattributt som den bør velge, slik at den implisitte prisen for et gitt attributt er lik grensekostnaden per bolig ved en partiell økning i mengden boligattributter. Ligning (8) forteller at bedriftene bør produsere det antallet boliger som gjør at grenseinntekten, gitt ved prisen på boligen, er lik grensekostnaden for å produsere boliger (Osland, 2001).

Offerfunksjonen

På tilbudssiden har vi offerfunksjonen: $\Phi = (Z, \pi, \beta)$.

Funksjonen illustrerer den minste prisen produsentene er villig til å akseptere, slik at de kan tilby boliger med ulike attributter, til et konstant profittnivå og gitt det optimale antall boliger som produseres (Osland, 2001). Ved å ta utgangspunkt i de optimale verdiene Z^* , M^* og π^* kan offerfunksjonen nå utledes. Vi får dermed profittfunksjonen:

$$(9) \pi^* = M^* * P(Z^*) - C(M^*, Z^*, \beta)$$

Dersom vi lar profittnivået være konstant lik π^* og setter offerfunksjonen inn i profittfunksjonen får vi:

$$(10) \pi^* = M^* * \Phi(Z^*, \pi^*, \beta) - C(M^*, Z^*, \beta)$$

Ved å derivere ligning (10) med hensyn på M og Z_i ($i=1, \dots, n$) får vi førsteordensbetingelsene (Osland, 2001):

$$(11) \quad \Phi(Z^*, \pi^*, \beta) = \frac{\partial C}{\partial M}$$

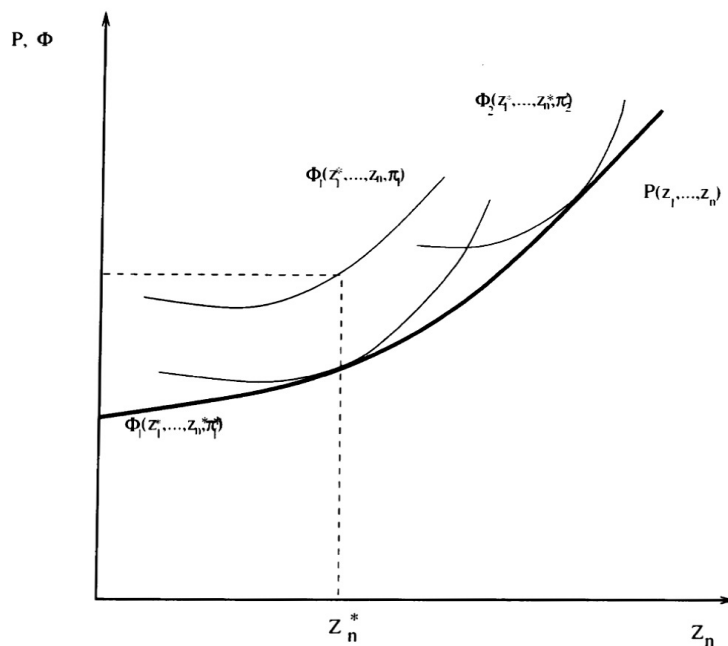
$$(12) \quad \frac{\partial \Phi}{\partial Z_i} = \frac{\partial C}{\frac{\partial Z_i}{M}} \quad i = 1, \dots, n$$

Løser vi ligning (11) med hensyn på M og deretter setter uttrykket inn i (12) forsvinner M . Profittfunksjonen viser nå forholdet mellom offerpriser og boligattributter (Osland, 2001):

$$(13) \quad \Phi = \Phi(Z, \pi^*, \beta)$$

I figuren nedenfor blir offerkurvene gitt en grafisk fremstilling, ved hjelp av et sett isoprofitkurver hvor det antas optimal tilpasning i alle attributter bortsett fra Z_n . Samtlige kurver er konvekse og profittnivået π stiger jo lenger man beveger seg oppover i diagrammet, som fører til $\frac{\partial \Phi}{\partial \pi} > 0$.

Mange produsenter vil ha ulik verdi på skiftparameteren β , og vil som følge av dette tilpasse seg ulike steder på prisfunksjonen. De produsentene som for eksempel tilbyr stort boligareal vil befinne seg lenger oppe langs prisfunksjonen.



Figur 3.5: Produsentenes offerfunksjon (Osland, 2001)

Betingelsen for likevekt på tilbudssiden får vi ved å sette ligning (7) og (12) sammen:

$$(14) \quad \frac{\partial \phi}{\partial Z_n} = \frac{\partial C}{\frac{\partial Z_n}{M}} = \frac{\partial P}{\partial Z_n}$$

I tillegg til dette kreves $\Phi(Z, \pi^*, \beta) = P(Z^*)$ for å oppnå likevekt. Offerprisen er da lik den eksogent gitte prisfunksjonen.

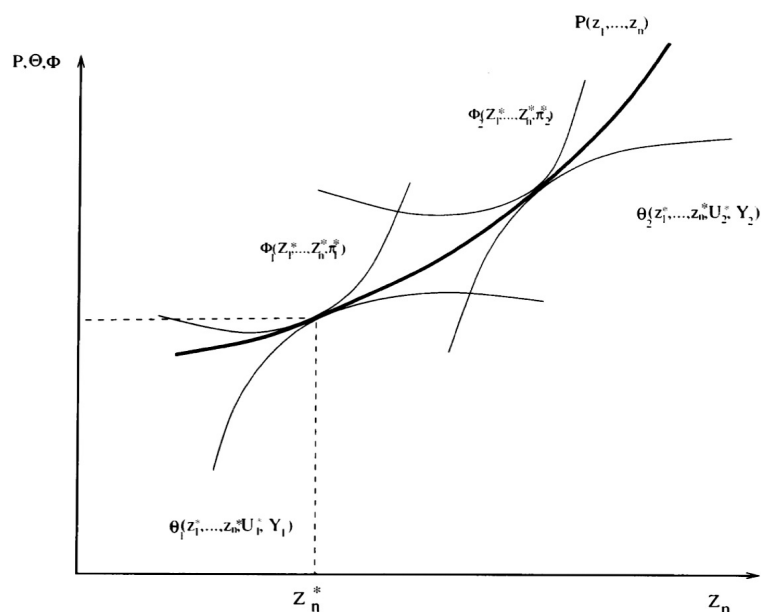
Markedslikevekt

For å oppnå markedslikevekt må husholdningens budfunksjon og produsentenes offerfunksjon tangere hverandre (Osland, 2001):

$$\frac{\partial \theta}{\partial Z_i} = \frac{\partial P}{\partial Z_i} = \frac{\partial C}{\frac{\partial Z_i}{M}} = \frac{\partial \phi}{\partial Z_i}$$

Denne fremstillingen viser at den hedonistiske prisfunksjonen er et samspill og likevekt mellom etterspørselen til husholdningen og tilbudet til produsentene.

Dette vises grafisk i figuren nedenfor:



Figur 3.6: Markedslikevekt (Osland, 2001)

I figuren er det to eksempler for likevekt mellom husholdningenes budfunksjon og produsentenes offerfunksjon. I virkeligheten vil det være flere skjæringspunkter, da det er mange ulike tilbud og husholdninger med ulik etterspørsel.

Videre vil jeg anvende teorien som er presentert til å belyse min problemstilling. For å kunne gjøre dette vil jeg å sette opp hypoteser jeg mener vil kunne bidra til dette.

3.5 Hypoteser

I denne oppgaven vil jeg undersøke om, og i så fall i hvilken grad husholdninger verdsetter ulike boligattributter i Stavanger og Haugesund. Ved hjelp av den hedonistiske metoden vil jeg estimere markedspris på ulike attributter i henholdsvis Stavanger og Haugesund. Videre vil jeg analysere om det er ulikheter i de estimerte prisene på attributtene i de to byene.

Det finnes en rekke empiriske studier av hvilke boligattributter som verdsettes av husholdninger. I følge artikkelen av Osland som har tatt utgangspunkt i den hedonistiske metoden, er det boligattributtene boligareal, tomteareal, boligalder, avstand til sentrum som har høyest forklaringskraft på boligprisen. Alonso-

Muth-Mills modellen ref. Kapittel 3.3.1 forklarte at pris på bolig og tomt varierer med avstand til sentrum.

De attributtene som inkluderes i min analyse antas å ha betydning for boligprisen. Jeg har benyttet meg av variabler som måler attributter ved selve boligene og tomtene samtidig. Videre har jeg benyttet meg av variabler som sier noe om attraktiviteten av området hvor boligene er lokalisert. Det antas at det vil være betalingsvillighet for disse attributtene. På bakgrunn av dette er de utvalgte variablene i min analyse:

- Boligareal.
- Avstand til sentrum.
- Avstand til sjø.
- Boligalder.
- Tomteareal.
- Boligtype.
- Må i bil til barnehage.

Boligareal

Det er rimelig å anta at jo større boligareal, jo høyere er boligprisen. Dette vil fortsette helt til boligarealet når et visst punkt. Etter det punktet er det rimelig å anta at boligprisen er avtagende. Det er interessant å se på boligarealets påvirkning på boligprisen. Boligareal sin innflytelse på boligprisen vil jeg anta at varierer fra by til by. Dermed vil det være av interesse undersøke hvor stor betydning boligarealet har for boligprisen i Haugesund i forhold til Stavanger.

Hypotesen jeg ønsker å teste er dermed:

H_0 : Boligarealet har lik betydning for boligprisen i Haugesund og Stavanger.

H_1 : Boligarealet har i Stavanger større betydning for boligprisen enn i Haugesund.

Avstand til sentrum

Denne hypotesen er basert på Alonso-Muth-Mills modellen som jeg tidligere presenterte som sa at jo kortere avstand til sentrum, jo høyere boligpris. Beliggenheten i forhold til sentrum er et boligattributt som ofte kan ha sterk forklaring for boligprisen. Stavanger og Haugesund har ulik bystørrelse og bygrense og vil derfor ha ulike avstander til sentrum. Innledningsvis i oppgaven nevnte jeg at preferanse i Stavanger i forhold avstand til sentrum ikke var like klare som andre steder.

Jeg ønsker derfor å teste:

H₂0: Avstand til sentrum har lik betydning for boligprisen i Haugesund og Stavanger.

H₂1: Avstand til sentrum har større betydning for boligprisen i Haugesund enn Stavanger.

Boligalder

Jeg vil anta at alderen på en bolig i de fleste tilfeller har betydning for boligprisen. Ofte kan alderen på en bolig gi en god indikasjon på boligstandarden. Eldre boliger bærer ofte preg av slitasje som kan bidra til en lavere verdi i forhold til nyere boliger. Oppussing kan derfor være nødvendig for en eventuell kjøper, noe som kan redusere betalingsvilligheten. Med dette tatt i betraktning er det rimelig å anta at jo høyere alderen er, jo lavere vil boligprisen være. Betalingsvilligheten for nyere eller eldre boliger kan variere mellom byer.

Jeg ønsker derfor å teste:

H₃0: Boligalder i Haugesund har lik betydning for boligprisen i Haugesund og Stavanger.

H₃1: Boligalder i Haugesund har større betydning for boligprisen enn i Stavanger.

Tomteareal til selveide eneboliger

Det er rimelig å anta at tomtestørrelser har betydning for boligprisen: Jo større tomt, jo høyere boligpris. Dette stemmer derimot ikke i alle tilfeller. Folk har ulike preferanser og en stor tomt med en mindre gunstig beliggenhet vil ikke

nødvendigvis føre til økt boligpris. Den topografiske utformingen på tomta vil også kunne spille inn på boligprisen. Ekstra arbeid på en tomt med for eksempel fjellknauser kan redusere betalingsvilligheten til en kjøper. Videre er det naturlig å anta at betalingsvilligheten for tomtestørrelse vil variere mellom de ulike boligtypene og eieformene. Dette tatt i betraktning vil jeg anta at tomtestørrelsen til en selveid enebolig har større påvirkning på boligprisen enn for eksempel ved en borettslagsleilighet hvor tomtearealene ofte er til felles bruk. Jeg ønsker å undersøke forskjellene mellom tomtearealets påvirkning på pris på enebolig i de to byene.

Med dette ønsker jeg å teste:

H₄₀: Tomtearealet til selveide eneboliger i Haugesund har lik betydning for boligprisen i Haugesund og Stavanger.

H₄₁: Tomtearealet til selveide enebolig i Haugesund har større betydning for boligprisen enn i Stavanger.

Boligtype

Både Stavanger og Haugesund har et variert marked for boligtyper. Boligtypene som er inkludert i denne analysen er eneboliger, rekkehus, leiligheter og tomannsbolig. Det vil alltid være forskjellige preferanser eller betalingsvillighet for de ulike typene. Geografisk plassering vil gjerne ha betydning ved valg av boligtype. I sentrum vil leilighetskomplekser være mer naturlig enn store eneboliger da tomtearealer i sentrum ofte er dyre. Andre kan ønske å ikke bo så tett og må da gjerne bevege seg ut av sentrum. Familiestørrelse og alder på kjøper må også tas i betraktning når det kommer til valg av boligtype. Andre forhold som kan påvirke boligprisen er eieform og fellesgjeld. Disse er derfor tatt med i analysen, men det blir ikke laget hypoteser av de.

Hypotesen jeg ønsker å teste er:

H₅₀: Boligtype har ikke betydning for boligprisen.

H₅₁: Boligtype har betydning for boligprisen.

Avstand til sjø

Både Haugesund og Stavanger er lokalisert på vestlandskysten og har derfor forholdsvis nær tilgang til sjøen. Innad i byene vil avstand til sjø likevel variere.

Mange vil verdsette en bolig med utsikt til havet eller nærhet til kysten. Dette vil derfor kunne å betydning for prisen på bolig.

Hypotesen jeg dermed ønsker å teste er:

H₃₀: Avstand til sjø har ikke betydning for boligprisen.

H₃₁: Avstand til sjø har betydning for boligprisen.

Må i bil for å komme til barnehage

Som jeg nevnte innledningsvis i oppgaven er folk bevisste på hvor nærmeste barnehage er lokalisert i forhold til den aktuelle boligen. Bevissthetsgraden vil likevel variere mellom de ulike aldersgruppene. Jeg antar at nærhet til barnehage vil ha betydning for boligprisen. Jeg ønsker å undersøke om det å slippe å kjøre til nærmeste barnehage vil gi et positivt utslag på boligprisen. Avgjørelsen om man slipper å kjøre vil jeg anta til å være en avstand på under 1 km.

Hypotesen blir dermed:

H₇₀: Må i bil for å komme til barnehage har ikke betydning for boligprisen.

H₇₁: Må i bil for å komme til barnehage har betydning for boligprisen

Med hensyn til hypotesene kan jeg sette opp en foreløpig forenklet funksjon for boligprisen. Boligprisen er derfor en funksjon av:

Boligpris = f (boligareal, avstand til sentrum, avstand til sjø, boligalder, tomteareal, boligtype, må i bil for å komme til barnehage).

Gitt min problemstilling er det nødvendig å innhente relevant data. Hvordan denne prosessen har foregått vil jeg gjøre rede for i neste kapittel.

4.0 Datainnsamling og beskrivelse av datamaterialet

I dette kapittelet vil jeg redegjøre for hvordan jeg samlet inn datamaterialet. Videre vil jeg presentere datamaterialet på en oversiktlig måte. Ved utgangspunkt i problemstillingen vil det være naturlig å foreta en kvantitativ undersøkelse. Kvantitativ metode er best egnet når man ønsker å foreta målinger av konsepter som direkte eller indirekte gir numeriske verdier. Disse verdiene kan enkelt bli behandlet av statistikkprogrammer. Metoden gjør det enklere å analysere flere enheter samtidig (Zikmund et al., 2013). Jeg vil benytte statistikkprogrammet Stata for gjennomføring av de kvantitative beregningene.

4.1 Datainnsamling

Ved bruk av kvantitativ metode er det nødvendig med et stort og bredt datamateriale for å kunne få et mest mulig grundig og nøyaktig resultat. Ved innsamlingen av datamaterialet har jeg benyttet sekundærdata. Innhenting av sekundærdata vil si at man baserer seg på noen andre sin datainnsamling.

Fordelene med denne typen datainnsamling kontra primærdata er at sekundærdata ofte er lett tilgjengelig og budsjettvennlig. Siden dette også er en tidsbegrenset oppgave vil benyttelse av sekundærdata bidra til spart tid. Jeg vil heller ikke ha muligheten til å påvirke respondentenes svar. Det man likevel må være oppmerksom på er at sekundærdata er opprinnelig skreddersydd til andre undersøkelser og formål. Det er derfor viktig å være kritisk og svært nøyaktig ved bruk av andres data (Zikmund et al., 2013).

En eiendomsdatabase utviklet av selskapet Eiendomsverdi ble benyttet for å samle inn nødvendig data for analysen. Denne eiendomsdatabasen fikk jeg tilgang til igjennom eiendomsverdi.no av Handelshøyskolen i Kristiansand. Databasen registrerer og overvåker all eiendomsaktivitet i Norge (Eiendomsverdi.no). Eiendomsverdi.no har et kraftig verktøy med mange filtreringsmuligheter. Her finnes tilgjengelig informasjon og opplysninger om adresse, omsetningspris, salgsdato, BOA/Primær-rom, BTA, eierform, tomteareal, prisantydning, registreringsdato, fellesgjeld, byggeår og boligtype.

Ved innhenting av data av avstander til sjøen, sentrum og barnehager har jeg brukt kartdatabaser utviklet av Google Maps, Gulesider og erikbostad.no.

4.2 Utvalgsprosess

For å få en god og representativ analyse om boligmarkedene i Haugesund og Stavanger trengte jeg et bredt datagrunnlag. Jeg valgte derfor å samle inn informasjon over en fem års periode, 1. januar, 2009 til 31. desember, 2013. Alle opplysninger om solgte og registrerte boliger i begge byene over denne perioden ble hentet fra eiendomsverdi.no.

I Haugesund kommune var det totalt 22 postnumre som hadde solgte og registrerte boliger de siste fem årene. Videre ble utvalget redusert til 21 postnumre da jeg valgte å utelate postnummeret tilhørende Røvær. Årsaken til dette er at det er en øy, lokalisert utenfor Haugesund og som derfor kunne bidra til å gi et feilaktig bilde av min analyse.

Stavanger hadde totalt 42 postnumre som tilhørte kommunen. Av samme årsak som ved Røvær, utelot jeg Vassøy i Stavanger og endte dermed opp med totalt 41 postnumre med solgte og registrerte boliger over min utvalgte tidsperiode.

Gjennom fem års perioden fikk jeg et representativt utvalg på 4667 boliger i Haugesund og 16596 i Stavanger. Det skal legges til at dette utvalget ble justert i etterkant når jeg foretok en rensing av dataene, ettersom det viste seg å være mangel på opplysninger på noen av boligene som er vesentlig for min analyse. I prinsippet hadde jeg muligheten på hver annonse å finne de manglende opplysninger, men på grunn av et omfattende datasett og at dette er en tidsbegrenset oppgave valgte jeg ikke å gjøre det. Jeg valgte dermed å utelate de boligene som manglet opplysningene jeg så som vesentlig for min analyse. Jeg endte dermed opp med et utvalg på 3 547 boliger i Haugesund og 13 565 boliger i Stavanger. En grundigere gjennomgang av datarensingen er presentert senere i oppgaven.

Analysen er rettet mot "vanlige" boliger og dermed er fritidsboliger utelatt da det er et annet markedssegment og vil derfor ikke vil være representativt for min analyse.

Det er viktig å merke seg at det er kun salg av boliger gjennom eiendomsmegler som er tatt med i analysen. Boligsalg uten eiendomsmegler og boliger som gaver, arv eller lignende vil ikke fanges opp i analysen. Det er både ulemper og fordeler med dette. Fordelene er at vi får mest sannsynlig korrekte markedssalg, mens ulempen vil være at vi mister noen omsetninger som også kunne hatt reelle markedsomsetninger.

Eieformene selveier og borettslag/andelseier er begge inkludert i analysen. Som følge av dette, var det nødvendig å ta med fellesgjeld i analysen da det antas at fellesgjeld kan ha en påvirkning på boligprisen, spesielt ved boliger som er en del av et borettslag.

For å kunne belyse attributter som avstand til sjøen, sentrum og barnehage brukte jeg internettsiden erikbolstad.no hvor man får tilgang til koordinatene og lokaliseringen til de ulike postnumrene i Haugesund og Stavanger. Videre ble gulesider.no brukt til å måle opp avstanden fra alle postnumrene til nærmeste sjøpunkt og sentrum. Sentrumpunktene for Haugesund og Stavanger er henholdsvis Haraldsgata og Kirkegata. For å måle avstanden til barnehage brukte jeg kombinasjonen av erikbolstad.no, haugesund-kommune.no, stavanger-kommune.no og internettværktøyet Google Maps for å lokalisere barnehagene i byene.

4.3 Presentasjon og operasjonalisering av variablene

En variabel kan defineres som hva som helst som varierer eller endres i et tidsrom. Variabler kan uttrykke forskjeller i verdi, vanligvis i form av størrelse eller styrke eller de kan vise en retning. Variabler kan deles inn i blant annet avhengige og uavhengige variabler (Zikmund et al., 2013).

