

En av fem leiligheter på Lund er solgt med tap i perioden 2010 - 2018

En analyse av prisutvikling på leiligheter i bydelen Lund.

.

HENRIETTE KRISTIANSEN

VEILEDER

Anne Wenche Emblem

Universitetet i Agder, 2019

Handelshøyskolen ved UIA

Forord

Denne masteroppgaven er skrevet våren 2019 og er den avsluttende delen av masterstudiet i økonomi og administrasjon. Formålet med oppgaven er å tilegne seg spesialisert kunnskap innenfor fagområdet og bruke teori og metode til å løse en problemstilling. Masteroppgaven utgjør 30 studiepoeng og går over ett semester.

Denne masteroppgaven er skrevet innenfor temaet eiendomsøkonomi og tar for seg leilighetsmarkedet i bydelen Lund i Kristiansand. Hensikten med denne oppgaven er å undersøke hva som bestemmer leilighetsprisen, samt finne ut om det er noen fellestrekk mellom de leilighetene som selges med tap i bydelen. Jeg valgte å undersøke bydelen Lund fordi det er en bydel hvor det har vært mye byggesaktivitet de siste årene, samt at jeg hadde kunnskap om at det fantes leiligheter i bydelen som ble solgt med tap. Dette motiverte meg til å undersøke om det var noen fellestrekk ved de leilighetene som selges med tap.

Arbeidet med masteroppgaven har vært en lærerik, men utfordrende og krevende prosess. Oppgaven har gitt meg mer kunnskap om eiendomsøkonomi som var et felt jeg var interessert i på forhånd, samt inngående innsikt i leilighetsmarkedet på Lund.

Jeg ønsker å takke veileder førsteamanuensis Anne Wenche Emblem for god oppfølging, nyttige tilbakemeldinger og gode forslag i løpet av prosessen. Jeg vil også rette en takk til professor Jochen Jungeilges og professor Valeriy Zakamulin ved Universitetet i Agder for emnene SE-506 og BE-506 som var lærerike på mange måter. I disse emnene fikk jeg både trening i å skrive oppgaver på egenhånd, samt kunnskap og god oppfølging om ulike statistiske program og analysemetoder. Disse emne la et veldig godt grunnlag for en kvantitativ analyse.

Til slutt ønsker jeg å takke kjæreste, familie og venner for motivasjon og råd underveis, samt korrekturlesing.

Henriette Kristiansen

Kristiansand, 28.mai 2019

Sammendrag

Leiligheter er heterogene goder som alle er unike. Det er sammensetninger av egenskaper ved boligen, samt forbrukernes betalingsvilje for de ulike sammensetningene av egenskaper som bestemmer boligprisen.

Hensikten med denne oppgaven er å undersøke hvilke egenskaper ved leiligheter i bydelen Lund i Kristiansand som påvirker leilighetens totalpris. I tillegg til å undersøke hvilke attributter som påvirker leilighetsprisen har jeg undersøkt om de leilighetene som selges med tap skiller seg ut fra de leilighetene som selges med gevinst. I dag er det mye fokus på prisveksten på boligmarkedet, derimot er det lite snakk om at det faktisk er en risiko for å selge bolig med tap. I denne analysen har jeg kun undersøkt leiligheter, og både leiligheter i hus og blokkleiligheter er inkludert.

Analysen er basert på sekundærdata fra Eiendomsverdi AS. Datagrunnlaget for å besvare problemstillingen består av leiligheter omsatt i bydelen Lund i perioden 2010 til 2018, og analysen er basert på 1871 observasjoner. En dobbellogaritmisk funksjonsform var foretrukket, og med denne modellen ble 86,18% av variasjonen i leilighetens totalpris forklart. I regresjonsanalysen finner jeg blant annet en positiv signifikant sammenheng mellom totalpris og leilighetens størrelse, og en negativ signifikant sammenheng mellom totalpris og avstand til sentrum, og totalpris og alder på leiligheten.

Del to av analysen baserer seg på observasjoner der leiligheter er omsatt mer enn en gang i løpet av perioden. I datasettet er det til sammen registrert 657 observasjoner som enten er solgt med gevinst eller med tap i løpet av perioden. Om lag 20% av leilighetene i bydelen, i den aktuelle perioden er solgt med tap. Analysen indikerer at leiligheter solgt med tap har høyere gjennomsnittlig omsetningshastighet, negativt gjennomsnittlig avvik mellom salgspris og prisantydning, og at det er kortere tid mellom to salg i de tilfeller hvor leiligheten er solgt med tap enn når leiligheten selges med gevinst. T-testen viste også at gjennomsnittlig leilighetsalder for leiligheter solgt med tap var lavere enn for leiligheter som selges med gevinst. Det kom også frem at andelen tap blant leiligheter som var yngre enn 10 år, var betydelig høyere enn andelen tap blant alle aldersgrupper. Hele en av fire leiligheter som var yngre enn 10 år har vært solgt med tap i perioden.

Innholdsfortegnelse

Forord	i
Sammendrag	ii
Figuroversikt	v
Tabelloversikt	vii
Liste over vedlegg	ix
1. Innledning	1
2. Bakgrunn	2
2.1. Kristiansand kommune	2
2.2. Bydel Lund	4
2.3. Boligmarkedet i Norge	7
2.4. Boligmarkedet i Kristiansand	8
2.4.1. Nye leiligheter	9
2.4.2. Nybygg og planlagte prosjekt	11
3. Teori	13
3.1. Prisdannelse i boligmarkedet	14
3.2. Den hedonistiske metoden	16
3.3. Alonso – Muth – Mills-modellen	20
3.4. Hypotesetesting	22
3.4.1. Hypoteser som omhandler leilighetsprisen	22
3.4.2. Hypoteser som omhandler tap ved salg	23
4. Innhenting og bearbeiding av datamaterialet	25
4.1. Sekundærdata	25
4.2. Innhenting av data	26
4.3. Bearbeiding av datamaterialet	27
4.3.1. Inndeling av områder i bydelen	28
4.3.2. Avstand til sentrum	28
4.3.3. Beregning av gevinst og tap	29
4.4. Datarensing	29
4.5. Deskriptiv statistikk for hele datasettet	32
4.5.1. Korrelasjon	36
4.6. Deskriptiv statistikk for leiligheter solgt mer enn en gang i perioden	37
4.7. Oppsummering av analyserte variabler	42
4.8. Utelatte variabler	42
5. Metode	44
5.1. Økonometrisk modell	44

5.1.1. Regresjonsanalyse	44
5.1.2. Spesifisering av funksjonsform	45
5.1.3. Forutsetninger for regresjonsanalyse	47
5.1.4. Modellens forklaringskraft	48
5.1.5. Hypotesetesting	48
5.2. T-test til sammenlikning av to grupper	50
6. Analyse ved hjelp av den hedonistiske metoden	51
6.1. Valg av forklaringsvariabler	51
6.2. Estimering av regresjonsmodellene	53
6.2.1. Lineær funksjonsform	53
6.2.2. Semilogaritmisk funksjonsform	57
6.2.3. Dobbellogaritmisk funksjonsform med totalpris som avhengig variabel	60
6.3. Valg av regresjonsmodell	63
6.4. Hypotesetesting	64
7. Analyse av leiligheter som er solgt mer enn en gang	66
7.1. Attributtenes påvirkning på leiligheter solgt mer enn en gang	68
7.2. Sammenlikning av regresjon for to grupper	69
7.3. T-test om signifikante forskjeller i gjennomsnittsverdier	71
7.4. Hypotesetesting	73
8. Videre analyse og diskusjon av resultater	75
8.1. Leilighetens størrelse	76
8.2. Virkningen av økt alder på leilighet	78
8.3. Avstand til sentrum	79
8.4. Betydningen av parkeringsplass og heis	80
8.5. Leiligheter som er solgt med tap i perioden 2010 – 2018.	81
8.5.1. Området Tobienborg	82
8.5.2. Leiligheter som er yngre enn 10 år	84
8.5.3. Leiligheter i første etasje	85
8.5.4. Omsetningshastighet og prisantydning	85
8.6. Tilbud og etterspørsel	86
8.7. Relevansen og viktigheten rundt dette temaet	88
8.8. Analysens svakheter	89
8.9. Forslag til videreføring av oppgaven	90
9. Konklusjon	91
10. Referanser	93
Vedlegg	96

Figuroversikt

Figur 2.1 Befolkningsutvikling til venstre og årlig prosentvis befolkningsendring i Kristiansand i perioden 1990 – 2018 til høyre (Statistisk Sentralbyrå, 2018b).	3
Figur 2.2 Andel sysselsatte i Kristiansand sammenlignet med andel sysselsatte på landsbasis (Statistisk Sentralbyrå, 2019d)	3
Figur 2.3 Bydelen Lund med postnummer. Kart: Google Maps (2019) redigert med utgangspunkt i Bolstad (2019).	5
Figur 2.4 Befolkningsutvikling for bydel Lund til venstre og årlig prosentvis befolkningsendring i bydelen i perioden 2010 -2017 til høyre (Kristiansand Kommune, 2018b).	6
Figur 2.5 Aldersfordeling i bydel Lund, sentrum(Kvadraturen), samt for kommunen totalt (Kristiansand Kommune, 2018c)	6
Figur 2.6 Utvikling i gjennomsnittlig kvadratmeterpris for leiligheter i perioden 2010-2018 (Eiendomsverdi AS, 2019b)	8
Figur 2.7 Utviklingen i gjennomsnittlig kvadratmeterpris til venstre og årlig prosentvis endring i denne til venstre (Eiendomsverdi AS, 2019b).	9
Figur 2.8 Oversikt over endringer i boligmassen i Kristiansand i perioden 2013 -2018. Figur laget med utgangspunkt i Kristiansand Kommune (2017).	9
Figur 2.9 Antall nybygde eneboliger og blokkleiligheter i Kristiansand, samt totalt antall nybygg bygd i perioden 1993 – 2016 (Statistisk Sentralbyrå, 2018a).	10
Figur 2.10 Store byggeprosjekt på Lund, redigert skjermdump fra Google Maps	13
Figur 3.1 Tilbud og etterspørsel i boligmarkedet	15
Figur 3.2 Effekt av økt tilbud og etterspørsel	15
Figur 3.3 Monosentrisk by	20
Figur 3.4 Husleien (DiPasquale & Wheaton, 1996)	21
Figur 4.1 Områdekart (redigert skjermdump fra Google Maps (2019))	28
Figur 4.2 Box plots: øverst fra venstre: totalpris, størrelse, boligalder, avstand til sentrum	32

Figur 4.3 Histogram for totalpris	34
Figur 4.4 Histogram for primærrom	35
Figur 4.5 Histogram for boligalder	35
Figur 4.6 Histogram for avstandsvariabelen	36
Figur 4.7 Histogram for omsetningshyppighet	36
Figur 4.8 Histogram over endring i totalpris mellom to salgstidspunkt	38
Figur 4.9 Histogram for totalpris, leiligheter solgt med gevinst til venstre og leiligheter solgt med tap til høyre	39
Figur 4.10 Primærrom, leiligheter solgt med gevinst til venstre og leiligheter solgt med tap til høyre	39
Figur 4.11 Antall soverom, leiligheter solgt med gevinst til venstre, leiligheter solgt med tap til høyre	40
Figur 4.12 Boligalder for leiligheter solgt med gevinst til venstre og solgt med tap til høyre	40
Figur 4.13 Avvik mellom salgspris og prisantydning, gevinst til venstre og tap til høyre	41
Figur 4.14 Omsetningshastighet, gevinst til venstre og tap til høyre	41
Figur 6.1 Histogram for feilledd	57
Figur 6.2 Normalfordelingsplott for feilledd i lineær funksjonsform	57
Figur 6.3 Normalfordelingsplott for feilledd i semilogaritmisk funksjonsform	60
Figur 6.4 Normalfordelingsplott for feilledd i dobbellogaritmisk modell	63
Figur 7.1 Andel leiligheter solgt med tap per område av totale salg per område	67
Figur 8.1 Virkningen av økt alder på leilighetens totalpris	78
Figur 8.2 Områdekart med estimerte prisnivåer	79

Tabelloversikt

Tabell 2.1 Leilighetsmasse og nye leiligheter per år (Eiendomsverdi AS, 2019b; Kristiansand Kommune, 2017)	11
Tabell 4.1 Endringer gjort for å oppnå et komplett datasett	31
Tabell 4.2 Deskriptiv statistikk for bydel Lund	33
Tabell 4.3 Korrelasjonsmatrise (n =1871 observasjoner)	37
Tabell 4.4 Deskriptiv statistikk for leiligheter som er solgt mer enn en gang i perioden 2010 - 2018 (n=657)	38
Tabell 4.5 Variabler med tilhørende Stata-variabel	42
Tabell 6.1 Modellens forklaringskraft ved ulike alternative variabler inkludert, n=1872	52
Tabell 6.2 Lineær multippel regresjon	54
Tabell 6.3 Multippel lineær regresjon med robuste standardfeil	56
Tabell 6.4 Regresjon med semilogaritmisk funksjonsform	58
Tabell 6.5 Robust regresjon med semilogaritmisk funksjonsform	59
Tabell 6.6 Regresjon med dobbellogaritmisk funksjonsform	61
Tabell 6.7 Regresjon med robuste standardfeil, dobbellogaritmisk funksjonsform	62
Tabell 7.1 Fordeling av tap, oppgitt i prosent av totale tap	67
Tabell 7.2 Andel leiligheter solgt med gevinst og tap per år, oppgitt i prosent av totale salg per år	68
Tabell 7.3 Antall år mellom omsetninger, oppgitt i prosent av totale observasjoner med respektivt antall	68
Tabell 7.4 Regresjonsmodell for leiligheter med registrert gevinst eller tap	69
Tabell 7.5 Interaksjonsmodell	70
Tabell 7.6 Resultat fra T-tester.	72
Tabell 8.1 Endring i pris ved økning på 1 kvm for gjennomsnittsleiligheten i hver gruppe	77
Tabell 8.2 Endring i pris ved økning på 1 kvm for medianleiligheten i hver gruppe	77
Tabell 8.3 Endring i pris ved økning på 1 kvm for gjennomsnittsleiligheten i hver gruppe	77

Tabell 8.4 Endring i pris ved økning på 1 kvm for medianleiligheten i hver gruppe	77
Tabell 8.5 Gjennomsnittlig totalpris for leiligheter som har primærrom mindre enn 100 m ²	87
Tabell 8.6 Lønnsomhet ved å eie leilighet (Smarte penger, 2018)	88

Liste over vedlegg

Vedlegg 1: Refleksjonsnotat	96
Vedlegg 2: Adresser på 4633 som tilhører bydelen Lund	99
Vedlegg 3: Utvikling i boligbygging i Kristiansand, sortert etter boligtype.	99
Vedlegg 4: Arbeidsledighet	99
Vedlegg 5: Regresjonsmodeller for små og store leiligheter	100
Vedlegg 6: Stata-kommandoer	105

1. Innledning

Et raskt søk på internett viser at boligpriser er noe som interesserer og engasjerer.

Fædrelandsvennen har månedlige artikler med overskrifter som «Boligprisene for januar: solgt rekordmange boliger», «Boligprisene sank så vidt i Desember: sterkere enn ventet», «Ferske mars-tall: så mye økte boligprisene i Kristiansand» og «Leiligheten kostet 520.000 kroner i 1998. Disse boligkjøperne kan le hele veien til banken». Men det er et viktig tema som uteblir – det er en mulighet, faktisk ganske stor, for at boligen selges med tap.

Bydelen Lund i Kristiansand er i stadig utvikling, og i løpet av noen få år har antall leiligheter økt betraktelig. Og flere skal det bli nå som store leilighetsprosjekt er igangsatt. Å vite hva som påvirker boligprisen i bydelen kan være viktig for både privatpersoner som ønsker å selge med gevinst og utbyggere som ønsker at alle leiligheter i deres leilighetsprosjekt skal bli solgt.

De fleste tenker at boligkjøp er en god investering fordi det er så stort fokus på prisveksten i dagens boligmarked. Det er derimot få som snakker om hvor mange boliger som faktisk selges med tap. Schultz og Kaspersen (2016) skriver at så mange som en tredjedel av boligene selges med tap. I et samfunn hvor det er økende fokus på å øke folks økonomiske forståelse, er det også viktig å snakke om at boliger kan selges med tap. Dagens boligmarked er karakterisert av høye boligpriser som betyr at de fleste som kjøper bolig tar opp et høyt lån. Derfor er det viktig å være klar over at man også kan tape penger ved boligsalg. I denne analysen skal jeg undersøke om det er noen egenskaper ved leiligheter og ved salgsprosessen av leiligheter som ser ut til å påvirke om leiligheten selges med tap.

I denne avhandlingen skal jeg gjøre en grundigere undersøkelse av bydelen Lund i Kristiansand og søker å finne svar på problemstillingen: **Hvilke attributter påvirker leilighetsprisen i bydel Lund i Kristiansand? Og finnes det noen fellestrekk mellom de leilighetene som er solgt med tap i perioden?** Med tap menes det her en leilighet som er solgt med lavere totalpris enn ved forrige omsetning.

For å svare på problemstillingen brukes sekundærdata om leiligheter i bydelen samlet inn av Eiendomsverdi til å først undersøke hvilke egenskaper som bestemmer leilighetsprisen i bydelen. Deretter er enkeltleiligheter i bydelen undersøkt grundigere og leiligheter solgt med tap er sammenliknet med leiligheter som er solgt med gevinst. Undersøkellesperioden er satt til å være 1.1.2010 til 31.12.2018.

For å bestemme hvilke egenskaper som påvirker leilighetsprisen i bydelen benyttes regresjon i henhold til den hedonistiske metoden. For å sammenlikne leiligheter som selges med tap og leiligheter som selges med gevinst brukes en t-test for å undersøke om det er signifikante forskjeller mellom gjennomsnittsverdiene til de to gruppene.

I neste kapittel vil nyttig bakgrunnsinformasjon om bydelen og boligmarkedet presenteres. Deretter presenteres relevant teori, og de modeller som skal brukes for å svare på problemstillingen. I kapittel 4 er datainnsamlingsprosessen beskrevet. Deretter følger et metodekapittel før resultatene presenteres og diskuteres. Tilslutt presenteres en konklusjon.

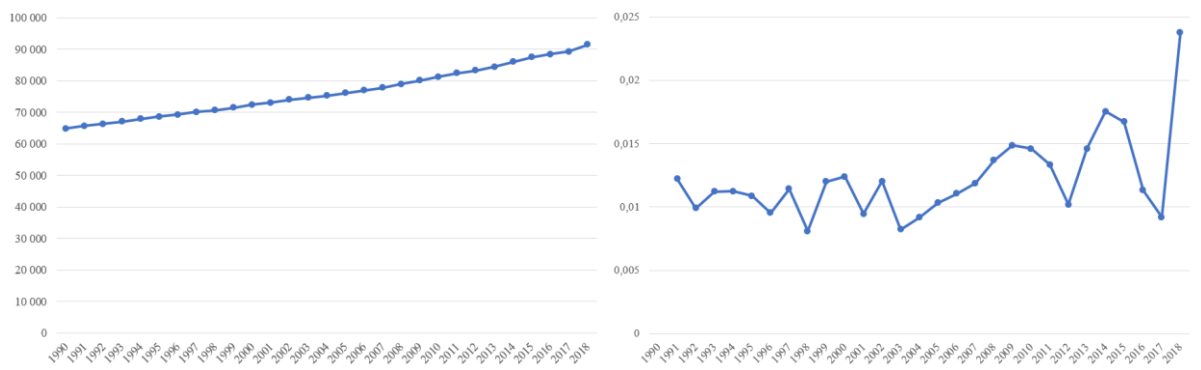
2. Bakgrunn

I dette kapitlet presenteres først litt bakgrunnsinformasjon om Kristiansand kommune, samt at bydelen Lund defineres og presenteres. Deretter gis en kort innføring i boligmarkedet i Norge, før det gis et innblikk i boligmarkedet og boligsammensetningen i Kristiansand og i bydel Lund.

2.1. Kristiansand kommune

Kristiansand kommune ligger i Vest-Agder og er den sjette største kommunen målt etter innbyggertall. Det bor om lag 92 000 mennesker i Kristiansand. (Statistisk Sentralbyrå, 2018b). 1.januar 2020 blir Kristiansand, Søgne og Songdalen slått sammen til en kommune. Det tas ikke hensyn til dette i denne analysen.

Kristiansand kommune har i perioden 1990 – 2018 hatt en relativt stabil befolkningsvekst med en årlig økning på mellom om lag 1% til 1,5%. Dette er illustrert i figur 2.1. Mellom 2017 og 2018 økte kommunens innbyggertall med 2,4%. Den høye befolkningsveksten skyldes en ordning der kommunen ga studenter som meldte flytting gratis busskort. 1930 studenter meldte flytting, mange av disse bodde allerede i byen men folkeregistrert i sin hjemkommune (Kjetil Reite, 2018). I følge Statistisk Sentralbyrå (2018b) var total befolkningsvekst i kommunen fra 2017 til 2018 på 2172. Det vil si at den reelle økningen i befolkningen var veldig lav sammenliknet med tidligere år.

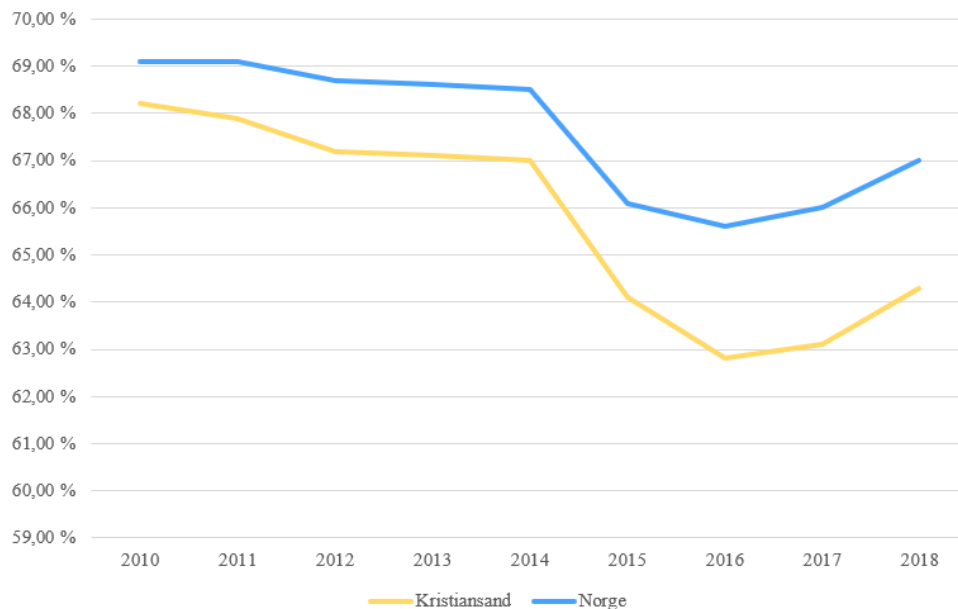


Figur 2.1 Befolkningsutvikling til venstre og årlig prosentvis befolkningsendring i Kristiansand i perioden 1990 – 2018 til høyre (Statistisk Sentralbyrå, 2018b).

Kristiansand Kommune (2018a) har i kommuneplanen antatt en årlig befolkningsvekst på ca. 1%, dette tilsvarer en økning på 800 – 1000 mennesker i året. Dette stemmer overens med det som er presentert i figur 2.1 som viser at befolkningsveksten i kommunen har vært relativt stabil og på mellom 1% og 1,5%.

Syssetting og arbeidsledighet

I 2018 utgjorde andelen sysselsatte i Kristiansand 64,3%. Dette er 2,7% lavere enn andelen sysselsatte på landsbasis. Andelen sysselsatte i Kristiansand i 2018 var også lavere enn andelen i andre store byer som Oslo, Stavanger, Bergen, Trondheim og Tromsø (Statistisk Sentralbyrå, 2019d).



Figur 2.2 Andel sysselsatte i Kristiansand sammenlignet med andel sysselsatte på landsbasis (Statistisk Sentralbyrå, 2019d)

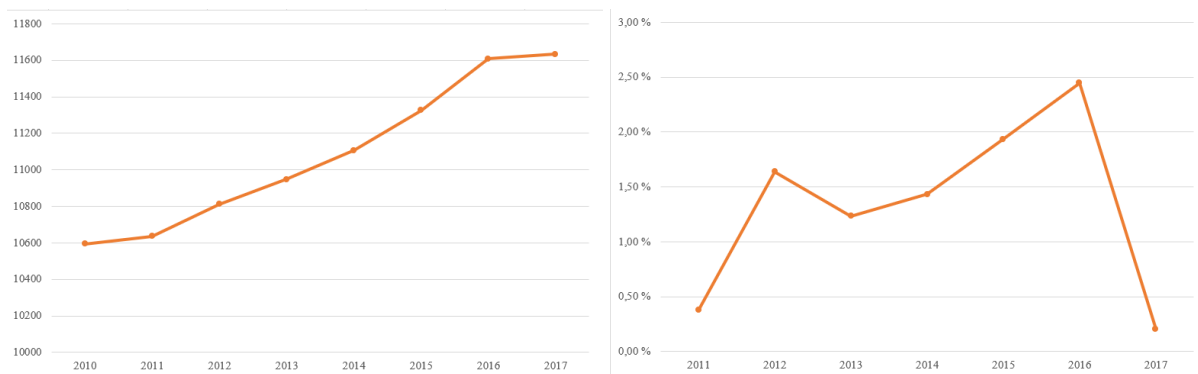
Figur 2.2 viser utviklingen i andelen sysselsatte blant personer i aldersgruppen 15-74år i Kristiansand og på landsbasis i perioden 2010 – 2018. Mellom 2014 og 2015 falt sysselsettingen både i Norge og i Kristiansand, dette skyldes antakelig at oljeprisen stupte og mange i oljebransjen mistet jobben. Sysselsettingsgraden i Kristiansand falt mer enn sysselsettingsgraden på landsbasis i perioden 2014-2016, dette kan skyldes at det var en stor andel sysselsatte i oljerelevante stillinger i Kristiansand. Både på landsbasis og i Kristiansand har sysselsettingen vært stigende siden den nådde bunnen i 2016.

2.2. Bydel Lund

Øst for Kristiansand sentrum ligger bydelen Lund og bydelen er definert i figur 2.3. Dette er en tettbebygd bydel og er den mest folkerike i Kristiansand.

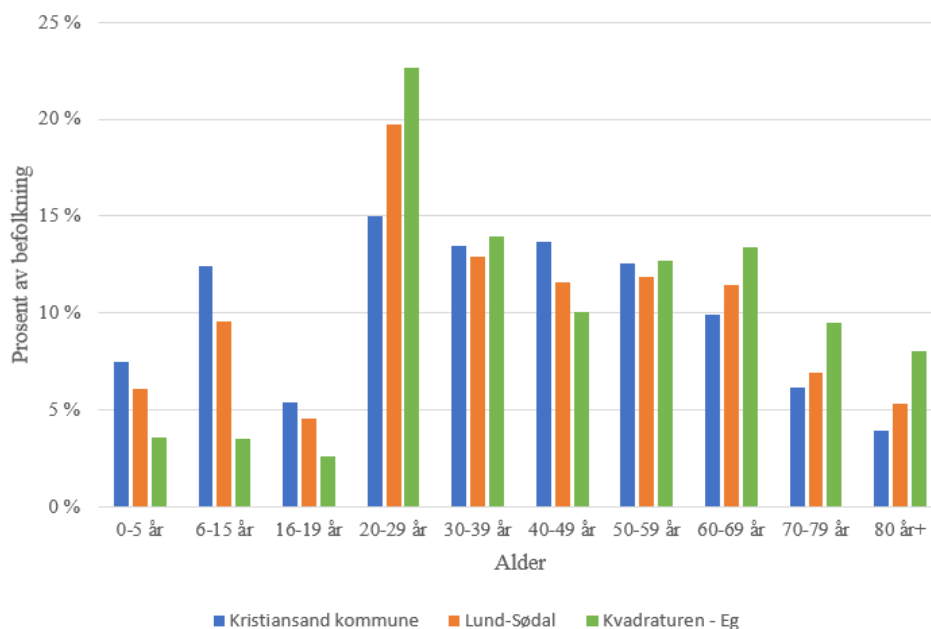
På Lund finner vi blant annet et godt utvalg dagligvarebutikker, Sør Arena og Universitetet i Agder. Sistnevnte fører til at mange studenter bor i området for en begrenset periode mens de studerer. Basert på en spørreundersøkelse blant bachelorstudenter, masterstudenter og studenter på årsstudium konkluderte en masteroppgave fra 2018 med at Lund er det mest attraktive boligområdet for studenter (Larsen & Nordheim, 2018).

I en undersøkelse utført av Kantar TNS vinteren 2018 var Lund-Sødal det mest prefererte området å bosette seg i. Dette er bydelen med høyest innbyggertall, og resultatet kan ha blitt påvirket av dette, da det ikke er gitt noen informasjon om hvor mange respondenter det er fra hver bydel.



Figur 2.4 Befolkningsutvikling for bydel Lund til venstre og årlig prosentvis befolkningsendring i bydelen i perioden 2010 - 2017 til høyre (Kristiansand Kommune, 2018b).

Andelen barn som bor i bydelen er lavere enn hva den er i kommunen totalt, jf. figur 2.5. Andelen av Lunds befolkning i aldersgruppen 20-69 år er 68%, og figur 2.5 viser også at andelen innbyggere over 60 år er høyere på Lund enn andelen i kommunen. Bydelen har en snittalder på 40 år. Det er ganske store forskjeller i snittalder mellom de forskjellige grunnkretsene på Lund. Høyest snittalder, 51,9 år, finner vi på Hamreheia som er et område med relativt dyre hus og leiligheter. Områdene Nedre Lund, Oddemarka, Solbygg og Tobienborg er karakterisert ved et stort antall boligblokker, og disse områdene har også en høyere snittalder enn gjennomsnittet på Lund (Kristiansand Kommune, 2018c).



Figur 2.5 Aldersfordeling i bydel Lund, sentrum(Kvadraturen), samt for kommunen totalt (Kristiansand Kommune, 2018c)

Andel av Kristiansand kommunes befolkning som bor på Lund har vært stabilt, og rundt 13% i perioden 2010 – 2017. Da undersøkelsen gjort for Kristiansand Kommune (2018a) fant at

det tydet på innbyggerne ønsket å bli boende i sin bydel samt at andelen innbyggere på Lund har vært stabil, er det rimelig å anta at Lund også i fremtiden vil utgjøre rundt 13% av befolkningen.

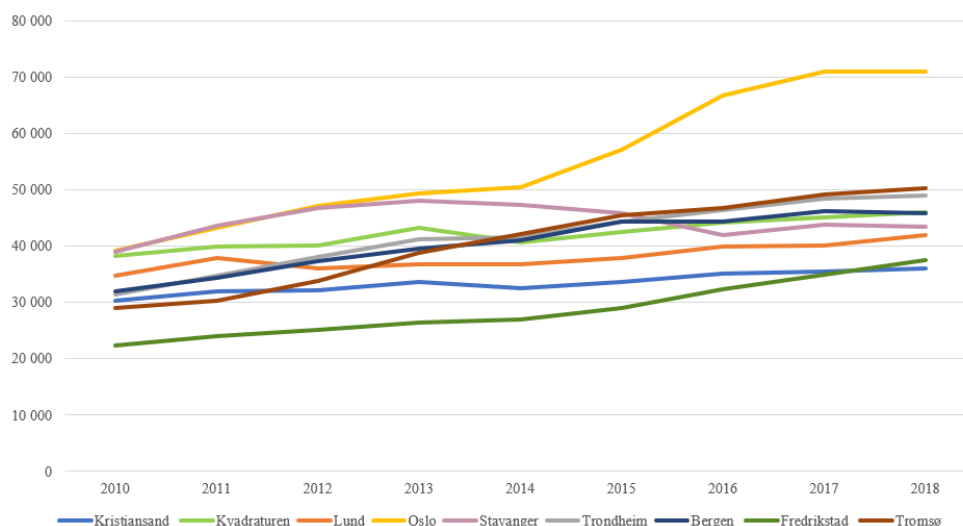
2.3. Boligmarkedet i Norge

I løpet av de siste tiårene har flere arbeidsplasser og studieplasser i storbyene ført til sentralisering og et presset boligmarked i byene. Antallet små husholdninger har også økt da dagens unge voksne studerer lenger og utsetter familieførøkelse mer enn tidligere. Dette har ført til det har vært en sterk prisstigning på de minste boligene, og boligene i byene (NOU 2002:2, 2002).

Statistisk Sentralbyrå (2018a) definerer en bolig som «en boenhet som består av ett eller flere rom, som er bygd eller ombygd om helårs privatbolig for en eller flere personer, og har egen atkomst uten at man må gå gjennom en annen bolig». I 2018 var det 2 547 732 boliger i Norge. Eneboliger er den mest vanlige boligformen, og om lag 50% av boligmassen i Norge er eneboliger. Om lag 24% av boenhetene er leiligheter. En leilighet kan defineres som «en bolig med minst ett rom og kjøkken» (Statistisk Sentralbyrå, 2018a). Det skilles mellom tre ulike eierformer: selveier, borettslag og aksjeleilighet.

I henhold til Statistisk Sentralbyrå (2019a) eier om lag 82% av den norske befolkningen sin egen bolig, de resterende 18% leier boligen de bor i. Andelen som eier sin egen bolig er mye høyere i Norge enn i våre naboland, i Finland, Sverige og Danmark er raten henholdsvis 71,4%, 66,20% og 62,20%, mens gjennomsnittet for Euroområdet er 66,4% (Trading Economics, 2019).

Boligprisene varierer mellom de ulike byene i Norge. I figur 2.6 er utviklingen i gjennomsnittlig kvadratmeterpris for leiligheter i perioden 2010 – 2018 presentert. Figuren viser årlige kvadratmeterpriser, og dermed vil ikke svingninger innenfor hvert år være synlige. Oslo har den klart høyeste leilighetsprisen, og har hatt det gjennom hele perioden. Kristiansand er blant kommunene med lavest leilighetspris og kurven er ganske slak. Bydelen Lund har litt høyere leilighetspris enn Kristiansand kommune. Figuren viser at kvadratmeterprisene i bydelen har gått fra fjerde til syvende plass, av de inkluderte byer og områder, i løpet av perioden.



Figur 2.6 Utvikling i gjennomsnittlig kvadratmeterpris for leiligheter i perioden 2010-2018 (Eiendomsverdi AS, 2019b)

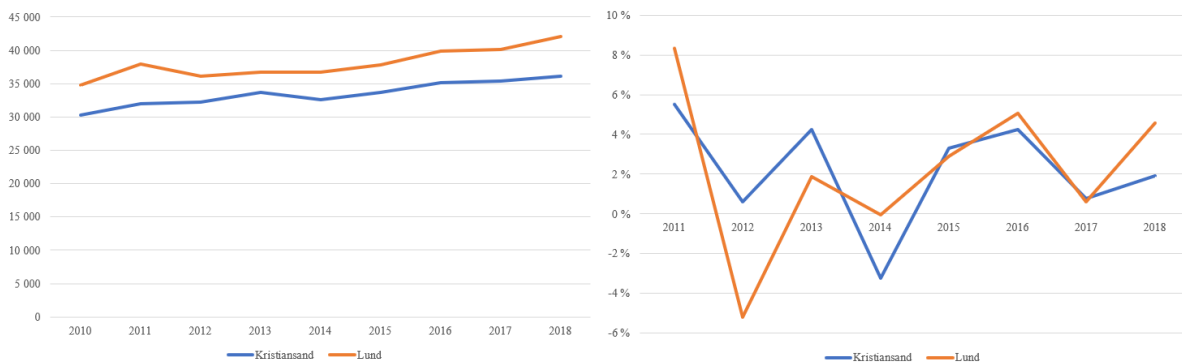
2.4. Boligmarkedet i Kristiansand

Figur 2.6 viser at gjennomsnittlig kvadratmeterpris i Kristiansand er lavere enn i andre store byer som Bergen, Stavanger, Trondheim og Oslo, og har vært det under hele undersøkelsesperioden. Gjennomsnittlig kvadratmeterpris i Kristiansand har også hatt en saktere prisvekst enn hva flere av de andre inkluderte byene i figuren har hatt. Dette kan skyldes kommunes boligpolitikk. Kristiansand kommunes boligpolitikk er å øke tilgangen på boliger slik at høy prisstigning unngås og «kjøpers marked» ivaretas. Krogsveen (2018) definerer «kjøpers marked» som en situasjonen der kjøper har flere enn en bolig å velge mellom, som er mulig å få kjøpt uten heftige budrunder. For å kunne ivareta denne politikken ønsker kommunen å tilrettelegge for at det bygges 700-800 nye boliger i året (Kristiansand Kommune, 2018a). I følge figur 2.6 ser det ut til at kommunen lykkes med sin boligpolitikk, da prisene i Kristiansand har en relativt jevn stigning i perioden.

I figuren under er gjennomsnittlig kvadratmeterpris for Kristiansand, samt bydel Lund trukket ut fra figur 2.6. Figur 2.7 viser at gjennomsnittlig kvadratmeterpris på Lund (oransje kurve) er høyere enn for hele kommunen (blå kurve) gjennom hele perioden. Lund er en sentrumsnær bydel, og den ligger innenfor bomringen som vil si at innbyggerne her kjører bomfritt til sentrum. Dette kan være blant årsakene til at gjennomsnittlig kvadratmeterpris er høyere på Lund enn det generelle nivået i kommunen.

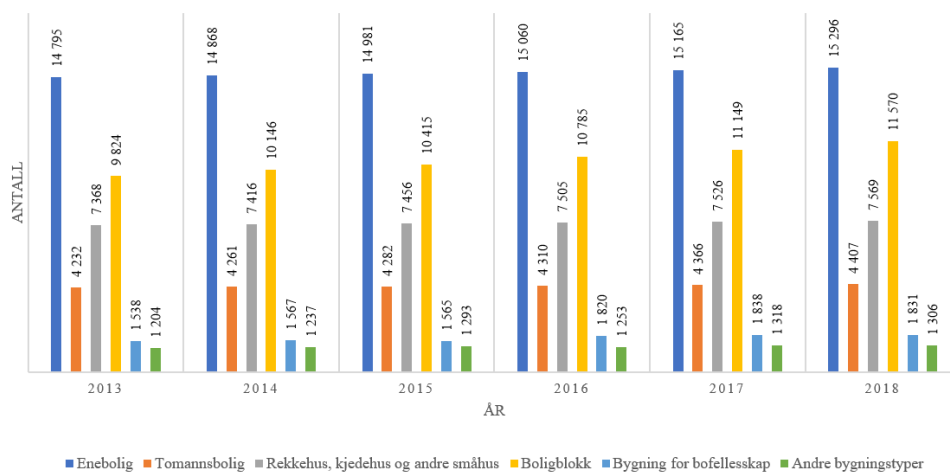
Mellom 2011 og 2012 sank gjennomsnittlig kvadratmeterpris både i kommunen og i bydelen, og nedgangen var størst på Lund. I 2014 var det igjen en liten nedgang, før prisen økte frem

til 2017 hvor den sank igjen. I 2018 nådde gjennomsnittlig kvadratpris i bydel Lund sitt høyeste punkt i løpet perioden.



Figur 2.7 Utviklingen i gjennomsnittlig kvadratmeterpris til venstre og årlig prosentvis endring i denne til venstre (Eiendomsverdi AS, 2019b).

Per 1.1.2018 er det 41 979 boliger i Kristiansand kommune. Av disse er 27,5% leiligheter i boligblokk. Figur 2.8 viser utviklingen i boligmassen i perioden 2013 til 2018. I denne perioden har total boligmasse økt med 7,2%. Boligtypen som har økt mest i denne perioden er leiligheter i boligblokk som har økt med 17,8%.

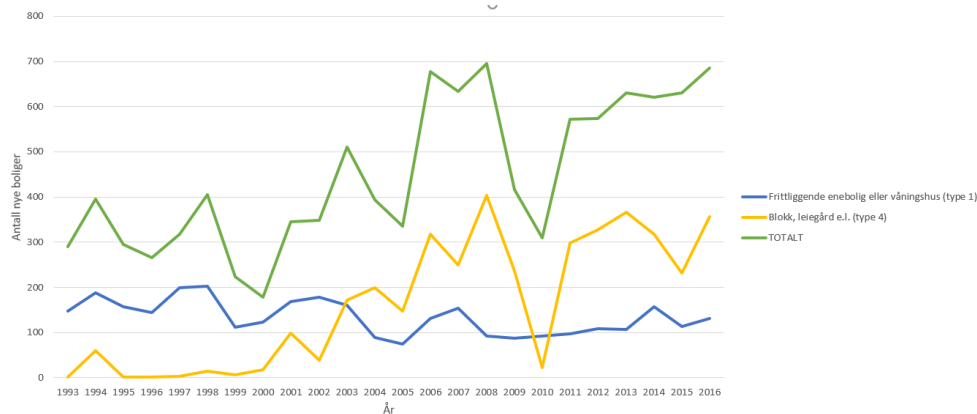


Figur 2.8 Oversikt over endringer i boligmassen i Kristiansand i perioden 2013 -2018. Figur laget med utgangspunkt i Kristiansand Kommune (2017).

I følge Kristiansand Kommune (2017) befinner 2468 av blokkleilighetene seg i bydelen Lund, det utgjør 36,1% av total boligmasse i bydelen. Etter Kvadraturen med 1,32 innbyggere per bolig, er Lund den bydelen med færrest innbyggere per bolig. I bydelen Lund er gjennomsnittlig innbyggere per bolig 1,70 per 1.1.2014.

2.4.1. Nye leiligheter

Figur 2.9 er basert på tall fra Kristiansand Kommune (2017) og viser utviklingen i byggingen av blokkleiligheter og eneboliger, samt totalt antall nybygg i Kristiansand i perioden 1993-2016. I figuren er rekkehus, småhus og forretningsgårder utelatt fordi utviklingen i boligbyggingen av disse boligtypene har vært relativt lav og stabil i perioden, og fremstår dermed som forstyrrende elementer i figuren. Figur med alle boligtyper inkludert finnes i vedlegg 3. Av figuren kan vi se at i perioden 1993-2002 ble det bygd flere eneboliger enn blokkleiligheter. De to kurvene krysser hverandre i 2003 og i perioden 2003 – 2009 er det bygd flere blokkleiligheter enn eneboliger. I 2010 var antall ferdigstilte blokkleiligheter på et veldig lavt nivå, det er naturlig å anta at dette var et resultat av nedgang i bygging i forbindelse med finanskrisen, da nivået av nye blokkleiligheter har vært stabilt høyt i perioden etter 2011. Det var også en liten nedgang i bygging av blokker i 2015, og det er rimelig å anta at dette har en sammenheng med oljeprisfallet.



Figur 2.9 Antall nybygde eneboliger og blokkleiligheter i Kristiansand, samt totalt antall nybygg bygd i perioden 1993 – 2016 (Statistisk Sentralbyrå, 2018a).

I henhold til Kristiansand Kommune (2018a) ønsker stadig flere innbyggere i Kristiansand å bosette seg i leilighet. Det søkes å tilrettelegge for at flere skal kunne bo i blokk, da bygging av blokkleiligheter har økt kraftig de siste årene. Rudberg (2018) skriver at flere og flere ønsker å bo tett, og figur 2.8 viser at mengden leiligheter i Kristiansand øker mer enn eneboliger.

Statistikk utarbeidet av Kristiansand Kommune (2017) viser mengden boligbygging i perioden 2010-2016 på Lund. For å finne antall nye leiligheter for årene 2017 og 2018 brukes Eiendomsverdi. For å kartlegge når de ulike nybyggene sto ferdig, benyttes omsetningsrapporter i Eiendomsverdi til å finne adresser og respektive byggeår.

Leilighetsprosjektene prospekt brukes til å bestemme antall leiligheter i byggene, da det kan være vanskelig å få eksakt informasjon om dette i databasen til Eiendomsverdi AS.

Statistikken presentert av Kristiansand Kommune (2017) sier at totalt antall blokkleiligheter på Lund 2468 per 1. januar 2017. Fra dette tallet beregnes leilighetsmassen på Lund i perioden 2010 – 2018 ved hjelp av den samme statistikken og informasjon fra Eiendomsverdi. Her forutsettes det at avgangen i leilighetsmarkedet har vært lik null, samt at statistikken er hentet fra ulike kilder. Dette kan føre til litt unøyaktige estimater, og dette må tas i betraktning.

Tabellen under er basert på Kristiansand Kommune (2017) og innsamlet data fra Eiendomsverdi og presenterer antall nybygg, samt total leilighetsmasse i perioden 2010-2018.

År	Nye leiligheter	Leilighetsmasse 31.12.201X	Nye i % av total leilighetsmasse
2010	16	2025	0,79%
2011	201	2041	9,84%
2012	53	2242	2,36%
2013	49	2295	2,14%
2014	134	2344	5,72%
2015	5	2478	0,20%
2016	99	2483	3,99%
2017	114*	2582	4,41%
2018	80*	2662	3,01%

Tabell 2.1 Leilighetsmasse og nye leiligheter per år (Eiendomsverdi AS, 2019b; Kristiansand Kommune, 2017)

* Tallet er hentet fra Eiendomsverdi

2.4.2. Nybygg og planlagte prosjekt

Figur 2.9 viser at antall ferdigstilte boliger varierer fra år til år, men at tilfanget av boliger har økt i perioden etter 2003. Figur 2.10 viser beliggenheten til de fem største nye leilighetsprosjektene på Lund, som enten er bygd, påbegynt eller hvor salget er påbegynt. I tillegg til disse fem store boligprosjektene er det også bygget og påbegynt flere mindre blokker og leilighetskompleks de siste årene.

Beliggenhet nr. 1 er bydel Bjørndalen hvor det er planlagt opp mot 500 leiligheter på sikt og salgsstart var høsten 2017. I henhold til prospektet skal første byggetrinn stå ferdig 1.kvartal 2021. Leilighetene har en høyere kvadratmeterpris enn gjennomsnittet i Kristiansand. I følge Eiendomsverdi AS (2019a) har det første byggetrinn som omfatter 30 leiligheter en

gjennomsnittlig kvadratpris på kr 52 586. Sammenliknet med gjennomsnittet for Kristiansand er disse leilighetene i snitt nesten kr 24 000 dyrere per kvadratmeter. Gjennomsnittlig kvadratmeterpris i bydelen Lund var i 2018 kr 42 037, det vil si at gjennomsnittlig pris for boligene i bydel Bjørndalen er om lag kr 10 000 mer enn gjennomsnittet på Lund. Her må det legges også til at leilighetene ikke skal stå ferdig før i 2021, men avviket til gjennomsnittet for kommunen er fortsatt stort (MittBjørndalen, 2019).

På beliggenhet nr.2 i figur 2.10 bygges Kongsgård Park, hvor første byggetrinn som omfatter 17 leiligheter skal stå ferdig 1.kvartal 2020. Her er 76% av leilighetene solgt. Andre byggetrinn skal også ferdigstilles 1.kvartal 2020, her er bare 40% av de 30 leilighetene solgt per 13.april 2019. Gjennomsnittlig kvadratmeterpris for begge byggetrinn er kr 53 145. Dette er også langt over snittet for både Kristiansand kommune og bydel Lund(Eiendomsverdi AS, 2019a).

Beliggenhet nr.3 er Marvika Torv som på sikt skal bestå av 147 leiligheter. Første byggetrinn omfatter 67 leiligheter som skal stå ferdig 4.kvartal 2020. Gjennomsnittlig kvadratmeterpris for leilighetene er kr 54 571. Per 13.april 2019 er 21% av leilighetene solgt.

Tobienborg Hage på beliggenhet nr. 4 består av to byggetrinn, hvor det første byggetrinnet var ferdigstilt høsten 2018. Det første byggetrinnet omfattet 36 leiligheter, og gjennomsnittlig kvadratmeterpris var kr 51 845. 78% av leilighetene er solgt per 6.februar 2019. På andre byggetrinn er 53% av de 34 leilighetene solgt. Her er kvadratmeterprisen litt høyere enn på første byggetrinn, kr 55 979. Andre byggetrinn ble ferdigstilt 1.kvartal 2019. Per 16.januar 2019 står 11 av de 17 toppleilighetene fortsatt usolgt og gjennomsnittlig kvadratmeterpris for de ledige toppleilighetene er kr 64 981 (Eiendomsverdi AS, 2019a; Tobienborg hage, 2019).

Kjøita Secret Garden, på beliggenhet nr. 5, ble ferdigstilt i 2017. Her er 1 av de 140 leilighetene fortsatt til salgs. Gjennomsnittlig kvadratmeterpris for leilighetene var kr 51 373 (Eiendomsverdi AS, 2019a).



Figur 2.10 Store byggeprosjekt på Lund, redigert skjermdump fra Google Maps

3. Teori

I dette kapittelet presenteres teorier og modeller som er relevante for å studere leilighetsprisutviklingen på Lund i perioden 2010 – 2018.