Videre kan vi skille mellom kontinuerlige og ikke kontinuerlige variabler. Kontinuerlige variabler kan ha hvilken som helst verdi. Eksempler på kontinuerlige variabler er boligareal, avstand til sentrum, avstand til sjø, boligpris etc. Ikke kontinuerlige variabler, også kalt dummyvariabler, har en verdi som er bestemt på forhånd og kan ha to mulig utfall. Eksempel på en dummyvariabel er eierform, som for eksempel selveier og borettslag. I en

regresjonsanalyse vil det være vil man derfor å kode disse alternativene med 0 og 1, hvor eksempelvis 1 representerer at det er selveier dersom det dreier seg om denne type eierform og 0 hvis det ikke er selveier (Zikmund et al., 2013).

4.3.1 Den avhengige variabelen

Den avhengige variabelen er en variabel som er forklart og/eller spådd av flere andre variabler (Zikmund et al., 2013).

- Boligpris

I min undersøkelse vil den avhengige variabelen være *boligpris*. Jeg ønsker å se på hvordan og hvor mye boligprisen blir påvirket av de ulike uavhengige variablene. Alle omsetningsprisene til boligene i Haugesund og Stavanger er hentet fra databasen til Eiendomsverdi.no.

4.3.2 Uavhengige variabler

Den uavhengige variabelen forventes å påvirke den avhengige variabelen på en eller annen måte (Zikmund et al., 2013). Uavhengige variabler kan også kalles forklaringsvariabler, da de antas å ha en forklarende effekt på den avhengige variabelen. Hvor stor effekt de vil ha på boligprisen vil følgende variere i ulik grad. I denne analysen skal jeg undersøke hvordan og hvor mye en rekke uavhengige variabler påvirker boligprisen. Videre vil jeg redegjør for hvordan datainnsamlingen har foregått:

- Boligareal

Boligareal er en uavhengig variabel og opplysningene er hentet fra eiendomsverdi.no En bolig kan inndeles i en primærdel og en sekundærdel. P-rom er bruksareal av primærdelen og måles til innsiden av omsluttende vegger for primærdelen. Det resterende arealet kalles s-rom som er bruksareal av sekundærdel. Typiske eksempler på sekundær rom er boder, garasje og tekniske rom. Jeg valgte å ta utgangspunkt i primærrom for min analyse, som er eksempelvis oppholdsrom, soverom, kjøkken, gang og bad²⁰ (Norges Takseringsforbund, 2007). Opplysningene om boligareal er registrert i

²⁰ <http://www.boligtakst.as/arealbegrep.html>

kvadratmeter og er hentet fra eiendomsverdi.no. De boligene som manglet opplysninger om boligareal ble utelatt fra datasettet.

- Avstand til Sentrum

Dette er en uavhengig variabel som antas å ha høy forklaringskraft på boligprisen ref. Alonso-Muth-Mills modellen tidligere presentert. Ved hjelp av internettsiden erikbolstad.no kunne jeg lokalisere og finne avgrensningene til alle postnumrene som jeg hadde hentet fra eiendomsverdi.no. Jeg definerte et sentralt punkt i hvert enkelt postnummer og brukte videre gulesider.no sitt kartverktøy til å måle avstanden inn til sentrum. Oppmålingene ble gjort i kilometer i bil. Jeg opererte med to ulike sentrum, henholdsvis Haraldsgata i Haugesund og Kirkegata i Stavanger. Optimalt sett skulle jeg målt avstanden fra hver enkelt husstand, men på grunn av svært stort utvalg ville dette blitt for tidkrevende.

- Avstand til sjø

Avstanden til sjø er en uavhengig variabel som antas å påvirke den avhengige variabelen. Mange verdsetter å ha nærhet til sjøen og sjøutsikt. Opplysninger om boligene har sjøutsikt eller om boligen ligger skjermet bak for eksempel et fjell som dekker utsikten har jeg ikke og er derfor følgende ikke tatt hensyn til. For å måle avstanden til sjøen brukte jeg samme metode som jeg gjorde for avstanden til sentrum. Jeg definerte et sentralt punkt i hvert enkelt postnummer. Forskjellen her er at jeg målte luftlinja til nærmeste sjøpunkt i antall kilometer. Optimalt sett skulle jeg målt avstanden fra hver enkelt husstand, men på grunn av omfattende datamateriale ville dette blitt for tidskrevende.

- Boligalder

Informasjonen om byggeåret og salgsåret til boligene er hentet fra eiendomsverdi.no og er en uavhengig variabel. For å finne alderen på boligen trakk jeg fra byggeåret fra årstallet i skrivende stund, nemlig 2014. Disse beregningene ble foretatt i Microsoft Excel. Det er viktig å merke seg at en eventuell oppussing eller utbygging av boligene etter at boligene ble bygget, ikke er tatt hensyn til i analysen.

- Tomteareal til selveide eneboliger

Tomteareal er en uavhengig variabel. Opplysningene om tomtenes areal ble også hentet fra Eiendomsverdi.no og er registrert i kvadratmeter. De boligene som ikke hadde opplysninger om tomteareal ble fjernet fra datasettet.

Tomteareal er en variabel som gjerne ikke er like interessant for alle boligtypene og alle eieformer. Derfor introduserte jeg en produktvariabel som beskriver kun tomtearealet til eneboliger som er selveid. Produktvariabelen har jeg gitt navnet "tomtenebolig". Grunnen til at jeg kun vil fokusere på tomtestørrelsen til selveier-eneboliger er fordi tomtestørrelsen ikke vil ha like stor betydning for en leilighetskjøper, som for en som kjøper en enebolig. Rekkehus, tomannsboliger, eneboliger og leiligheter som er deler av et borettslag har ofte store felles tomteareal og det ville derfor vært vanskelig å gi et riktig bilde av analysen.

- Fellesgjeld

Opplysningene om fellesgjeld er hentet fra eiendomsverdi.no

Om en bolig har fellesgjeld eller ikke kan påvirke totalprisen på en bolig. Høy fellesgjeld vil ofte kunne gi en gunstig omsetningspris. Fellesgjeld blir derfor tatt med i analysen som en uavhengig variabel fordi det antas at denne har påvirkning på boligprisen. Dette støttes av Robertsen og Theisen (2011) som har gjennom sin forskning funnet ut at dersom fellesgjelden øker med en krone, reduseres prisen med 90 øre.

4.3.3 Dummyvariabler

I en regresjonsanalyse hvor man har dummyvariabler, koder man disse med alternative 0 og 1, hvor eksempelvis 1 representerer at det er enebolig dersom det er det dreier seg om denne type bolig, og 0 hvis det ikke er en enebolig. (Zikmund et al., 2013).

- Boligtype

Opplysningene om boligtype er hentet fra eiendomsverdi.no.

Boligtype har som hensikt å skille mellom ulike boligtyper. De boligtypene som er inkludert er enebolig, leilighet, rekkehus og tomannsbolig. Jeg har gitt hver boligtype en verdi fra 1-4 i analysen i Stata.

- Må i bil for å komme til barnehage

Jeg ønsker å undersøke om nærhet til barnehage har påvirkning på boligprisen. Operasjonaliseringen av denne variabelen ble gjort ved å kartlegge hvor mange barnehager som fantes i byene og hvor de var lokalisert. Jeg lokaliserte barnehagene i Stavanger og Haugesund ved hjelp av søkemotoren til Googlemaps.no. Jeg dobbeltsjekket med barnehageoversikten på Stavanger- og Haugesund Kommune sine hjemmesider at dette stemte overens. Videre definerte jeg et sentralt punkt i hvert enkelt postnummer hentet fra erikbolstand.no. Ved hjelp av dette kunne jeg si hvorvidt man måtte ta bilen fatt for å komme til nærmeste barnehage. Avgjørelsen om man måtte i bil eller ikke for å komme til nærmeste barnehage ble gjort ved et 1 km skille. Dersom det var over 1 km til nærmeste til barnehage, indikerer det at man sannsynligvis må i bil for å komme til barnehage, mens under 1 km viser at det trolig er gåavstand. I flere tilfeller kan det antas at man må kjøre forbi den nærmeste barnehagen da den ikke har tilgjengelig kapasitet, dette er ikke hensyntatt i analysen. Variabelen er en dummyvariabel, hvor 1 viser at man trolig må i bil for å komme til bil, mens verdien 0 indikerer at man slipper å kjøre for å komme til barnehage.

- Eieform

Opplysningene om eieform ble hentet fra eiendomsverdi.no. Jeg valgte å inkludere eieformene selveier og borettslag/andelseier. Eierforholdet kan antas å ha betydning for boligprisen, men er ikke hovedfokuset i oppgaven.

- Salgsår

Det kan være høy aktivitet i boligmarkedet ett år, mens det i andre år kan være preg av at kjøpere sitter på gjerdet i frykt for prisfall. Gode år i boligmarkedet kan slå positivt ut på boligprisen, mens dårlige år kan slå negativt ut. Jeg utarbeidet dummyvariabler for salgsårene 2009-2013. Når en bolig er solgt er ikke hovedfokuset i oppgaven, men er tatt med da salgstidspunkt kan ha betydning for boligprisen. Opplysningene fra når boligene ble solgt er hentet i fra eiendomsverdi.no.

4.4 Koding

Etter å ha samlet datamaterialet fra eiendomsverdi.no inn i Microsoft Excel hentet jeg det videre inn i statistikkprogrammet Stata. Ved behandling og analyse av datamaterialet i Stata tok jeg i bruk koder av hensiktsmessige årsaker.

Kodene bidro til et ryddig og et mer oversiktlig bilde av datamaterialet som ble bearbeidet. I tabellen under er variablene med tilhørende koder presentert.

Tabell 4.1: Koding

Variabel	Koding
Boligpris	I hele kroner
Fellesgjeld	I hele kroner
Boligareal	I hele kvadratmeter
Tomtestørrelse	I hele kvadratmeter
Avstand til sentrum	I kilometer
Avstand til sjøen	I kilometer
Boligalder	Antall år (2014 minus byggeåret)
Må i bil til barnehage	1 hvis må i bil, 0 hvis ikke
Eieform:	
Selveier	1 hvis selveier, 0 hvis ikke
Borettslag	1 hvis borettslag, 0 hvis ikke
Boligtype:	
Leilighet	1 hvis leilighet, 0 hvis ikke
Enebolig	1 hvis enebolig, 0 hvis ikke
Rekkehus	1 hvis rekkehus, 0 hvis ikke
Tomannsbolig	1 hvis tomannsbolig, 0 hvis ikke
Salgsår:	
Salgsår 2009	1 hvis salgsår 2009, 0 hvis ikke
Salgsår 2010	1 hvis salgsår 2010, 0 hvis ikke
Salgsår 2011	1 hvis salgsår 2011, 0 hvis ikke
Salgsår 2012	1 hvis salgsår 2012, 0 hvis ikke
Salgsår 2013	1 hvis salgsår 2013, 0 hvis ikke

4.5 Datarensing

Rensing av dataene som er samlet inn er veldig viktig for troverdigheten av analysen. Ved innsamling av store mengder data, vil det ofte være manglende eller ufullstendige opplysninger i dataene. I mitt datasett var noen observasjoner som manglet opplysninger om boligpris, boligareal, byggeår og tomteareal. Dersom jeg ikke hadde tatt hensyn til de manglende opplysningene ville de kunne påvirke resultatene og troverdigheten av resultatene ville blitt lavere enn ønsket.

Det finnes forskjellige måter å løse utfordringene med mangelfulle opplysninger. I mitt tilfelle valgte jeg å utelate de boligene som hadde manglende opplysninger om boligpris og de som har pris lik 0. De boligene som ikke hadde oppgitt

byggeår eller byggeår lik 0 valgte jeg også å fjerne. Videre valgte jeg å se bort i fra de observasjonene som manglet boligareal og tomteareal. Til slutt valgte jeg å fjerne boliger som hadde fellesgjeld under 1000 kroner. Dette ble gjort for å unngå eventuelle innlastningsfeil.

Før datarensningen ble gjort hadde jeg et utgangspunkt på 4 667 og 16 596 observasjoner for henholdsvis Haugesund og Stavanger. Etter datarensningen var unnagjort stod jeg igjen med 3 547 observasjoner i Haugesund og 13 565 i Stavanger. I tabell 4.2 er det gitt en oversikt over gjennomføringen av datarensningen.

Tabell 4.2: Oversikt over gjennomføringen av datarensning

	Haugesund	Stavanger	Totalt
Observasjoner totalt før datarensning	4667	16596	21263
- Sletter observasjoner med manglende pris	166	110	276
- Sletter observasjoner med manglende boligareal	465	1997	2462
- Sletter observasjoner med manglende tomteareal	316	867	1183
- Sletter observasjoner med manglende byggeår	173	30	203
- Sletter observasjoner med fellesgjeld <1000	0	27	27
Observasjoner totalt etter datarensning	3547	13565	17112

Videre skal den deskriptive statistikken presenteres før jeg ser på de ulike variablene nærmere og korrelasjonen mellom disse.

4.6 Deskriptiv statistikk

Etter at datamaterialet er kodet og rensset, kan datamaterialet presenteres ved deskriptiv statistikk. Deskriptiv statistikk er også omtalt som beskrivende statistikk, som har som mål å summere og beskrive datamaterialet på en enkel og forståelig måte. Den deskriptive statistikken blir presentert ved hjelp av figurer og tabeller som bidrar til ryddighet og et mer oversiktlig bilde av datamaterialet. Dette vil igjen gjøre den lettere å tolke og analysere dataene. (Zikmund et al., 2013). I tabellen under er den deskriptive statistikken illustrert med alle ulike variablene jeg har med i undersøkelsen:

Tabell 4.3: Deskriptiv statistikk, Haugesund & Stavanger

Haugesund					
Variabel	N	Gjennomsnitt	Standardavvik	Minimum	Maksimum
Boligpris	3547	2005544	938994.2	250000	7500000
Fellesgjeld	905	213564.4	213564.4	1068	7000000
Boligareal	3547	114.7296	56.5563	24	420
Tomtenebolig	1271	738.8293	428.71	25	7224
Boligalder	3547	45.36482	29.13305	1	174
Avstand til sentrum	3547	2.62808	1.543833	0.5	9.6
Avstand til sjøen	3547	.8647871	.6284503	0.2	3
Må i bil til barnehage	3547	.0969834	.2959768	0	1
Selveier	3547	.7454187	.435687	0	1
Borettslag	3547	.2545813	.435687	0	1
Leilighet	3547	.4324782	.4954897	0	1
Enebolig	3547	.3645334	.4813669	0	1
Rekkehus	3547	.0919087	.2889377	0	1
Tomannsbolig	3547	.1110798	.3142752	0	1
Salgsår09	3547	.1807161	.384837	0	1
Salgsår10	3547	.2004511	.4003944	0	1
Salgsår11	3547	.2207499	.4148107	0	1
Salgsår12	3547	.2049619	.4037307	0	1
Salgsår13	3547	.1931209	.3948027	0	1
Stavanger					
Variabel	N	Gjennomsnitt	Standardavvik	Minimum	Maksimum
Boligpris	13565	3333433	121582.7	45000	19500000
Fellesgjeld	4090	398426.4	592609	1014	5580000
Boligareal	13565	104.219	55.86838	18	478
Tomtenebolig	2705	580.8998	1159.432	40	55014
Boligalder	13565	47.0895	36.06517	1	246
Avstand til sentrum	13565	4.33366	2.933602	0.4	22
Avstand til sjøen	13565	.7998305	.6443428	0	2.4
Må i bil til barnehage	13565	.0505713	.2191287	0	1
Selveier	13565	.7023959	.4572213	0	1
Borettslag	13565	.2976041	.4572213	0	1
Leilighet	13565	.4911906	.4999408	0	1
Enebolig	13565	.2075931	.4072368	0	1
Rekkehus	13565	.1223738	.2075931	0	1
Tomannsbolig	13565	.1788426	.3832346	0	1
Salgsår09	13565	.1906377	.3928184	0	1
Salgsår10	13565	.2030225	.4022639	0	1
Salgsår11	13565	.2047917	.4035642	0	1
Salgsår12	13565	.209436	.4069211	0	1
Salgsår13	13565	.1921121	.3939752	0	1

I første kolonne i tabellen er det gitt en oversikt over de ulike variablene. N angir antall observasjoner som var med i undersøkelsen. Minimum- og maksimumsverdiene angir den laveste og høyeste verdien som ble undersøkt. Gjennomsnittet viser den gjennomsnittlige verdien til hver enkelt variabel som ble observert igjennom tidsperioden 2009-2013. Dette ble gjort ved å summere alle verdiene til den enkelte variabel og dividere på antall respektive observasjoner.

Videre er standardavviket illustrert. Standardavviket har som mål å vise spredningen eller variabiliteten fra gjennomsnittet. Hvilket betyr at dersom vi har et høyt standardavvik er spredningen fra gjennomsnittet høyt og omvendt. (Zikmund et al., 2013).

4.7 Deskriptiv statistikk av hver enkelt variabel

Etter å ha presentert den deskriptive statistikken med hensikt å gi et klarere oversiktsbilde av datasettet er det naturlig å gå nærmere inn på hver enkelt variabel, og det er nettopp det jeg skal gjøre videre i dette avsnittet. Variablene er presentert med tilhørende tabeller og diagrammer.

4.7.1 Boligpris

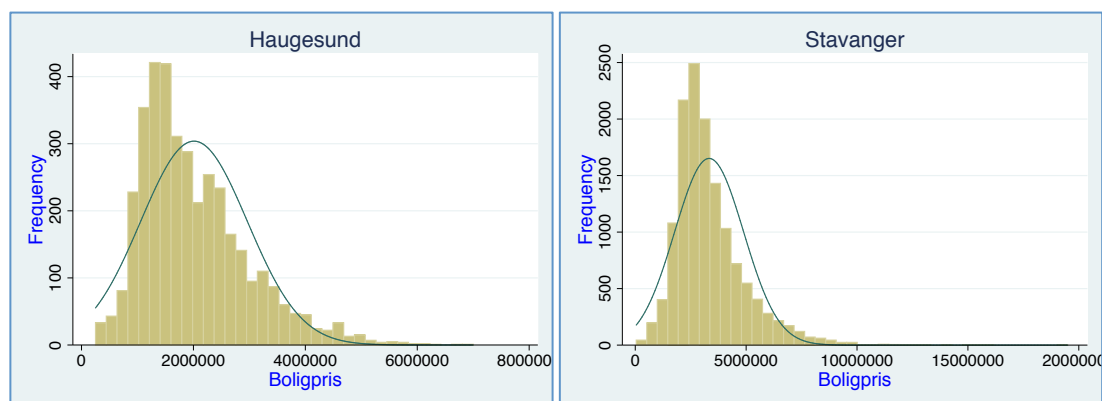
Den første variabelen som skal presenteres er den avhengige variabelen, boligpris. I denne oppgaven ønskes det å se på forklaringskraften fra de forskjellige attributtene på boligprisen.

Totalt var det 3 547 registrerte solgte boliger i Haugesund og 13 565 i Stavanger. Gjennomsnittlig boligpris var i overkant av 1,3 millioner høyere i Stavanger i forhold til Haugesund. Standardavviket indikerer at de fleste boligprisene i Haugesund ligger mellom 1 000 000 og 3 000 000 kroner, som indikerer en relativ stor spredning rundt gjennomsnittet. I de fleste tilfeller ligger boligprisene i Stavanger mellom 3 200 000 og 3 400 000 kroner, som indikerer en relativt lav spredning rundt gjennomsnittet. Den høyeste observerte prisen i Stavanger var over dobbelt så høy i Stavanger enn hva den var i Haugesund.

Tabell 4.4: Boligpris for Haugesund og Stavanger (01.01.09-31.12.13)

Salgspris	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	3547	2005544	938994.2	250000	7500000
Stavanger	13565	3333433	121582.7	45000	19500000

I figur 4.1 er det illustrert et histogram for begge byene med boligpris. I begge figurene har jeg illustrert normalfordelingen og ved å se på dens toppunkt kan vi se gjennomsnittsprisen på en bolig. Videre er det mulig å se at det er en del boliger i Haugesund som er solgt for underkant av 2 millioner, noe som vil dra ned gjennomsnittet. I Stavanger har man flere boliger som er solgt for over 4 millioner, som vil bidra til å øke gjennomsnittlig boligpris.



Figur 4.1: Fordeling av boliger etter boligpris i perioden 2009-2013

4.7.2 Boligtype

Tabellen under gir en oversikt over de forskjellige boligtypene i de respektive byene. Både i Haugesund og Stavanger er det leiligheter som er registrert mest omsatt de 5 siste årene. Prosentvis har det vært omsatt flest leiligheter i begge byer, noe mer i Stavanger enn i Haugesund. Videre har det prosentvis vært omsatt flere eneboliger i Haugesund enn hva som er tilfellet i Stavanger. De to siste boligtypene, rekkehus og tomannsbolig har begge en høyere prosentandel i Stavanger enn i Haugesund.