Hvilke kilder som brukes som utgangspunkt i denne teoridelen vil påvirke hvordan analysen gjennomføres og dermed kvaliteten på avhandlingen. Derfor er det viktig å være å ha et bevisst forhold til de kilder som benyttes. I dette kapittelet skal jeg først gi en introduksjon av prisdannelsen i boligmarkedet. Dette delkapittelet er basert på en NOU (Norsk Offentlig Utredning) utarbeidet av boligutvalget og noen utvalgte medlemmer fra kommunal- og regionaldepartementet i 2002. Boligutvalget bestod av blant annet en professor, en forskningssjef og administrerende direktør i Norges boligbyggerlag.

Deretter presenteres den hedonistiske metoden. Til dette har jeg valgt å i hovedsak bruke to artikler. Den første artikkelen tar for seg boligprisen i Haugesund og er utgitt i Norsk Økonomisk Tidsskrift i 2001. Den andre artikkelen som er brukt som teoretisk utgangspunkt er Rosen (1974), denne kan betraktes som en kjerneartikkel innenfor hedonistisk prissetting med sine over 11 000 siteringer.

Videre presenteres Alonso-Muth-Mills modellen. Det teoretiske grunnlaget som er brukt til å forklare denne modellen er en bok skrevet av DiPasquale og Wheaton (1996) som er sitert mer enn 1 000 ganger.

I tillegg til artiklene nevnt ovenfor er forelesningsmateriale presentert av Emblem (2018) og Theisen (2018) benyttet for å supplere og hente inspirasjon til artikler og modeller.

3.1. Prisdannelse i boligmarkedet

NOU 2002:2 (2002) beskriver boliger som både konsumgoder og kapitalobjekter, det betyr at når man kjøper en bolig kjøpes både bolig tjenesten, samt at boligen er et investeringsobjekt. Boligmarkedet er et stort og komplekst marked med mange delmarked og mange ulike aktører. Blant aktørene i markedet finner vi etterspørrere og tilbydere.

Etterspørrerne er de som ønsker å kjøpe eller leie bolig, mens tilbyderne er alle som eier boliger. Etterspørrerne har ulik betalingsvilje, og betalingsviljen avhenger av betalingsevne og hvordan den enkelte etterspørrer verdsetter bolig fremfor andre konsumgoder og investeringer. Etterspørselskurven i boligmarkedet er en fallende kurve som viser alle etterspørrere for bolig til hvert nivå av prisen. Desto lavere prisnivå, desto flere etterspør bolig som vist i figur 3.1 (NOU 2002:2, 2002).

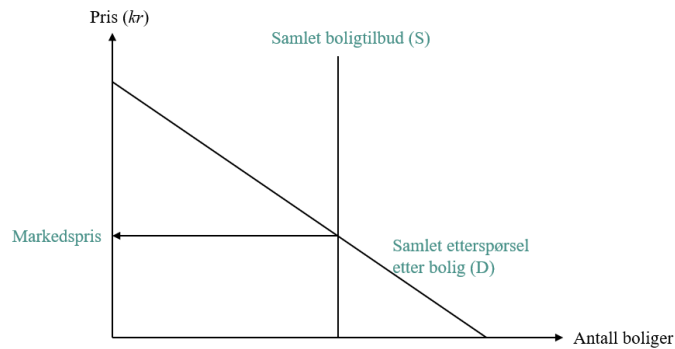
Prisantydningen er en antakelse om hvor mye man minst forventer å få solgt en leilighet for. Den prisantydning som er satt på en bolig påvirker både den endelige salgsprisen og hvor lenge boligen er tilgjengelig i markedet før den blir solgt. En prisantydning som er satt for lav kan resultere i raskt salg, men til en lavere pris enn man egentlig kunne oppnådd. Dersom prisantydningen er for høy reduseres mengden potensielle kjøpere, og resultatet av dette kan være at boligen er tilgjengelig i markedet i en lang periode (Knight, 2002).

Tilbyderne er de som eier bolig, og det er en forutsetning at tilbyderne ønsker å selge de boligene de ikke selv skal bruke. Tilbudet av boliger øker når det bygges nye boliger, og reduseres ved rivning, brann og omreguleringer. Å bygge boliger tar tid, derfor er tilbudet av boliger på kort sikt gitt som en loddrett kurve (NOU 2002:2, 2002).

Boligprisene på nybygg følger prisene på liknende brukte boliger. Dette er fordi etterspørrerne ønsker å betale så lite som mulig for bolig, og dermed vil ikke nye boliger bli solgt dersom prisen på de er mye høyere enn prisen for en brukt bolig. Det kan kanskje påregne mindre vedlikehold på en nyere bolig, og dette er en faktor som kan være med å trekke prisen på en ny bolig opp i forhold til en brukt bolig. Økt bolig etterspørsel og boligpris

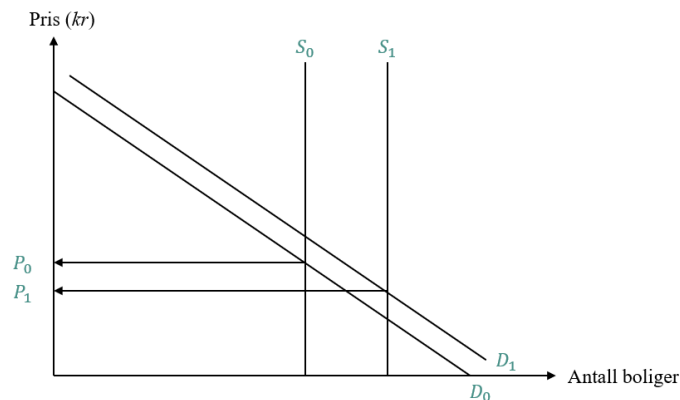
gir incentiv til å øke boligbyggingen. Når tilbudet av boliger øker vil prisen etter hvert reduseres, dersom etterspørselen er konstant (NOU 2002:2, 2002).

Markedsprisen i boligmarkedet bestemmes av skjæringspunktet mellom tilbudskurven og etterspørselskurven. De etterspørerne som har høyere betalingsvilje enn markedsprisen vil få kjøpe bolig.



Figur 3.1 Tilbud og etterspørsel i boligmarkedet

Når mengden nybygg er større enn antall boliger som forsvinner fra markedet øker boligmassen og tilbudskurven skifter utover i diagrammet. Over tid endres også etterspørselen, den kan for eksempel øke på grunn av befolkningsvekst, inntektsvekst eller redusert arbeidsledighet. Da vil etterspørselskurven også skifte utover i diagrammet. Hvordan likevektsprisen påvirkes av skift i tilbud og etterspørsel avhenger av om tilbudsveksten eller etterspørselsveksten har vært størst. Figur 3.2 viser en situasjon der økningen i tilbudet er større enn etterspørselsveksten. Dette fører til at markedsprisen reduseres fra P_0 til P_1 . Tilsvarende vil markedsprisen øke dersom etterspørselsveksten er større enn tilbudsveksten (NOU 2002:2, 2002).



Figur 3.2 Effekt av økt tilbud og etterspørsel

3.2. Den hedonistiske metoden

En sentral del av boligprisutviklingen er hvilke faktorer som bestemmer hvor mye en bolig blir solgt for, og hvilke faktorer som ser ut til å være de viktigste. Den hedonistiske metoden er en modell som kan brukes til å analysere heterogene goder, i dette tilfellet leiligheter, som har ulike egenskaper eller attributter. Modellen søker å forklare hvordan den hedonistiske prisfunksjonen bestemmes av samspillet mellom tilbydere og etterspørrere i boligmarkedet (Osland, 2001).

Det finnes mange artikler som tar for seg den hedonistiske metoden og hvordan ulike attributter påvirker boligprisen. Dette kapitlet er basert på teori presentert av Osland (2001) og Rosen (1974), samt forelesningsmateriale presentert av Theisen (2018).

Attributtene kan deles i to hovedgrupper: attributter som er knyttet til selve leiligheten som størrelse og antall soverom, og attributter knyttet til leilighetens plassering som avstand til sentrum og solforhold. Disse attributtene genererer nytte for konsumentene, og markedsprisen er et resultat av i hvilken grad attributtene verdsettes. Den hedonistiske metoden kan brukes til å finne ut hvor mye de ulike attributtene påvirker markedsprisen for leiligheten (Osland, 2001; Rosen, 1974).

En studie fra 2001 gjort i Haugesund kommune viste at boligens totalpris økte ved en økning i boareal og tomteareal. Økt avstand til sentrum reduserte boligens totalpris, det gjorde også økt alder på boligen (Osland, 2001).

Rosen (1974) definerer attributtvektoren som er formulert i funksjon 3.1. Attributtvektoren sier at et gode kan sees på som en vektor som består av n attributter som er objektivt målt og hvor z_i måler mengden av attributt i som er representert i godet. Totalprisen for boligen, Z , er dermed et resultat av mengden og kombinasjonen av attributter og deres implisitte pris. Denne modellen forutsetter at det er til enhver tid finnes et stort antall boliger på markedet, slik at konsumenten har valget mellom ulike attributtvektorer. Det antas også at markedet er karakterisert av fri konkurranse der det er mange små aktører som hver for seg ikke har noen innflytelse på pris og markedsforhold deltar. En tredje forutsetning for modellen er at alle aktører har full informasjon om priser og egenskaper ved de ulike boligene (Osland, 2001).

$$Z = (z_1, z_2, \dots, z_n) \quad (3.1)$$

Attributtene og kombinasjoner av disse kan gi nytte for husholdningene, og det er vanlig å betegne attributtprisen som implisitte, indirekte eller hedonistiske priser. Den hedonistiske

prisfunksjonen er definert som $P(Z)$, det betyr at boligens totalpris er en funksjon av mengden attributter og deres indirekte pris (Osland, 2001).

I den hedonistiske modellen har vi en etterspørselsside og en tilbudsside. Etterspørselssiden består av husholdninger og tilbudssiden består av mange små bedrifter. Både etterspørrere og tilbydere søker å maksimere sin nytte.

Husholdningenes nyttefunksjon avhenger av mengden attributter ved boligen (Z), alle andre goder enn bolig (X) og husholdningens preferanser (α_j) og er strengt konkav. Husholdningene bruker inntekten de har til disposisjon (Y_j) på andre konsumgoder enn bolig (X) og boligkostnader ($P(Z)$), og posisjonerer seg slik at deres nytte maksimeres.

$$\text{Maksimer } U_j = (Z, X, \alpha_j) \text{ gitt } Y_j = X + P(Z) \quad (3.2)$$

For å finne optimum deriveres nyttefunksjonen med hensyn på Z . I henhold til Osland (2001) svarer høyresiden av funksjon 3.3 til hedonistiske priser for attributt i . Det tilsvarer den økningen eller reduksjonen i leilighetsprisen attributtet gir.

$$\frac{\frac{\partial U_j}{\partial Z_j}}{\frac{\partial U_j}{\partial X}} = \frac{\partial P}{\partial Z_i} \quad (3.3)$$

Husholdningens budfunksjon defineres som maksimal betalingsvillighet for ulike sammensetninger av attributter når inntektsnivå og nyttenivå holdes konstant. Budfunksjonen vil opptre som en indifferenskurve, som er en kurve i et diagram som viser ulike kombinasjoner av to goder som gir forbrukeren samme nytte.

I henhold til Osland (2001) utledes budfunksjonen ved å ta utgangspunkt i de optimale verdiene for boligvektoren (Z^*) og vektoren som består av mengden konsumenten ønsker å bruke på andre goder enn bolig (X^*).

$$X^* = Y_j - P(Z^*) \quad (3.4)$$

Budfunksjonen settes inn i nyttefunksjonen, 3.1, og man finner optimalt nyttenivå.

$$U_j = (Z^*, Y_j - P(Z^*), \alpha_j) = U_j^* \quad (3.5)$$

Videre forutsettes det at konsumentens maksimale betalingsvillighet er lik prisen man betaler for boligen, det vil si $\theta = P(Z^*)$, der θ representerer maksimal betalingsvilje. Nyttefunksjon kan dermed uttrykkes på følgende måte:

$$U_j = (Z^*, Y_j - P(Z^*), \alpha_j) = U_j^* = U(Z, Y_j - \theta_j, \alpha_j) \quad (3.6)$$

Uttrykk 3.6 definerer i henhold til Osland (2001) «implisitt en relasjon for maksimal betalingsvillighet ved andre sammensetninger enn den optimale, samtidig som husholdningene oppfatter disse kombinasjonene som likeverdige». Budfunksjonen varierer med konsumentens nyttenivå og inntektsnivå, og kan mer generelt uttrykkes ved:

$$\theta_j = \theta(Z, Y_j, U_j, \alpha_j) \quad (3.7)$$

For å finne maksimal betalingsvillighet for en partiell økning i et boligattributt, brukes implisitt derivasjon på uttrykket 3.7.

$$\frac{\partial \theta_j}{\partial z_i} = \frac{\frac{\partial U_j}{\partial z_i}}{\frac{\partial U_j}{\partial x}} > 0 \quad (3.8)$$

I optimum er maksimal betalingsvilje for leilighet lik hva konsumenten faktisk betalte for leiligheten. Dette er vist i uttrykk 3.9. I optimum er marginal betalingsvilje for et attributt lik marginalprisen til attributtet – økningen eller reduksjonen i leilighetspris ved å legge til en enhet eller 1% av attributtet.

$$\theta(Z, Y_j, U_j, \alpha_j) = P(Z^*) \quad (3.9)$$

På tilbudssiden av markedet finnes det mange små bedrifter som ønsker å maksimere sin fortjeneste. På kort sikt har bedriftene to mulige tilpasningsmuligheter, enten endre produsert antall enheter eller endre antall og sammensetninger av attributter. Bedriftens profittfunksjon er som følger:

$$\pi = M \times P(Z) - C(M, Z, \beta) \quad (3.10)$$

Hvor M er bedriftens tilbud av boliger. C er kostnadsfunksjonen, som er konveks og øker med antall boliger produsert. β representerer en vektor av skiftparametre, som materialpriser og produksjonsteknologi.

Bedriften bør velge den sammensetningen av boligattributter som gjør at den implisitte attributtprisen er lik marginalkostnad per bolig ved en partiell økning i mengden boligattributter. Dette fremkommer av førsteordensbetingelsen for maksimal fortjeneste.

$$\frac{\partial P}{\partial z_i} = \frac{\frac{\partial C}{\partial z_i}}{M} = P_i = \frac{C_i}{M} \Leftrightarrow P(Z) = C_M \quad (3.11)$$

I henhold til (Osland, 2001) defineres offerfunksjonen som «det minste beløpet eller prisen produsentene er villig til å akseptere for å kunne tilby boliger med ulike attributter, til et konstant profittnivå og gitt det optimale antall boliger som produseres».

$$\Phi = (Z, \pi, \beta) \quad (3.12)$$

Ved å ta utgangspunkt i de optimale verdiene, M^* , Z^* og π^* , kan profittfunksjonen utledes:

$$\pi^* = M^* \times P(Z^*) - C(M^*, Z^*, \beta) \quad (3.13)$$

Ved konstant profittnivå kan uttrykket ovenfor skrives som:

$$\pi^* = M^* \times \Phi(Z^*, \pi^*, \beta) - C(M^*, Z^*, \beta) \quad (3.14)$$

Funksjon 3.14 deriveres med hensyn på M og Z og gir følgende to førsteordensbetingelser.

$$\Phi(Z^*, \pi^*, \beta) = \frac{\partial C}{\partial M} \Leftrightarrow \Phi(Z^*, \pi^*, \beta) = C_M \quad (3.15)$$

$$\frac{\partial \Phi}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} \Leftrightarrow \frac{C_i}{M} \quad (3.16)$$

Dette løses med hensyn på M og settes inn i uttrykk 3.14, dette gir

$$\Phi = \Phi(Z, \pi^*, \beta) \quad (3.17)$$

Uttrykket 3.3 definerer implisitt en relasjon mellom offerprisene og boligattributtene (Osland, 2001). Når uttrykk 3.3 er likt uttrykk 3.15 er det likevekt på tilbudssiden. Da tangerer offerkurvene for hver produsent prisfunksjonen som er eksogent gitt.

$$\frac{\partial \Phi}{\partial Z_n} = \frac{\frac{\partial C}{\partial Z_n}}{M} = \frac{\partial P}{\partial Z_n} \quad (3.18)$$

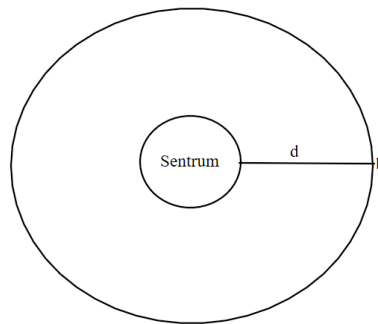
Når det er likevekt på tilbudssiden er offerprisen lik marginalkostnadene ved å produsere ytterligere en bolig. Bedriften velger attributtvektoren der den implisitte prisen til hvert attributt er lik marginalkostnadene per bolig ved en partiell økning i attributtet. I likevekt vil også offerprisen være lik den faktiske prisen bedriften får for salg av boligen.

Når konsumentenes budfunksjon og bedriftenes offerfunksjon tangerer hverandre oppnås markedsliekevt.

$$\frac{\partial \theta}{\partial Z_i} = \frac{\partial P}{\partial Z_i} = \frac{\frac{\partial C}{\partial Z_i}}{M} = \frac{\partial \Phi}{\partial Z_i} \quad (3.19)$$

3.3. Alonso – Muth – Mills-modellen

Avstanden til sentrum er en sentral attributt i den hedonistiske modellen, og i henhold til DiPasquale og Wheaton (1996) er et fundamentalt trekk ved det urbane bolig og tomtemarkedet at boliger med bedre plassering er dyrere enn de boligene som har en dårligere plassering. Alonso-Muth-Mills-modellen er en forenklet modell som forutsetter en monosentrisk by, dette er illustrert i figur 3.3.



Figur 3.3 Monosentrisk by

I tillegg til forutsetningen om at det er et arbeidssentrum presenterer DiPasquale og Wheaton (1996) følgende forutsetninger for modellen.

1. Byens innbyggere reiser til byens sentrum langs en rett linje. Kostnaden forbundet med å pendle til sentrum er k kr per kilometer.
2. Boligens beliggenhet refererer til den lineære distansen fra boligen til bykjernen, d .
3. Utenfor bygrensen er jordbruk den alternative bruken av areal. Tomtene på bygrensen har en tomteleie på $r^a q$.
4. Husholdningene er identiske og antall pendlere i hver husholdning er gitt.
5. Husholdningene bruker sin inntekt (y) på pendling (kd), andre goder (x) og bolig.
6. Husleien uttrykkes ved $R(d)$, og denne varierer for de ulike beliggenhetene.
7. Boligene er gitt ved en fast tomtestørrelse (q) og en fast byggekostnad © . Dermed er boligtettheten gitt ved $1/q$.
8. Boligene okkuperes av de husholdningene som er villige til å betale mest.

Det forutsettes at alle husholdningene er identiske, dermed vil også nivået av konsum av andre goder (x^0) være det samme for alle lokasjoner ($x = x^0$). Dermed kan husleien uttrykkes slik:

$$R(d) = y - kd - x^0 \quad (3.20)$$

Innbyggere som bor i sentrum, der $d=0$, vil ha pendlekostnader $k=0$. Dermed kan husleien i sentrum uttrykkes som $R(0) = y - x^0$. Ved å bevege seg lenger fra sentrum, vil husleien, dette er for å kompensere for økte pendlerkostnader. Ved bygrensa (b) vil husleien ha det laveste nivået, samtidig som pendlekostnadene har sitt høyeste nivå. Det er forutsatt at tomteleien på bygrensen er $r^a q$, dermed vil husleien på bygrensen være lik $r^a q$ pluss de årlige konstruksjonskostnadene (c).

$$R(b) = r^a q + c \quad (3.21)$$

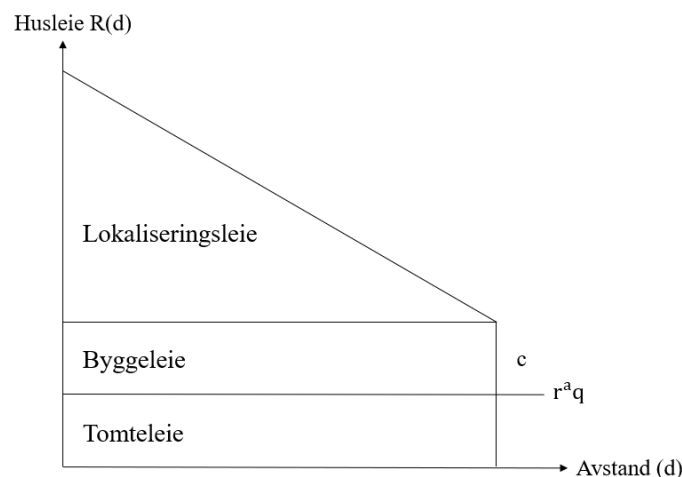
Konsumnivået av andre goder for de husholdningene som bor på bygrensen bestemmes ved å sette $R(d)=R(b)$, og $d=b$ og kombinere likning 3.20 og 3.21 og løse med hensyn på x^0 .

$$x^0 = y - kb - (r^a q + c) \quad (3.22)$$

Uttrykket 3.21 definerer hvordan husleien må øke når pendlekostnadene reduseres for å opprettholde velferdnivået definert i 3.22. Ved å kombinere disse to uttrykkene finner vi uttrykket for husleien uansett beliggenhet.

$$R(d) = (r^a q + c) + k(b - d) \quad (3.23)$$

Husleien består av tre komponenter: lokaliseringsleie, byggeleie og tomteleie. Byggeleien og tomteleien er gitte størrelser som er uavhengig av beliggenhet. Husleien er illustrert i figur 3.4.



Figur 3.4 Husleien (DiPasquale & Wheaton, 1996)

Modellen er bygd opp slik at det er forutsatt at alle innbyggerne i byen har samme mengde kostnader relatert til bolig og pendlekostnader. Det er kun lokaliseringsleien som endrer seg med beliggenhet, og desto lenger fra sentrum en bolig ligger, desto lavere er

lokaliseringsrenten, slik at det kompenseres for de økte pendlekostnader. Ved å derivere uttrykk 3.23 med hensyn på d finner vi at stigningstallet til husleiefunksjonen er $-k$. Det betyr at husleien reduseres med k for hver kilometer en beveger seg ut fra bysentrum.

3.4. Hypotesetesting

Basert på teorien som er presentert, samt antakelser skal det nå formuleres forskningsspørsmål og tilhørende testbare hypoteser. Dette skal hjelpe til med å besvare problemsstillingen. I kapittel 6 og 7 vil relevante analyser gjennomføres for å kunne konkludere for hver av hypotesene.

3.4.1. Hypoteser som omhandler leilighetsprisen

Forskningsspørsmål 1: Har antall m² primærrrom betydning for leilighetens totalpris?

I en større leilighet er det plass til flere mennesker, og av den grunn er det naturlig å tenke at en større leilighet har en høyere totalpris enn en mindre leilighet. Tidligere studier viser også at økt boareal fører til økt boligpris. Her brukes en ensidig test, da antakelsen er at økt størrelse fører til økt leilighetspris.

H₀: Det er ingen korrelasjon mellom leilighetens størrelse og leilighetspris

H₁: Det er en positiv sammenheng mellom leilighetens størrelse og leilighetspris

Forskningsspørsmål 2: Har leilighetens alder betydning for totalprisen?

Nye leiligheter krever som regel mindre vedlikehold og oppussing enn eldre leiligheter. I tillegg er nye leiligheter som regel bedre isolert enn eldre leiligheter, samt at de ofte har energibesparende løsninger. Dette fører til at kostnadene forbundet med leiligheten er lavere ved en ny leilighet, derfor er det naturlig å tenke at forbrukernes betalingsvilje synker desto eldre leiligheten er. Dette legger grunnlaget for den ensidige hypotesen:

H₀: Det er ingen korrelasjon mellom leilighetens alder og leilighetspris

H₁: Det er en negativ sammenheng mellom leilighetens alder og leilighetspris

Forskningsspørsmål 3: Har avstand til sentrum betydning for totalprisen?

Modellen presentert av DiPasquale og Wheaton (1996) sier at desto lenger fra bykjernen en bolig er plassert, desto lavere er husleien. Basert på denne teorien utarbeides følgende hypotese. Dette gir grunnlaget for følgende hypotese.