Tabell 4.5: Boligtype i Haugesund og Stavanger i antall og prosent

Boligtype:	Haugesund		Stavanger	
	Frekvens	Prosent	Frekvens	Prosent
Enebolig	1293	36,5 %	2816	20,76 %
Leilighet	1534	43,3 %	6663	49,12 %
Rekkehus	326	9,2 %	1660	12,24 %
Tomannsbolig	394	11,1 %	2426	17,88 %
Sum	3547	100 %	13565	100 %

4.7.3 Eieform

Denne variabelen viser hvilken eieform de ulike boligene har. Det skilles her ved mellom selveier og borettslag. I tabellen under viser man at selveier er den klart mest vanlige eieformen begge steder, med en noe større borettslagandel i Stavanger enn i Haugesund.

Tabell 4.5: Eieform i antall og prosent for Haugesund og Stavanger

Eieform	Haugesund		Stavanger	
	Frekvens	Prosent	Frekvens	Prosent
Borettslag	903	25,5 %	4037	29,8 %
Selveier	2644	74,5 %	9528	70,2 %
Sum	3547	100,0 %	13565	100,0 %

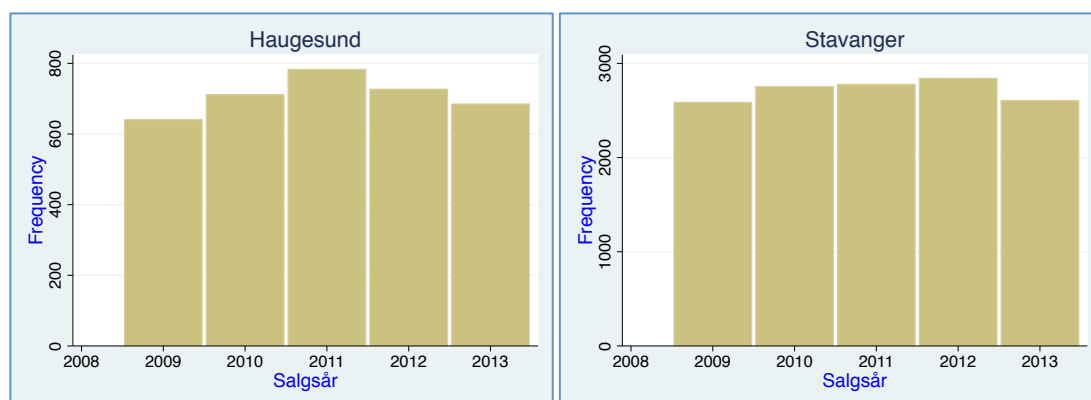
4.7.4 Salgsår

I tabellen under er det gitt en oversikt over hvor mange boliger som er solgt i hvert enkelt år i Haugesund og Stavanger. I 2011 ble det solgt flest boliger i Haugesund, mens i Stavanger var 2012 året hvor det var omsatt flest boliger.

Tabell 4.6: Antall boliger solgt hvert år (2009-2013)

År	Haugesund		Stavanger	
	Frekvens	Prosent	Frekvens	Prosent
2009	641	18,1 %	2586	19,1 %
2010	711	20,1 %	2754	20,3 %
2011	783	22,1 %	2778	20,5 %
2012	727	20,5 %	2841	20,9 %
2013	685	19,3 %	2606	19,2 %
Totalt	3547	100 %	13565	100 %

Nedenfor er utviklingen gjennom de fem siste årene illustrert. Ut i fra histogrammene kan man se at Stavanger har en flatere utvikling i antall solgte boliger, mens man i Haugesund har man en klarere "salgstopp" i 2011. Videre er det omsatt flere boliger i 2010, 2011 og 2012 enn i 2013 i begge byene.



Figur 4.2: Histogrammer for salgsårene 2009-2013

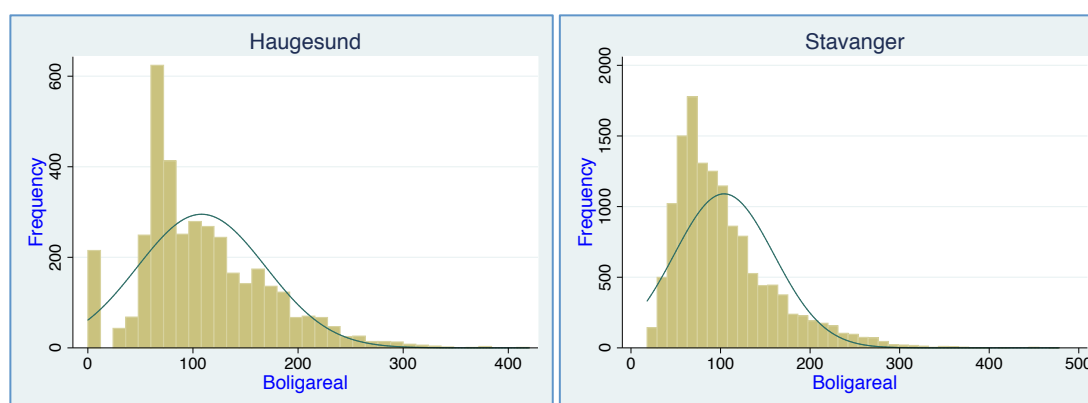
4.7.5 Boligareal

Boligareal antas å ha stor betydning for boligprisen. Gjennomsnittlig boligareal i Haugesund er 10 kvm² høyere enn i Stavanger. I Stavanger er det observert en bolig på kun 18 kvm². I Stavanger er også boligen med det høyeste boligareal observert.

Tabell 4.7: Boligareal i kvm² for Haugesund og Stavanger

Boareal	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	3547	114.7296	56.5563	24	420
Stavanger	13565	104.219	55.86838	18	478

Figurene under viser at Haugesund har en del boliger i 70 kvm² siktet som drar ned gjennomsnittet noe. I Stavanger ser man at det er en jevnere fordeling av boligareal størrelsene i forhold til hva som er tilfellet i Haugesund.



Figur 4.3: Fordeling av boliger etter boligareal i kvadratmeter

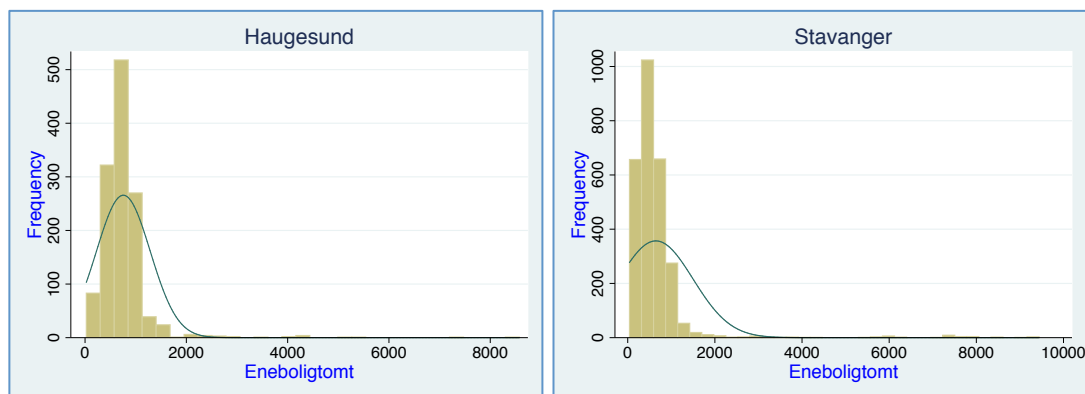
4.7.6 Tomteareal til selveide eneboliger

Totalt i Haugesund og Stavanger var det henholdsvis 1271 og 2705 registrerte tomtearealer til selveide eneboliger. Den minste tomten ble observert i Haugesund på 25 kvm², mens den største var observert i Stavanger på 55014 kvm². Gjennomsnittlig tomtestørrelse er mindre i Stavanger enn i Haugesund.

Tabell 4.8: Tomteareal til enebolig i Haugesund og Stavanger i kvm²

Tomtenebolig	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	1271	738.8293	428.71	25	7224
Stavanger	2705	580.8998	1159.432	40	55014

I figurene nedenfor kommer det tydelig frem at tomtestørrelsene som regel er under 2000 kvm² i begge byer.



Figur 4.4: Fordeling av tomteareal til eneboliger i Haugesund og Stavanger

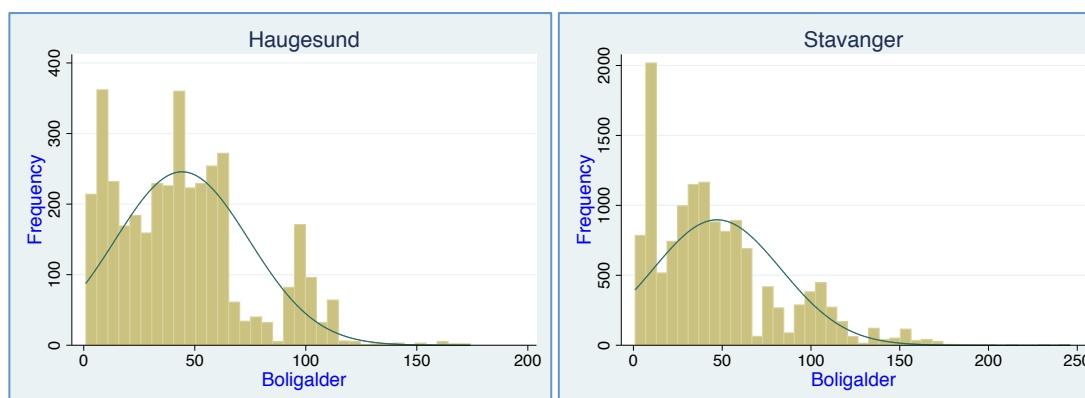
4.7.7 Boligalder

Den laveste observerte boligalderen er 1 år i begge byene, da jeg har operert med 2014 utgangspunkt og trukket fra når boligene ble bygget. Den eldste boligen er observert i Stavanger og man ser av tabell 4.9 at Stavanger har i gjennomsnitt eldre boliger enn Haugesund. Gjennomsnittsalderen er 45 og 47 år for henholdsvis Haugesund og Stavanger.

Tabell 4.9: Boligalder etter antall år

Boligalder	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	3547	45.36482	29.13305	1	174
Stavanger	13565	47.0895	36.06517	1	246

Figurene under gir en fordeling av boligalder for observasjonene i Haugesund og Stavanger, og man kan se at de fleste boliger ligger under gjennomsnittsalderen i begge byene. I Stavanger har man noen boliger av det eldre slaget, som gjør at gjennomsnittsalderen blir trukket opp.



Figur 4.5: Fordeling av boligalder etter år for Haugesund og Stavanger

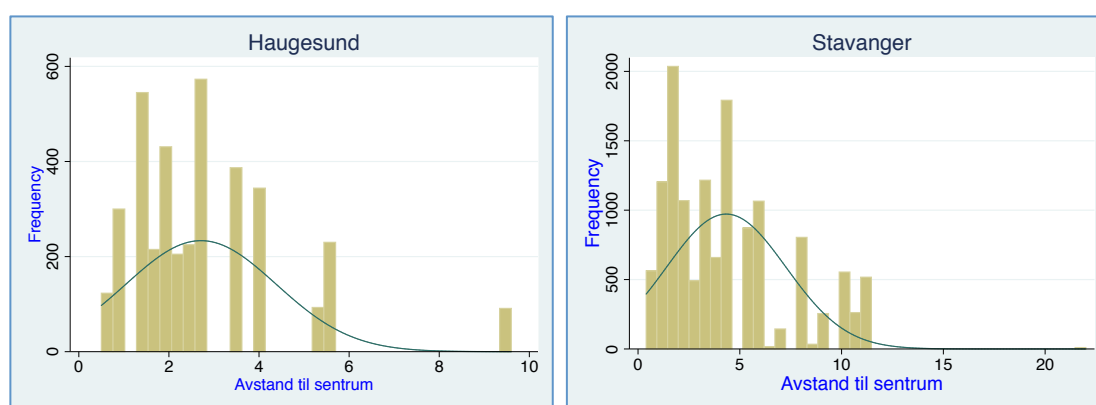
4.7.8 Avstand til sentrum

Avstanden til sentrum antas å ha stor forklaringskraft på boligprisen. Jeg målte avstanden fra hvert enkelt postnummer til sentrum i både Haugesund og Stavanger. Sentrum er som tidligere nevnt i oppgaven definert som Haraldsgata og Kirkegata for henholdsvis Haugesund og Stavanger. Den lengste avstanden fra sentrum var målt i Stavanger på 22 km, mens den "kun" var 9,6 km i Haugesund. Den lengste observerte avstanden i Stavanger er observert til å være plassen som heter Åmøy. Minimum avstanden til sentrum er relativt lik i begge byene. Gjennomsnittlig avstand til sentrum i Stavanger og Haugesund er henholdsvis 4,3 km og 2,6 km. Dette tilsvarer en gjennomsnittlig differanselengde på 1,7 km.

Tabell 4.10: Avstand til sentrum i kilometer

Avstand til sentrum	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	3547	2.62808	1.543833	0.5	9.6
Stavanger	13565	4.33366	2.933602	0.4	22

Figurene under illustrerer forskjellige avstander til sentrum for de observerte boligene i begge byene. Det kommer frem at de fleste observerte avstandene ligger rundt gjennomsnittet i Haugesund, mens det er flere observasjoner under gjennomsnittet i Stavanger.



Figur 4.6: Antall boliger fordelt etter avstand til sentrum i Haugesund og Stavanger

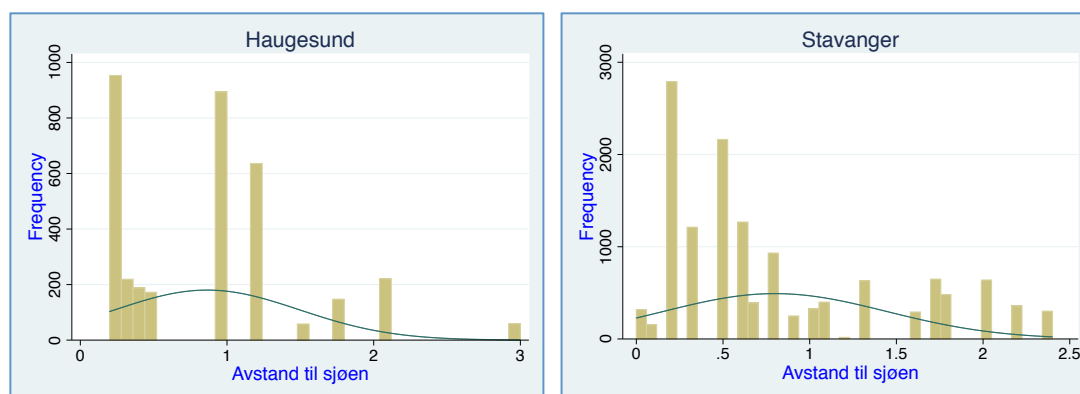
4.7.9 Avstand til sjø

Avstanden til sjøen ble målt i luftlinje fra hvert enkelt postnummer i byene og til nærmeste sjøpunkt. Gjennomsnittlig er avstanden til sjø så og si identisk for begge, med en avstand i underkant av 1 km. Videre ser man at både minimum og maksimum avstander er forholdsvis lik begge steder, noe som ikke er overraskende med tanke på at begge byene er lokalisert langs kysten.

Tabell 4.11: Avstand til sjø i kilometer

Avstand til sjøen	N	Gjennomsnitt	Standardavvik	Minimum	Maximum
Haugesund	3547	.8647871	.6284503	0.2	3
Stavanger	13565	.7998305	.6443428	0	2.4

Nedenfor er fordelingen avstanden til sjøen for byene illustrert. Mange observasjoner er samlet til venstre i figurene som indikerer svært kort avstand til sjøen.



Figur 4.7: Antall boliger fordelt etter avstand til sjøen i Haugesund og Stavanger

4.7.10 Må i bil for å komme til barnehage

Her betyr 0 at man slipper bil og 1 at man må i bil for å komme til barnehagen. I 804 av tilfellene var det rimelig å anta at man måtte i bilen i Haugesund for å komme til nærmeste barnehage, mens i Stavanger var det totalt 686. Tabellen viser at i Haugesund er sannsynligheten større for at man må i bil enn i Stavanger.

Tabell 4.12: Må i bil for å komme til barnehage (antall og prosent)

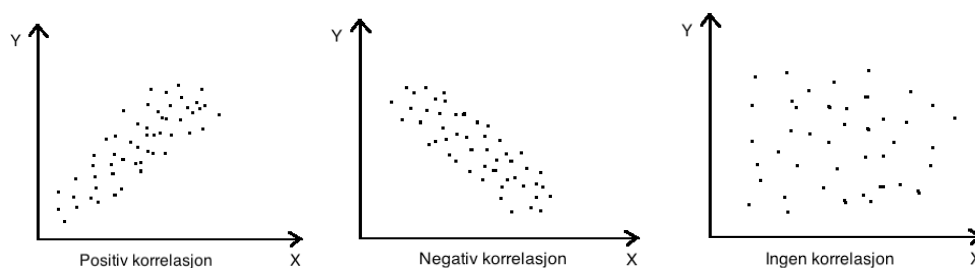
Haugesund	Frekvens	Prosent
0	2743	77,33 %
1	804	22,67 %
Totalt	3547	100 %
Stavanger	Frekvens	Prosent
0	12879	94,94 %
1	686	5,06 %
Totalt	13565	100 %

4.8 Korrelasjon mellom variablene

Det vil være interessant å se på om det foreligger korrelasjon mellom de variablene jeg har inkludert i analysen. En korrelasjonskoeffisient er et statistisk

mål på hvor sterk samvariasjon det er mellom to variabler.

Korrelasjonskoeffisienten kan være negativ og positiv og ligger mellom -1.0 til +1.0 En korrelasjon på + 1.0 indikerer at det er et perfekt positivt forhold mellom variablene. En korrelasjon på -1.0 tilsier at det er et perfekt negativt forhold mellom variablene. Dersom det ikke er noen form for korrelasjon mellom variablene indikeres korrelasjonen med verdien 0. I figur 4.8 er de ulike korrelasjonssammenhengene illustrert (Zikmund et al., 2013).



Figur 4.8: Korrelasjonsmønster (Zikmund et al., 2013)

Alle korrelasjoner som er signifikante på et 95% signifikansnivå er markert med en stjerne (*) i korrelasjonsmatrisen. Dette indikerer at det kan trygt sies at det er korrelasjon mellom variablene (Midtbø, 2012).

Når man har med mange uavhengige variabler kan det oppstå multikollinearitet. Det vil si at det er to uavhengige variabler som korrelerer med hverandre. Korrelasjonskoeffisientene vil følgende nærme seg verdiene 1 eller -1. Dette kan føre til at den avhengige variabelen ikke klarer å skille mellom effekten fra de ulike uavhengige variablene. Dette medfører at man kan ende opp med feil verdier og avvik i analysen. Dette problemet kan imidlertid løses ved å utelukke en av de uavhengige variablene (Thrane, 2003).

Jeg har valgt å ta med en korrelasjonsmatrise som inkluderer alle variabler unntatt tomteareal for borettslag. Dette fører til at cellen som i utgangspunktet ville gitt en korrelasjonsverdi mellom "tomtenebolig" og fellesgjeld er tom.