H₀: Det er ingen sammenheng mellom avstand til sentrum og leilighetspris

H₁: Det er en negativ sammenheng mellom avstand til sentrum og leilighetspris

Forskningsspørsmål 4: Har de leilighetene som ligger i første etasje en lavere totalpris enn leiligheter i øvrige etasjer?

Det er naturlig å tenke at leiligheter som ligger i første etasje er mer utsatt for innsyn og støy. Dessuten ligger ofte de største og dyreste leilighetene på toppen i leilighetsbyggene. Dette er grunnlaget for følgende hypotese.

H₀: Det er ingen korrelasjon mellom avstand til leilighetens etasjeplassering og totalpris

H₁: Det er en positiv korrelasjon mellom leilighetens etasjeplassering og totalprisen

Forskningsspørsmål 5: Har parkeringsplass en positiv effekt på leilighetens totalpris?

Utbyggingen på Lund fører til at området blir tettere bebyggt enn tidligere, dette kan føre til at det blir vanskeligere å finne parkeringsplass dersom man ikke har en fast plass. Lund ligger relativt sentralt og det finnes mange dagligvarebutikker i bydelen. Men det er fortsatt naturlig å tenke at mange likevel trenger og ønsker å ha bil. Det finnes også kollektivtransporttilbud i bydelen, men busstilbudet er varierende. Derfor antar jeg at tilgang på parkeringsplass vil øke leilighetens totalpris. Hypotesen som skal testes er dermed:

H₀: Det er ingen forskjell i leilighetsstørrelse mellom leiligheter solgt med tap og leiligheter solgt med gevinst

H₁: Gruppen med leiligheter som selges med tap har lavere gjennomsnittsstørrelse enn leilighetene som selges med gevinst.

3.4.2. Hypoteser som omhandler tap ved salg

Forskningsspørsmål 6: Er det de små leilighetene som selges med tap?

Det finnes mange små leiligheter, på rundt 50 m², disse har ofte kun ett soverom. Prisen på disse leilighetene skiller seg ikke betydelig fra en leilighet på rundt 60 m² med to soverom. En leilighet med to soverom kan oppleves som mer attraktiv og permanent enn en leilighet med ett soverom. Dermed er min antakelse at det er flere små leiligheter selges med tap.

H₀: Det er ingen forskjell i leilighetsstørrelse mellom leiligheter solgt med tap og leiligheter solgt med gevinst

H₁: Gruppen med leiligheter som selges med tap har lavere gjennomsnittsstørrelse enn leilighetene som selges med gevinst.

Forskningsspørsmål 7: Er det flest nye leiligheter som selges med tap?

Desto eldre leiligheten er, desto større er behovet for vedlikehold og rehabilitering. I denne analysen kan boligalderen maksimalt ha økt med 8 år, og det er begrenset hvor mye vedlikeholdskostnadene har økt på kun 8 år. En eldre bolig kan være dårligere isolert og ha en dårligere energimerking enn nyere leiligheter. Dette trekker i retning av at økt alder gir redusert pris.

På den andre siden kan vi ha situasjonen der leilighetens første eier har vært villig til å betale mer for et nybygg enn hva den neste eieren er villig til å betale for en brukt, men relativt ny bolig. Dette trekker i retning av at gjennomsnittlig alder for leiligheter solgt med tap er lavere enn for leiligheter solgt med gevinst. Disse kreftene trekker i hver sin retning, og danner grunnlaget for den tosidige hypotesen.

H₀: Det er ingen forskjell i leilighetens alder mellom leiligheter solgt med tap og leiligheter solgt med gevinst

H₁: De leilighetene som er solgt med tap har lavere gjennomsnittsalder enn de leilighetene som er solgt med gevinst.

Forskningsspørsmål 8: Vil leiligheter som omsettes hyppig være mer utsatt for salg med tap?

I henhold til Schultz og Kaspersen (2016) er antall tap ved salg av bolig størst når eieren har eid boligen i en kort periode.

H₀: Det er ingen sammenheng mellom hvor mange år det er mellom to salg av samme leilighet og om den selges med tap

H₁: Leiligheter som selges med tap omsettes hyppigere enn de som selges med gevinst.

Forskningsspørsmål 9: Øker risikoen for salg med tap desto lenger leiligheten ligger i markedet?

Tidligere studier viser at å sette en for høy prisantydning kan føre til at leiligheten er tilgjengelig i markedet over en lengere periode. Jeg antar at selger har et økende ønske om å selge leiligheten etter hvert som salgsperioden øker. Det kan da føre til at leiligheten selges med tap, fordi selges ønske om å få solgt leiligheten blir stort. Det gir grunnlag for å teste følgende hypotese:

H₀: Det er ingen forskjell i omsetningshastighet mellom de to gruppene

H₁: De leiligheter som selges med tap har høyere gjennomsnittlig omsetningshastighet

Forskningsspørsmål 10: Har de leilighetene som er solgt med tap en høyere prisantydning enn salgspris?

Prisantydningen er en antakelse om hvor mye en minst forventer å få solgt leilighetene for. I dagens marked med god prisvekst på boliger, settes gjerne prisantydningen ved til et høyere prisnivå enn det boligen ble kjøpt for da de som skal selge kjøpte den. I henhold til Schultz og Kaspersen (2016) var det flest tap blant de leiligheter som ble solgt igjen etter kort tid. Hvis prisantydningen settes lik eller høyere enn det forrige salgspris var, og leiligheten selges med tap vil vi få et negativt avvik. Dette gir grunnlag for å undersøke om avviket mellom salgspris og prisantydning er størst, i negativ retning, i den gruppen leiligheter som selges med tap.

H₀: Gjennomsnittlig avvike mellom salgspris og prisantydning er likt for de to gruppene

H₁: Avviket mellom salgspris og prisantydning er mest negativt for de leilighetene som er solgt med tap

4. Innhenting og bearbeiding av datamaterialet

I dette kapittelet vil jeg først beskrive datakilden og undersøkelsesområdet. Deretter vil jeg beskrive innsamlingsprosessen og prosessen med å rense innsamlet data. I kapittelet gis også avslutningsvis en oversikt de variabler som skal brukes i undersøkelsen samt deskriptiv statistikk for de to datasettene som benyttes i analysen og en vurdering rundt de variabler som er utelatt.

4.1. Sekundærdata

Ved kvalitativ metode undersøkes et stort antall enheter og formålet er ofte å teste en hypotese, en antakelse om virkeligheten. I en slik type analyse er dataene på tallform og analyseres ved hjelp av statistiske og økonometriske metoder (Dahlum, 2018).

For å svare på problemstillingen brukes sekundærdata innsamlet av Eiendomsverdi AS. Eiendomsverdi AS er Norges største boligdatabase og eies av DNB, Nordea, Sparebank1 og Eika Boligkreditt (Eiendomsverdi AS, 2019). Fordelen ved å bruke data innsamlet av andre er at det er tidsbesparende, samt at store deler av datamaterialet allerede er samlet inn og systematisert. I databasen finnes en oversikt over boliger som er i markedet og boliger som er solgt. Ulempen ved å bruke sekundærdata er at jeg ikke har kunnskap om innsamlingsprosessen utover det de selv skriver og at jeg må stole på at informasjonen jeg får

av databasen er riktig. Eiendomsverdi AS (2019) skriver at de «følger og registrerer aktiviteten i det norske boligmarkedet daglig» og at de «leverer markedets mest komplette eiendomsdata i realtid».

Databasen leverer omsetningsrapporter hvor blant annet adresse, primærrom, salgspris, byggeår og omsetningshastighet inkludert. Omsetningsrapporten lages utfra valgt periode, valgt område og valgt boligtype. Jeg valgte områder på bakgrunn av postnumre. Lund omfatter postnumrene 4630, 4631 og 4632 samt deler av 4633 jf. figur 2.2. I denne analysen har jeg valgt å kun inkludere leiligheter og både selveier og borettslagsleiligheter er inkludert. Denne analysen omfatter både leiligheter i boligblokk og leiligheter som ikke er i blokk. Aksjeleiligheter utelates på grunn av måten databasen er bygd opp på.

I tillegg til omsetningsrapportene inneholder databasen ytterligere informasjon om de enkelte leilighetene. Funksjonen *enkelteiendom* i Eiendomsverdi AS gir informasjon om hvilken etasje leiligheten ligger i, antall soverom, om det er heis i blokken og om leiligheten har egen parkeringsplass. Her er det også en oversikt over hvor mange ganger leiligheten er solgt og pris og prisantydning ved hver omsetning. Dette gjør det mulig å si noe om leilighetens prisutvikling.

4.2. Innhenting av data

Fra omsetningsrapporten i Eiendomsverdi AS innhentes det data om følgende variabler: Adresse, eierform, boligtype, primærrom, bruksareal, registreringsdato, salgsdato, omsetningshastighet, prisantydning, omsetningspris, fellesgjeld, kvadratmeterpris, tomtestørrelse og byggeår.

Inntrykket jeg satt igjen med etter å ha lastet ned og sett litt nærmere på omsetningsrapport for Lund var at prisene nærmest sto stille. Dette støttes også av figur 2.6 som viser at kvadratmeterprisen i bydelen har en relativt slak kurve i perioden. Derfor er mitt formål å undersøke prisutviklingen i bydelen ytterligere og til dette trenger jeg mer data enn hva omsetningsrapportene gir.

Å laste ned omsetningsrapportene var relativt ukomplisert og lite tidkrevende. Å hente inn ytterligere informasjon om hver enkelt eiendom i datasettet visste seg å bli mer krevende enn først antatt. Data om enkeltleiligheter må samles manuelt for hver enkelt leilighet og legges til riktig observasjon i omsetningsrapporten. Eiendomsverdi AS har en funksjon hvor man kan søke på selveierleiligheter og en hvor man kan søke på borettslagsleiligheter. I datasettet for

bydelen Lund er ca 8% av observasjonene aksjeleiligheter. Da det ikke finnes en funksjon for aksjeleiligheter, vil innsamlingsprosessen for disse leilighetene bli tungvint og unøyaktig. Det ser heller ikke ut som prisene eller leilighetstypene skiller seg betydelig fra de andre eierformene. Derfor velger jeg å utelate aksjeleiligheter fra denne analysen, og dette er dermed en undersøkelse av selveier- og borettslagsleiligheter i bydelen Lund i Kristiansand.

For å kunne svare på problemstillingen og hypotesene trengs i tillegg informasjon om følgende variabler: antall soverom, hvilken etasje leiligheten ligger i, hvor mange ganger leiligheten er solgt i perioden 2010 til 2018, om det er heis i bygget, om leiligheten har parkeringsplass og om leiligheten ligger i blokk. De tre sistnevnte variablene kan bare ha variablene 0 og 1, hvor 1 svarer til «ja». Informasjon om disse variablene finnes ved å bruke Enkelteiendom-funksjonen.

Arbeidet med å samle informasjon om flere egenskaper ved leilighetene krevde stor grad av nøyaktighet. Å koble informasjonen fra enkelteidommene med omsetningsrapporten viste seg å være utfordrende. Det finnes flere leiligheter som har samme gatenavn og husnummer. I mange tilfeller kan leilighetene skilles på størrelse, men det finnes også leiligheter med samme størrelse og samme adresse. Jeg har også oppdaget tilfeller hvor det skilte et par kvadratmeter på forskjellige observasjoner av samme leilighet. For å være sikker på at jeg har koblet informasjon om samme leilighet, har jeg brukt registreringsdato, salgsdato og omsetningspris og sjekket at de stemmer overens for alle observasjonene.

Formålet med denne analysen er å undersøke hvilke faktorer som påvirker prisen på leiligheter i bydelen, og deretter undersøke om det finnes noen fellestrekk ved de leilighetene som er solgt med tap i tidsperioden 2010 – 2018. Dette krever at det også inkluderes et identifikasjonsnummer. Hver enkelt leilighet får sitt unike identifikasjonsnummer, og flere observasjoner av samme leilighet får følgelig samme identifikasjonsnummer.

Dataen som er samlet inn om leiligheter i bydelen Lund lagres kun på min pc. Når de nødvendige analysene er gjennomført slettes alt datamaterialet.

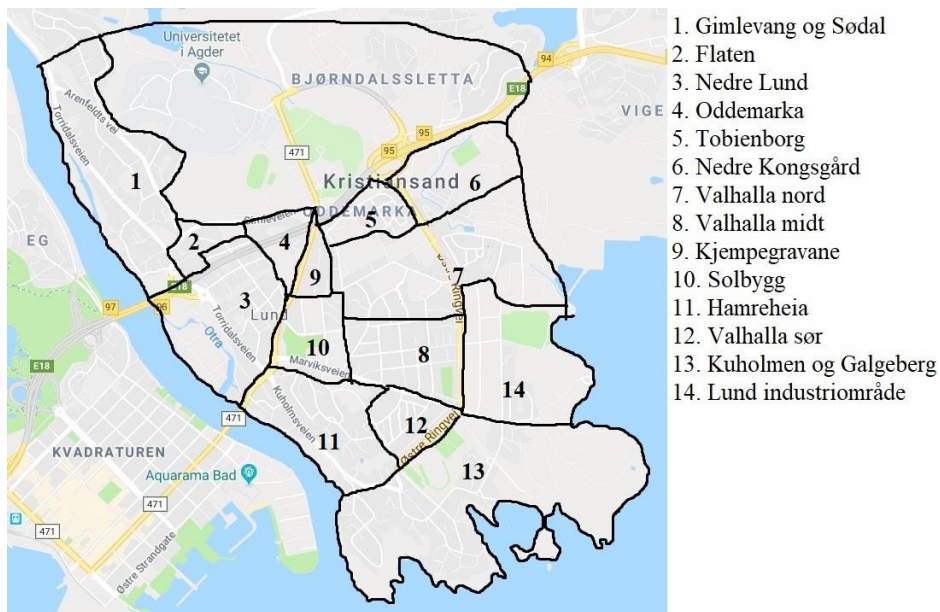
4.3. Bearbeiding av datamaterialet

Datasettet som inkluderer alle de nødvendige variablene importeres til Stata, og første rad importeres som variabelnavn. Variabelen omsetningshastighet leses som tekst, og må derfor gjøres numerisk.

Leilighetens eierform er oppgitt å enten være selveier eller borettslag. Det lages en dummyvariabel som har to verdier, 0 og 1, hvor 1=selveierleilighet, det betyr at observasjoner hvor denne variabelen har verdien 0 er borettslagsleiligheter.

4.3.1. Inndeling av områder i bydelen

For å undersøke om det er noen forskjeller mellom de ulike områdene i bydelen, deles den opp i mindre områder. I Eiendomsverdi er Kristiansand delt opp basert på postnumre eller områder. Jeg velger å ta utgangspunkt i disse områdene og finner de adressene som tilhører hvert område. Områdene er presentert på kartet under.



Figur 4.1 Områdekart (redigert skjermdump fra Google Maps (2019))

Disse 14 områdene brukes som utgangspunkt for å lage dummyvariabler for hvert område i bydelen. Variabelen L_n genereres, hvor $n=1, \dots, 14$. Hver variabel kan to verdier, 0 og 1, hvor verdien 1 angir at leiligheten ligger i det gitte området.

Området Solbygg skiller seg ut ved at det har klart flest observasjoner, mens Lund industriområde og Nedre Kongsgård har færrest. Solbygg er et område som nesten utelukkende består av blokker, og dermed blir det mange observasjoner fra området.

4.3.2. Avstand til sentrum

Avstanden til sentrum er et sentralt attributt i den hedonistiske metoden, og denne må bergenes. Bykjernen er satt til et sentralt punkt i gågaten i Kristiansand sentrum, *Markens gate 19, 4611 Kristiansand*.

I første omgang beregnet jeg avstanden fra et punkt i hvert av områdene definert i figur 4.1, slik at alle leiligheter i samme område fikk samme verdi på avstand til sentrum. Dette ble gjort ved hjelp av veibeskrivelsesfunksjonen i Google Maps. Senere ble jeg gjort oppmerksom på at det fantes et tillegg til Stata som beregnet avstand fra en startadresse til en sluttadresse. Ved hjelp av *georoute* funksjonen i Stata ble avstand fra de enkelte leiligheter til det bestemte punktet i sentrum beregnet.

4.3.3. Beregning av gevinst og tap

Fra datasettet som nå er laget, skilles det ut et mindre deldatasett som kun inneholder de observasjonene som er omsatt mer enn en gang. Dette datasettet skal brukes til å undersøke om det er noen forskjeller mellom de leilighetene som er solgt med gevinst og de som er solgt med tap.

For å kunne si noe om hvor mange leiligheter som er solgt med tap lages en variabel, *endring*. Endringen berignes ved å trekke totalprisen siste gang solgt fra totalprisen forrige gang leiligheten ble solgt. Ved å bruke totalprisen tas det også hensyn til om fellesgjelden har økt eller blitt redusert mellom salgene. Anta en leilighet som er solgt tre ganger, for eksempel i 2011, 2013 og 2018. Da berignes *endring* ved å trekke totalpris i 2013 fra 2011, og salgspris i 2018 minus totalpris i 2013. Verdien for gevinst eller tap legges da til observasjonene med salgsår 2013 og 2018. Det inkluderes også en variabel som kan ha to ulike verdier: 0 og 1. Hvor 1=tap og 0=gevinst.

Dette deldatasettet importeres til Stata, og de observasjonene der leiligheten er solgt første gang, altså der det ikke er registrert gevinst eller tap registrert, utelates. Ferdig rensset består datasettet av 657 observasjoner.

Det kan være interessant å se om avviket mellom omsetningspris og prisantydning er ulik for de leilighetene som er solgt med tap og de leilighetene som er solgt med gevinst. Derfor skapes en variabel, *avvik*, som er beregnet ved å trekke prisantydningen fra salgspris.

4.4. Datarensing

Før datasettet kan brukes til analyse, må det undersøkes om det finnes manglende verdier, feilkodinger og sære verdier. Byggeår og primærom lik 0, samt veldig lave omsetningspriser kan påvirke resultatene av analysen og derfor er det viktig forsikre seg om at dataene er logiske.

Data for variablene adresse, eierform, boligtype, primærrrom, registreringsdato, salgsdato, omsetningshastighet, prisantydning, omsetningspris, fellesgjeld, kvadratmeterpris og byggeår er kopiert fra Eiendomsverdi AS, og feil i datasettet kan komme av at jeg har limt inn dataene på feil plass. En annen feilkilde er at det er registrert feil tall, for eksempel at det mangler eller er lagt til et siffer.

Verdier for variablene antall soverom, leilighetens etasjeplassering, antall ganger leiligheten er omsatt i perioden 2010 – 2018, om det er parkeringsplass og om det er heis i bygget, samt identifikasjonsnummer er fylt inn manuelt. En stor mengde data ble fylt inn, og en mulig feilkilde er at feil verdi er lagt til for enkelte observasjoner. Derfor er det viktig å inspisere datamaterialet, for å forsikre meg om at det ikke er noen variabler som har ulogiske eller sære verdier.

Undersøkellesperioden er satt til salg i perioden fra 1.1.2010 til 31.12.2018. Derfor sjekkes det at salgsår bare har verdier mellom 2010 og 2018, da inntasting av salgsår er gjort manuelt, er det en risiko for tastefeil. Selv om variablene bare har logiske verdier kan det finnes observasjoner hvor feil år er registrert.

Det visste seg at flere av eiendommene som var registrert som leilighet ikke var leiligheter. Disse observasjonene utelates. Ytterligere en leilighet fjernes. Denne leiligheten er omsatt to ganger, den ene gangen for 5 050 000 og to år senere for 2 050 000. Begge observasjonene står oppført med primærrrom lik 60 m². Den første salgsprisen virker urealistisk for størrelsen på leiligheten, og Finn-annonsen viser at det i dette tilfellet er solgt en leilighet på 147 m² med mulighet for oppdeling.

Rådata i datasettet utgjør 2 656 observasjoner. Databasen er noe mangelfull og observasjoner som mangler verdi for primærrrom, omsetningspris, eller byggeår fjernes. I tillegg fjernes de observasjonene hvor det var mangelfull informasjon om enkelte eiendommen. Etter disse observasjonene er fjernet står jeg igjen med 1 871 observasjoner.

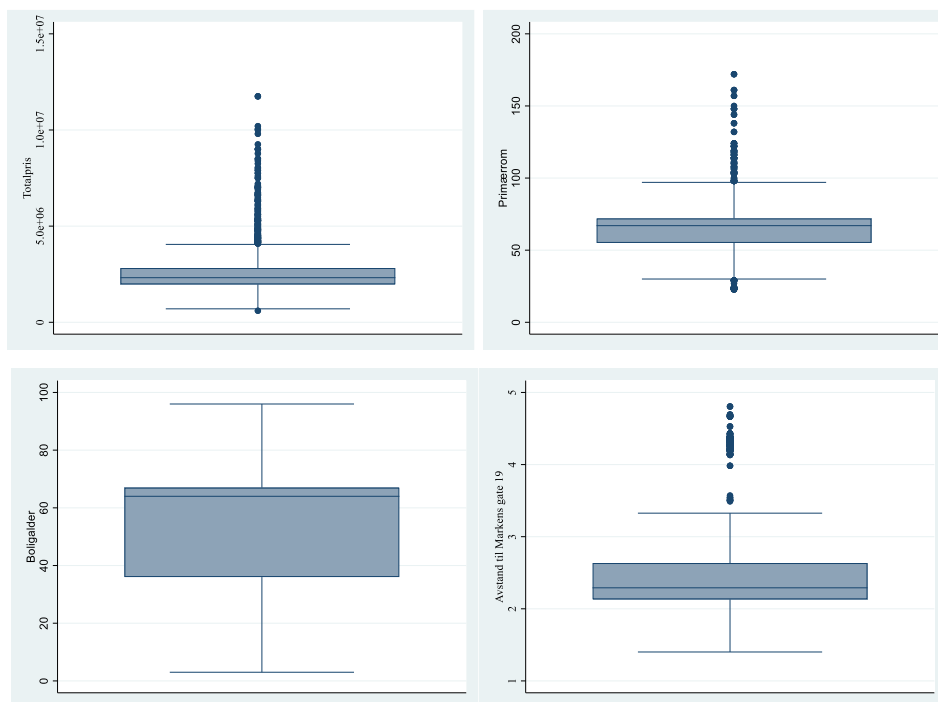
Opprinnelig antall observasjoner	2 656
Observasjoner fjernet på grunn av manglende P-rom	452
Observasjoner fjernet på grunn av manglende eller unormal omsetningspris	16
Observasjoner fjernet på grunn av manglende byggeår	11
Observasjoner fjernet på grunn av manglende data om enkelteiendom	306
Observasjoner etter datarensing	1871

Tabell 4.1 Endringer gjort for å oppnå et komplett datasett

Før datamaterialet brukes i analyse forsikrer jeg meg også om at ingen dummyvariabler har andre verdier enn 0 og 1. Observasjoner der det finnes dummyvariabler registrert med andre verdier enn 0 og 1 undersøkes, og rettes opp.

En outlier, heretter kalt avvik, er en observasjon som betydelig skiller seg fra de andre observasjonene. Observasjoner som er veldig forskjellig fra de andre observasjonene vil påvirke resultatet, og det er derfor viktig å sjekke at verdiene er korrekt registrert (Sekaran & Bougie, 2016).

Det er flere muligheter for å identifisere avvik, blant annet ved hjelp av frekvenstabeller, tabeller med maksimum, minimum og gjennomsnitt eller box plots. I box plottet er medianen vist som en linje inne i den fargede boksen. Medianen er den midterste verdien når observasjonene er organisert kronologisk. Ved å organisere observasjonene i kronologisk rekkefølge og dele på fire, finner vi de fire kvartilene. Bunnpunktet i den fargede boksen er 1.kvartil, mens toppunktet er 3.kvartil, dermed viser boksen de midterste 50% av observasjonene når de er ordnet stigende rekkefølge. Box plottet viser også variasjonen og de individuelle punktene kan være avvik (Sekaran & Bougie, 2016).



Figur 4.2 Box plots: øverst fra venstre: totalpris, størrelse, boligalder, avstand til sentrum

Figur 4.2 viser at det er mange individuelle punkter for de tre variablene totalpris, primærrom og avstand til bysentrum. Disse punktene skiller seg betydelig fra resten av populasjonen, og kan påvirke resultatene. For totalpris og primærrom ser vi at det finnes noen leiligheter som skiller seg ut ved at de er mindre enn det flesteparten av verdiene i datasettet er. Det finnes også noen observasjoner som har avstand til sentrum som avviker fra majoriteten av observasjonene ved at de er lokalisert lenger fra sentrum. Observasjoner til de tilhørende avvik inspiseres nærmere, og all tilgjengelig informasjon tilsier at registreringen er korrekt. Hvordan disse observasjonene som skiller seg betydelig fra de andre observasjonene skal behandles, kommer jeg tilbake til i forbindelse med analyse i kapittel 6. For boligalder er det ingen observasjoner som skiller seg betydelig ut.

At det ikke finnes observasjoner med ulogiske eller sære verdier er ingen garanti for at det ikke er feilkodinger. Det kan fortsatt finnes tilfeller hvor det er verdien «0» er registrert i stedet for «1» eller at feil antall soverom eller etasje er oppgitt. Men ved å undersøke datamaterialet ytterligere er risikoen for store feil i observasjonene minimert.

4.5. Deskriptiv statistikk for hele datasettet

Før jeg begynner på analysen skal de ulike variablene som skal undersøkes beskrives. Først presenteres en tabell med oppsummerende statistikk for hele datasettet. Deretter vil de variablene som krever ytterligere forklaring beskrives. Disse variablene vil forklares både ved

hjelp av ord, statistikk og figurer. Hensikten med denne delen er å beskrive datamaterialet ved hjelp av noen beskrivende verdier samt histogram.

I tabell 4.2 er gjennomsnittsverdier, standardavvik, minimumsverdier og maksimumsverdier, samt median for de variabler som antas å påvirke leilighetsprisen presentert. Standardavvik er definert som kvadratroten av variansen, og er det vanligste spredningsmålet. Standardavviket brukes til å måle hvor tett observasjonene er samlet rundt middelverdien og forteller noe om hvor mye observasjonene i gjennomsnitt avviker fra middelverdien (Gripsrud, Olsson & Silkoset, 2017). Median er et mål som viser verdien av den midterste observasjonen når observasjonene er sortert i kronologisk rekkefølge.

Variabel	Gjennomsnitt	Standardavvik	Minimumsverdi	Maksimumsverdi	Median
Totalpris	2 602 480	1 227 072	600 000	11 800 000	2320000
Primærrom	68,65	24,09	16	214	68
Boligalder	52,53	25,54	1	219	63
Antall soverom	1,83	0,78	0	6	2
Etasjeplassering	2,45	1,70	1	10	2
Avstand	2,54	0,69	1,401	4,805	2,291
Omsetningshyppighet	1,89	0,82	1	5	2
Parkeringsplass	0,37	0,48	0	1	
Heis	0,30	0,46	0	1	
Selveier	0,57	0,49	0	1	
Blokk	0,72	0,45	0	1	

Tabell 4.2 Deskriptiv statistikk for bydel Lund

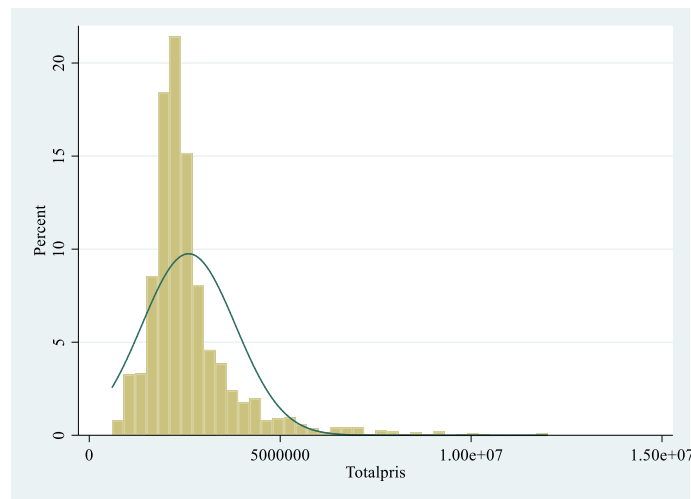
Totalpris

I tillegg til den omsetningsprisen en kjøper leilighet for, følger det i mange tilfeller med en fellesgjeld. Dette er gjeld et borettslag eller et sameie har tatt opp, ofte i forbindelse med vedlikehold og bygging. Når en kjøper en leilighet i et sameie eller borettslag hvor det er fellesgjeld betaler man en salgssum, samt at man overtar leilighetens andel av fellesgjelden. Fellesgjelden blir en del av de månedlige felleskostnadene (Aktiv Eiendomsmegling, 2019). Jeg antar at fellesgjeld er lik 0 når det ikke er oppgitt fellesgjeld i databasen. Fellesgjelden antas å påvirke boligprisen og variabelen tas derfor med i betraktning.

For å finne totalprisen legges omsetningspris og fellesgjeld sammen. I henhold til Eretveit og Theisen (2016) som studerte boligmarkedet i Kristiansand, legges ikke hele fellesgjelden til, men det brukes en faktor, γ . Dermed kan leilighetens totalpris uttrykkes på følgende måte:

$$\text{Totalpris} = \text{Omsetningspris} \times \gamma \text{Fellesgjeld} \quad (4.1)$$

I henhold til resultatene til Eretveit og Theisen (2016) er $\gamma=0,87$, det vil si at 87% av fellesgjelden legges til omsetningsprisen for å finne totalpris.



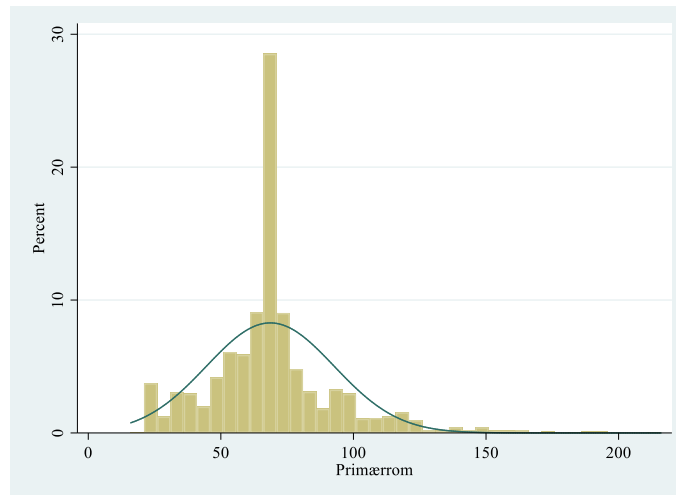
Figur 4.3 Histogram for totalpris

Figur 4.3 viser fordelingen av leilighetens totalpris. Hver søyle har en bredde på 300 000. Histogrammet viser at majoriteten av de solgte leilighetene i bydelen har en totalpris på mellom 1,8 millioner og 2,5 millioner kroner.

Leilighetens størrelse

Leilighetens størrelse omfattes av variabelen primærrom som er et størrelsesmål og omfatter alle rom i leiligheten med unntak av boden. Verdien av variabelen er oppgitt i kvadratmeter. Størrelsen til denne variabelen er bruksarealet, målt på innsiden av veggene (Anderssen, 2019a). For å beregne leilighetens kvadratmeterpris divideres omsetningsprisen på leilighetens størrelse.

Figur 4.4 viser hvor fordelingen av størrelsen på leiligheter i datasettet. Hver søyle har en bredde på 5. Figuren viser at det er flest observasjoner med primærrom på mellom 70 m² og 75 m², hvor 28,5% av observasjonene befinner seg.

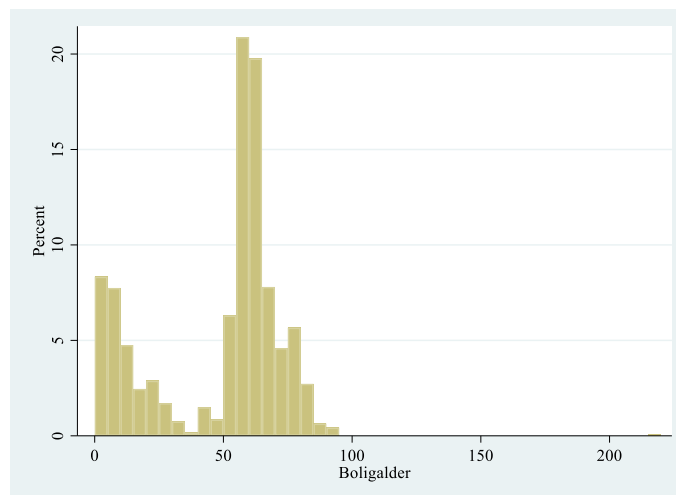


Figur 4.4 Histogram for primærom

Om lag 22,5% av leilighetene i bydelen er mindre enn 55 m². Av de leilighetene som er inkludert i analysen er nesten 30% av leilighetene ett- eller to-roms leiligheter.

Leilighetens alder

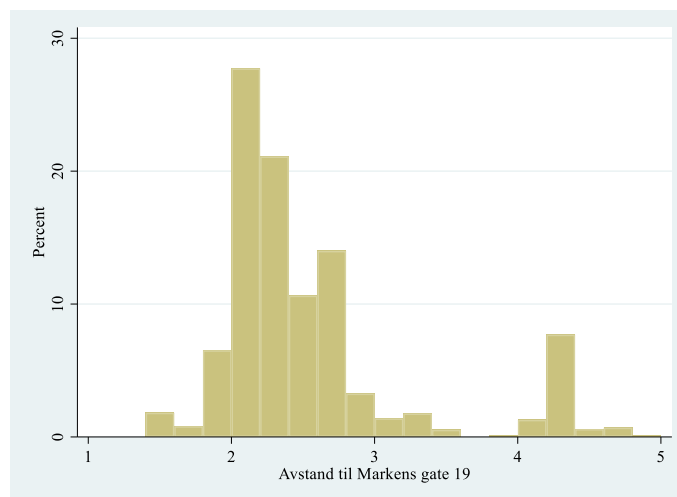
Boligalderen er beregnet ved å trekke leilighetens byggeår fra salgsår. Histogrammet viser at det er flest leiligheter som er mer enn 50 år gamle, men at det i løpet av de siste tiårene er bygget en del nye leiligheter.



Figur 4.5 Histogram for boligalder

Avstand til bykjernen

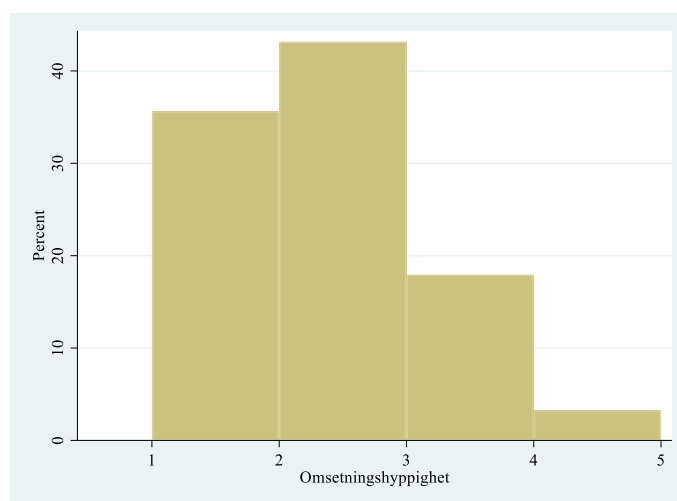
Variabelen *avstand* måler antall kilometer fra leilighetens adresse til Markens gate 19, som er et sentralt punkt i sentrum. Om lag 77% av observasjonene har en avstand på mellom 2 km og 3 km til sentrum, jf. figur 4.6.



Figur 4.6 Histogram for avstandsvariabelen

Omsetningshyppighet

Histogrammet for variabelen som sier hvor mange ganger en leilighet er omsatt i perioden 2010-2018 er gjengitt nedenfor. Histogrammet viser hvor stor prosent av utvalget som er registrert med de ulike omsetningshyppighetene.



Figur 4.7 Histogram for omsetningshyppighet

4.5.1. Korrelasjon

Korrelasjon mellom to variabler måler i hvilken grad de to variablene har en lineær sammenheng, altså i hvilken grad to variabler henger sammen. Korrelasjonen sier derimot ingenting om endring i den ene variabelen forårsaker endringer i den andre variabelen (Brooks, 2008).

Korrelasjonsmatrisen brukes til å undersøke forhold mellom variabler på intervall eller forholdstallsnivå. I matrisen kommer det frem om det er et positivt eller negativt forhold mellom to variabler, og hvor sterkt dette forholdet er. Korrelasjonskoeffisienten varierer mellom -1 og 1, hvor -1 er perfekt negativt forhold og 1 er perfekt positivt forhold (Sekaran & Bougie, 2016).

I korrelasjonsmatrisen er variablene presentert ved hjelp av symboler, forklaring på disse symbolene finnes i tabell 4.5 i kapittel 4.7.

	P	M	S	Y	km	E	G	H	BD	SD
P	1,000									
M	0,765	1,000								
S	0,133	0,028	1,000							
Y	-0,472	-0,196	-0,078	1,000						
km	-0,134	0,017	-0,016	-0,034	1,000					
E	-0,001	-0,107	0,015	-0,010	-0,175	1,000				
G	0,478	0,313	0,045	-0,617	0,040	0,037	1,000			
H	0,461	0,205	0,071	-0,670	-0,076	0,267	0,677	1,000		
BD	0,095	-0,078	0,059	-0,376	-0,142	0,322	0,161	0,401	1,000	
SD	0,369	0,290	-0,004	-0,331	0,159	-0,318	0,458	0,208	-0,347	1,000

Tabell 4.3 Korrelasjonsmatrise (n = 1871 observasjoner)

Korrelasjonsmatrisen presentert i tabell 4.5 viser at det er en relativt sterk, positiv korrelasjon mellom totalpris og størrelse på leilighet. Dette indikerer at desto større leiligheten er, desto høyere er totalprisen.

Det er en middels sterk, negativ korrelasjon mellom leilighetens totalpris og alderen på leiligheten. Dette indikerer at eldre leiligheter har lavere totalpris enn nyere leiligheter.

Korrelasjonsmatrisen viser også at det er en middels sterk positiv korrelasjon mellom leilighetens totalpris og egenskapene parkeringsplass og heis.

4.6. Deskriptiv statistikk for leiligheter solgt mer enn en gang i perioden

Det er 657 observasjoner i deldatasettet som er omsatt for andre, tredje, fjerde eller femte gang i løpet av perioden. Til disse observasjonene er det lagt til en variabel som indikerer om leiligheten er solgt med gevinst eller tap. Nedenfor presenteres deskriptiv statistikk for dette deldatasettet.

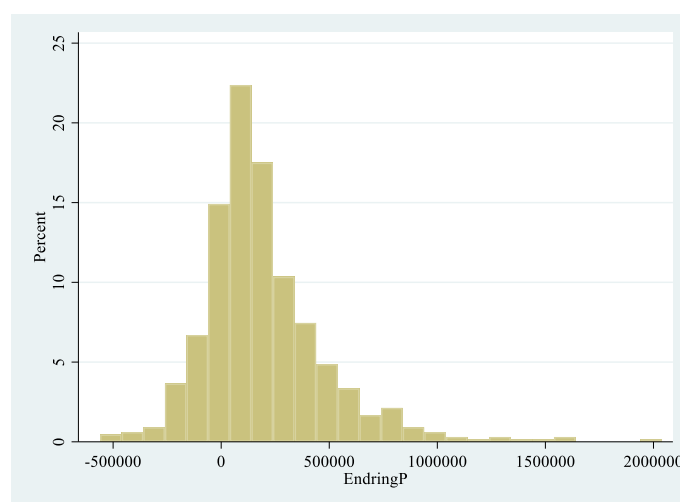
At undersøkelsesperioden er satt til perioden 2010 – 2018, medfører at en leilighet som for eksempel er solgt i 2009 og 2016 ikke vil være med i utvalget.

Variabel	Gjennomsnitt	Standardavvik	Minimumsverdi	Maksimumsverdi	Median
Totalpris	2 527 457	951 933	700 000	9 000 000	2 350 000
Endring	192 125	288 676	-560 359	2 000 000	140 000
Avvik	1 153	138 892	-800 000	900 000	0
Primærrom	66,19	22,89	23	172	67
Boligalder	54,84	24,65	3	96	64
Antall soverom	1,77	0,80	0	6	2
Etasjeplassering	2,44	1,75	1	10	2
Avstand	2,54	0,72	1,401	4,805	2,29
Omsetningshyppighet	2,47	0,64	2	5	2
Omsetningshastighet	42,03	72,15	0	887	14
Parkeringsplass	0,35	0,47	0	1	
Heis	0,25	0,43	0	1	
Selveier	0,55	0,50	0	1	
Blokk	0,68	0,47	0	1	
Tap	0,1948	0,39	0	1	

Tabell 4.4 Deskriptiv statistikk for leiligheter som er solgt mer enn en gang i perioden 2010 - 2018 (n=657)

Endring

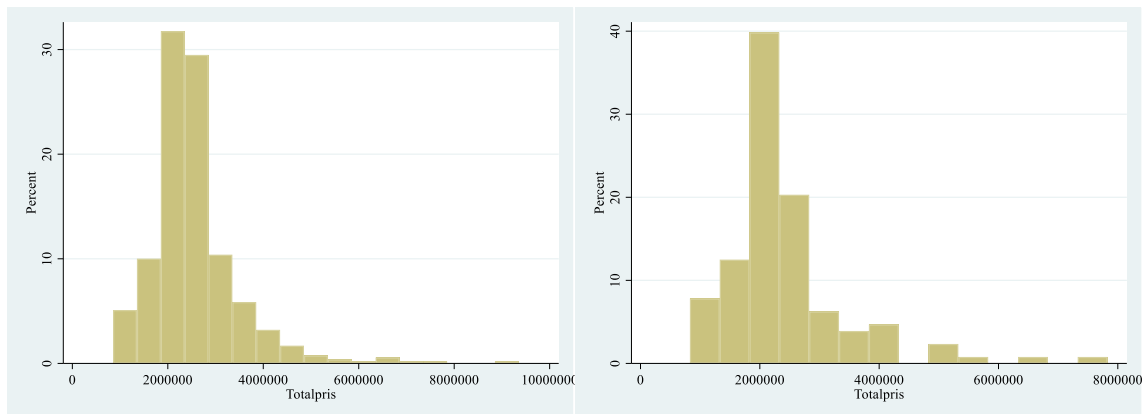
Variabelen «Endring» er definert som totalpris ved forrige omsetning minus totalpris ved nåværende omsetning. Det vil si at dersom endringen er negativ er leiligheten solgt med tap. Histogrammet ovenfor viser at om lag 20% av leilighetene i deldatasettet er solgt med tap.



Figur 4.8 Histogram over endring i totalpris mellom to salgstidspunkt

Videre følger en grafisk sammenlikning av noen utvalgte variabler mellom de leiligheter som selges med gevinst og de leiligheter som selges med tap.

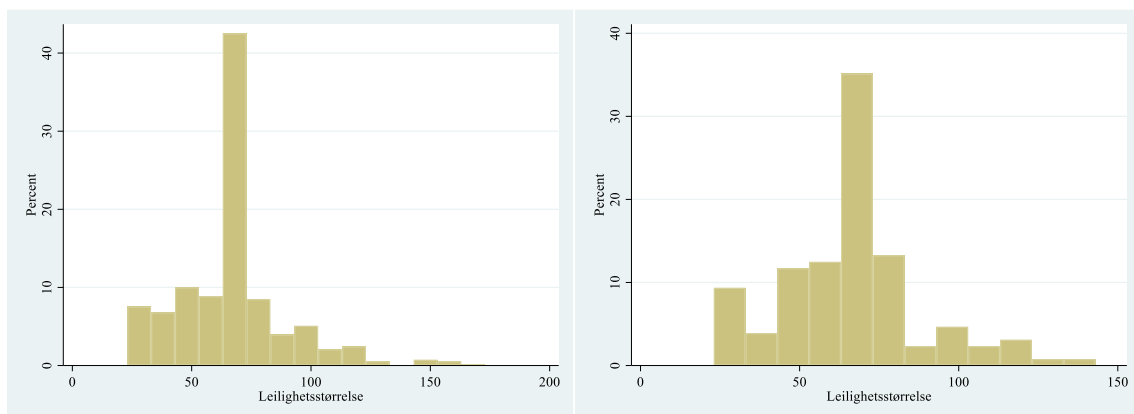
Totalpris



Figur 4.9 Histogram for totalpris, leiligheter solgt med gevinst til venstre og leiligheter solgt med tap til høyre

Histogrammene viser at det er en liten forskjell i totalpris mellom gruppene. Hver søyle har en bredde på kr 500 000. Som vi ser av figur 4.9 er det omtrent 60% av leilighetene som har en totalpris på mellom kr 2 000 000 og kr 3 000 000 i begge gruppene. Men i gruppen med leiligheter solgt med tap er det om lag 10% flere leiligheter som har en totalpris på mellom kr 2 000 000 og kr 2 500 000.

Leilighetens størrelse

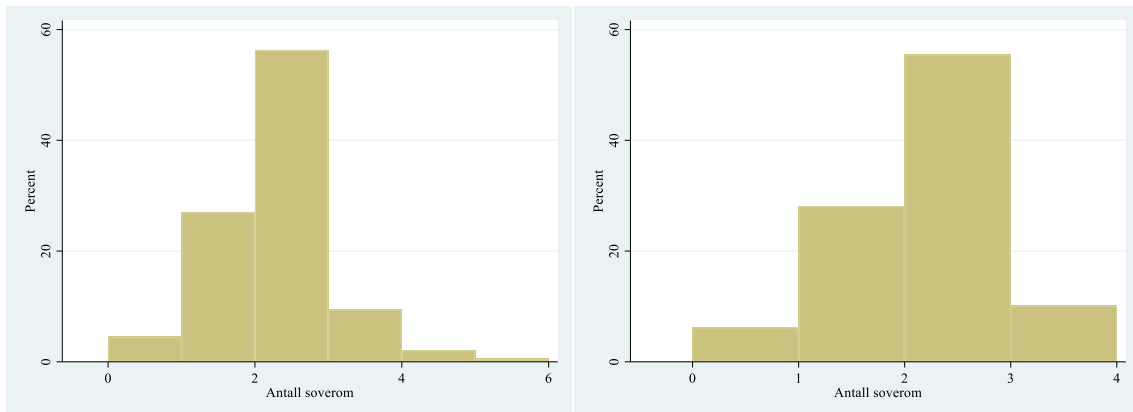


Figur 4.10 Primærrom, leiligheter solgt med gevinst til venstre og leiligheter solgt med tap til høyre

Begge histogrammene starter på 23 m², og hver stolpe har bredde på 10. Av figuren ser vi at det er flest leiligheter med størrelse på mellom 63 m² og 73 m² for både leiligheter solgt med gevinst og leiligheter solgt med tap, det er henholdsvis 42,5% og 35,1%.

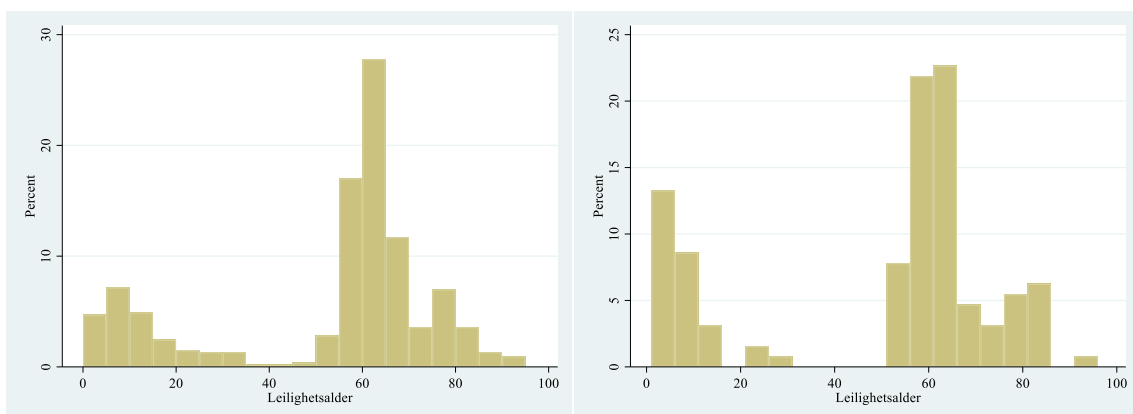
Det ser ikke ut som om det er noen betydelige forskjeller i leilighetens størrelse mellom de to gruppene. Figur 4.10 indikerer at de største leilighetene i bydelen har blitt solgt med gevinst i denne perioden.

Av figur 4.11 ser vi at det ikke finnes noen leiligheter med fem eller seks soverom som er solgt med tap. Men utover dette er fordelingen av antall soverom i de to gruppene relativt lik.