Videre er korrelasjonsmatrisene med alle variablene for Haugesund og Stavanger presentert:

Tabell 4.13: Korrelasjonsmatrise, Haugesund

Haugesund	Boligpris	Leilighet	Enebolig	Rekkehus	Tomannsbolig	Boligareal	Boligalder	Sjøen	Sentrum	Barnehage	Selveier	Borettslag	Fellesgjeld	Tomtenebolig	Salgsår09	Salgsår10	Salgsår11	Salgsår12	Salgsår13
Boligpris	1.0000																		
Leilighet	-0.5009*	1.0000																	
Enebolig	0.5389*	-0.6612*	1.0000																
Rekkehus	-0.0469*	-0.2777*	-0.2410*	1.0000															
Tomannsbolig	0.0075	-0.3086*	-0.2677*	-0.1125*	1.0000														
Boligareal	0.7848*	-0.6374*	0.6241*	-0.0439*	0.0893*	1.0000													
Boligalder	-0.1896*	-0.1276*	0.0615*	-0.0780*	0.1787*	0.0179	1.0000												
Sjøen	0.0556*	-0.0540*	0.0751*	-0.0052	-0.0252	0.0560*	-0.3172*	1.0000											
Sentrum	0.0974*	-0.1896*	0.1874*	0.1303*	-0.1079*	0.1320*	-0.3607*	0.5763*	1.0000										
Barnehage	0.0010	0.0004	0.0041	-0.1023*	0.0872*	-0.0072	0.2048*	-0.0230	-0.0477*	1.0000									
Selveier	0.5115*	-0.5663*	0.4130*	0.0627*	0.2025*	0.4369*	-0.0050	-0.0792*	0.0233	0.1495*	1.0000								
Borettslag	-0.5115*	0.5663*	-0.4130*	-0.0627*	-0.2025*	-0.4369*	0.0050	0.0792*	-0.0233	-0.1495*	-1.0000	1.0000							
Fellesgjeld	-0.3435*	0.0448	0.0099	-0.0568	-0.0137	0.0768*	-0.5567*	-0.1627*	-0.0778*	-0.0353	-0.0606	0.0606	1.0000						
Tomtenebolig	0.5294*	-0.5284*	0.7991*	-0.1926*	-0.2140*	0.5868*	0.0296	0.0842*	0.2424*	0.0292	0.3537*	-0.3537*	-	1.0000					
Salgsår09	-0.1144*	-0.0122	0.0020	0.0129	0.0042	-0.0056	0.0008	0.0309	0.0336*	-0.0005	0.0205	-0.0205	0.0016	0.0027	1.0000				
Salgsår10	-0.0693*	0.0050	-0.0281	0.0113	0.0247	-0.0053	0.0138	-0.0446*	-0.0247	0.0048	-0.0178	0.0178	0.0376	-0.0225	-0.2352*	1.0000			
Salgsår11	0.0095	0.0142	0.0220	-0.0423*	-0.0173	0.0035	-0.0130	0.0089	-0.0058	-0.0073	-0.0026	0.0026	-0.0360	0.0128	-0.2500*	-0.2665*	1.0000		
Salgsår12	0.0556*	-0.0048	-0.0247	0.0319	0.0161	-0.0063	0.0035	-0.0303	-0.0143	0.0070	0.0146	-0.0146	-0.0056	-0.0227	-0.2385*	-0.2542*	-0.2702*	1.0000	
Salgsår13	0.1150*	-0.0032	0.0286	-0.0123	-0.0275	0.0137	-0.0048	0.0367*	0.0129	-0.0039	-0.0141	0.0141	0.0034	0.0300	-0.2298*	-0.2450*	-0.2604*	-0.2484*	1.0000

Tabell 4.14: Korrelasjonsmatrise, Stavanger

Stavanger	Boligpris	Leilighet	Enebolig	Rekkehus	Tomannsbolig	Boligareal	Boligalder	Sjøen	Sentrum	Barnehage	Selveier	Borettslag	Fellesgjeld	Tomtenebolig	Salgsår09	Salgsår10	Salgsår11	Salgsår12	Salgsår13
Boligpris	1.0000																		
Leilighet	-0.4914*	1.0000																	
Enebolig	0.5809*	-0.5029*	1.0000																
Rekkehus	-0.0073	-0.3669*	-0.1911*	1.0000															
Tomannsbolig	0.0326*	-0.4585*	-0.2389*	-0.1743*	1.0000														
Boligareal	0.8057*	-0.6129*	0.6211*	0.0596*	0.0912*	1.0000													
Boligalder	0.0001	-0.2731*	0.0886*	-0.0865*	0.3364*	0.0276*	1.0000												
Sjøen	0.1065*	-0.0062	0.0227*	0.1025*	-0.1035*	0.1127*	0.2023*	1.0000											
Sentrum	0.1028*	-0.1055*	0.1369*	0.1995*	-0.1779*	0.2069*	-0.4218*	0.0463*	1.0000										
Barnehage	0.0958*	-0.0928*	0.0768*	-0.0800*	0.1083*	0.0840*	0.1605*	-0.0742*	-0.0573*	1.0000									
Selveier	0.4015*	-0.3742*	0.2890*	-0.1358*	0.2983*	0.2974*	0.2271*	-0.2243*	-0.1876*	0.1502*	1.0000								
Borettslag	-0.4015*	0.3742*	-0.2890*	0.1358*	-0.2983*	-0.2974*	-0.2271*	0.2243*	0.1876*	-0.1502*	-1.0000	1.0000							
Fellesgjeld	-0.5121*	0.1064*	-0.0163	-0.1024*	-0.0253	-0.0648*	-0.6663*	-0.1065*	-0.0337*	-0.0176	-0.1476*	0.1476*	1.0000						
Tomtenebolig	0.3150*	-0.2006*	0.3989*	-0.0762*	-0.0953*	0.2963*	0.0009	0.0199*	0.0958*	0.0479*	0.1329*	-0.1329*	-	1.0000					
Salgsår09	-0.1881*	-0.0095	0.0172*	-0.0232*	0.0140	0.0035	0.0113	-0.0081	0.0113	-0.0109	0.0134	-0.0134	-0.0088	-0.0022	1.0000				
Salgsår10	-0.0795*	0.0118	-0.0012	-0.0078	-0.0074	-0.0026	0.0056	-0.0058	-0.0128	0.0040	-0.0010	0.0010	0.0293	0.0149	-0.2450*	1.0000			
Salgsår11	0.0163	0.0002	-0.0003	0.0067	-0.0056	-0.0002	0.0001	0.0007	0.0055	0.0029	-0.0045	0.0045	0.0364*	0.0019	-0.2463*	-0.2561*	1.0000		
Salgsår12	0.1028*	0.0118	-0.0146	-0.0048	0.0042	-0.0149	-0.0105	-0.0104	-0.0036	-0.0006	0.0085	-0.0085	-0.0198	-0.0121	-0.2498*	-0.2598*	-0.2612*	1.0000	
Salgsår13	0.1458*	-0.0150	-0.0005	0.0292*	-0.0049	0.0148	-0.0062	0.0241*	-0.0000	0.0045	-0.0166	0.0166	-0.0376*	-0.0024	-0.2367*	-0.2461*	-0.2475*	-0.2510*	1.0000

Korrelasjonskoeffisienten er sterkt positiv mellom boligpris og boligareal, med henholdsvis 0,8057 og 0,7848 for Stavanger og Haugesund. Dette indikerer at en økning i boligareal gir høyere boligpris, noe som er rimelig. Videre kan man se at korrelasjonen mellom enebolig og boligpris er positiv, mens det er en negativ sammenheng mellom boligpris og leilighet. Man kan også observere at enebolig og boligareal har en positiv korrelasjon, mens den er negativ mellom leilighet og boligareal. Dette virker fornuftig ettersom leiligheter i mange tilfeller har et mindre boligareal enn eneboliger.

Korrelasjonskoeffisienten mellom boligpris og fellesgjeld er negativ og da den også er signifikant kan jeg med 95% sannsynlighet si at dersom en bolig har fellesgjeld vil dette ha en negativ innvirkning på boligprisen. Dette er begrunnet med at boliger med fellesgjeld ofte har en lavere boligpris.

”Tomtenebolig” har en positiv korrelasjon med boligpris i Haugesund, som betyr at ved økning av tomtestørrelsen til enebolig gir høyere boligpris. Denne er også signifikant ved et 95% signifikansnivå.

Hovedforskjellene mellom byene er at boligalder har en negativ innvirkning på boligprisen i Haugesund, noe som vil si at økt boligalder vil gi lavere boligpris. I Stavanger har ikke boligalderen noen korrelasjon med boligprisen.

Det er perfekt negativ korrelasjon mellom borettslag og selveier, ettersom jeg opererer med to ulike eierformer. Eierformen er derfor enten borettslag eller selveier. For å unngå multikollinearitet vil jeg utelate borettslag i regresjonsanalysen. Hvordan man forebygger og tar hensyn til multikollinearitet vil jeg forklare nærmere i neste kapittel.

5. Analyse

I analysen vil jeg anvende informasjonen jeg har fremskaffet til å estimere en funksjon som kan predikere boligpris. For å forklare effekten de ulike attributtene har på boligprisen vil jeg benytte meg av regresjonsanalyse. På grunnlag av resultatene vil jeg få empirisk støtte til kunne beholde eller forkaste hypotesene.

5.1 Kort om regresjonsanalyse

Regresjonsanalyser har som formål å forklare sammenhengen mellom den avhengige variabelen og en eller flere uavhengige variabler.

Korrelasjonsmatrisene jeg presenterte tidligere viste kun samvariasjon mellom variablene og fungerer som et første spadestik. En regresjonsanalyse tar steget langt videre med hensyn til tolkningsmulighetene, modellspesifikasjoner og typer av data som kan undersøkes. Analysen vil også skille klart mellom variabler som forklarer og variabler som blir forklart. Videre vil regresjonsanalysen flytte fokuset fra størrelsen på sammenhengen til styrken på årsakseffekten (Midtbø, 2007).

For at vi kan stole på at resultatene regresjonsanalysen gir er riktige, må seks forutsetninger ligge i bunn (Thrane, 2003):

1. Linearitet – Avhengig variabel er en lineær funksjon av alle uavhengige variabler og restleddet.
2. Ukorrelerte restledd – restledd skal være ukorrelerte med andre restledd.
3. Homoskedastisitet – variansen til restleddet må være konstant.
4. Fravær av multikollinearitet.
5. De uavhengige variablene er ukorrelert med restleddet.
6. Normalfordelt restledd.

Dersom forutsetningene ikke holder, holder ikke resultatene. Dette er mulig å teste flere av forutsetningene, noe jeg også skal gjøre.

Til å begynne med vil jeg presentere en enkel lineær regresjon som inkluderer kun en avhengig variabel. Videre vil jeg utarbeide en multivariat lineær regresjonsanalyse som inkluderer alle variablene, for å så presentere to ulike logaritmiske regresjonsanalyser, henholdsvis semilogaritmisk- og dobbellogaritmisk regresjonsanalyse. Avslutningsvis vil jeg presentere en interaksjonsmodell.

5.2 Bivariat regresjonsanalyse

Bivariat regresjonsanalyse er en enkel lineær regresjonsanalyse som inkluderer kun en uavhengig variabel og dens påvirkning på den avhengige variabelen.

Den bivariante regresjonsmodellen kan skrives slik:

$$P = \beta_0 + \beta_1 Z + \varepsilon$$

P er den avhengige variabelen boligpris, mens Z er attributtet og den uavhengige variabelen. Z forklarer variasjonen i P, men den kan ikke forklare alt. Dette gjør at regresjonsmodellen også inkluderer et restledd ε . Restleddet representerer alt utenom Z som påvirker P. Dette betyr at jo mer som forklares av Z, jo bedre forklaringskraft har modellen. β_0 er konstantleddet som angir størrelsen til P når Z er lik null, mens β_1 helningsleddet som sier hvor mye P øker når Z øker med én enhet (Midtbø, 2012).

5.2.1 Eksempel på bivariat regresjonsanalyse

I dette eksempelet vil jeg ha boligpris som den avhengige variabelen og boligareal som den uavhengige variabelen. Den uavhengige variabelen ble valgt på grunnlag av høyest korrelasjon med boligpris. Videre er bivariat regresjonsanalyse for Haugesund og Stavanger presentert i tabell 5.1 og 5.2.

Tabell 5.1: Bivariat regresjonsanalyse, Haugesund

Kilde	SS	df	MS	N	3547
				F(1, 3545)	5684.39
Modell	1.9256e+15	1	1.9256e+15	Prob > F	0.000
Residual	1.2009e+15	3545	3.3876e+11	R2	0.6159
Total	3.1265e+15	3546	8.8171e+11	Justert R2	0.6158
				Root MSE	5.8e+05

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	13029.79	172.8206	75.39	0.000	12690.95	13368.63
Konstant	510640.6	22105.23	23.10	0.000	467300.4	553980.9

Tabell 5.2: Bivariat regresjonsanalyse, Stavanger

Kilde	SS	df	MS	N	13565
				F(1, 3545)	25100.01
Modell	2.1304e+16	1	2.1304e+16	Prob > F	0.000
Residual	1.1512e+16	13563	8.4878e+11	R2	0.6492
Total	3.2816e+16	13564	2.4194e+12	Justert R2	0.6492
				Root MSE	9.2e+05

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	22503.38	142.04	158.43	0.000	22224.96	22781.8
Konstant	992055.4	16749.04	59.23	0.000	959225	1024886

Øverst til høyre viser N antall observasjoner som er med i analysen. Prob > F viser om modellen er signifikant. Dersom den er under 5% kan man si med 95% at modellen er signifikant (Midtbø, 2012).

R² er forklart varians som sier noe om hvor mye av variasjonen i den avhengige variabelen kan tillegges den uavhengige variabelen. Jo større andelen forklart av variasjon utgjør av den totale variasjonen, desto større er R². R² vil ligge et sted mellom 0 og 1. Dersom R² ligger tett opp til 1 forklarer den uavhengige variabelen mye av variasjonen i den avhengige variabelen. Er R² lik null forklarer den uavhengige variabelen ingenting. Utfordringen med R² er at det ikke er mulig å gi et klart svar på hva som er lav og høy R². Løsningen på dette kan være å sammenligne R² med forskjellige uavhengige variabler og samtidig beholde den samme avhengige variabelen. Den modellen som har høyest R² vil da ha størst forklaringskraft (Midtbø, 2007).

I kolonnen nedenfor har vi justert forklart varians R² som aldri kan være større enn R². Dette kommer av at denne foretar en justering av frihetsgradene. Justert R² har som fordel at den tar hensyn til flere uavhengige variabler om det inkluderes i modellen (Midtbø, 2012). Root MSE er presentert på den nederste linjen og vil si noe om standardfeilen til regresjonen. Root MSE måler spredningen rundt regresjonslinjen, og forteller hvor god eller dårlig modellen er. Jo lavere standardfeil, desto mer forklarer modellen (Midtbø, 2007).

Den viktigste informasjonen finner man i den nederste tabellen. Her finner vi koeffisientene til både standardfeilen, t-verdien, og p-verdien for konstantleddet og helningskoeffisienten (Thrane, 2003). Koeffisienten til konstantleddet er til

høyre for konstanten, mens helningskoeffisienten er til høyre for boligareal (Midtbø, 2012).

Videre er standardfeilen standardavviket i sannsynlighetsfordelingen. Dersom utvalget er stort, avvikene mellom faktiske og predikerte verdier er små og spredningen hos den uavhengige variabelen stor, vil standardfeilen være lav og tillitten til modellen høy. Når koeffisienten er stor og standardfeilen liten, blir t-verdien høy (Midtbø, 2007). $P > |t|$ og t-verdiene indikerer om variablene er signifikante. En $P > |t|$ verdi på 0,05 sammen med en t-verdi større eller lik 1.96 i en to-halet test forteller oss at variabelen er signifikant på et 5% signifikansnivå (Midtbø, 2007).

For Haugesund er R^2 0.6159, noe som tilsier at boligarealet forklarer hele 61,59% av variasjonen i boligprisen. De resterende 38,41% forklares av feilleddet ϵ . I Stavanger forklarer boligarealet litt mer med 64,92%. Den bivariate regresjonsanalysen indikerer en t-verdi på 75,39 i Haugesund og 158,43 i Stavanger, samt en $P > |t|$ lik 0. Boligareal er en signifikant variabel og jeg kan med 95% sannsynlighet si at det er en sammenheng mellom boligpris og boligareal.

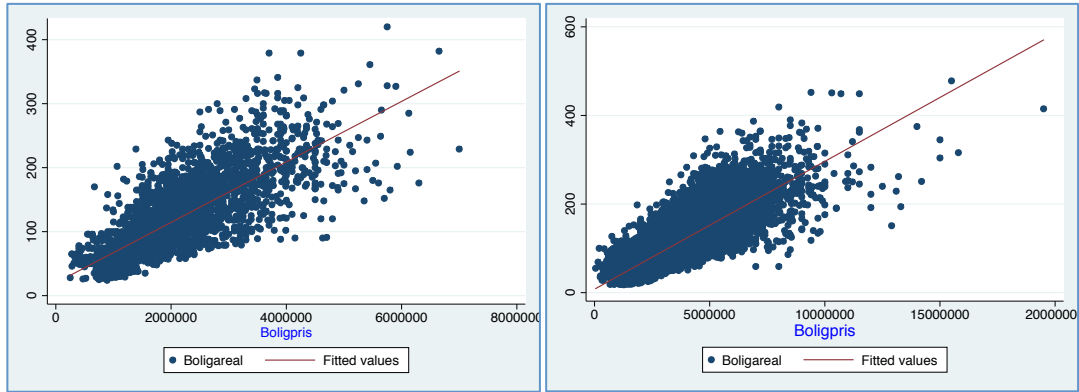
Den bivariate modellen kan derfor skrives slik for de to byene (Midtbø, 2012):

1. Boligpris (Haugesund) = $510640,6 + 13029,79 * \text{Boligareal}$
2. Boligpris (Stavanger) = $992055,4 + 22503,38 * \text{Boligareal}$

Konstantleddet viser at dersom boligareal er lik null, vil boligprisen være 510 641 kroner i Haugesund og 992 055 kroner i Stavanger. Helningskoeffisienten til boligarealet viser at for hver ekstra kvadratmeter en bolig har, vil føre til en økning på 13 030 kroner i boligprisen i Haugesund, mens økningen vil være på 22 503 kroner i Stavanger da jeg antar linearitet.

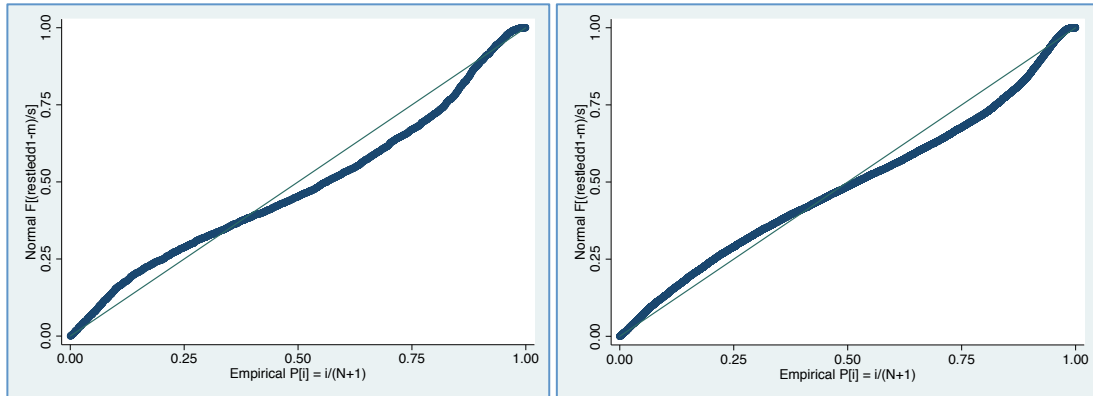
5.2.2 Vurdering av feilleddene

Figur 5.1 nedenfor viser regresjonslinjen for boligpris og boligareal for alle dataene. Vi ser at boligprisen øker med antall kvadratmeter. De prikkene som avviker med en større avstand fra regresjonslinjen har enten svært høy boligpris eller boligareal. De fleste observasjonene ligger derimot tett langs den lineære linjen, noe som gjør at jeg kan si at regresjonsanalysen gir gode estimater.



Figur 5.1: Korrelasjonsplott mellom boligpris og boligareal, Haugesund & Stavanger

Figur 5.2 illustrerer et normalskråplott for boligprisen og boligarealet. Et slikt normalskråplott ønsker å si noe om feileddene og om de er normalfordelt eller ikke. Dette er som tidligere nevnt en viktig forutsetning for at vi kan anta at det som er kommet frem i regresjonsanalysen er riktig. Dersom residualene er perfekt normalfordelt, vil de følge den lineære linjen og dermed ikke ha avvik mellom forventende og de faktiske verdiene (Midtbø, 2012).



Figur 5.2: Normalskråplott for den enkle lineære modellen, Haugesund & Stavanger

I en god modell bør sannsynligheten for å underestimere og overestimere en verdi være forholdsvis lik. Sagt på en annen måte bør fordelingen til restleddet være symmetrisk (Midtbø, 2012). Vi kan se her at residualene ikke er perfekt normalfordelt ettersom de svinger i noe grad fra den lineære linjen. På tross av dette kan jeg likevel si at residualene ligger tilnærmet normalfordelt langs den lineære linjen.

5.3 Multivariat regresjonsanalyse

En multivariat regresjonsanalyse, også kalt multipel regresjonsanalyse, knytter en avhengig variabel til to eller flere forklaringsvariabler. Å ta skrittet fra bivariat til multivariat regresjonsanalyse er helt nødvendig. Dette er fordi multivariat regresjon har flere fordeler i forhold til regresjonsanalyse med kun en uavhengig variabel. Videre er fordelene presentert (Midtbø, 2007):

Den første fordelen er at den gir et mer helhetlig bilde av fenomenene som studeres (Midtbø, 2007).

Den andre fordelen som kanskje er enda viktigere, er at beskrivelsene av årsakssammenhengene er mer presise og blir dermed mer troverdig. Modeller med bare en forklaringsvariabel fremstår som underspesifiserte og kan dermed gi resultater som er misvisende. Videre vil den multivariate regresjonsanalysen identifisere effekten av en variabel samtidig som de andre variablene holdes konstant. Dette gjør at forklaringer isoleres fra andre forklaringer, og effektene tolkes uavhengig av andre effekter (Midtbø, 2007).

En tredje fordel med multivariat regresjonsanalyse er at den ikke bare gir svar på om det er en god forklaring eller ikke. Modellen vil også si noe om hvor god denne forklaringen er i forhold til andre forklaringer (Midtbø, 2007).