Figur 4.11 Antall soverom, leiligheter solgt med gevinst til venstre, leiligheter solgt med tap til høyre

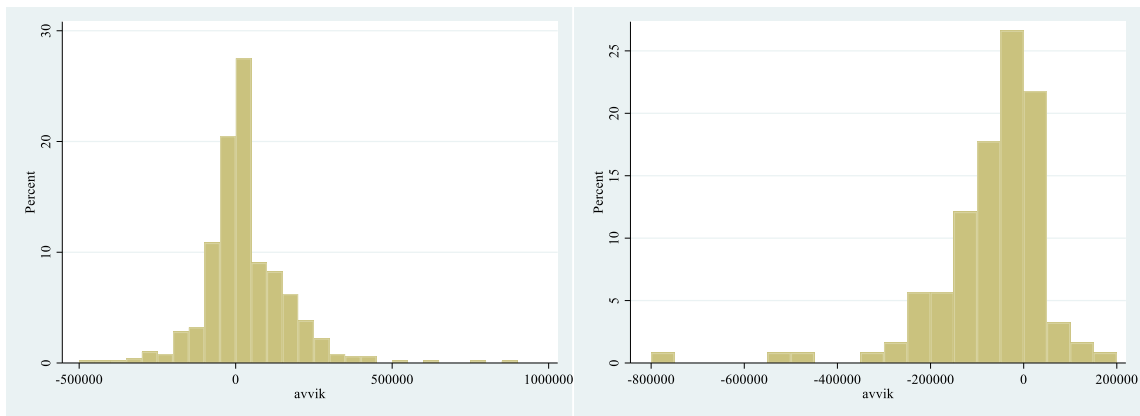
Leilighetens alder



Figur 4.12 Boligalder for leiligheter solgt med gevinst til venstre og solgt med tap til høyre

Figur 4.12 indikerer at det er en større andel leiligheter med en alder på mellom 0 og 20 år som selges med tap enn med gevinst.

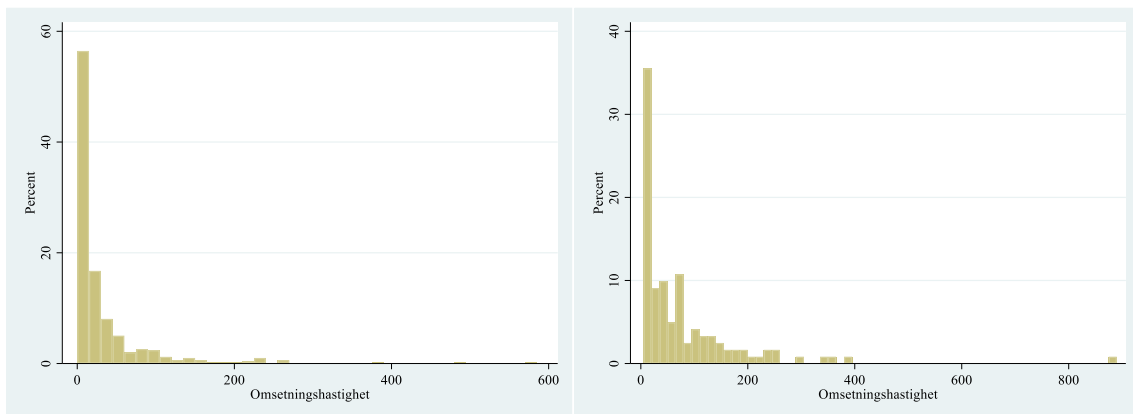
Avvik mellom salgspris og prisantydning



Figur 4.13 Avvik mellom salgspris og prisantydning, gevinst til venstre og tap til høyre

Variabelen «Avvik» er definert som salgspris minus prisantydning. Det vil at dersom denne verdien er negativ er leiligheten solgt til en pris som er lavere enn prisantydningen. Av figur 4.13 er det tydelig at avviket mellom salgspris og prisantydning er forskjellig for de to gruppene. I gruppen for gevinst er det flest leiligheter som selges til en høyere pris enn prisantydning, mens det i gruppen for tap er flest leiligheter med negativt avvik. Mindre enn 30% av leilighetene som selges med tap har høyere salgspris enn prisantydning. Av de leilighetene som selges med gevinst er omtrent 40% solgt til en pris lavere enn prisantydning.

Omsetningshastighet



Figur 4.14 Omsetningshastighet, gevinst til venstre og tap til høyre

Det ser ut som om en større andel av de leilighetene som selges med gevinst selges i løpet av kort tid. Om lag 56% av de leilighetene som selges med gevinst selges i løpet av de første 15 dagene, det er nesten 20% flere enn i gruppen som består av leiligheter solgt med tap.

4.7. Oppsummering av analyserte variabler

Tabellen nedenfor oppsummer de variabler som vil bli brukt i videre analyse. I tabeller med mange variabler og kolonner, slik at det er liten plass, vil symbolene under «Stata-variabel» benyttes gjennom hele avhandlingen.

Variabelnavn	Stata-variabel	Forklaring
Primærrrom	M	Målt i m ²
Omsetningspris	p	Målt i norske kroner
Totalpris	P	Omsetningspris + 0,87*Fellesgjeld.
Boligalder	Y	2019 – boligens byggeår.
Soverom	R	Målt i antall
Leilighetens etasjeplassering	E	
Omsetningshyppighet	Z	Måles i antall ganger leiligheten er omsatt i perioden
Parkeringsplass	G	1 = har parkeringsplass 0 = har ikke parkeringsplass
Heis	H	1 = heis i blokken 0 = ikke heis i blokken
Avstand	km	Avstand til bykjerne, målt i km.
Selveier	SD	1 = selveier 0 = borettslag
Blokk	BD	1 = selveier 0 = borettslag
År201X	år2010,...,år2018	Salgsår År2010: leilighet solgt i 2010 får verdien 1 År2011: leilighet solgt i 2011 får verdien 1.
L _n	L1, L2,...,L14	Områdedummy. Dersom variabelen har verdien 1 indikerer det at observasjonen ligger i gitt område
Endring	C	Endring mellom forrige og nåværende salgspris
Avvik	d	Avvik mellom omsetningspris og prisantydning

Tabell 4.5 Variabler med tilhørende Stata-variabel

4.8. Utelatte variabler

Hvilke variabler som inkluderes i analysen påvirker resultatet. På bakgrunn av den informasjonen jeg har tilgjengelig har jeg valgt å bruke de variablene som er presentert i tabell 4.5. I denne delen følger noen variabler som kunne vært nyttige å inkludere i analysen, men som jeg på nåværende tidspunkt ikke har inkludert.

Et alternativ til totalpris var å se på kvadratmeterpris for å se på utviklingen i denne. Men det visste seg at mange av de samme leilighetene var registrert med ulik størrelse på primærrrom,

selv om det er helt sikkert at det var samme leilighet. Det største avviket jeg merket meg var på 8 m². I Eiendomsverdi er det opplyst om kvadratmeterpris, men stikkprøver visste at denne ikke alltid stemte overens med de opplysninger som var gitt. Det var også tilfeller der kvadratmeterpris i omsetningsrapporten avviker fra oppgitt kvadratmeterpris under enkelteiendom, selv om alle andre opplysninger var like. Et eksempel er en leilighet som er solgt to ganger, til henholdsvis kr 2 325 000 og kr 2 250 000. De tilhørende kvadratmeterprisene er kr 38 680 og kr 37 463. Den første gangen er leiligheten registrert med et primærrom på 76 m², og beregnet kvadratmeterpris blir da kr 30 592, mot kr 38 680 som var registrert. Den andre gangen leiligheten ble solgt, var primærrom opplyst å være 68 m², som gir en kvadratmeterpris på kr 33 088 mot kr 37 463 som var registret. Begge de registrerte kvadratmeterprisene er gjort med utgangspunkt i ca. 60,1 m², men dette stemmer ikke overens med de andre opplysningene.

Da leilighetene er registrert med forskjellig primærrom, samt at den oppgitte kvadratmeterprisen ikke stemmer overens med leilighetens størrelse vil det bli unøyaktig å bruke denne variabelen for å se på prisutvikling. Et alternativ hadde vært å endre de observasjonene det gjaldt slik at de hadde den samme størrelsen, men dette vil bli veldig tidkrevende, samt at jeg ikke har kunnskap om hvilken størrelse som er den rette.

I analysen er variablene støy og utsikt, som antas å påvirke leilighetsprisen utelatt. Å samle inn disse variablene ville vært et tidkrevende arbeid, samt at faren for unøyaktigheter er stor. Det er også forskjellig type utsikt, og det antas at havutsikt øker leilighetens pris mer enn utsikt til en gate gjør. Heller ikke informasjon om tomtepriser og materialpriser er inkludert. Dette antas også å ha en påvirkning på leilighetens totalpris.

Makroøkonomiske faktorer som arbeidsledighet og inntektsnivå er ikke inkludert i modellen, selv om det kan påvirke boligsalg og prisen på boligen. Det antas at i perioder med høy arbeidsledighet, vil leilighetsprisen være lavere enn i gode tider. På bakgrunn av modellen jeg har lagd er det ikke mulig å bevise at det er en årsakssammenheng mellom leilighetspris og makroøkonomiske faktorer.

På bakgrunn av at sentrale variabler ikke er tatt med i modellen, er det begrenset hvor egnet modellen er til å predikere fremtidig boligpris. Derimot sier modellen noe om i hvor stor grad de inkluderte variablene påvirker leilighetsprisen.

5. Metode

I dette kapitlet er de metoder som er brukt for å besvare problemstillingen presentert. Den første delen av analysen vil være å undersøke hvilke faktorer som påvirker leilighetsprisene i bydelen. I denne delen brukes regresjonsanalyse og teorien om den hedonistiske modellen som er presentert i kapittel 3.2. I den andre delen av analysen er formålet å undersøke om det finnes noen betydelige forskjeller mellom leiligheter som selges med tap og leiligheter som selges med gevinst. Til dette skal t-test som sammenlikner gjennomsnitt mellom to grupper brukes.

5.1. Økonometrisk modell

5.1.1. Regresjonsanalyse

I henhold til Brooks (2008) brukes regresjon til å beskrive og evaluere forholdet mellom en avhengig variabel og en eller flere uavhengige variabler. Med regresjonen søker vi å forklare endringer i den avhengige variabelen gitt endringer i de uavhengige variablene.

Regresjonsanalyse kan brukes til å teste om det er en signifikant sammenheng mellom en avhengig variabel og en eller flere uavhengige variabler. Årsakssammenhengen kan derimot ikke bevises, men ved hjelp av regresjonsanalysen kan vi teste om sammenhenger er signifikant forskjellige fra null. Analysen kan brukes til å undersøke hvordan endringer i de uavhengige variablene kan forklare endringer i den avhengige variabelen (Gripsrud et al., 2017).

Enkel og multippel regresjon

Det skilles mellom enkel og multippel regresjonsanalyse. I enkel regresjon søker vi å forklare hvordan en uavhengig variabel påvirker en avhengig variabel. Den generelle enkle regresjonsmodellen er presentert i funksjon 5.1.

$$y = \alpha + \beta_1 x_1 + \varepsilon \quad (5.1)$$

I modellen ovenfor representerer α konstantleddet og β_1 er stigningstallet.

I multippel regresjonsanalyse er mer enn en uavhengig variabel inkludert for å forklare endringene i den avhengige variabelen. Den generelle regresjonsmodellen med en avhengig variabel og flere uavhengige variabler er presentert som funksjon 5.2. Det er denne regresjonsmodellen som vil bli brukt videre i analysen, da det antas at omsetningsprisen påvirkes av flere forskjellige variabler.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p + \varepsilon \quad (5.2)$$

5.1.2. Spesifisering av funksjonsform

Den hedonistiske metoden brukes til å undersøke hvordan ulike attributter eller egenskaper ved boligen påvirker leilighetsprisen. Den hedonistiske prisfunksjonen ser slik ut:

$$P = f(z_1, z_2, \dots, z_n) = P(Z) \quad (5.3)$$

Den hedonistiske prisfunksjonen (5.3) brukes som utgangspunkt for analysen og det må med utgangspunkt i denne formuleres en funksjon som kan brukes i regresjon. Det finnes mange varianter av denne funksjonen som kan brukes og i henhold til Osland (2001) er et sentralt spørsmål hvilken funksjonsform som skal velges. I analyser som omhandler eiendomspriser brukes ofte lineær, semilogaritmisk og dobbellogaritmisk funksjonsform eller Box-Cox transformasjon. Slik som blant annet Schnare og Struyk (1976) velger jeg å bruke følgende funksjonsformer:

$$P = \alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon \quad (5.4)$$

$$P = e^{\alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon} \quad (5.5)$$

$$P = \alpha z_1^{\beta_1} z_2^{\beta_2} z_3^{\beta_3} e^{z_4^{\beta_4} + z_5^{\beta_5} + \varepsilon} \quad (5.6)$$

Lineær funksjonsform

Funksjon 5.4 er en lineær funksjon som kan brukes i regresjonsanalyse uten at den må transformeres. De kategoriske variablene inkluderes som dummyvariabler.

$$P = \alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon \quad (5.4)$$

Regresjonskoeffisientene, β_i , forteller hvordan en endring i den uavhengige variabelen påvirker den avhengige variabelen. De estimerte β -verdiene gir uttrykk for hvor mye, og i hvilken retning de ulike attributtene påvirker leilighetsprisen.

Semilogaritmisk funksjonsform

Funksjon 5.5 er en eksponentialfunksjon hvor høyresiden er på formen e^x , hvor e er en matematisk konstant med verdi tilnærmet lik 2,71828.

$$P = e^{\alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon} \quad (5.5)$$

Ved hjelp av den naturlige logaritmen kan eksponentialfunksjonen få en lineær funksjonsform. Funksjonen transformeres ved å legge til den naturlige logaritmen, \ln , på

begge sider av likhetstegnet. En logaritmisk transformasjon av funksjonen fører til at høyresiden av funksjonen blir på samme form som den generelle regresjonsmodellen presentert i funksjon 5.2. På venstresiden vil jeg stå igjen med den naturlige logaritmen av totalprisen, $\ln P$.

Etter transformering av funksjon 5.5 får funksjonen en logaritmisk venstreside og en ikke-logaritmisk høyreside, dette kommer frem av funksjon 5.7. Dette kalles en semilogaritmisk funksjon eller en log-lineær funksjon.

$$\ln P = \alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon \quad (5.7)$$

Hver av de estimerte regresjonskoeffisientene, β_i , tolkes som den partielt deriverte av den naturlige logaritmen til den avhengige variabelen med hensyn på den tilhørende uavhengige variabelen. Det vil si at en økning på en enhet i uavhengige variabelen fører til en $100\% \times \beta_i$ økning eller reduksjon i den avhengige variabelen (Brooks, 2008).

Dobbellogaritmisk funksjonsform

Den tredje funksjonsformen har attributtene som «grunntall» som opphøyes i regresjonskoeffisientene. De ulike «grunntallene» med tilhørende potens multipliseres med hverandre. For at funksjon 5.6 skal kunne brukes i lineær regresjon må den transformeres ved hjelp av den naturlige logaritmen.

$$P = \alpha z_1^{\beta_1} z_2^{\beta_2} z_3^{\beta_3} e^{z_4^{\beta_4} + z_5^{\beta_5} + \varepsilon} \quad (5.6)$$

I funksjon 5.6 er variabel 4 og 5 dummyvariabler som bare har verdiene 0 eller 1.

Dummyvariabler er inkludert fordi variabler på nominalnivå og ordinalnivå kan ha en forklaringskraft, men kan ikke brukes i regresjon dersom de ikke er bearbeidet.

Dummyvariablene står på formen e^x , som medfører at jeg etter transformasjon med den naturlige logaritmen kun står igjen med x .

$$\ln P = \ln \alpha + \ln \beta_1 z_1 + \ln \beta_2 z_2 + \ln \beta_3 z_3 + \beta_4 z_4 + \beta_5 z_5 + \varepsilon \quad (5.8)$$

Regresjonskoeffisientene fra den dobbellogaritmiske funksjonsformen tolkes som elastisiteten til den avhengige variabelen med hensyn på den uavhengige variabelen. En betakoeffisient på for eksempel $\beta=5,00$, betyr at en økning på 1% i den uavhengige variabelen fører til at den avhengige variabelen øker med 5% dersom alt annet er konstant (Brooks, 2008).

I kapittel 6 vil det gjennomføres regresjonsanalyser med de tre modellene beskrevet ovenfor.

Den funksjonsformen som best forklarer omsetningsprisen, samt oppfyller forutsetningene for

regresjon vil bli valgt. Utgangspunktet for de tre regresjonsmodellene som skal brukes i lineær regresjon er presentert under.

$$P = \alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon \quad (5.4)$$

$$\ln P = \alpha + \beta_1 z_1 + \beta_2 z_2 + \dots + \beta_p z_p + \varepsilon \quad (5.7)$$

$$\ln P = \ln \alpha + \ln \beta_1 z_1 + \ln \beta_2 z_2 + \ln \beta_3 z_3 + \beta_4 z_4 + \beta_5 z_5 + \varepsilon \quad (5.8)$$

5.1.3. Forutsetninger for regresjonsanalyse

For at regresjonsanalysen skal gi troverdige estimater er det viktig at disse syv forutsetningene oppfylles. Forutsetningene presentert under er basert på Brooks (2008), Verbeek (2008) og Sekaran og Bougie (2016).

1. Det må være et lineært forhold mellom den avhengige variabelen og de uavhengige variablene.
2. $\mathbb{E}[\varepsilon_i] = 0$, $i = 1, \dots, N$. Dette betyr at forventet verdi av feilleddene, også kalt residualer eller restledd, er null.
3. Multikollinearitet innebærer at to eller flere uavhengige variabler er sterkt korrelert og dette bør unngås. Dersom det er en situasjon der to eller flere uavhengige variabler er sterkt korrelert, vil estimering av regresjonskoeffisienter være umulig. Estimerer gjort basert på regresjonskoeffisientene vil ikke være troverdige. En korrelasjon på mer enn 0,70 regnes vanligvis som høy (Sekaran & Bougie, 2016).
4. $Cov[\varepsilon_i, x_i] = 0$, $i, j = 1, \dots, N, i \neq j$. Residual og x-verdi skal ikke være korrelert.
5. $V[\varepsilon_i] = \sigma^2$, $i = 1, \dots, N$, hvor $\sigma^2 < \infty$. Variansen til feilleddene skal være konstant, dette kalles homoskedastisitet.
6. $\varepsilon_{(N,1)} \sim MVN_N(\mathbb{E}[\varepsilon]V[\varepsilon]) = MVN_N(0_{(N,1)}, \sigma^2 I_{(N,N)})$. Dette betyr at det forutsettes at feilleddene følger en normalfordeling.

Dersom forutsetningene er oppfylt vil regresjonsparameterne $\hat{\alpha}$ og $\hat{\beta}$ være BLUE (Best linear unbiased estimators). Akronymet BLUE, som står for best linear unbiased estimator, kan forklares på følgende måte. Med *best* menes det at OLS estimatoren $\hat{\beta}$ har den minste variansen sammenliknet med andre lineære objektive estimatorer. *Estimator* står for at parameterne $\hat{\alpha}$ og $\hat{\beta}$ er estimatorer på den ekte verdien av α og β . Med *linear* menes det at $\hat{\alpha}$ og $\hat{\beta}$ er lineære estimatorer. Hvis regresjonen gjennomføres mange ganger, vil verdiene av $\hat{\alpha}$

og $\hat{\beta}$ gjennomsnittlig være lik deres ekte verdier, dette er hva som menes med *unbiased* (Brooks, 2008).

5.1.4. Modellens forklaringskraft

Modellens forklaringskraft sier noe om hvor godt regresjonsligningen er tilpasset datamaterialet. Dette målet kan brukes til å sammenlikne modeller. I dette tilfellet har jeg tre alternative regresjonsmodeller, presentert i kapittel 5.1, hvor den som best forklarer datamaterialet, samt oppfyller forutsetningene i størst grad skal brukes til videre analyse.

Den totale variasjonen kan deles i forklart og uforklart variasjon.

$$R^2 = \frac{RSS}{TSS} \Leftrightarrow R^2 = 1 - \frac{ESS}{TSS} \quad (5.9)$$

Total variasjon er i funksjon 5.9 representert av TSS som står for *total sum of squares*.

Forholdet $\frac{RSS}{TSS}$ er andelen forklart variasjon, og RSS står for *regression sum of squares*. R^2 forteller stor andel av variasjonen i den avhengige variabelen som beskrives av de uavhengige variablene. Den siste komponenten i funksjonen, ESS (*error sum of squares*) er et mål på den uforklarte variasjonen i modellen (Gripsrud et al., 2017).

Verdien av R^2 er mellom 0 og 1, og desto høyere verdi, desto bedre er modellen tilpasset datamaterialet. Dersom $R^2 = 1$ betyr det at den estimerte verdien av den avhengige variabelen er lik den ekte verdien av den avhengige variabelen for alle observasjoner.

En høy R^2 betyr ikke nødvendigvis at det er en god modell. Jo flere uavhengige variabler som inkluderes i modellen, desto høyere vil R^2 bli. Det betyr at dersom en variabel som ikke er relevant inkluderes, vil likevel R^2 øke. Derfor bør man justere for frihetsgrader, $n-k-1$, hvor n er antall observasjoner og k er antall uavhengige variabler som er inkludert. R^2 som er justert for frihetsgrader er preferert ovenfor R^2 når målet skal brukes som et kriterium for å bestemme hvilken modell som skal brukes.

$$\overline{R^2} = 1 - \frac{\frac{ESS}{(n-k-1)}}{\frac{TSS}{(n-1)}} \quad (5.10)$$

Når det velges modell på bakgrunn av $\overline{R^2}$, velges den modellen som minimerer *mean squared error* (Jungeilges, 2018).

5.1.5. Hypotesetesting

I henhold til Hellevik (2011) kan en hypotese defineres som «en presis påstand om virkeligheten som enda ikke er bekreftet». Når sammenhenger skal undersøkes formuleres to hypoteser. En nullhypotese og en alternativ hypotese, hvor den alternative hypotesen (H_1) representerer teorien. Nullhypotesen (H_0) sier at det ikke er noen sammenheng mellom variablene som testes. Det forutsettes at nullhypotesen er sann, inntil det motsatte er bevist. Alternativhypotesen sier at det er en sammenheng mellom variablene eller som indikerer at det er forskjeller mellom grupper (Gripsrud et al., 2017; Sekaran & Bougie, 2016).

Ved regresjon i Stata beregnes en p-verdi, og det er det er denne jeg vil ta utgangspunkt i når jeg skal bestemme om nullhypotesene skal forkastes eller ikke. Det er vanlig å benytte signifikansnivå på 1%, 5% og 10%. Ved et forhåndsbestemt signifikansnivå på 5% forkastes nullhypotesen dersom $\alpha < 0,05$. Dersom nullhypotesen forkastes gis alternative hypotesen støtte.

I hypotesetesting er det viktig å være klar over at feil kan oppstå. Det skilles mellom type I-feil og type II-feil. Type I-feil er sannsynligheten for at en sann nullhypotese forkastes, og dette er den mest alvorlige feilen. Sannsynligheten for at type I-feil oppstår er det forhåndsbestemte signifikansnivået, det vil at dersom $\alpha=0,05$ er det 5% sannsynlighet for at en sann nullhypotese forkastes. Type II-feil er sannsynligheten for å beholde en usann nullhypotese.

F-test

F-testen brukes for å finne ut om det er en sammenheng mellom den avhengige variabelen og de uavhengige variablene som er inkludert i modellen, altså en test av hele regresjonsmodellen. Hypotesen som testes med denne testen er

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

$$H_1: \text{minst en } \beta_i \neq 0$$

Dersom p-verdien relatert til modellen, vist som *Prob > F* i Stata, er mindre enn 0,05 forkastes nullhypotesen med et signifikansnivå på 5%. Når nullhypotesen forkastes gis det støtte til at det er et forhold mellom den avhengige variabelen og minst en av de uavhengige variablene (Gripsrud et al., 2017).

T-test

T-testen brukes til å sikre seg mot at en sann nullhypotese forkastes, og dette gjøres ved å teste signifikansen til hver av regresjonskoeffisientene. Ved hjelp av T-testen testes det om korrelasjonskoeffisienten, β , er signifikant forskjellig fra null. Dette betyr, man tester om det er en sammenheng mellom den avhengige variabelen og hver av de uavhengige variablene. Hypotesen relatert til denne testen er:

$$H_0: \beta_i = 0$$

$$H_1: \beta_i \neq 0 \text{ (tosidig test) eller } H_1: \beta_i < (>) 0 \text{ (ensidig test)}$$

Dersom p-verdien som er relatert til den gitte β -koeffisient er mindre enn 0,05 (5%), forkastes nullhypotesen, og det gis støtte til at β -koeffisienten er signifikant ulik null. Dersom konklusjonen fra testen er at β -koeffisienten er lik null, er det en indikasjon på at den avhengige variabelen og den uavhengige variabelen er uavhengige fra hverandre (Gripsrud et al., 2017).

5.2. T-test til sammenlikning av to grupper

I den totale populasjonen er det 529 leiligheter som er solgt med gevinst og 128 som er solgt med tap. Disse to gruppene, leiligheter solgt med gevinst og leiligheter solgt med tap, skal undersøkes ytterligere og sammenliknes.

De to gruppene, leiligheter solgt med gevinst og leiligheter solgt med tap, skal sammenliknes for å undersøke om det finnes noen vesentlige forskjeller mellom de to gruppene. Til denne undersøkelsen velger jeg å bruke en «independent sample t-test». I denne testen brukes en kategorisk variabel til å dele utvalget i to grupper, og gjennomsnittsverdier fra de to gruppene sammenliknes. Testen kan være både ensidig og tosidig. Nullhypotesen er den samme i begge testene. Den ensidige testen brukes dersom man på forhånd har en klar formening om hvilken gjennomsnittsverdi som er høyest. I den ensidige testen er den alternative hypotesen at den ene gruppen har høyere eller lavere gjennomsnittsverdi enn den andre gruppen.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2 \text{ eller } H_1: \mu_1 < \mu_2$$

Den tosidige testen brukes når vi ikke har noen klar formening om hvilken gruppe som har den høyeste gjennomsnittsverdien på forhånd.

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Denne testen er en t-test som har samme handlingsregel som t-testen som er relatert til regresjonen. Det vil si at dersom testen returnerer en p-verdi mindre enn 0,05 forkastes nullhypotesen til et signifikansnivå på 5%.

6. Analyse ved hjelp av den hedonistiske metoden

I dette kapittelet skal jeg bruke regresjon for å undersøke hvordan de ulike attributtene påvirker leilighetsprisen i bydelen. I kapittel 5 er de tre ulike funksjonsformene som er utgangspunkt for utarbeiding av regresjonsfunksjoner presentert. I regresjonsfunksjonene brukes symbolene som er presentert i tabell 4.6.

De to funksjonsformene semilogaritmisk og dobbeltlogaritmisk skal brukes og det krever at både avhengige og uavhengige variabler transformeres ved hjelp av den naturlige logaritmen. I Stata transformeres de aktuelle variablene ved at det lages nye variabler som inneholder den naturlige logaritmen av de aktuelle variablene. Variablene $\ln P$, $\ln M$, $\ln N$, $\ln R$, $\ln E$ og $\ln Y$ lages. Som nevnt skal ikke dummyvariablene transformeres ved hjelp av den naturlige logaritmen.

6.1. Valg av forklaringsvariabler

Hvilke attributter som skal inkluderes i regresjonsmodellen er et sentralt spørsmål. I henhold til Gripsrud et al. (2017) er det viktig å bare inkludere relevante forklaringsvariabler i modellen. Men man kan ikke med sikkerhet på forhånd si hvilke variabler som er relevante, samt at det ikke er mulig å få tak i alle variabler man ønsker å inkludere. Osland (2001) har inkludert følgende attributter i sin analyse: boligareal, tomteareal, antall soverom, boligtype, boligens alder, utsikt, parkett i stue, peis, antall toaletter, hybel, kjeller, avstand til sentrum og dummy-variabler for omsetningsår.

I denne avhandlingen er variabelen totalpris er den avhengige variabelen i modellen. De uavhengige variablene som inkluderes i modellen er leilighetens alder, avstand til sentrum, etasje, om det er parkeringsplass, heis, eierform og om leiligheten ligger i blokk, samt dummyvariablene som sier noe om hvilket område leiligheten ligger i. I tillegg skal det inkluderes en variabel som sier noe om størrelsen på leiligheten, enten leilighetens størrelse eller antall soverom. Det skal også inkluderes en variabel som sier noe om hvilket år leiligheten er solgt.

Valg av variabler der det finnes alternative variabler

Med dataen samlet inn, finnes det flere alternative regresjonsmodeller og det må vurderes hvilke variabler som skal inkluderes i den endelige modellen. Primærrom og soverom er høyt korrelert, med en korrelasjonskoeffisient på 0,763, og bør derfor ikke inngå i samme regresjonsmodell.

I datasettet finnes både en variabel som angir salgsår og dummyvariabler for hvert salgsår. Her må det vurderes om det skal brukes dummy-variabler for hvert år eller om salgsår-variabelen skal brukes.

Med utgangspunkt i funksjonsformene presentert i kapittel 5 gjennomføres regresjon med de uavhengige variablene det ikke finnes alternativer til og hver av de alternative variablene. I tabellen nedenfor er forklaringskraften til de ulike modellene med de forskjellige alternativene presentert.

Variabler inkludert	Lineær	Semilogaritmisk	Dobbellogaritmisk
Primærrom og salgsår	0,8031	0,8311	0,8609
Primærrom og dummyvariabler for salgsår	0,8071	0,8315	0,8618
Soverom og salgsår	0,6503	0,6496	0,6948
Soverom og dummyvariabler for salgsår	0,6507	0,7239	0,6953

Tabell 6.1 Modellens forklaringskraft ved ulike alternative variabler inkludert, n=1872

Fra tabell 6.1 er det tydelig at inkludering variabelen primærrom gir modellen en betydelig høyere forklaringskraft enn det variabelen antall soverom gjør. Forklaringskraften i alle funksjonsformer er også litt høyere ved bruk av dummy-variabler for salgsårene istedenfor variabelen salgsår. Andre analyser gjort ved bruk av den hedonistiske metoden, blant annet Osland (2001) bruker dummyvariabler for salgsår, derfor er det også det jeg vil bruke i den endelige modellen. Det er også mest hensiktsmessig å bruke salgsår-dummyer, da det gir muligheten til å undersøke hvordan leilighetsprisen er påvirket av salgsåret.

Behandling av verdier som skiller seg betydelig fra resten av verdiene

I kapittel 4 ble klart at flere av variablene hadde noen verdier som skiller seg betydelig fra resten av verdiene. Disse verdiene som enten er betydelig lavere eller betydelig høyere enn majoriteten av verdiene vil påvirke resultatet. Derfor prøver jeg å utelate de mest ekstreme tilfellene av verdier som skiller seg ut for å se om dette øker modellens forklaringskraft og

oppfyller forutsetningene bedre. Det ser ikke ut til at modellen blir bedre av å utelate de mest ekstreme avvikene. Derfor vil regresjon med hele datasettet brukes som utgangspunkt.

Dummyvariabler

Dummyvariablene for salgsår og områder inneholder mer enn to alternativer. Dersom alle dummyvariablene innenfor en kategori inkluderes får vi en situasjon med multikollinearitet. Løsningen på dette er å droppe en av dummyvariablene. Dummyvariabel for det første salgsåret som er inkludert i analysen, 2010, utelates. Dette er fordi det er året som er tidligst i undersøkelsesperioden og det antas at totalprisen på leiligheter var lavest i 2010.

I tillegg må et av områdene utelates. Utgangspunktet mitt var å utelate område nummer 1, Gimlevang/Sødal. Det viste seg at vif-verdien som indikerer multikollinearitet var ganske høy for variabelen L10, som svarer til området Solbygg, i denne modellen. Solbygg er det området med flest observasjoner og Solbygg er også det området med lavest gjennomsnittlig totalpris. Dersom Solbygg utelates og brukes som referansekategori er alle vif-verdiene lave og dette er årsaken til at området Solbygg brukes som referansekategori.

Standardleiligheten er da en borettslagsleilighet i småhus med en alder på null år, som er omsatt i 2010. Leiligheten er lokalisert i området Solbygg og har ikke heis eller parkeringsplass.

6.2. Estimering av regresjonsmodellene

På bakgrunn av vurderingen gjort ovenfor utarbeides de endelige regresjonsmodellene som skal brukes til analysen av boligmarkedet i bydel Lund i Kristiansand. I regresjonsresultatet presentert i dette kapitlet består variabelen totalpris av omsetningspris og 87% av fellesgjelden brukt som avhengig variabel¹.

6.2.1. Lineær funksjonsform

$$P = \alpha + \beta_1 M + \beta_2 Y + \beta_3 km + \beta_4 E + \beta_5 H + \beta_6 G + \beta_7 SD + \beta_8 BD + \beta_9 \text{år}2011 + \dots + \beta_{15} \text{år}2018 + \beta_{16} L1 + \dots + \beta_{28} L14 + \varepsilon \quad (6.1)$$

¹ Jeg har også prøvd med en avhengig variabel der 100% av fellesgjelden er inkludert, men denne gav lavere forklaringskraft. Dermed legges en fellesgjeldsgrad på 0,87 til grunn, i henhold til resultatene til Eretveit, S. & Theisen, T. (2016). Efficiency and justice in the market for cooperative dwellings. *International Real Estate Review*, 19(3), 297-326. .

Source	SS	df	MS	Number of obs	=	1,871
Model	2.2734e+15	29	7.8392e+13	F(29, 1841)	=	265.56
Residual	5.4345e+14	1,841	2.9519e+11	Prob > F	=	0.0000
				R-squared	=	0.8071
				Adj R-squared	=	0.8040
Total	2.8168e+15	1,870	1.5063e+12	Root MSE	=	5.4e+05

P	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
M	33913.75	588.1937	57.66	0.000	32760.15 35067.34
Y1	-9775.869	816.4961	-11.97	0.000	-11377.22 -8174.513
km	-124380	25825.71	-4.82	0.000	-175030.8 -73729.22
E	41821.88	8880.937	4.71	0.000	24404.11 59239.64
G	-149654.8	43950.11	-3.41	0.001	-235852.1 -63457.54
H	251787.5	54377.75	4.63	0.000	145139 358436.1
BD	133467.4	46509.03	2.87	0.004	42251.39 224683.4
SD	154406	38725.82	3.99	0.000	78454.84 230357.1
år2011	36415.11	56264.06	0.65	0.518	-73932.96 146763.2
år2012	56852.91	54916.43	1.04	0.301	-50852.13 164557.9
år2013	85737.72	55997.04	1.53	0.126	-24086.66 195562.1
år2014	73785.18	55878.51	1.32	0.187	-35806.75 183377.1
år2015	182812.5	56372.03	3.24	0.001	72252.67 293372.3
år2016	345304.5	53559.97	6.45	0.000	240259.8 450349.1
år2017	318969	55148.3	5.78	0.000	210809.2 427128.8
år2018	442741	53776.22	8.23	0.000	337272.2 548209.8
L1	-417351.7	76303.83	-5.47	0.000	-567002.9 -267700.6
L2	57783.36	98533.87	0.59	0.558	-135466.5 251033.3
L3	670911.6	70117.13	9.57	0.000	533394.1 808429
L4	13027.54	53910.42	0.24	0.809	-92704.47 118759.5
L5	133364.6	98150.85	1.36	0.174	-59134.1 325863.3
L6	295204.7	120184.9	2.46	0.014	59491.61 530917.7
L7	91018.63	56585.91	1.61	0.108	-19960.68 201997.9
L8	98492.14	65711.97	1.50	0.134	-30385.69 227370
L9	-377088.8	66065.22	-5.71	0.000	-506659.4 -247518.1
L11	928907.4	66438.78	13.98	0.000	798604.1 1059211
L12	73052.03	70420.47	1.04	0.300	-65060.34 211164.4
L13	548492.5	85946.56	6.38	0.000	379929.5 717055.5
L14	-275734.7	103548.2	-2.66	0.008	-478819 -72650.38
_cons	447684.3	106437.2	4.21	0.000	238933.9 656434.6

Tabell 6.2 Lineær multippel regresjon

Modellen har høy forklaringskraft da de uavhengige variablene i modellen forklarer 80,40% av variasjonen i totalprisen. Men for å være sikker på at den leverer troverdige resultater må jeg undersøke om den oppfyller de bakenforliggende forutsetningene for lineær regresjon.

Variance inflation factor (vif) er en indikator på multikollinearitet. En vif-verdi på mer enn 10 indikerer at to eller flere av de uavhengige variablene i modellen er høyt korrelert. Den høyeste vif-verdien som returneres for den lineære modellen er 3,93, og gjennomsnittlig vif-verdi er 1,94. Dette resultatet indikerer at det ikke er noe problem med multikollinearitet.

En annen viktig forutsetning er at feilleddene har konstant varians, denne forutsetningen kan undersøkes ved hjelp av en Breusch-Pagan test. Hypotesen som testes i Breusch-Pagan testen er:

H₀: Konstant varians i feilleddene (homoskedastisitet)

H₁: Feilleddene har ikke konstant varians (heteroskedastisitet)

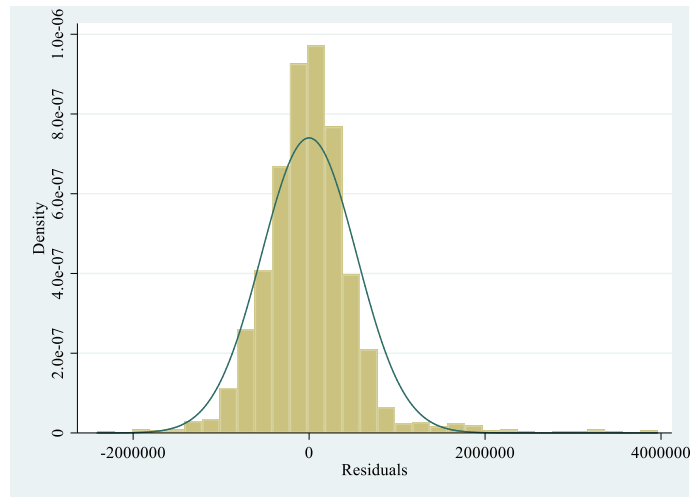
Handlingsregelen er at nullhypotesen forkastes dersom p-verdien testen returnerer er mindre enn et bestemt signifikansnivå. Dersom nullhypotesen forkastes indikerer det at feilleddene ikke har konstant varians. Testen returnerer en p-verdi på 0,000, det vil si at nullhypotesen forkastes både til et signifikansnivå på 1%. Det betyr at det er 99% sikkert at den rette konklusjonen er å forkaste nullhypotesen. Konsekvensen av å at feilleddene ikke har konstant varians vil føre til at estimatorene fra regresjonen ikke er best linear unbiased estimators. I henhold til Brooks (2008) finnes det to løsninger på heteroskedastisitet: 1) bruke logaritmetransformasjon og 2) bruke «heteroscedasticity-consistent standard error estimates», dette gjøres ved en robust regresjon. I dette tilfellet velger jeg å bruke alternativ 2 og estimerer en robust regresjonsmodell.

Linear regression		Number of obs = 1,871				
		F(29, 1841) = 106.67				
		Prob > F = 0.0000				
		R-squared = 0.8071				
		Root MSE = 5.4e+05				
P	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
M	33913.75	1199.618	28.27	0.000	31560.99	36266.5
Y1	-9775.869	1173.261	-8.33	0.000	-12076.93	-7474.807
km	-124380	22821.43	-5.45	0.000	-169138.6	-79621.38
E	41821.88	9527.379	4.39	0.000	23136.27	60507.48
G	-149654.8	48829.14	-3.06	0.002	-245421.1	-53888.52
H	251787.5	51127.24	4.92	0.000	151514	352061
BD	133467.4	45251.01	2.95	0.003	44718.69	222216.1
SD	154406	31315.39	4.93	0.000	92988.58	215823.4
år2011	36415.11	50839.08	0.72	0.474	-63293.21	136123.4
år2012	56852.91	48805.87	1.16	0.244	-38867.77	152573.6
år2013	85737.72	49950.5	1.72	0.086	-12227.88	183703.3
år2014	73785.18	52732.33	1.40	0.162	-29636.28	177206.6
år2015	182812.5	54457.42	3.36	0.001	76007.72	289617.3
år2016	345304.5	56979.58	6.06	0.000	233553.1	457055.9
år2017	318969	49877.82	6.40	0.000	221146	416792.1
år2018	442741	55067.45	8.04	0.000	334739.8	550742.2
L1	-417351.7	80264.3	-5.20	0.000	-574770.4	-259933.1
L2	57783.36	112173.7	0.52	0.607	-162217.7	277784.4
L3	670911.6	96919.49	6.92	0.000	480827.9	860995.2
L4	13027.54	40846.05	0.32	0.750	-67081.92	93136.99
L5	133364.6	104115.9	1.28	0.200	-70833.03	337562.2
L6	295204.7	90221.05	3.27	0.001	118258.3	472151
L7	91018.63	43990.31	2.07	0.039	4742.489	177294.8
L8	98492.14	59038.91	1.67	0.095	-17298.12	214282.4
L9	-377088.8	61538.44	-6.13	0.000	-497781.2	-256396.3
L11	928907.4	84298.67	11.02	0.000	763576.3	1094238
L12	73052.03	55039.63	1.33	0.185	-34894.64	180998.7
L13	548492.5	141891.1	3.87	0.000	270208.2	826776.8
L14	-275734.7	68380.65	-4.03	0.000	-409846.5	-141622.9
_cons	447684.3	126461.8	3.54	0.000	199660.7	695707.9

Tabell 6.3 Multipl linear regresjon med robuste standardfeil

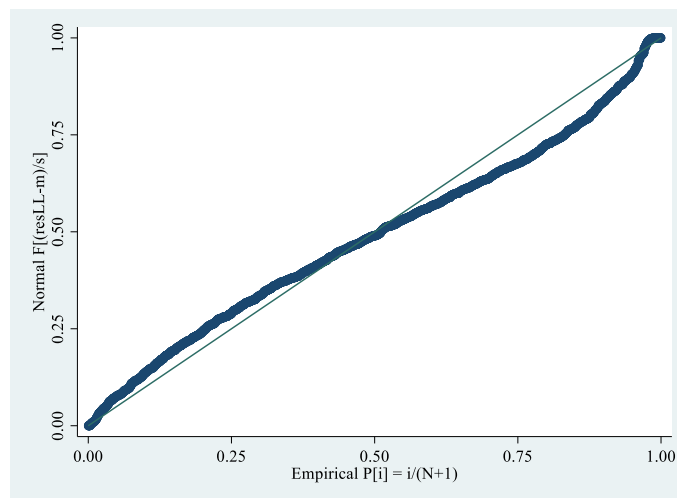
Konstantleddet i modellen angir prisen på standardleiligheten, som er definert som en borettslagsleilighet med alder på null år, i småhus lokalisert i området Solbygg omsatt i 2010, som ikke har heis og parkeringsplass. Prisen på standardleiligheten er estimert til å være kr 447 684,30.

En siste forutsetning er normalfordelte feilledd. I figur 6.1 er et histogram over feilleddene presentert. Histogrammet er relativt symmetrisk og det ser ut til at feilleddene følger en tilnærmet normalfordeling.



Figur 6.1 Histogram for feilledd

Figur 6.2 viser et normalfordelingsplott for feilledd i den lineære funksjonsformen, denne typen plott viser tydeligere om feilleddene er normalfordelt. I figuren viser den tynne diagonale kurven situasjonen med perfekt normalfordeling, det vil si at ved perfekt normalfordeling skal den tykke blå kurven være lineær. Figuren indikerer at forutsetning om normalfordeling i feilleddene ikke er tilstrekkelig oppfylt.



Figur 6.2 Normalfordelingsplott for feilledd i lineær funksjonsform

6.2.2. Semilogaritmisk funksjonsform

Før den semilogaritmiske modellen kan brukes i regresjon må variabelen $\ln P$ genereres. Dette er den naturlige logaritmen til variabelen totalpris. Deretter estimeres regresjonsmodell 6.2.

$$\ln P = \alpha + \beta_1 M + \beta_2 Y + \beta_3 km + \beta_4 E + \beta_5 H + \beta_6 G + \beta_7 SD + \beta_8 BD + \beta_9 \text{år}2011 + \dots + \beta_{15} \text{år}2018 + \beta_{16} L1 + \dots + \beta_{28} L14 \varepsilon \quad (6.2)$$

Source	SS	df	MS	Number of obs	=	1,871
Model	231.639713	29	7.98757631	F(29, 1841)	=	319.17
Residual	46.072922	1,841	.02502603	Prob > F	=	0.0000
				R-squared	=	0.8341
				Adj R-squared	=	0.8315
Total	277.712635	1,870	.14850943	Root MSE	=	.1582

lnP	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
M	.0111417	.0001713	65.06	0.000	.0108058 .0114776
Y1	-.002457	.0002377	-10.33	0.000	-.0029233 -.0019907
km	-.0383923	.0075196	-5.11	0.000	-.0531402 -.0236444
E	-.0087406	.0025858	-3.38	0.001	-.0138121 -.0036691
G	-.0166243	.0127969	-1.30	0.194	-.0417222 .0084736
H	.0842896	.0158331	5.32	0.000	.0532369 .1153422
BD	.027261	.0135419	2.01	0.044	.0007019 .0538202
SD	.0437757	.0112757	3.88	0.000	.0216611 .0658902
år2011	.0445163	.0163823	2.72	0.007	.0123865 .0766461
år2012	.0452517	.0159899	2.83	0.005	.0138915 .076612
år2013	.0507542	.0163045	3.11	0.002	.0187768 .0827315
år2014	.0557004	.01627	3.42	0.001	.0237907 .08761
år2015	.0978384	.0164137	5.96	0.000	.065647 .1300299
år2016	.1228031	.0155949	7.87	0.000	.0922175 .1533887
år2017	.1483508	.0160574	9.24	0.000	.1168581 .1798435
år2018	.1718371	.0156579	10.97	0.000	.1411279 .2025462
L1	-.0886567	.0222172	-3.99	0.000	-.1322304 -.0450831
L2	.0062575	.0286899	0.22	0.827	-.0500106 .0625257
L3	.165285	.0204159	8.10	0.000	.1252443 .2053257
L4	.0585748	.015697	3.73	0.000	.027789 .0893605
L5	.1193016	.0285784	4.17	0.000	.0632522 .1753511
L6	.0930153	.034994	2.66	0.008	.0243832 .1616474
L7	.0473047	.016476	2.87	0.004	.0149911 .0796183
L8	.046567	.0191332	2.43	0.015	.0090419 .0840921
L9	-.0292356	.0192361	-1.52	0.129	-.0669624 .0084912
L11	.2734735	.0193448	14.14	0.000	.2355333 .3114136
L12	.0749834	.0205042	3.66	0.000	.0347696 .1151973
L13	.100644	.0250249	4.02	0.000	.0515639 .1497241
L14	-.0182568	.0301499	-0.61	0.545	-.0773885 .0408748
_cons	13.96036	.0309911	450.46	0.000	13.89958 14.02115

Tabell 6.4 Regresjon med semilogaritmisk funksjonsform

Den semilogaritmiske modellen har høyere forklaringskraft enn den lineære funksjonsformen gav, og her er 83,15% av variasjonen i totalprisen forklart av de uavhengige variablene som er inkludert i modellen.

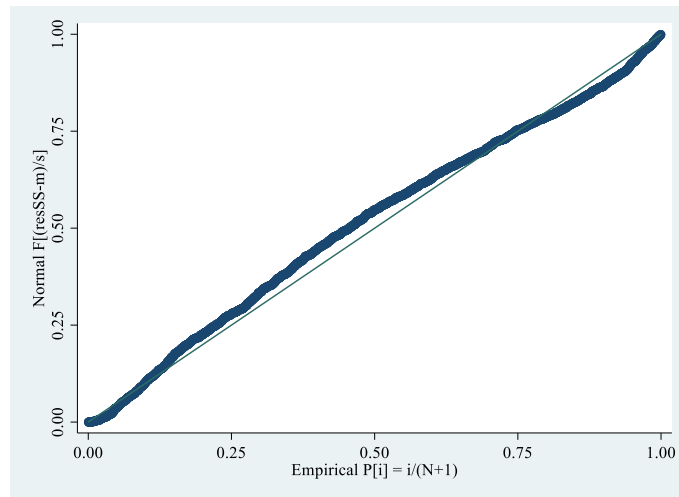
De uavhengige variablene i modellen med lineær funksjonsform og modellen med den semilogaritmiske funksjonsformen er de samme. Dermed er det heller ikke i denne modellen noe problem med multikollinearitet i modellen presentert i tabell 6.4. Breusch-Pagan returnerer også her en p-verdi på 0,000, derfor gjennomføres robust regresjon.