Den siste fordelen som er verdt å nevne er at multivariat regresjonsanalyse kan sette opp grupper av forklaringsvariabler opp mot hverandre. Den vil kunne identifisere den samlede påvirkningen av gruppene som helhet, men også se på den unike effekten på enkeltvariablene innenfor hver av gruppene (Midtbø, 2007).

Den generelle multivariate regresjonsmodellen kan skrives slik (Midtbø, 2012):

$$P = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_n Z_n + \varepsilon$$

P er boligpris og er den avhengige variabelen. Konstantleddet, β_0 angir den forventende verdien til P når de uavhengige variablene Z er lik null. β er helningskoeffisient som angir effekten av sin forklaringsvariabel kontrollert for

effekten av den andre forklaringsvariabelen. Det betyr at β_1 sier hvor mye P forventes å endres seg når Z_1 øker med en enhet, når de andre uavhengige variablene holdes konstant. Restleddet ε representerer alt utenom Z som påvirker P. Korrelasjonen mellom restleddet ε og Z må være lik null. (Midtbø, 2012).

I analyser er det ikke gunstig å bruke for komplekse funksjonsformer. Hovedbegrunnelsen for dette er at ulike transformasjonsformer av de uavhengige variablene kan gi upresise estimater på individuelle attributtpriser. Tidligere analyser viser at enkle funksjonsformer gir mer presise estimat på individuelle koeffisienter, når all relevant informasjon ikke er inkludert i funksjonen eller ved bruk av dummyvariabler (Osland, 2001). Jeg vil utarbeide funksjonsformer for lineær regresjon, semilogaritmisk og dobbeltlogaritmisk regresjonsmodell.

Ved utarbeidelse av den lineære regresjonsmodellen måtte jeg foreta noen endringer på forhånd. For å unngå problemer i Stata og multikollinearitet ble en borettslagstomannsbolig med fellesgjeld solgt i 2009 min basisbolig. Valget av basisbolig ble valgt fordi det ville vært problematisk om selveier var en del av basisboligen ettersom den er inkludert i produktvariabelen "tomtenebolig".

5.3 Lineær regresjon med relevante variabler (Multivariat)

I tabell 5.3 og 5.4 er lineær regresjon med alle variablene for både Haugesund og Stavanger presentert.

Tabell 5.3: Lineær regresjon med alle relevante variabler, Haugesund

Kilde	SS	df	MS	N	3547
				F(1, 3545)	691.42
Modell	2.2907e+15	14	1.6362e+14	Prob > F	0.000
Residual	8.3584e+14	3532	2.3665e+11	R2	0.7327
Total	3.12.65e+15	3546	8.171e+11	Justert R2	0.7316
				Root MSE	4.9e+05

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	11328.95	208.0835	54.44	0.000	10920.98	11736.93
Leilighet	180987.9	33190.17	5.45	0.000	115914.1	246061.7
Enebolig	78739.41	36239.38	2.17	0.000	7687.19	149791.6
Rekkehus	80041.2	38113.94	2.10	0.036	5313.638	154768.8
Boligalder	-7317.916	328.6311	-22.27	0.000	-7962.242	-6673.59
Selveier	461651.4	24169.25	19.10	0.000	414264.3	509038.5
Avstand til sjøen	22593.37	16406.3	1.38	0.169	-9573.399	54760.15
Avstand til sentrum	-63175.96	7083.204	-8.92	0.000	-77063.54	-49288.37
Må i bil til barnehage	34276.75	20545.73	1.67	0.095	-6005.951	74559.44
Tomtenebolig	223.205	32.3338	6.90	0.000	159.8102	286.5998
Salgsår 2010	113577.9	26550.53	4.28	0.000	61521.96	165633.8
Salgsår 2011	225178.2	25947.07	8.68	0.000	174305.4	276050.9
Salgsår 2012	332333.8	26388.05	12.59	0.000	280596.5	384071.2
Salgsår 2013	423376.2	26762.95	15.82	0.000	370903.8	475848.6
Konstant	436573.7	58038.28	7.52	0.000	322781.8	550365.7

Tabell 5.4: Lineær regresjon med alle relevante variabler, Stavanger

Kilde	SS	df	MS	N	13565
				F(1, 3545)	3083.01
Modell	2.4976e+16	14	1.7840e+15	Prob > F	0.000
Residual	7.8407e+15	13550	5.7865e+11	R2	0.7611
Total	3.2816e+16	23564	2.4194e+12	Justert R2	0.7608
				Root MSE	7.6e+05

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	19675.43	173.1756	113.62	0.000	19335.99	20014.88
Leilighet	226547.3	22475.04	10.08	0.000	182493.1	270601.5
Enebolig	469291.3	24129.63	19.45	0.000	421993.9	516588.7
Rekkehus	125420.6	26321.00	4.77	0.000	73827.82	177013.5
Boligalder	-2980.052	220.5064	-13.51	0.000	-3412.275	-2547.829
Selveier	678426.9	17051.95	39.79	0.000	645002.7	711851.1
Avstand til sjøen	158360.8	10719.71	14.77	0.000	137348.6	179372.9
Avstand til sentrum	-34728.69	2647.335	-13.12	0.000	-39917.84	-29539.55
Må i bil til barnehage	98537.50	30543.64	3.23	0.001	38667.72	158407.3
Tomtenebolig	147.5983	12.60902	11.71	0.000	122.8828	172.3137
Salgsår 2010	371262.9	20843.99	17.81	0.000	330405.7	412120
Salgsår 2011	667578.5	20796.14	32.10	0.000	626815.2	708341.9
Salgsår 2012	956059.0	20685.35	46.22	0.000	915512.8	996605.2
Salgsår 2013	1051684	21141.27	49.75	0.000	1010244	1093124
Konstant	113141.6	40537.06	2.79	0.000	33683.29	192599.8

R² er på 0,7327 og 0,7611 som betyr at de 14 variablene forklarer henholdsvis 73,27% og 76,11% av variasjonen i boligprisen i Haugesund og Stavanger. Jeg kan dermed si at modellene forklarer relativt mye, noe som er bra.

I modellen er det sentralt å se på hvilke variabler som er signifikante. De variablene som har en p-verdi $<0,05$ og en t-verdi større eller lik 1,96 anses å være signifikante på et 5% nivå. De variablene som ikke er signifikante vil ikke kunne forklare variasjonen i boligprisen.

Totalt sett er det 12 av 14 variabler signifikante i modellen for Haugesund. Variablene avstand til sjøen og må i bil til barnehage er ikke signifikante. I modellen til Stavanger er alle signifikante, men variablene avstand til sjøen og "må i bil til barnehage" har positive koeffisienter, noe som er uforventet.

5.3.1 Test på multikollinearitet

Som tidligere nevnt kan multikollinearitet oppstå når man inkluderer flere uavhengige variabler i en regresjonsanalyse. En av forutsetningene til regresjonsanalyser var at det skulle være fravær av multikollinearitet. I statistikkprogrammet Stata kunne jeg kjøre en "VIF"-test, som står for "Variance Inflation Score". Denne testen vil fortelle noe om i hvilken grad de uavhengige variablene i modellen korrelerer med de andre uavhengige variablene. De uavhengige variablene bør ikke ha en VIF-skåre på mer enn 10 og gjennomsnittlig VIF for alle de uavhengige variablene bør ikke være så mye større enn 1. Dersom noen avviker fra dette bør det vurderes om de aktuelle variablene bør utelates fra regresjonsmodellen. Nedenfor er testene for multikollinearitet presentert (Thrane, 2003).

Tabell 5.5: VIF-test, Haugesund & Stavanger

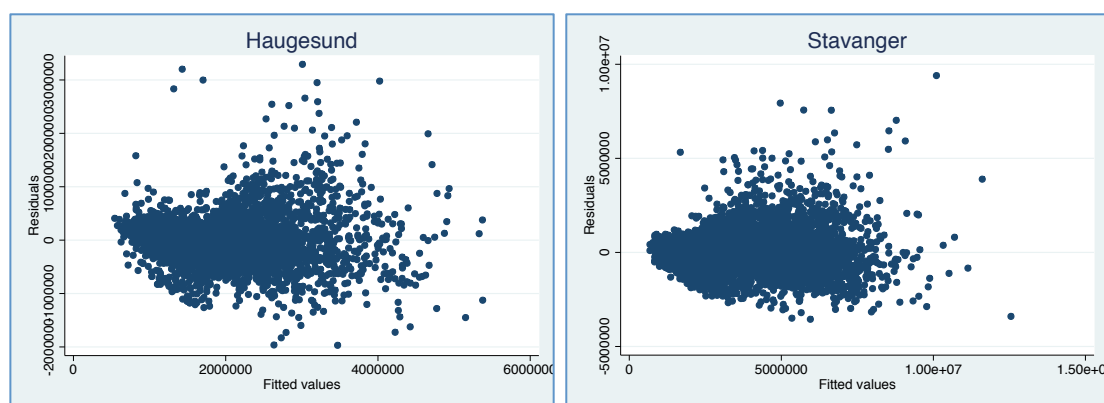
Variabel	Haugesund		Stavanger	
	VIF	1/VIF	VIF	1/VIF
Boligareal	2.08	0.481864	2.18	0.458636
Enebolig	4.56	0.219304	2.25	0.445380
Rekkehus	1.82	0.550282	1.74	0.573311
Leilighet	4.05	0.246759	2.96	0.337899
Boligalder	1.37	0.728068	1.48	0.674539
Selveier	1.66	0.601847	1.42	0.701818
Sjøen	1.59	0.627767	1.12	0.894181
Sentrum	1.79	0.558087	1.41	0.707302
Barnehage	1.11	0.901647	1.05	0.952325
Tomtenebolig	3.00	0.333556	1.20	0.833674
Salgsår10	1.69	0.590526	1.65	0.606794
Salgsår11	1.74	0.576083	1.65	0.605668
Salgsår12	1.70	0.587981	1.66	0.602115
Salgsår13	1.67	0.597769	1.63	0.614930
Gj.snitt VIF	2.13		1.67	

Ut ifra VIF-testene ovenfor er det er ingen av de uavhengige variablene som har en VIF-skåre høyere enn 10.

5.3.2 Vurdering av feilleddene

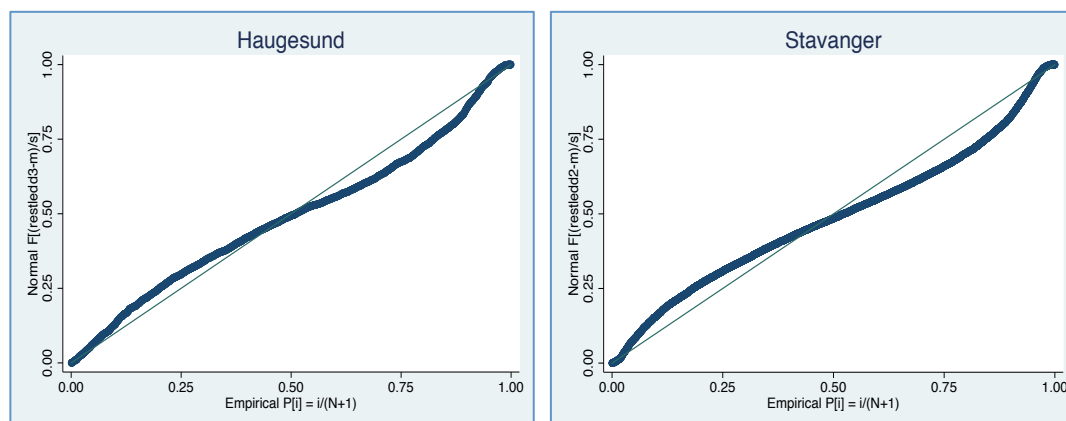
For å vurdere residualene til den lineære regresjonsmodellen vil jeg først presentere et residualplot og deretter et normalskråplott for dataene i analysen.

En av forutsetningene til regresjonsanalyse er som nevnt homoskedastisitet. Restleddet skal av den grunn ikke være heteroskedastisk. Identifisering av heteroskedastisitet kan gjøres grafisk. Dette gjøres ved å plote residualene opp mot de predikerte verdiene. Et mønster i spredningen, for eksempel i form av en sky av prikker som enten blir trangere eller bredere når de predikerte verdiene blir større, indikerer uønsket heteroskedastisk (Thrane, 2003).



Figur 5.3: Residualplott til den lineære modellen, Haugesund & Stavanger

I figurene ovenfor er det ingen klare tendenser av heteroskedasitet å spore, selv om det er noen ensomme plasseringer.



Figur 5.4: Normalskråplott til den lineære modellen, Haugesund & Stavanger

I figurene ovenfor ser vi at residualene ikke er perfekt normalfordelt ettersom de svinger i noe grad fra den lineære linjen. Residualene svinger noe mer i Stavanger enn i Haugesund. På tross av dette kan vi si at residualene er tilnærmet normalfordelte.

5.4 Logistiske modeller

Grunnen til at lineære modeller ofte blir brukt i statistiske analyser er fordi det er lettere rent matematisk, ikke fordi man tror at alle interessante sammenhenger er eksakt rettlinjede og lineære. I mange tilfeller vil lineære modeller likevel vise gode resultater. For å finne den lineære sammenhengen kan det være naturlig, og vel så nødvendig, å utarbeide en logistisk modell. Jeg vil derfor utarbeide to typer logistiske modeller. Det skiller herved mellom semi- og dobbeltlogaritmisk modell (Skog, 1998).

Logaritmer kan også oppfattes som en omkoding, herved under en ikke-lineær omkoding (Skog, 2004). Ved omkoding av en variabel i logaritmisk modell vil regresjonsparameteren vise prosentvise endringer istedenfor absolutte endringer. I den lineære regresjonsmodellen viste regresjonsparameteren hvor mange enheter den avhengige variabelen vil øke (eller minke) når den uavhengige variabelen øker med én enhet (Skog, 1998).

Det er viktig å være oppmerksom på at logaritmiske omkodinger bare kan brukes på variabler som har positive verdier. Grunnen til dette er at logaritmen til null er lik minus uendelig, mens logaritmen til negative tall ikke er definert (Skog, 2004).

5.4.1 Semilogaritmisk modell

I en semilogaritmisk modell kan vi kun omkode den avhengige variabelen, som i min analyse er boligpris. Regresjonsparameteren vil nå fortelle oss hvor stor relativ prosentvis endring den avhengige variabelen får, når den uavhengige variabelen øker med én enhet (Skog, 1998).

Funksjonsformen til denne modellen kan uttrykkes ved (Stock & Watson, 2003):

$$\ln(P_i) = \beta_0 + \beta_1 Z_i + \varepsilon_i$$

Nedenfor er resultatene fra den semilogaritmisk regresjonsanalysen for Haugesund og Stavanger presentert.

Tabell 5.6: Semilogaritmisk regresjonsanalyse, Haugesund

Kilde	SS	df	MS	N	3547
				F(14, 3532)	610.26
Modell	548.63711	14	39.188365	Prob > F	0.000
Residual	226.809159	3532	.064215504	R2	0.7075
Total	775.446269	3546	0.218681971	Justert R2	0.7064
				Root MSE	.25341

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	.0048907	.0001084	45.12	0.000	.0046782	.0051032
Leilighet	.036295	.0172894	2.10	0.036	.0023968	.0701931
Enebolig	.0430771	.0188778	2.28	0.023	.0060647	.0800895
Rekkehus	.0567207	.0198543	2.86	0.004	.0177938	.0956477
Boligalder	-.0027524	.0001712	-16.08	0.000	-.0030881	-.0024168
Selveier	.3423819	.0125902	27.19	0.000	.317697	.3670667
Avtand til sjøen	.0120986	.0085463	1.42	0.157	-.0046577	.0288548
Avstand til sentrum	-.0198392	.0036898	-5.38	0.000	-.0270735	-.0126049
Avstand til barnehage	-.0192469	.0107027	-1.80	0.072	-.0402309	.0017371
Tomtenebolig	.0000582	.0000168	3.45	0.001	.0000252	.0000912
Salgsår 2010	.0566689	.0138307	4.10	0.000	.029552	.0837858
Salgsår 2011	.1133625	.0135163	8.39	0.000	.0868619	.1398631
Salgsår 2012	.1556153	.013746	11.32	0.000	.1286644	.1825663
Salgsår 2013	.2020352	.0139413	14.49	0.000	.1747013	.229369
Konstant	13.60112	.0302332	449.87	0.000	13.54184	13.66039

Tabell 5.7: Semilogaritmisk regresjonsanalyse, Stavanger

Kilde	SS	df	MS	N	13565
				F(14, 13550)	2318.71
Modell	1989.69378	14	142.120985	Prob > F	0.000
Residual	830.52246	13550	.06129317	R2	0.7055
Total	2820.21624	13564	.207919216	Justert R2	0.7052
				Root MSE	.24757

Boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
Boligareal	.0049823	.0000564	88.40	0.000	.0048718	.0050928
Leilighet	.0136939	.0073148	1.87	0.061	-.000644	.0280318
Enebolig	.0889101	.0078533	11.32	0.000	.0735166	.1043036
Rekkehus	.0772656	.0085665	9.02	0.000	.0604741	.0940571
Boligalder	-.0002669	.0000718	-3.72	0.000	-.0004075	-.0001262
Selveier	.2775591	.0055498	50.01	0.000	.2666808	.2884374
Avtand til sjøen	.0590902	.0034889	16.94	0.000	.0522516	.0659288
Avstand til sentrum	-.0034039	.0008616	-3.95	0.478	-.0050928	-.0017151
Avstand til barnehage	.0070613	.0099408	0.71	0.001	-.012424	.0265466
Tomtenebolig	.0000197	4.10e-06	4.80	0.000	.0000116	.0000277
Salgsår 2010	.1305278	.0067839	19.24	0.000	.1172304	.1438252
Salgsår 2011	.2307519	.0067683	34.09	0.000	.217485	.2440188
Salgsår 2012	.3244201	.0067323	48.19	0.000	.3112239	.3376163
Salgsår 2013	.3524477	.0068807	51.22	0.000	.3389606	.3659348
Konstant	13.93928	.0131932	1056.55	0.000	13.91342	13.96514

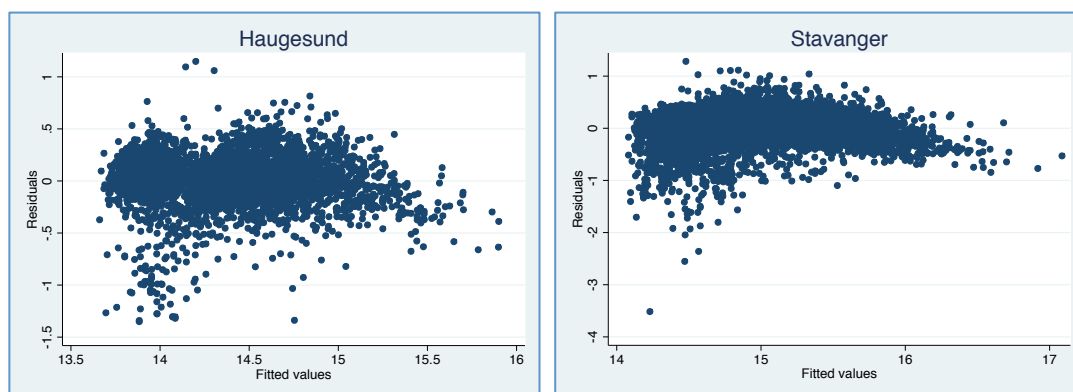
Forklaringskraften R^2 til semilogaritmisk regresjonsmodell er noe lavere sammenlignet med lineære regresjon med henholdsvis 70,75% og 70,55% for Haugesund og Stavanger. Det er utfordrende å si om dette gir grunnlag for å antyde at det er en svakere modell ettersom det er to ulike modeller.

Det som er interessant å legge merke til er at dummyvariabelen "må i bil til barnehage" i Haugesund har skiftet fortegn til minus. Denne er likevel ikke signifikant på et 5% signifikansnivå. Den ville derimot vært signifikant om jeg opererte med et 10% signifikansnivå. Det andre er som er viktig å legge merke til er at variabelen "avstand til sentrum" for Stavanger ikke er signifikant, noe den var i den lineære regresjonen.

Ut i fra koeffisientene kan vi si at dersom boligarealet øker med en kvadratmeter for en bolig i Haugesund, vil boligprisen økes med 0,489%, mens den vil øke med 0,498% i Stavanger.

5.4.2 Vurdering av feilleddene

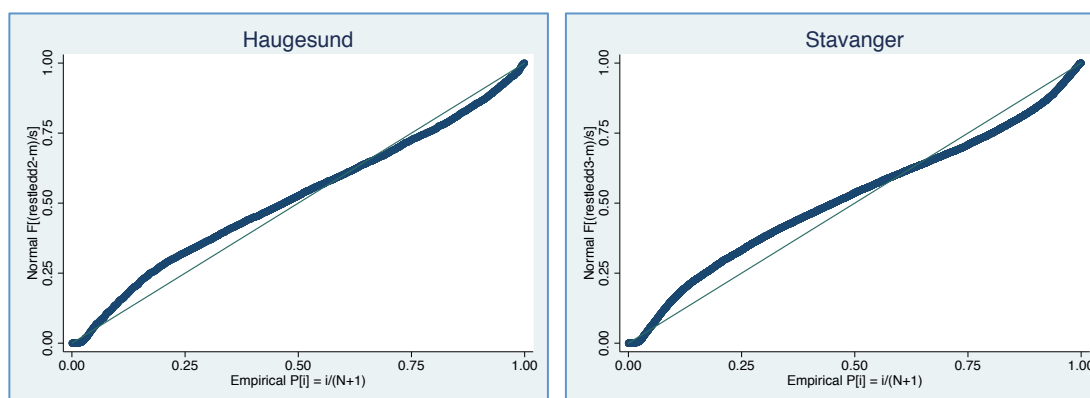
Residualplottene for den semilogaritmiske modellen er illustrert nedenfor.



Figur 5.5: Residualplott til den semilogaritmiske modellen, Haugesund & Stavanger

Residualene er ganske konsentrerte i begge figurene, noe som tyder på at variansen til restleddet er konstant.

Videre er normalskråplottene for den semilogaritmiske regresjonen illustrert i figurene nedenfor.



Figur 5.6: Normalskråplott til den semilogaritmiske modellen, Haugesund & Stavanger

Verdiene følger ikke den lineære linja optimalt, men man kan si at restleddene er tilnærmet normalfordelt for begge regresjonene.

5.4.3 Dobbellogaritmisk modell

I en dobbellogaritmisk modell omkoder man både den avhengige og den uavhengige variabelen.

Funksjonsformen til denne modellen kan uttrykkes ved (Stock & Watson 2003):

$$\ln(P_i) = \beta_0 + \beta_1 \ln(Z_i) + \varepsilon_i$$

For å unngå multikollinearitet måtte jeg fjerne variabelen "tomtenebolig". Hadde denne blitt inkludert ville jeg ikke kunnet vise resultatene for de forskjellige boligtypene og selveier. Det totale utvalget observasjoner ble for Stavanger redusert til 13248 fordi avstand til sjøen var 0 kilometer for noen av observasjonene. Jeg velger likevel å vise resultatene i tabell 5.9 og 5.10.

Tabell 5.8: Dobbellogaritmisk regresjon, Haugesund

Kilde	SS	df	MS	N	3547
				F(14, 3532)	784.70
Modell	575.968898	13	44.3052998	Prob > F	0.0000
Residual	199.477371	3533	.056461186	R2	0.7428
Total	775.446269	3546	0.218681971	Justert R2	0.7418
				Root MSE	.23762

In_boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
In_boligareal	.7067711	.0125978	56.10	0.000	.6820715	.7314708
In_boligalder	-.0846471	.0052764	-16.04	0.000	-.0949922	-.074302
In_sentrum	-.0353946	.008621	-4.11	0.000	-.0522972	-.018492
In_sjøen	.0001202	.005511	0.02	0.983	-.0106847	.0109252
Barnehage	-.0158369	.0100068	-1.58	0.114	-.0354565	.0037827
Selveier	.3107018	.0124949	24.87	0.000	.2862039	.3351998
Leilighet	.1455054	.0167512	8.69	0.000	.1126624	.1783484
Enebolig	.0921799	.0142116	6.49	0.000	.064316	.1200437
Rekkehus	.0803048	.018449	4.35	0.000	.0441331	.1164765
Salgsår 2010	.0517942	.0129675	3.99	0.000	.0263698	.0772187
Salgsår 2011	.1134928	.0126714	8.96	0.000	.0886488	.1383368
Salgsår 2012	.1500305	.0128962	11.63	0.000	.1247458	.1753153
Salgsår 2013	.1991261	.0130749	15.23	0.000	.173491	.2247613
Konstant	11.02478	.0680513	162.01	0.000	10.89136	11.1582

Tabell 5.9: Dobbellogaritmisk regresjon, Stavanger

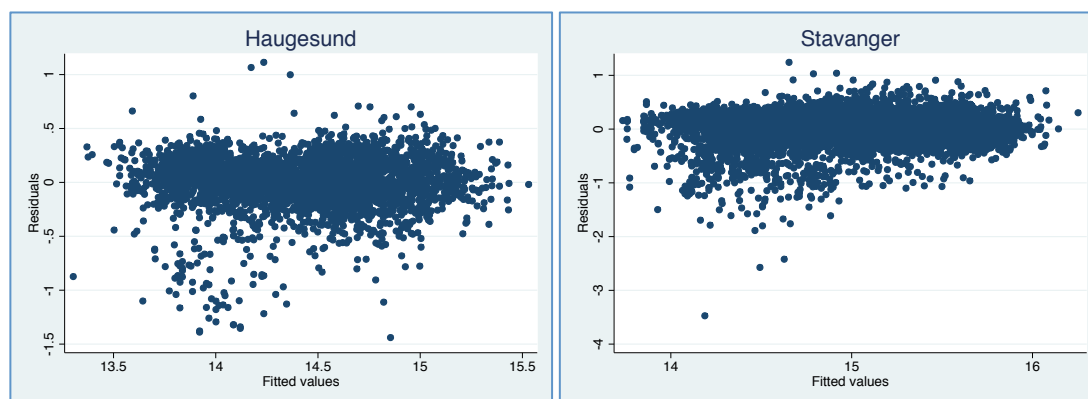
Kilde	SS	df	MS	N	13248
				F(14, 3532)	2928.36
Modell	2059.07363	13	158.390279	Prob > F	0.0000
Residual	715.805875	13234	.054088399	R2	0.7420
Total	2774.8795	13247	.209472296	Justert R2	0.7418
				Root MSE	.23257

In_boligpris	Koeffisient	Std.feil	t	P> t	[95% Konf. Intervall]	
In_boligareal	.6157422	.0060127	102.41	0.000	.6039565	.627528
In_boligalder	-.0005866	.002473	-0.24	0.812	-.0054341	.0042609
In_sentrum	-.0277236	.0034728	-7.98	0.000	-.0345307	-.0209164
In_sjøen	.0368389	.0026825	13.73	0.000	.0315809	.042097
Barnehage	.0099251	.0093975	1.06	0.291	-.0084953	.0283455
Selveier	.2881254	.0053478	53.88	0.000	.2776429	.2986078
Leilighet	.1187016	.007204	16.48	0.000	.1045806	.1328225
Enebolig	.1535395	.0070938	21.64	0.000	.1396347	.1674444
Rekkehus	.0843501	.0080505	10.48	0.000	.06857	.1001303
Salgsår 2010	.1325058	.0064552	20.53	0.000	.1198526	.145159
Salgsår 2011	.2369272	.0064397	36.79	0.000	.2243044	.24955
Salgsår 2012	.3280529	.0064104	51.17	0.000	.3154875	.3406183
Salgsår 2013	.3549233	.0065446	54.23	0.000	.342095	.3677517
Konstant	11.68016	.0314664	371.20	0.000	11.61848	11.74184

Forklaringskraften er lavere i denne modellen enn hva den var hos den lineære modellen for Stavanger, mens den er litt høyere for Haugesund. Totalt er 11 av 13 variabler signifikante for både Haugesund og Stavanger.

5.4.4 Vurdering av feilleddene

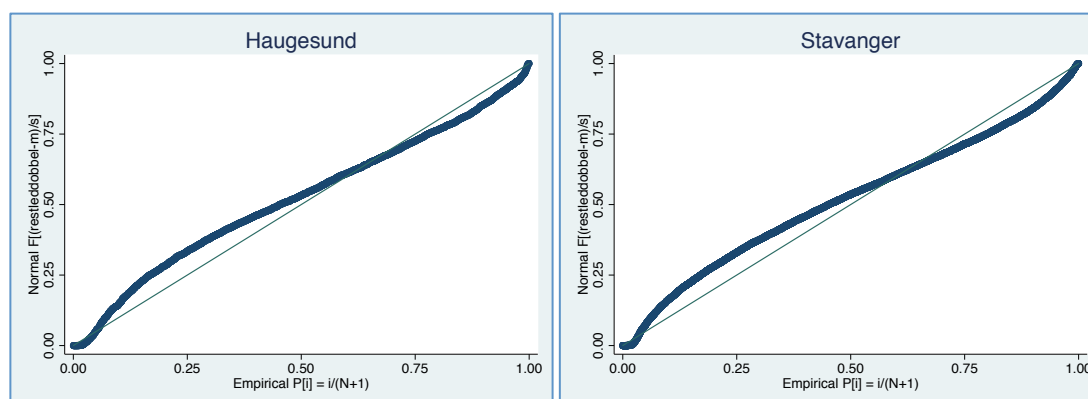
Videre er det gitt en vurdering av feilleddene.



Figur 5.7: Residualplott til den dobbellogaritmiske modellen, Haugesund & Stavanger

Residualene er mer konsentrert for Stavanger enn hos Haugesund. Grunnen til at de er sentrert høyere i figuren for Stavanger er fordi y-aksen har lavere verdier enn hva den har i figuren for Haugesund.

I figurene nedenfor er normalskråplott for den dobbeltlogaritmiske modellen illustrert.



Figur 5.8: Normalskråplott til den dobbellogaritmiske modellen, Haugesund & Stavanger

Residualene avviker i en viss grad fra den lineære linja. Likevel kan vi si at restleddet er tilnærmet normalfordelt.

5.5 Vurdering av modellene - Valg av modell

Når jeg nå vil gå videre i analysen er det viktig å velge den funksjonsformen som gir de mest troverdige resultatene for min analyse. Hvilken funksjonsform som er valgt, er valgt på grunnlag av forklaringskraft, vurdering av feilleddene, det jeg anser som riktige fortegn og det jeg vil anse som eventuelle svakheter ved modellene. Videre er det gitt en vurdering av de ulike regresjonsmodellene for begge byene.

5.5.1 Haugesund

Den dobbeltlogaritmiske modellen gav den høyeste forklaringskraften med 74,28%, noe som er relativt bra. Videre var forutsetningen om et normalfordelt restledd oppfylt. På den andre siden måtte jeg utelate den uavhengige variabelen "tomtenebolig" med hensyn til multikollinearitet, noe jeg anser som mindre gunstig. Videre var variablene "må i bil til barnehage" og avstand til sjøen ikke statistisk signifikant. Totalt sett gjorde dette at denne funksjonsformen ble uaktuell.

Den semilogaritmiske modellen viste en R^2 på 70,75% som var den laveste av samtlige modeller. Variabelen "må i bil til barnehage" hadde et negativt fortegn, noe som var forventet, da jeg antar at å slippe å kjøre til barnehage vil gi en dyrere boligpris. Videre har variabelen "avstand til sentrum" en lavere t-verdi enn hos den lineære modellen. Forutsetningen om et normalfordelt restledd var oppfylt. Alt dette tatt i betraktning gjorde at heller ikke denne modellen ble valgt.

Den tredje og siste, nemlig den lineære modellen hadde en R^2 på 73,27%, som var høyere enn hos den semilogaritmiske modellen, men litt lavere enn ved den dobbeltlogaritmiske modellen. Restleddet var tilnærmet normalfordelt. Alle variablene utenom "må i bil til barnehage" og avstand til sjøen var statistisk signifikant. På grunnlag av dette ble denne modellen valgt til videre analyse.

5.5.2 Stavanger

Den dobbeltlogaritmiske modellen gav den nest høyeste forklaringskraften med 74,20%. Som for ved modellen for Haugesund måtte jeg fjerne den uavhengige variabelen "tomtenebolig" for å unngå multikollinearitet. Resultatet av det var at

det samtidig ble frafall av noen observasjoner ettersom logistiske modeller ikke kan behandle nullverdier. Videre var ikke boligalder signifikant. På bakgrunn av dette var ikke den modellen aktuell for min analyse.

Den semilogaritmiske modellen hadde den laveste forklaringskraften med R^2 70,55%. Variabelen avstand til sentrum var ikke signifikant. Dette gjorde det enkelt å utelukke denne modellen.

Den lineære modellen med alle relevante variabler hadde en forklaringskraft på 76,11% noe som er høyere enn samtlige modeller. Alle variablene var signifikante og det var det tilnærmet normalfordelt restledd. Selv om variablene "må i bil til barnehage" og avstand til sjøen har positive fortegn ble denne valgt.

Kort oppsummert har jeg valgt den lineære modellen for begge byer. Det var den modellen som hadde høyest forklaringskraft, flest signifikante variabler, antagelse om forventende fortegn og færrest svakheter. Videre var vurderingen av feilleddene ved de ulike modellene var utfordrende ettersom de var tilnærmet likt normalfordelt.

Ved estimering av separate modeller for hver enkelt gruppe, er det vanskelig å si noe om forskjellene mellom gruppene²¹. Videre har jeg derfor kjørt de to lineære regresjonene i en og samme modell for å se om attributter har forskjellig betydning for boligprisen i de to byene.

5.6 Sammenligning av to regresjoner (Statistisk regresjon)

Statistisk interaksjon, også kalt samspill, er en uavhengig variabels påvirkning på en avhengig variabel som er betinget av (verdien på) en annen uavhengig variabel. Dette er et vanlig fenomen i samfunnsvitenskapene. Regresjon kan benyttes for å finne ut om det forekommer samspill mellom variablene (Thrane, 2003).

Interaksjonseffekter er empirisk relevante, men kan være vanskelige å tolke. Vi er ikke på jakt etter effekten av X på Y kontrollert for Z, men effekten av X på Y

²¹ <http://www3.nd.edu/~rwilliam/stats2/l51.pdf>

som en funksjon av Z. En interaksjonseffekt forutsetter at koeffisienten for interaksjonsleddet er signifikant forskjellig fra null (Midtbø, 2012).

Jeg vil se om attributter har større/mindre eller lik betydning for boligprisen i de to byene. Det vil det derfor være gunstig å gjennomføre en interaksjonsanalyse. Ved å inkludere for mange variabler er det vanskelig å oppdage effekten mellom gruppene. Videre skal man ikke inkludere for mange kategoriske variabler i interaksjonsmodeller. Jeg antar at effekten av noen av variablene varierer mellom Haugesund og Stavanger, men ikke alle²². Jeg vil derfor se på om effekten av de signifikante kontinuerlige variablene fra den lineære modellen er betinget av om det er Stavanger eller Haugesund. Disse er som følger:

- Boligareal.
- Boligalder.
- Avstand til sentrum.
- Tomteareal til selveid enebolig.

Ved gjennomføringen måtte jeg tilføre en dummy variabel som har verdien 1 hvis Stavanger 1 og 0 hvis Haugesund. Videre konstruerte jeg en interaksjonsvariabel for hver av variablene, for eksempel produktet av Stavanger og boligareal (dvs. Stavanger * Boligareal) som jeg navngav "Stvgareal". Hvordan beregningene i Stata er gjort er hentet fra "An Introduction to Modern Econometrics Using Stata", Baum. C.F (2006).

Funksjonen for sammenligning av to regresjoner vil se slik ut (Gujarati & Porter, 2010):

$$1. P_i = B_1 + B_2D_i + B_3Z_1 + B_4(D_iZ_1) + B_5Z_2 + B_6(D_iZ_2) + B_7Z_3 + B_8(D_iZ_3) + B_9Z_4 + B_{10}(D_iZ_4) + \varepsilon$$

Der Z_1 = Boligareal, Z_2 = Boligalder, Z_3 = Avstand til sentrum, Z_4 = Tomteareal til selveid enebolig.

Haugesund funksjonen for boligpris ($D_i=0$) blir (Gujarati & Porter, 2010):

$$2. E(P_i | D_i = 0, Z_i) = B_1 + B_3Z_1 + B_5Z_2 + B_7Z_3 + B_9Z_4$$

²² <http://www3.nd.edu/~rwilliam/stats2/l51.pdf>

Stavanger funksjonen for boligpris ($D_i=1$) blir (Gujarati & Porter, 2010):

$$3. E(P_i | D_i = 1, Z_i) = B_1 + B_2D_i + (B_3 + B_4D_i)Z_1 + (B_5 + B_6D_i)Z_2 + (B_7 + B_8D_i)Z_3 + (B_9 + B_{10}D_i)Z_4$$

$$= (B_1 + B_2) + (B_3 + B_4)Z_1 + (B_5 + B_6)Z_2 + (B_7 + B_8)Z_3 + (B_9 + B_{10})Z_4 \text{ fordi } D_i = 1$$

Statistikkprogrammet Stata antar homoskedasitet standardfeil, som følge av dette bør man justere modellen for heteroskedasitet. En tommelfingerregel er at man skal alltid anta at det er heteroskedasitet i modellen (Stock & Watson, 2003). Jeg har derfor tatt i bruk kommandoen "robust" i Stata ved utarbeidelsen av regresjonen.

Resultatene fra interaksjonsmodellen er vist i tabellen nedenfor:

Tabell 5.10: Interaksjonsmodell med relevante variabler.

N	17112
F(14, 3532)	3286.11
Prob > F	0.0000
R2	0.7053
Root MSE	8.4e+05

Boligpris	Koeffisient	Robust std.feil	t	P> t 	[95% Konf. Intervall]	
Stavanger	216259.1	52755.77	4.10	0.000	112852.4	319665.8
Stvgareal	10423.84	475.1656	21.94	0.000	9492.463	11355.21
Boligareal	11978.61	256.8536	46.64	0.000	11475.15	12482.07
Stvgsentrum	22744.9	7557.601	3.01	0.003	7931.223	37558.57
Avstand til sentrum	-74892.37	6848.453	-10.94	0.000	-88316.04	-61468.7
Stvgalder	5351.555	491.1091	10.90	0.000	4388.931	6314.179
Boligalder	-8095.923	418.8354	-19.33	0.000	-8916.884	-7274.963
Stvgtomt	-69.46955	117.5757	-0.59	0.555	-299.9301	160.991
Tomtenebolig	307.4416	28.99125	10.60	0.000	250.6158	364.2674
Konstant	1113942	38737.77	28.76	0.000	1038012	1189872

Totalt er det 17 112 (13 565 + 3 547) observasjoner for Haugesund og Stavanger. Når man sammenligner grupper, bør man utarbeide en statistisk test. Nedenfor er resultatene fra en kollektiv test presentert (Midtbø, 2012):

- (1) Stavanger = 0
- (2) Stvgareal = 0
- (3) Stvgsentrum = 0
- (4) Stvgalder = 0
- (5) Stvgtomt = 0

$$F(5, 17102) = 3458.58$$

$$\text{Prob} > F = 0.0000$$

Interaksjonsvariablene "Stvgareal", "Stvgsentrum", "Stvgalder" er signifikante ettersom de har t-verdier forskjellig fra null. Interaksjonsvariabelen "Stvgtomt" er ikke signifikant. Videre er dummyvariabelen "Stavanger" også signifikant, som betyr at regresjonslinjene til Stavanger og Haugesund er forskjellig. Jeg ønsker å finne den marginale endringen til hvert enkelt attributt. Dette vil jeg gjøre ved å derivere med hensyn på attributtet, gitt alt annet holdes likt.

- **Boligareal Haugesund:**

$$\frac{\partial P}{\partial Z_1} = \frac{\partial \text{Boligpris}}{\partial \text{Boligareal}} = 11\,978,61$$

- **Boligareal Stavanger:**

$$\frac{\partial P}{\partial Z_1} = \frac{\partial \text{Boligpris}}{\partial \text{Boligareal}} = 22\,402,45$$

Interaksjonseffekt forutsetningen er oppfylt ettersom "Stvgareal" har en t-verdi på 21,94. Jeg kan si at det er en betalingsvillighet på 22 402 kr per kvadratmeter boligareal i Stavanger, mens betalingsvilligheten per kvadratmeter i Haugesund er 11 979 kr. Estimert er det en større betalingsvillighet for attributtet boligareal i Stavanger enn i Haugesund. Med andre ord er effekten av boligareal større i Stavanger enn hva den er i Haugesund.

- **Boligalder (Haugesund):**

$$\frac{\partial P}{\partial Z_2} = \frac{\partial \text{Boligpris}}{\partial \text{Boligalder}} = -8095,923$$

- **Boligalder (Stavanger):**

$$\frac{\partial P}{\partial Z_2} = \frac{\partial \text{Boligpris}}{\partial \text{Boligalder}} = -2744,368$$

Vi kan se av interaksjonsmodellen at det foreligger en interaksjonseffekt for "Stvgalder". Ved å se på helningskoeffisientene til Haugesund og Stavanger kan vi se at eldre boliger i Haugesund gir lavere boligpris enn hva som er tilfellet i Stavanger. Ved en alders økning på et år reduseres betalingsvilligheten med 8

096 kr i Haugesund, mens den reduseres med 2 744 kr i Stavanger. Vi kan derfor si at effekten av boligalder er større for Haugesund enn hva den er for Stavanger.

- **Avstand til sentrum (Haugesund):**

$$\frac{\partial P}{\partial Z_3} = \frac{\partial \text{Boligpris}}{\partial \text{Avstand til sentrum}} = -74892,37$$

- **Avstand til sentrum (Stavanger):**

$$\frac{\partial P}{\partial Z_3} = \frac{\partial \text{Boligpris}}{\partial \text{Avstand til sentrum}} = -52147,47$$

Forutsetningen om interaksjonseffekt er oppfylt da interaksjonsvariabelen "Stvgsentrum" er signifikant. Videre kan vi se at boligprisen varierer med avstand til sentrum i større grad i Haugesund enn Stavanger. For hver kilometer bort ifra sentrum i Haugesund vil betalingsvilligheten reduseres med 74 892 kr. I Stavanger vil betalingsvilligheten reduseres med 52 147 kr. Vi kan dermed si at effekten av avstand til sentrum er større for Haugesund enn hva den er for Stavanger.

- **Tomteareal til selveide eneboliger (Haugesund):**

$$\frac{\partial P}{\partial Z_4} = \frac{\partial \text{Boligpris}}{\partial \text{Tomteareal til selveide eneboliger}} = 207,4416$$

- **Tomteareal til selveide eneboliger (Stavanger):**

$$\frac{\partial P}{\partial Z_4} = \frac{\partial \text{Boligpris}}{\partial \text{Tomteareal til selveide eneboliger}} = 137,97205$$

Av interaksjonsmodellen ser vi at interaksjonsvariabelen "Stvgtomt" ikke er signifikant. Jeg kan dermed ikke si at det foreligger en interaksjonseffekt.

Kort oppsummert er det en interaksjonseffekt for tre av fire interaksjonsvariabler. Effekten av boligareal er større i Stavanger enn i Haugesund, mens effekten av boligalder og avstand til sentrum er større i Haugesund enn i Stavanger.

5.7 Eksempel ved bruk av lineær modell

Jeg vil nå bruke den lineære modellen til å vise et eksempel for å se hvilken boligpris den gir ut ifra mine signifikante variabler. Først vil jeg vise for den lineære modellen for Haugesund, og deretter for Stavanger. Jeg har valgt å ta utgangspunkt i en registrert solgt bolig i Haugesund av eiendomsverdi.no. Boligen ble omsatt for 2 900 000 kr i 2013. Egenskapene til boligen er som følger:

- * Selveid enebolig med selveid tomt på 885kvm².
- * Boligareal på 161kvm².
- * Bygget i 1974.
- * Boligen ligger 3 km fra Haraldsgata i Haugesund.

Ved beregning av boligpris tar jeg utgangspunkt i funksjonen for den lineære modellen:

$$P = \beta_0 + \beta_1 Z_1 + \beta_2 Z_2 + \dots + \beta_n Z_n + \varepsilon$$

$$\text{Boligpris}_{2013} = 436573.7 + (11328.95 * 161) + (78739.41 * 1) - (7317.916 * 39) + (461651.4 * 1) - (63175.96 * 3) + (223.205 * 885) + (423376.2 * 1) = \underline{\underline{2\ 946\ 911}}$$

Ut ifra beregningene gir den lineære modellen en 46 911 kr høyere boligpris i forhold til hva den var registrert omsatt for, noe jeg vil si er et tilfredsstillende estimat.

Videre vil jeg se på hvordan estimatet til den lineære modellen for Stavanger. Jeg har tatt utgangspunkt i en bolig omsatt for 4 400 000 kr i 2013. Egenskapene til boligen er som følger:

- * Selveid enebolig med selveid tomt på 124kvm².
- * Boligareal på 135 kvm².
- * Bygget i 1850.
- * Boligen ligger 1,4 km fra Kirkegata i Stavanger.

$$\text{Boligpris}_{2013} = 113141,6 + (19675,43 \cdot 135) + (469291,3 \cdot 1) - (2980,052 \cdot 162) + (678426,9 \cdot 1) - (34728,69 \cdot 1,4) + (147,5983 \cdot 124) + (1051684 \cdot 1) = \underline{\underline{4\ 455\ 640\ \text{kr.}}}$$

Den lineære modellen for Stavanger gir et estimat som er 55 640 kr høyere boligpris enn hva den var registrert omsatt for, noe som er et relativt nøyaktig estimat.

5.8 Resultater - Hypotesetesting

I dette avsnittet vil jeg presentere mine resultater og trekke konklusjoner. Hypotesetesting er en prosedyre for å avgjøre hvorvidt forskjellen mellom utvalgsverdien og populasjonsverdien er reell, eller om den skyldes tilfeldigheter. For å avgjøre om jeg kan forkaste eller beholde hypotesene vil jeg se på om de har empirisk støtte fra analysen. Dette vil jeg gjøre ved hjelp av de lineære regresjonsmodellene og interaksjonsmodellen. For at jeg kan forkaste nullhypotesene må variablene være statistisk signifikant (Midtbø, 2007).

Hypotese 1: Boligareal

Den første hypotesen jeg skal teste er om en økning i boligarealet gir en høyere boligpris i Stavanger enn i Haugesund.

Hypotesen var som følger:

H₁₀: Boligareal har lik betydning for boligprisen i Haugesund og Stavanger.

H₁₁: Boligareal har i Stavanger større betydning for boligprisen enn i Haugesund.

Den lineære regresjonen viser at boligareal har en t-verdi på 113,62 i Stavanger og 54,44 i Haugesund. Begge byene har p-verdier lik null. Jeg kan dermed si med 95% sannsynlighet at boligareal har betydning for boligprisen i både Haugesund og Stavanger. Resultatene fra interaksjonsmodellen sier oss at helningskoeffisienten for boligareal er 10423,84 (22402,45 - 11978,61) høyere i Stavanger enn hva den er for Haugesund. For hver økte kvadratmeter øker boligprisen med 10 423,84 kr mer i Stavanger enn i Haugesund. Vi kan dermed forkaste nullhypotesen.

Konklusjon: Boligareal i Stavanger har større betydning for boligprisen enn i Haugesund.

Hypotese 2: Avstand til sentrum

Jeg ville se om avstand til sentrum påvirker boligprisen i de to byene forskjellig.

Hypotesen var som følger:

H₂₀: Avstand til sentrum har lik betydning for boligprisen i Haugesund og Stavanger.

H₂₁: Avstand til sentrum har større betydning for boligprisen i Haugesund enn Stavanger.

Variabelen avstand til sentrum er ved begge de lineære modellene signifikant ved et 5% signifikansnivå, med t-verdiene -8,92 i Haugesund og -13,12 i Stavanger. Vi kan dermed med 95% sikkerhet si at beliggenheten til en bolig i Haugesund og Stavanger i forhold til sentrum har betydning for boligprisen. Dersom vi ser på helningskoeffisientene til avstand til sentrum i interaksjonsmodellen, vil boligprisen reduseres med henholdsvis 74 892,37 kr og 52 147,47 kr for hver økte kilometer bort ifra Haraldsgata i Haugesund og Kirkegata i Stavanger. Vi kan på bakgrunn av dette forkaste nullhypotesen.

Konklusjon: Avstand til sentrum har større betydning for boligprisen i Haugesund enn Stavanger.

Hypotese 3. Boligalder

I hypotesen for boligalder antok jeg at en eldre bolig vil kunne gi en lavere boligpris og vice versa. Jeg ønsket å se på om boligalder har ulik effekt på boligprisen i Haugesund i forhold til Stavanger

Hypotesen var som følger:

H₃₀: Boligalder i Haugesund har lik betydning for boligprisen i Haugesund og Stavanger.

H₃₁: Boligalder i Haugesund har større betydning for boligprisen enn i Stavanger.

I den lineære regresjonen er boligalder signifikant for begge byer representert med t-verdier -22,27 og -13,51 for henholdsvis Haugesund og Stavanger. Jeg kan dermed si at boligalder har betydning for boligprisen. For å avgjøre om alderen på en bolig har større påvirkning på boligprisen i Haugesund enn i Stavanger må vi se på interaksjonsmodellen. Helningskoeffisienten i Haugesund er negativt med 8 096 kr, mens for Stavanger er helningskoeffisienten 2 744 kr. Boligprisen vil derfor falle med 8 096 kr for hvert eldre år en bolig er i Haugesund, mens den falle med 2 744 kr for hvert år i Stavanger. På grunnlag av dette kan vi forkaste nullhypotesen.

Konklusjon: Boligalder har større betydning for boligprisen i Haugesund enn i Stavanger.

Hypotese 4. Tomteareal til selveide eneboliger

Jeg ønsket å se på om tomtearealet til selveide eneboliger har ulik effekt på boligprisen i Haugesund i forhold til Stavanger.

Hypotesen var som følger:

H₄₀: Tomtearealet til selveide eneboliger i Haugesund har lik betydning for boligprisen i Haugesund og Stavanger.

H₄₁: Tomtearealet til selveide enebolig i Haugesund har større betydning for boligprisen enn i Stavanger.

Den lineære modellen viser t-verdier på 11,71 og 6,90 for henholdsvis Stavanger og Haugesund. Begge byene har signifikans lik null. Vi kan dermed si med 95% sannsynlighet at tomteareal til selveide eneboliger har betydning for boligprisen i Haugesund og Stavanger. Dersom vi ser på sammenligningen av regresjonene ser vi at det ikke foreligger en interaksjonseffekt

Konklusjon: Jeg kan ikke påstå at tomtearealet til selveide eneboliger har større betydning for boligprisen i Haugesund i forhold til Stavanger.

Videre er resultatene for variablene som ikke ble inkludert i interaksjonsmodellen presentert.

Hypotese 5: Boligtype

Jeg har samlet inn data for fire ulike boligtyper. Som følge av min basisbolig er tomannsbolig er inkludert i konstantleddet. Det er naturlig å anta at enhver boligtype har ulike egenskaper og at det derfor vil være ulik betalingsvillighet for boligtypene.

Hypotesen var som følger:

H_{50} : Boligtype har ikke betydning for boligprisen.

H_{51} : Boligtype har betydning for boligprisen.

Resultatene fra den lineære modellen viser at samtlige boligtyper har høye t-verdier og p-verdier under 0,05 som forteller at de er signifikante. Jeg kan dermed si at boligtypen har betydning for boligprisen. Koeffisienten for enebolig i Haugesund sier oss at en enebolig er 78 739,41 kr dyrere enn en tomannsbolig. I Stavanger er koeffisienten for enebolig 469 291,3, som sier at en enebolig er 469 291, 3 kr dyrere enn en tomannsbolig. Vi kan dermed forkaste nullhypotesen.

Konklusjon: Boligtype har betydning for boligprisen i både Haugesund og Stavanger.

Hypotese 6: Avstand til sjø

Begge byene har forholdsvis nær tilgang til sjøen. Jeg ønsket å teste om nærhet til sjøen verdsettes innad i de to byene og om dette i så fall vil ha betydning for boligprisen.

Hypotesen jeg ønsket å teste var:

H_{60} : Avstand til sjø har ikke betydning for boligprisen.

H_{61} : Avstand til sjø har betydning for boligprisen.

I den lineære modellen for Haugesund er variabelen avstand til sjø ikke signifikant. Jeg kan derfor ikke si at avstand til sjø har betydning for boligprisen i Haugesund. Variabelen avstand til sjø har en positiv koeffisient i den lineære modellen for Stavanger, noe som ikke var forventet. Med andre ord betyr en positiv koeffisient at økt avstand til sjø gir lavere boligpris.

Konklusjon: Jeg kan ikke påstå at avstand til sjø har betydning for boligprisen verken i Haugesund eller i Stavanger

Hypotese 7: Må i bil til barnehage

Jeg antok at dersom man fra boligen må i bil for å komme til nærmeste barnehage, vil dette gi en lavere boligpris.

Hypotesen var som følger:

H₇₀: Må i bil for å komme til barnehage har ikke betydning for boligprisen.

H₇₁: Må i bil for å komme til barnehage har betydning for boligprisen.

I den lineære modellen for Haugesund er ikke variabelen signifikant. Jeg kan derfor ikke si at om man må i bil for å komme til barnehage har betydning for boligprisen i Haugesund. I den lineære modellen for Stavanger er variabelen signifikant. Koeffisienten til denne variabelen er derimot positiv, noe som ikke er et forventet fortegn. På bakgrunn av dette kan jeg ikke forkaste nullhypotesen.

Konklusjon: Jeg kan ikke påstå at om man må i bil for å komme til barnehage har betydning for boligprisen verken i Haugesund eller Stavanger.

6. Drøfting av resultatene

I dette kapittelet vil jeg drøfte resultatene og gi en oppsummering av det som kom frem av analysen. Dette blir gjort med formål å svare på min problemstilling. Avslutningsvis vil jeg gi noen kritiske vurderinger av oppgaven.

6.1 Drøfting av problemstillingen

Problemstillingen min gikk ut på å undersøke hvilken betydning ulike boligattributter har på boligprisen i Haugesund og Stavanger i perioden 2009 til 2013. I drøftingen blir det tatt utgangspunkt fra resultatene fra den lineære modellen for begge byene og sammenligning av de lineære modellene.

I den lineære modellen for Stavanger var totalt 12 av 14 variabler signifikante på et 5% signifikansnivå og som hadde forventende fortegn. I den lineære modellen

for Haugesund var 12 av 14 variabler signifikante på et 5% signifikansnivå. Boligareal, leilighet, enebolig rekkehus, boligalder, selveier, avstand til sentrum, "tomtenebolig" og salgsårene 2009-2013 var de variablene som forklaringskraft på boligprisen.

Jeg vil først se nærmere på de variablene som er inkludert i interaksjonsmodellen. Den første signifikante variabelen boligareal er et attributt som uten tvil har betydning for boligprisen i begge byene. Av interaksjonsmodellen kunne jeg med sikkerhet si at det var en interaksjonseffekt for boligareal, noe som gjør at jeg med sikkerhet kan si at boligareal har større betydning for boligprisen i Stavanger enn i Haugesund, slik jeg antok det ville være. I Haugesund er det en betalingsvillighet på 11 979 kr per kvm², mens den er i Stavanger 22 202 kr per kvm². Økt antall kvadratmeter vil derfor gi økt pris. Slik som den hedonistiske prisfunksjonen indikerte, vil en økt mengde av attributtet gi høyere pris.

Avstand til sentrum er et attributt som har forklaringskraft på boligprisen da den er statistisk signifikant. De lineære modellene for begge byer viser negative koeffisienter som betyr at ved bevegelse bort i fra sentrum vil følgende boligprisene reduseres. Dette stemmer overens med teorien om tomtemarkedet ved Alonso-Muth-Mills modellen jeg presenterte i kapittel 3. Videre kunne jeg si av interaksjonsmodellen at boligprisen øker mer i Haugesund med nærhet til sentrum i forhold til Stavanger, da det foreligger en interaksjonseffekt. For hver ekstra kilometer bort ifra Haraldsgata i Haugesund vil boligprisen reduseres med 74 892 kr. I Stavanger vil boligprisen reduseres med 52 147 kr per kilometer fra Kirkegata i Stavanger. Dette samsvarer med det jeg antydte i tidligere i oppgaven, at skillet i boligpreferanse om man bor i indre eller ytre delen av byen ikke er så klart i Stavanger i forhold til andre byer.

Variabelen boligalder er statistisk signifikant i den lineære modellen for både Haugesund og Stavanger. Interaksjonsmodellen viser en interaksjonseffekt som betyr at alderen på en bolig har større påvirkning på boligprisen i Haugesund enn i Stavanger. For hvert år eldre en bolig er i Stavanger vil boligprisen reduseres med 2 744 kr, mens i Haugesund vil prisen reduseres med 8 096 kr

per år. Jeg kan dermed konkludere med at betalingsvilligheten i Haugesund og Stavanger reduseres jo eldre boligene er.

Tomtestørrelsen til selveid enebolig er statistisk signifikant ved den lineære regresjonen for begge byer. Koeffisientene er positive som betyr at for hver økning i kvadratmeter tomt gir høyere boligpris. Ved sammenslåingen av regresjonene for Haugesund og Stavanger var interaksjonsvariabelen ikke signifikant.

Når det gjelder dummyvariablene for boligtypene enebolig, leilighet og rekkehus er alle signifikante, og alle har positive koeffisienter for både Haugesund og Stavanger. Dette betyr at alle boligtypene er dyrere enn tomannsbolig, ettersom tomannsbolig er inkludert i basisboligen. Enebolig, leilighet og rekkehus har dermed en positiv effekt på boligprisen. Av de lineære modellene er det interessant å legge merke til at i Haugesund er leilighet dyrest i forhold til tomannsbolig, mens i Stavanger er det enebolig som er den dyreste i forhold til tomannsbolig. I Haugesund er leilighet 180 987,9 kr dyrere enn en tomannsbolig, mens leilighet i Stavanger er 226 547,3 kr dyrere. En enebolig i Stavanger er 469 291,3 kr dyrere enn en tomannsbolig, mens den bare er 78 739,41 kr dyrere i Haugesund.

Variabelen avstand til sjøen er signifikant for Stavanger med en positiv betakoeffisient. Dette indikerer at desto større avstand fra sjøen boligen er lokalisert, jo høyere boligpris, noe som ikke var forventet. Avstand til sjø for Haugesund er ikke signifikant på et 5% signifikansnivå. På bakgrunn av dette vil jeg ikke kunne si at nærhet til sjø vil ha forklaringskraft på boligprisen.

Dummyvariabelen "må i bil til barnehage" ser vi at ikke er signifikant i Haugesund. I Stavanger har variabelen en positiv koeffisient som vil si at dersom man må bruke bil for å komme til nærmeste barnehage vil det gi økt boligpris, noe som ikke er rimelig etter mine antagelser.

Salgsårene 2009-2013 ble behandlet som dummyvariabler, men ble ikke testet i en egen hypotese. Salgsåret 2009 er en del av basisboligen. De andre salgsårene hadde positive koeffisienter og er statistisk signifikante med veldig høye t-

verdier, noe som sier oss at det har vært en positiv utvikling i boligprisen siden 2009 i både Haugesund og Stavanger. Vi kan dermed si at boliger har blitt dyrere og dyrere i begge byer.

Dummyvariabelen for selveier ble ikke testet i en egen hypotese. Likevel kan vi se at betakoeffisienten er positiv og signifikant. Dette kommer av at borrettslag er en del av basisboligen. Selveier har dermed en positiv effekt på boligprisen. Ser vi på de lineære modellene til byene kan vi se at dette er et attributt som verdsettes høyt. Dette er naturlig ettersom borettslag ofte har fellesgjeld som vil gi kunne gi en lavere boligpris.

6.2 Kritiske vurderinger

Ved gjennomføringer av slike analyser vil det alltid være rom for ting som kunne eller burde vært gjort for å få en høyere validitet. Det vil i så måte være nødvendig å foreta noen kritiske vurderinger av det som er gjort.

For det første ble en del observasjoner utelatt fra analysen på grunn av mangelfulle opplysninger. Hadde disse blitt inkludert kunne de vært med og påvirket resultatene.

Avgjørelsen av hvilke attributter som jeg anså å ha størst forklaringskraft ble hentet fra teori og de som jeg anså som interessante å undersøke. Det vil da være slik at vi utelater andre variabler som også vil ha betydning for boligprisen. For å styrke analysen vil det vært bedre om jeg inkluderte flere variabler.

Variabelen avstand til sjøen var en variabel som ikke gav ønskede resultater. Noe av årsaken til dette er at jeg ikke hadde informasjon om egenskapene til boligene i forhold til sjø, når det gjelder sjøutsikt eller hvordan tilgangen til sjøen er. Det ville nok gitt mer riktige og forventende resultater om jeg hadde studert avstand til sjø for fritidsboliger istedenfor ordinære boliger.

Boligprisene er ikke blitt justert etter boligprisindeksen. Hadde jeg gjort dette kunne det gitt andre resultater.

Vurderingen om man må i bil for å komme til barnehage var krevende og vanskelig å undersøke. Det var noen tvilstilfeller om avgjørelse om man måtte sette seg i bil for å komme til barnehage eller ikke. Jeg satte denne grensen på 1 kilometer. Hadde jeg økt eller endret denne grensen litt kunne det fått store utslag for analysen. Dette kan være noe av grunnen til denne variabelens svake resultater. Videre er det uavklart og ikke blitt tatt hensyn til muligheten for å få barnehageplass hos den nærmeste barnehagen. Samtidig verdsetter aldersgrupper nærhet til barnehage ulikt. Jeg kunne dermed fått en høyere grad av representative estimater om jeg for eksempel hadde undersøkt hvordan husholdninger med barn i barnehagealder, verdsetter nærhet til barnehage.

7. Konklusjon

Jeg ønsket å analysere hvilken betydning utvalgte attributter har for boligprisen i Haugesund og Stavanger. Her vil jeg oppsummere hovedresultatene før jeg avslutningsvis vil gi noen forslag til videre forskning.

Analysen viser at vi med 95% sannsynlighet kan si at boligattributtene boligareal, boligalder, avstand til sentrum, tomteareal til selveide eneboliger har betydning for boligprisen i begge byene. Vi kan med høy sikkerhet si at boligalder har større betydning på boligprisen i Haugesund enn i Stavanger. Eldre boliger i Stavanger holder seg dermed bedre i verdi enn hva som er tilfellet i Haugesund.

Boligareal har større betydning for boligprisen i Stavanger enn i Haugesund, noe som ikke var overraskende. Videre kan vi konkludere med at boligenes beliggenhet i forhold til sentrum varierer med boligprisen og at avstand til sentrum har større betydning for boligprisen i Haugesund enn i Stavanger.

Videre kan vi konkludere med at de ulike boligtypene har betydning for boligprisen i begge byene.

Til slutt kan vi ikke si at tomtearealet til selveide eneboliger har større betydning for boligprisen i Haugesund enn Stavanger. Videre kan vi ikke påstå at variablene avstand til sjø og "må i bil til barnehage" har betydning for boligprisen verken i Haugesund eller Stavanger.

7.1 Forslag til videre forskning

Under samtale med eiendomsutvikler Olav Askeland i Haugesund og Stavanger, utdypet han at boligfinansieringsformen husbankfinansiering var spesielt viktig for boligkjøpere i Haugesund. Husbankfinansieringen kan ses på som attraktiv og gunstig for mange. Det kunne vært interessant å studere hvilken effekt dette har for boligprisen og boligkjøpere.

En annen ting som kunne vært interessant å undersøke er hvor de forskjellige aldersgruppene velger å bosette seg i byene og hvilke boligtyper de etterspør. Det pressede boligmarkedet i Stavanger gjør at eksisterende sentrumsnære boliger ofte vil gå i arv eller i familien ved eierskifte.

På litt lenger sikt vil Haugesund og Stavanger etter alt å dømme bli knyttet nærmere hverandre som følge av de store veiprosjektene i regionen. Det kan derfor være interessant å undersøke hvordan boligmarkedet både i Haugesund og Stavanger reagerer når veiprosjektene er ferdigstilt.

Litteraturliste

- Baum. C. F (2006). An Introduction to Modern Econometrics Using Stata.
- Byggenæringens landsforening, BNL (2010). Byggenæringen i Rogaland. Hentet 01.03.2014 fra,
http://www.bnl.no/getfile.php/Filer/Publikasjoner/BNL_Rogaland_web%282%29.pdf
- Byggenæringens landsforening, BNL (2014). Byggenæringen i Rogaland. Hentet 01.03.2014 fra,
<http://www.bnl.no/article.php?articleID=2379&categoryID=6>
- Dinside.no (2014). Fall i boligprisene. 15.01.2014 hentet fra,
<http://www.dinside.no/926037/boligprisene-ned-i-desember>
- Dinside.no. (2013). Stavanger dyreste by å bo i. Hentet 15.02.2014 fra,
<http://www.dinside.no/921053/stavanger-er-den-dyreste-byen-aa-etablere-seg-i>
- Eiendom Norge (2014) Boligprisutvikling. Hentet 09.03.2014 fra,
http://www.eff.no/filer/Boligpris%202014_januar_org01.pdf
- Eiendomsverdi.no (2014)
- Eiendomsverdi.no (2014) Fakta boligmarkedet. Hentet 20.02.2014 fra,
http://eiendomsverdi.no/app/Factsheets/Fakta_januar_2014.pdf
- Erik Bolstad. Postnummer i Haugesund og Stavanger. Hentet 01.02.2014 fra,
<http://www.erikbolstad.no/postnummer-koordinatar/kommune.php?kommunennummer=1106>
- Estate Nyheter (2013). Om Haugesund. Hentet 30.02.2014 fra,
<http://www.estatenyheter.no/index.php/component/content/article/5-nyheter/nyheter/1006-full-rulle-pa-haugalandet>
- Farsund (2012). Scenarier 2029. Boligmarkedet – drivkrefter, utviklingstrekk, og fremtidsperspektiver. IRIS.
- Google Maps (2014). Avstand, Haugesund-Stavanger. Hentet 16.02.2014 fra,
<https://www.google.no/maps/dir/Stavanger+Sentrum,+Stavanger/Haugesund/@59.189242,5.2306882,10z/data=!3m1!4b1!4m13!4m12!1m5!1m1!1s0x463a3535ddb0ef0f:0xa1e547ad9160937!2m2!1d5.7397907!2d58.9706621!1m5!1m1!1s0x463b9e2daa5b77ff:0xc124394a88e63a66!2m2!1d5.2679869!2d59.413581>
- Haugesund kommune (2014) Fakta om kommunen. Hentet 15.01.2014 fra,
<http://www.haugesund.kommune.no/om-kommunen>
http://statistikk.stavanger.kommune.no/generelt_01s.html
<http://www.nrk.no/rogaland/barnehageplass-utenfor-egen-bydel-1.8146537>
- Institute for digital research and education. Sammenligning av to regresjonskoeffisienter. Hentet 01.04.2014 fra,
<http://www.ats.ucla.edu/stat/stata/faq/compreg2.htm>

International institute of Stavanger. Boligmarkedet i Haugesund og Stavanger.

Hentet 05.02.2014 fra,

<http://www.scenarier2029.no/filer/bakgrunnsnotater/Boligmarkedet2012.pdf>

Knudsen J.P og Sødal. S (2010). *Økonomi og tid, 18 essays i Pufendorf-tradisjon.*

Midtbø. T (2012). Stata – en entusiastisk innføring. Universitetsforlaget

Midtbø. T. (2007). Regresjonsanalyse for samfunnsvitere. Universitetsforlaget

N. Gujarati og C. Porter (2010). *Essentials of Econometrics.* 4th edition.

Norges Taksteringsforbund. Arealbegrep. Hentet 16.02.2014 fra,

<http://www.boligtakst.as/arealbegrep.html>

Nrk.no (2012). Barnehageplass. Hentet 05.02.2014 fra,

<http://www.nrk.no/rogaland/barnehageplass-utenfor-egen-bydel-1.8146537>

Nrk.no (2012). Prisforskjeller. Hentet 20.01.2014 fra,

<http://www.nrk.no/rogaland/enorme-prisforskjeller-pa-bolig-1.8109099>

Nrk.no (2012). Prisforskjeller. Hentet 20.01.2014 fra,

<http://www.nrk.no/rogaland/store-prisforskjeller-pa-bolig-1.8268358>

Omdømmebarometeret utarbeidet av Ordkraft og utgitt av Hovedorganisasjonen VIRKE (2014)

Regjeringen.no (2002). Boligmarkedets virkemåte. Hentet 28.02.2014 fra,

<http://www.regjeringen.no/nb/dep/kmd/dok/nouer/2002/nou-2002-2/4.html?id=366185>

Robertsen, K (2013) Forelesningsnotater BE-409, Eiendomsverdi. Universitetet i Agder

Robertsen, K & Theisen, T (2011). The Impact of Financial Arrangements and Institutional Form on Housing Prices.

Skog.O (1998). Å forklare sosiale fenomener. En regresjonsbasert tilnærming. 1.Utgave.

Skog O. (2004). Å forklare sosiale fenomener. En regresjonsbasert tilnærming. 2.Utgave.

Statistisk Sentralbyrå (2014a). Boliger, etter bygningstype og antall rom (K).

Hentet 18.01.2014 fra,

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

Statistisk Sentralbyrå (2014b). Selveierboliger. Gjennomsnittlig

kvadratmeterpris og antall omsetninger (K). Hentet 20.01.2014 fra,

<https://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=SelveierKVp2&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=priser-og-prisindekser&KortNavnWeb=bpi&StatVariant=&checked=true>

Statistisk Sentralbyrå (2014c). Hovedposter fra ligninga, etter størrelsen på skattepliktig nettoformue. Bosatte personer 17 år og eldre. Gjennomsnitt (kr) (K). Hentet 05.05.2014 fra,

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

Statistisk Sentralbyrå (2014d). Virksomheter, etter næring (SN2007) og antall ansatte (K). Hentet 25.01.2014 fra,

<https://www.ssb.no/statistikkbanken/selectvarval/Define.asp?subjectcode=&ProductId=&MainTable=NYBedrifter&nvl=&PLanguage=0&nyTmpVar=true&CMSSubjectArea=virksomheter-foretak-og-regnskap&KortNavnWeb=bedrifter&StatVariant=&checked=true>

Stavanger kommune (2014). Kommuneplanen. Hentet 03.03.2014 fra,

<https://www.stavanger.kommune.no/PageFiles/4409/kommplanen2010-2025.pdf>

Stavanger Kommune (2014). Stavanger-statistikken. Hentet 10.02.2014 fra,

<http://statistikk.stavanger.kommune.no>

Stock. H & Watson.W (2003). *Introduction to Econometrics*

Thrane. C (2003). Regresjonsanalyse i praksis. Høyskoleforlaget

TV-Haugaland (2013). Boligmarked i Haugesund. Hentet 18.01.2014 fra,

<http://tvh.no/serier/episode/horison-2-oktober-2013-boligmarkedet-og-meglerbransjen>

Zikmund, Babin, Carr, Griffin (2013). *Business Research Methods*. 9th edition.

Vedlegg

// Do-fil Masteroppgave Haugesund

*Lager navn til variablene

```
rename prom boligareal
rename salgsdato salgsår
rename bygger byggeår
rename alderbolig boligalder
rename avstandtilsentrum sentrum
rename avstandtilsj sjøen
rename mibiltilbarnehage barnehage
rename tomt toa
rename salgspris boligpris
```

*Fjerner variabler

```
drop adresse
drop prisant
drop regdata
drop bta
drop m2prom
drop omshast
drop v1
```

*Renser dataene

```
drop if byggeår == 0
drop if byggeår == .
drop if boligpris == .
drop if boligpris == 0
drop if boligareal == 0
drop if boligareal == .
drop if toa == 0
drop if toa == .
drop if fellesgjeld < 1000
```

*Lager variabler numeriske

```
encode eieform, gen(eieform1)
encode boligtype, gen(boligtype1)
```

*Lager dummyvariabler for eieform

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

```
replace selveier = 0
replace selveier = 1 if eieform1 == 2
replace borettslag = 1 if eieform1 == 1
```

*Lager dummyvariabler for boligtyper

```
gen leilighet = 0
gen enebolig = 0
generate rekkehus = 0
gen tomannsbolig = 0
```

```
replace leilighet = 1 if boligtype1 == 2
replace enebolig = 1 if boligtype1 == 1
replace rekkehus = 1 if boligtype1 == 3
replace tomannsbolig = 1 if boligtype1 == 4
```

*Årsdummyer

```
gen salgsår09 = 0
replace salgsår09 = 1 if salgsår == 2009
gen salgsår10 = 0
replace salgsår10 = 1 if salgsår == 2010
gen salgsår11 = 0
replace salgsår11 = 1 if salgsår == 2011
gen salgsår12 = 0
replace salgsår12 = 1 if salgsår == 2012
gen salgsår13 = 0
replace salgsår13 = 1 if salgsår == 2013
```

*Lager produktvariabel for enebolig, selveier og tomt med boligpris

```
gen tomtenebolig = toa*enebolig*selveier
```

*Deskriptiv statistikk

```
sum boligpris fellesgjeld boligareal boligalder sentrum sjøen barnehage selveier
borettslag leilighet enebolig rekkehus tomannsbolig salgsår09 salgsår10
salgsår11 salgsår12 salgsår13
```

```
sum tomt if enebolig == 1 & selveier == 1
```

*Histogrammer

```
histogram boligpris, normal freq
histogram boligareal, normal freq
histogram tomt, normal freq
histogram boligalder, normal freq
histogram sentrum, normal freq
histogram sjøen, normal freq
histogram tomtenebolig, normal freq
```

histogram salgsår, normal freq

*Lager frekvenstabell for dummyvariabler

tabulate barnehage

tabulate boligtype1

tabulate eieform1

tabulate salgsår

*Korr matrise

pwcorr boligpris leilighet enebolig rekkehus tomannsbolig boligareal boligalder
sjøen sentrum barnehage selveier borettslag fellesgjeld tomtenebolig salgsår09
salgsår10 salgsår11 salgsår12 salgsår13, sig star(5)

*Regresjon med en variabel (boligareal)

reg boligpris boligareal

*Normalskråplott for bivariat reg.analyse

predict restledd1, resid

pnorm restledd1

*Plott av data med regresjonslinjen

twoway (scatter boligareal boligpris)(lfit boligareal boligpris)

*Regresjon med alle variabler untatt (borettslag salgsår09, fellesgjeld,
tomannsbolig)

reg boligpris boligareal leilighet enebolig rekkehus boligalder selveier sjøen
sentrum barnehage tomtenebolig salgsår10 salgsår11 salgsår12 salgsår13

*Tester for multikollinearitet

vif

predict restledd2, resid

pnorm restledd2 (normalskråplott)

rvfplot (residualplott)

*Semilogaritmisk regresjonsanalyse alle variabler

gen ln_boligpris = ln(boligpris)

reg ln_boligpris boligareal leilighet enebolig rekkehus boligalder selveier sjøen
sentrum barnehage tomtenebolig salgsår10 salgsår11 salgsår12 salgsår13

predict restledd3, resid

pnorm restledd3 (normalskråplott)

rvfplot (residualplott)

*Dobbeltlogaritmisk regresjonsanalyse uten tomtenebolig

```
gen ln_boligpris =ln(boligpris)
gen ln_boligareal =ln(boligareal)
gen ln_boligalder =ln(boligalder)
gen ln_sentrum =ln(sentrum)
gen ln_sjøen =ln(sjøen)
```

```
reg ln_boligpris ln_boligareal ln_boligalder ln_sentrum ln_sjøen barnehage
selveier leilighet enebolig rekkehus salgsår10 salgsår11 salgsår12 salgsår13
```

```
predict restleddobbel, resid
pnorm restleddobbel
rvfplot
```

// Do-fil Masteroppgave Stavanger

```
*Lager navn til variablene
rename prom boligareal
rename salgsdato salgsår
rename bygger byggeår
rename alderbolig boligalder
rename avstandtilsentrum sentrum
rename avstandtilsj sjøen
rename mibiltilbarnehage barnehage
rename tomt toa
rename salgspris boligpris
```

```
*Fjerner variabler
drop adresse
drop prisand
drop regdata
drop bta
drop m2prom
drop omshast
drop antall
```

```
*Renser dataene
drop if byggeår == 0
drop if byggeår == .
drop if boligpris == .
drop if boligpris == 0
drop if boligareal == 0
drop if boligareal == .
drop if toa == 0
drop if toa == .
```

drop if fellesgjeld <1000

*Lager variabler numeriske
encode eieform, gen(eieform1)
encode boligtype, gen(boligtype1)

*Lager dummyvariabler for eieform
gen selveier = 0
gen borettslag = 0

replace selveier = 0
replace selveier = 1 if eieform1 == 2
replace borettslag = 1 if eieform1 == 1

*Lager dummyvariabler for boligtyper
gen leilighet = 0
gen enebolig = 0
generate rekkehus = 0
gen tomannsbolig = 0

replace leilighet = 1 if boligtype1 == 2
replace enebolig = 1 if boligtype1 == 1
replace rekkehus = 1 if boligtype1 == 3
replace tomannsbolig = 1 if boligtype1 == 4

*Årsdummyer
gen salgsår09 = 0
replace salgsår09 = 1 if salgsår == 2009
gen salgsår10 = 0
replace salgsår10 = 1 if salgsår == 2010
gen salgsår11 = 0
replace salgsår11 = 1 if salgsår == 2011
gen salgsår12 = 0
replace salgsår12 = 1 if salgsår == 2012
gen salgsår13 = 0
replace salgsår13 = 1 if salgsår == 2013

*Lager produktvariabel for enebolig, selveier og tomt med boligpris
gen tomtenebolig = toa*enebolig*selveier

*Deskriptiv statistikk
sum boligpris fellesgjeld boligareal boligalder sentrum sjøen barnehage selveier
borettslag leilighet enebolig rekkehus tomannsbolig salgsår09 salgsår10
salgsår11 salgsår12 salgsår13

sum tomt if enebolig == 1 & selveier == 1

*Histogrammer

histogram boligpris, normal freq
histogram boligareal, normal freq
histogram tomt, normal freq
histogram boligalder, normal freq
histogram sentrum, normal freq
histogram sjøen, normal freq
histogram tomtenebolig normal freq
histogram salgsår, normal freq

*Lager frekvenstabell for dummyvariabler

tabulate barnehage
tabulate eieform1
tabulate boligtype1
tabulate salgsår

*Korr matrise

pwcorr boligpris leilighet enebolig rekkehus tomannsbolig boligareal boligalder
sjøen sentrum barnehage selveier borettslag fellesgjeld tomtenebolig salgsår09
salgsår10 salgsår11 salgsår12 salgsår13, sig star(5)

*Regresjon med en variabel (boligareal)

reg boligpris boligareal

*Normalskråplott for bivariat reg.analyse

predict restledd1, resid
pnorm restledd1

*Plott av data med regresjonslinjen

twoway (scatter boligareal boligpris)(lfit boligareal boligpris)

*Regresjon med alle variabler untatt (borettslag salgsår09, fellesgjeld,
tomannsbolig)

reg boligpris boligareal leilighet enebolig rekkehus boligalder selveier sjøen
sentrum barnehage tomtenebolig salgsår10 salgsår11 salgsår12 salgsår13

*Tester for multikollinearitet

vif

predict restledd2, resid
pnorm restledd2 (normalskråplott)
rvfplot (residualplott)


```
*Semilogaritmisk regresjonsanalyse alle variabler unntatt toa
gen ln_boligpris = ln(boligpris)
reg ln_boligpris boligareal leilighet enebolig rekkehus boligalder selveier sjøen
sentrum barnehage tomtenebolig salgsår10 salgsår11 salgsår12 salgsår13
```

```
predict restledd3, resid
pnorm restledd3 (normalskråplott)
rvfplot (residualplott)
```

```
*Dobbeltlogaritmisk regresjonsanalyse uten tomtenebolig
```

```
gen ln_boligpris =ln(boligpris)
gen ln_boligareal =ln(boligareal)
gen ln_boligalder =ln(boligalder)
gen ln_sentrum =ln(sentrum)
gen ln_sjøen =ln(sjøen)
```

```
reg ln_boligpris ln_boligareal ln_boligalder ln_sentrum ln_sjøen barnehage
selveier leilighet enebolig rekkehus salgsår10 salgsår11 salgsår12 salgsår13
```

```
predict restleddobbel, resid
pnorm restleddobbel
rvfplot
```

// Do-fil Masteroppgave, sammenligning av regresjonene Haugesund og Stavanger

```
*Lager navn til variablene
rename prom boligareal
rename salgsdato salgsår
rename bygger byggeår
rename alderbolig boligalder
rename avstandtilsentrum sentrum
rename avstandtilsj sjøen
rename mibiltilbarnehage barnehage
rename tomt toa
rename salgspris boligpris
```

```
*Fjerner variabler
drop prisand
drop regdato
drop bta
drop m2prom
drop omshast
drop antall
```

```
*Renser dataene
drop if byggeår == 0
drop if byggeår == .
drop if boligpris == .
drop if boligpris == 0
drop if boligareal == 0
drop if boligareal == .
drop if toa == 0
drop if toa == .
drop if fellesgjeld <1000
```

```
*Lager variabler numeriske
encode eieform, gen(eieform1)
encode boligtype, gen(boligtype1)
```

```
*Lager dummyvariabler for eieform
gen selveier = 0
gen borettslag = 0
```

```
replace selveier = 0
replace selveier = 1 if eieform1 ==2
replace borettslag = 1 if eieform1 == 1
```

```
*Lager dummyvariabler for boligtyper
gen leilighet = 0
gen enebolig = 0
gen rekkehus = 0
gen tomannsbolig = 0
```

```
replace leilighet = 1 if boligtype1 == 2
replace enebolig = 1 if boligtype1 == 1
replace rekkehus = 1 if boligtype1 == 3
replace tomannsbolig = 1 if boligtype1 == 4
```

```
*Årsdummyer
gen salgsår09 = 0
replace salgsår09 = 1 if salgsår == 2009
gen salgsår10 = 0
replace salgsår10 = 1 if salgsår == 2010
gen salgsår11 = 0
replace salgsår11 =1 if salgsår == 2011
gen salgsår12 = 0
replace salgsår12 = 1 if salgsår == 2012
gen salgsår13 = 0
```

replace salgsår13 = 1 if salgsår == 2013

*Lager produktvariabel for enebolig, selveier og tomt med boligpris
gen tomtenebolig = toa*enebolig*selveier

*Lager variabler numeriske
encode adresse, gen(adresse1)

*Lager dummyvariabler for byene
gen Stavanger = 0
gen Haugesund = 0

replace Stavanger = 1 if adresse1 == 2
replace Haugesund = 1 if adresse1 == 1

*Produktvariabel boligareal
replace Stavanger = 1 if adresse == "Stavanger"
replace Stavanger = 0 if adresse == "Haugesund"

*Genererer produktvariabler
generate stvgareal = Stavanger*boligareal
generate stvgsentrum = Stavanger*sentrum
generate stvgalder = Stavanger*boligalder
generate stvgtomt = Stavanger*tomtenebolig

*Interaksjonsmodell
reg boligpris Stavanger stvgareal boligareal stvgsentrum sentrum stvgalder
boligalder stvgtomt tomtenebolig, robust

* F-test
testparm Stavanger stvgareal stvgsentrum stvgalder stvgtomt