Linear regression		Number of obs = 1,871				
		F(29, 1841) = 202.15				
		Prob > F = 0.0000				
		R-squared = 0.8341				
		Root MSE = .1582				
lnP	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
M	.0111417	.0002632	42.33	0.000	.0106255	.011658
Y1	-.002457	.0002657	-9.25	0.000	-.0029781	-.0019359
km	-.0383923	.0068255	-5.62	0.000	-.0517788	-.0250058
E	-.0087406	.0028667	-3.05	0.002	-.0143629	-.0031184
G	-.0166243	.0134179	-1.24	0.216	-.0429403	.0096917
H	.0842896	.0159331	5.29	0.000	.0530408	.1155384
BD	.027261	.0128384	2.12	0.034	.0020817	.0524404
SD	.0437757	.0104654	4.18	0.000	.0232503	.0643011
år2011	.0445163	.015726	2.83	0.005	.0136736	.075359
år2012	.0452517	.0150129	3.01	0.003	.0158076	.0746959
år2013	.0507542	.0170814	2.97	0.003	.0172533	.0842551
år2014	.0557004	.0151997	3.66	0.000	.0258899	.0855109
år2015	.0978384	.0153401	6.38	0.000	.0677527	.1279242
år2016	.1228031	.0149796	8.20	0.000	.0934244	.1521818
år2017	.1483508	.0143556	10.33	0.000	.1201957	.1765059
år2018	.1718371	.0163787	10.49	0.000	.1397143	.2039598
L1	-.0886567	.0243032	-3.65	0.000	-.1363214	-.0409921
L2	.0062575	.0351633	0.18	0.859	-.0627067	.0752217
L3	.165285	.0228258	7.24	0.000	.1205178	.2100522
L4	.0585748	.0133345	4.39	0.000	.0324224	.0847272
L5	.1193016	.0250511	4.76	0.000	.07017	.1684332
L6	.0930153	.0341185	2.73	0.006	.0261003	.1599303
L7	.0473047	.0148831	3.18	0.002	.0181152	.0764942
L8	.046567	.018082	2.58	0.010	.0111036	.0820304
L9	-.0292356	.0175395	-1.67	0.096	-.0636349	.0051638
L11	.2734735	.0195986	13.95	0.000	.2350356	.3119113
L12	.0749834	.0172103	4.36	0.000	.0412297	.1087372
L13	.100644	.0272668	3.69	0.000	.0471669	.1541211
L14	-.0182568	.0236729	-0.77	0.441	-.0646854	.0281717
_cons	13.96036	.0342047	408.14	0.000	13.89328	14.02745

Tabell 6.5 Robust regresjon med semilogaritmisk funksjonsform

I motsetning til den lineære funksjonsformen, hvor bare dummyvariablene for årene 2015-2018 var signifikante, er alle salgsårsdummyene signifikante i denne modellen.

Figur 6.3 viser normalfordelingsplottet til feilleddene i modellen basert på den semilogaritmiske funksjonsformen. Det ser ut som om feilleddenes tilnærming til normalfordeling er bedre i denne modellen enn i modellen med lineære parametere.



Figur 6.3 Normalfordelingsplott for feilledd i semilogaritmisk funksjonsform

6.2.3. Dobbellogaritmisk funksjonsform med totalpris som avhengig variabel

Den dobbellogaritmiske funksjonen er presentert i funksjon 6.3. I denne modellen er også de uavhengige variablene logaritmiske. Den naturlige logaritmen av 0 er udefinert, og den naturlige logaritmen til 1 er 0. Derfor er standardleiligheten i denne regresjonsmodellen enten null eller et år gammel.

$$\ln P = \alpha + \beta_1 \ln M + \beta_2 \ln Y + \beta_3 \ln km + \beta_4 \ln E + \beta_5 H + \beta_6 G + \beta_7 SD + \beta_8 B D_s + \beta_9 \text{år}2011 + \dots + \beta_{15} \text{år}2018 + \beta_{16} L1 + \dots + \beta_{28} L14 + \varepsilon \quad (6.3)$$

Source	SS	df	MS	Number of obs = 1,831		
				F(29, 1801) = 394.65		
Model	233.466416	29	8.05056606	Prob > F = 0.0000		
Residual	36.7389506	1,801	.020399195	R-squared = 0.8640		
				Adj R-squared = 0.8618		
Total	270.205366	1,830	.147653206	Root MSE = .14283		

lnP	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnM	.7606096	.010188	74.66	0.000	.7406281	.7805912
lnY1	-.0656473	.0050199	-13.08	0.000	-.0754928	-.0558019
lnkm	-.079708	.0207225	-3.85	0.000	-.1203508	-.0390653
lnE	.0354206	.0064705	5.47	0.000	.0227302	.048111
G	-.0109829	.0116098	-0.95	0.344	-.033753	.0117871
H	.0533727	.0143742	3.71	0.000	.0251808	.0815646
BD	.0062497	.0121437	0.51	0.607	-.0175676	.0300669
SD	.0527817	.0102131	5.17	0.000	.032751	.0728124
år2011	.0383719	.0148521	2.58	0.010	.0092428	.067501
år2012	.050772	.01451	3.50	0.000	.0223138	.0792302
år2013	.0544177	.014821	3.67	0.000	.0253496	.0834857
år2014	.0419722	.0148759	2.82	0.005	.0127963	.0711482
år2015	.0962814	.0148074	6.50	0.000	.06724	.1253228
år2016	.1407462	.0140939	9.99	0.000	.1131041	.1683883
år2017	.1460009	.0145174	10.06	0.000	.1175283	.1744736
år2018	.1717101	.0141681	12.12	0.000	.1439224	.1994977
L1	-.1052874	.0207099	-5.08	0.000	-.1459054	-.0646694
L2	-.0045662	.0258871	-0.18	0.860	-.055338	.0462056
L3	.1512	.0186954	8.09	0.000	.114533	.1878669
L4	.0128601	.0145217	0.89	0.376	-.0156211	.0413413
L5	.0538938	.029912	1.80	0.072	-.0047721	.1125597
L6	.0319872	.0326622	0.98	0.328	-.0320726	.096047
L7	.0156283	.015305	1.02	0.307	-.0143892	.0456457
L8	.0212099	.0174451	1.22	0.224	-.0130049	.0554248
L9	-.074565	.0175054	-4.26	0.000	-.108898	-.0402321
L11	.2781499	.0177295	15.69	0.000	.2433774	.3129224
L12	.0608906	.0187527	3.25	0.001	.0241112	.09767
L13	.1314802	.0227722	5.77	0.000	.0868176	.1761429
L14	-.0421396	.0276348	-1.52	0.127	-.0963392	.0120601
_cons	11.6298	.0501884	231.72	0.000	11.53136	11.72823

Tabell 6.6 Regresjon med dobbellogaritmisk funksjonsform

Denne modellen har den høyeste forklaringskraften av de tre estimerte modellene, og 86,18% av variasjonen i totalprisen forklares av modellen.

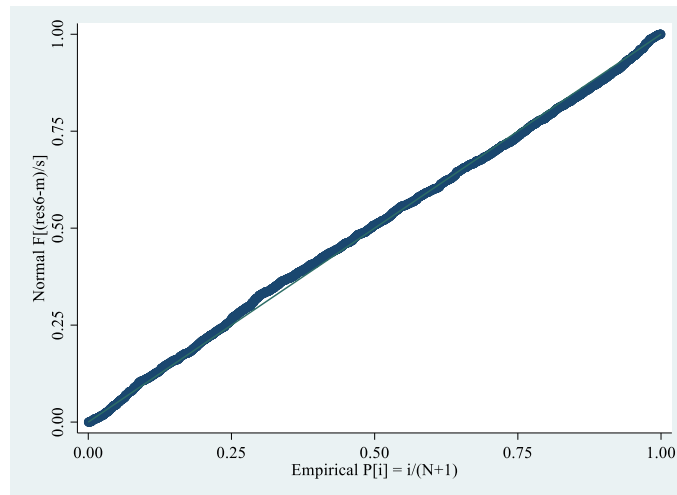
Den høyeste vif-verdien som returneres fra vif-testen er på 3,88, og gjennomsnittlig vif-verdi er 1,95. Det indikerer at det ikke er for høy korrelasjon mellom de uavhengige variablene i modellen.

For at estimatene modellen gir skal være troverdige er en forutsetning at feilleddene har konstant varians. Breusch-Pagan gir en p-verdi på 0,000 som betyr at nullhypotesen om homoskedastisitet forkastes. Som for de to foregående modellene, gjennomføres robust regresjon.

Linear regression		Number of obs		=		1,831	
		F(29, 1801)		=		280.33	
		Prob > F		=		0.0000	
		R-squared		=		0.8640	
		Root MSE		=		.14283	
lnP	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]		
lnM	.7606096	.0120639	63.05	0.000	.7369489	.7842703	
lnY1	-.0656473	.0064135	-10.24	0.000	-.078226	-.0530687	
lnkm	-.079708	.0214146	-3.72	0.000	-.121708	-.037708	
lnE	.0354206	.0065713	5.39	0.000	.0225324	.0483088	
G	-.0109829	.0123797	-0.89	0.375	-.0352631	.0132972	
H	.0533727	.0151579	3.52	0.000	.0236438	.0831016	
BD	.0062497	.0129071	0.48	0.628	-.0190648	.0315642	
SD	.0527817	.009886	5.34	0.000	.0333925	.0721709	
år2011	.0383719	.0150698	2.55	0.011	.0088158	.0679281	
år2012	.050772	.0143323	3.54	0.000	.0226623	.0788817	
år2013	.0544177	.0150276	3.62	0.000	.0249444	.083891	
år2014	.0419722	.0143771	2.92	0.004	.0137747	.0701698	
år2015	.0962814	.0148801	6.47	0.000	.0670974	.1254654	
år2016	.1407462	.0144063	9.77	0.000	.1124914	.169001	
år2017	.1460009	.0142577	10.24	0.000	.1180375	.1739643	
år2018	.1717101	.015183	11.31	0.000	.1419319	.2014883	
L1	-.1052874	.0219693	-4.79	0.000	-.1483755	-.0621993	
L2	-.0045662	.0290663	-0.16	0.875	-.0615734	.052441	
L3	.1512	.0245399	6.16	0.000	.1030703	.1993296	
L4	.0128601	.0120781	1.06	0.287	-.0108284	.0365486	
L5	.0538938	.029956	1.80	0.072	-.0048583	.1126459	
L6	.0319872	.0277587	1.15	0.249	-.0224555	.0864299	
L7	.0156283	.0143851	1.09	0.277	-.012585	.0438415	
L8	.0212099	.0174186	1.22	0.224	-.0129528	.0553727	
L9	-.074565	.0179956	-4.14	0.000	-.1098596	-.0392705	
L11	.2781499	.0192491	14.45	0.000	.2403969	.3159029	
L12	.0608906	.016172	3.77	0.000	.0291727	.0926084	
L13	.1314802	.0265535	4.95	0.000	.0794014	.183559	
L14	-.0421396	.0214169	-1.97	0.049	-.0841441	-.0001351	
_cons	11.6298	.0594202	195.72	0.000	11.51326	11.74634	

Tabell 6.7 Regresjon med robuste standardfeil, dobbellogaritmisk funksjonsform

Normalfordelingsplottet for feilleddene i den dobbellogaritmiske modellen er presentert i figur 6.4 viser at feilleddene har en veldig god tilnærming til normalfordelingen.



Figur 6.4 Normalfordelingsplott for feilledd i dobbellogaritmisk modell

6.3. Valg av regresjonsmodell

Alle de tre funksjonsformene gir modeller med høy forklaringskraft. Det er den dobbellogaritmiske funksjonsformen som gir den høyeste forklaringskraften, men de to andre modellene gir også en tilfredsstillende forklaringskraft.

Basert på forutsetningene som ligger bak regresjon er det som skiller de tre modellene feilleddets tilnærming til normalfordeling. Det er den dobbellogaritmiske modellen som oppfyller denne forutsetningen best. Dermed er det den dobbellogaritmiske modellen som best oppfyller de bakenforliggende forutsetninger for regresjon.

I alle de seks regresjonene utført i kapittel 6.2 er $Prob > F = 0,000$. Det vil si at nullhypotesen i F-testen, forklart i kapittel 5.1.5, forkastes, og det gis støtte til at minst en av koeffisientene er ulik fra null og at det er en sammenheng mellom avhengig og minst en uavhengig variabel.

Regresjonsanalysen gir også et mål på standardavviket til feilleddene, denne er uttrykt som *root MSE*. RMSE sier da noe om spredningen til feilleddene. Denne verdien er lavest i den dobbellogaritmiske modellen.

Basert på oppfyllelse av forutsetninger, forklaringsgrad, samt at det er hensiktsmessig å bruke modellen som gir prosentvis endring i totalpris ved en 1% økning i de uavhengige variablene brukes den dobbellogaritmiske modellen til hypotesetesting. β_i representerer i denne modellen den prosentvise endringen i totalpris ved en økning på 1% i mengden attributt.

6.4. Hypotesetesting

Regresjonsresultatene fra tabell 6.7 skal brukes til å vurdere hypotesene som omhandler hvilke attributter som påvirker leilighetsprisen, presentert i kapittel 3. Til å besvare hypotesene gjennomføres t-testen som beskrevet i under hypotesetesting i metodekapittelet.

Forskningsspørsmål 1: Har antall m² primærrom betydning for leilighetens totalpris?

H₀: Det er ingen korrelasjon mellom størrelse på leilighet og leilighetspris

H₁: Det er en positiv sammenheng mellom leilighetens størrelse og leilighetspris

Følgende testbare hypotese utarbeides:

$$H_0: \beta_M = 0$$

$$H_1: \beta_M > 0$$

Nullhypotesen om at koeffisienten til leilighetens størrelse forkastes til et signifikansnivå på 1%, da p-verdien testen returner er 0,000. Det betyr at sannsynligheten for at feil nullhypotese forkastes er 1%.

Når nullhypotesen forkastes gis det støtte til at alternativhypotesen som sier at leilighetens størrelse har en betydning for leilighetens totalpris. Ifølge denne modellen gir en økning på 1% i leilighetens størrelse en økning på 0,76% i totalprisen.

Forskningsspørsmål 2: Har leilighetens alder betydning for totalpris?

H₀: Det er ingen korrelasjon mellom leilighetens alder og leilighetspris

H₁: Det er en negativ sammenheng mellom leilighetens alder og leilighetspris

Dette kan uttrykkes som

$$H_0: \beta_Y = 0$$

$$H_1: \beta_Y < 0$$

P-verdien relatert til boligalder er 0,000, og dermed forkastes nullhypotesen med et signifikansnivå på 1%. Alternativhypotesen som sier at det er en negativ sammenheng mellom boligalderen og leilighetsprisen gis støtte. Betaverdien til boligalder er -0,0656 som vil si at en 1% økning i boligalder fører til en reduksjon i leilighetspris på 0,0656%.

Forskningsspørsmål 3: Har avstand til sentrum betydning for totalprisen?

H₀: Det er ingen sammenheng mellom avstand til sentrum og leilighetspris

H₁: Det er en negativ sammenheng mellom avstand til sentrum og leilighetspris

Dette kan uttrykkes som

$$H_0: \beta_{km} = 0$$

$$H_1: \beta_{km} < 0$$

Tabell 6.7 viser at koeffisienten relatert til avstand fra leilighet til sentrum har en p-verdi på 0,000, det betyr at denne nullhypotesen forkastes med et signifikansnivå på 1%. Dermed gis det støtte til at dersom avstanden til sentrum øker med 1%, reduseres leilighetsprisen med 0,0797%. At leiligheter nærmest sentrum har høyere pris er i tråd med økonomisk teori og Alonso-Muth-Mills modellen.

Forskningsspørsmål 4: Har de leilighetene som ligger første etasje en lavere totalpris enn leiligheter i øvrige etasjer?

H₀: Det er ingen korrelasjon mellom avstand til leilighetens etasjeplassering og totalpris

H₁: Det er en positiv korrelasjon mellom leilighetens etasjeplassering og totalprisen

Dette kan uttrykkes som

$$H_0: \beta_E = 0$$

$$H_1: \beta_E > 0$$

P-verdien relatert til koeffisienten til leilighetens etasjeplassering er av Stata beregnet til å være 0,000. Det vil si at vi beholder den alternative hypotesen som sier at leilighetsprisen øker desto høyere etasje leiligheten er lokalisert i. En økning på 1% i etasje fører til at leilighetens totalpris øker med 0,0354%.

Forskningsspørsmål 5: Har parkeringsplass en positiv effekt på leilighetens totalpris?

H₀: Det er ingen sammenheng mellom leilighetsprisen og om leiligheten har parkeringsplass

H₁: Det er en positiv sammenheng mellom leilighetsprisen og om leiligheten har parkeringsplass

Dette kan uttrykkes som

$$H_0: \beta_G = 0$$

$$H_1: \beta_G > 0$$

Regresjonen returnerer en p-verdi på 0,295 for koeffisienten parkeringsplass. Denne p-verdien er høyere enn et signifikansnivå på 5% og nullhypotesen kan dermed ikke forkastes. Dette indikerer at i denne modellen har ikke parkeringsplass betydning for leilighetens totalpris.

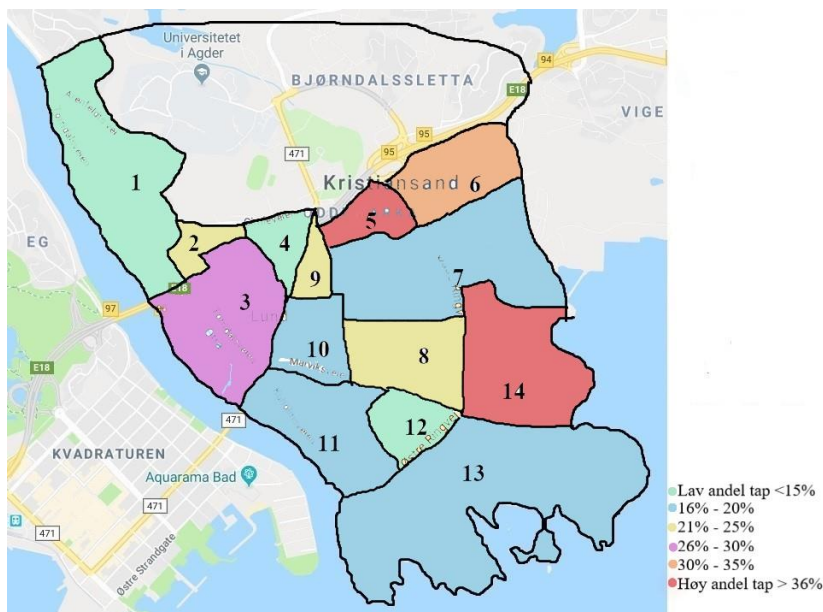
7. Analyse av leiligheter som er solgt mer enn en gang

I dette kapitlet vil jeg se nærmere på de leilighetene som er solgt mer enn en gang. For å undersøke forskjeller mellom de leilighetene som selges med gevinst og de leilighetene som selges med tap gjennomføres en interaksjonsmodell, samt t-tester for å sammenlikne gjennomsnittsverdiene til de to gruppene.

I mitt datasett som dekker tidsperioden 2010 – 2018 viste det seg at nesten 20% av de leilighetene som er solgt mer enn en gang i perioden er solgt med tap. I dette kapitlet presenteres først noen tabeller som beskriver hvordan gevinst og tap ved leilighets salg fordeler seg i bydelen i perioden 2010-2018. Deretter vil jeg sammenlikne leiligheter solgt med gevinst med leiligheter solgt med tap ved hjelp av t-tester.

Figur 7.1 viser hvor stor andel av leilighetene i de ulike områdene som er solgt med gevinst og tap. Det er varierende hvor mange leiligheter som er solgt igjen i området. Solbygg og Valhalla Nord er de områdene hvor flest leiligheter er omsatt mer enn en gang, og de største andelen tap oppgitt i prosent av totalt antall tap finner vi i disse to områdene. Derfor er det i dette tilfellet mest hensiktsmessig å oppgi antall gevinst og tap i prosent av totale salg per område. Det selges leiligheter med tap i alle områdene, men andelen av totale salg som er solgt med tap varierer mellom områdene.

Områdene Gimlevang, Oddemarka og Valhalla sør er de områdene hvor lavest andel av de solgte leilighetene i det respektive område er solgt med tap. I bydelene Tobienborg, Lund industriområde og Nedre Kongsgård har størst andel av de solgte leilighetene blitt solgt med tap.



Figur 7.1 Andel leiligheter solgt med tap per område av totale salg per område

Det er verdt å merke seg at områdene Oddemarka (4) og Tobienborg (5) ligger ganske nær hverandre og ser ut til å ha de samme fasilitetene. Samtidig selges en mye høyere andel av de leilighetene som selges i området Tobienborg med tap. Området Tobienborg består i hovedsak av nyere leiligheter, og gjennomsnittlig leilighetsalder i dette området er 6,53 år, mens gjennomsnittlig leilighetsalder i området Oddemarka er 42,21 år.

På grunn av at undersøkelsesperioden kun strekker seg fra 2010 til 2018 er de leilighetene som er solgt i 2011 kun eid i et år før de er solgt igjen. Dermed er det naturlig at lavest andel tap, målt i prosent av total mengde tap i perioden er lavest de første årene etter undersøkelsesperioden startet. Tabell 7.1 viser hvordan den totale mengden leiligheter solgt med tap fordeler seg over undersøkelsesperioden. Tabellen viser at det er størst andel leiligheter solgt med tap i 2015, 2016 og 2018.

	2011	2012	2013	2014	2015	2016	2017	2018
Tap	2,34	7,81	12,50	14,84	18,75	17,97	8,59	16,40

Tabell 7.1 Fordeling av tap, oppgitt i prosent av totale tap

I tabell 7.2 er andel gevinst og tap per år, oppgitt i prosent av totale salg per år presentert. Her ser vi at andelen tap i prosent av totale salg per år er relativt stabil på mellom 20% og 30%. 2017 er et unntak, der kun 8,53% av de solgte leilighetene ble solgt med tap.

	2011	2012	2013	2014	2015	2016	2017	2018
Gevinst	76,92	73,68	71,43	67,80	74,19	81,60	91,47	85,31
Tap	23,08	26,32	28,57	32,20	25,81	18,40	8,53	14,69

Tabell 7.2 Andel leiligheter solgt med gevinst og tap per år, oppgitt i prosent av totale salg per år

En studie av leiligheter omsatt to eller flere ganger i Norge, presentert av Schultz og Kaspersen (2016) fant at sannsynligheten for tap er større når man eier leiligheten i en kort periode. Derfor inkluderes det her en tabell som presenterer andel solgt med gevinst og tap, oppgitt i prosent av totale antall salg med det respektive antall år mellom omsetningene.

	1	2	3	4	5	6	7	8
Gevinst	70,29	85,81	77,40	85,09	79,01	82,98	100,00	80,00
Tap	29,70	14,18	22,60	14,91	20,99	17,02	0,00	20,00

Tabell 7.3 Antall år mellom omsetninger, oppgitt i prosent av totale observasjoner med respektive antall

7.1. Attributtenes påvirkning på leiligheter solgt mer enn en gang

En regresjonsmodell basert på funksjon 6.3, men med kun de observasjonene som er solgt mer enn en gang er presentert i tabell 7.4. Det er undersøkt at forutsetningene for regresjon er oppfylt. I denne modellen har ikke avstand til sentrum signifikant koeffisient, det betyr at det antas at den ikke har betydning for leilighetens totalpris.

I denne regresjonsmodellen har området Tobienborg, i tillegg til Kjempegravane, Gimlevang og Lund industriområde et lavere prisnivå enn referansekategori Solbygg. I motsetning til regresjonsmodellen som inkluderte alle observasjoner er det i denne regresjonsmodellen bare områdene Hamreheia og Nedre Lund som har høyere prisnivå enn Solbygg.

F(29, 1160) = 191.00						
Prob > F = 0.0000						
R-squared = 0.8698						
Root MSE = .12681						
lnP	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
lnM	.7152994	.0130767	54.70	0.000	.6896428	.740956
lnY	-.1010836	.0099616	-10.15	0.000	-.1206283	-.0815389
lnkm	-.0121512	.0208682	-0.58	0.560	-.0530949	.0287925
lnE	.0214527	.0075056	2.86	0.004	.0067265	.0361788
G	-.0343965	.0147656	-2.33	0.020	-.0633668	-.0054263
H	.0608613	.0179909	3.38	0.001	.025563	.0961597
BD	-.0171611	.0148835	-1.15	0.249	-.0463627	.0120406
SD	.0379418	.0111291	3.41	0.001	.0161065	.0597772
år2011	.0514105	.0172519	2.98	0.003	.0175621	.085259
år2012	.0519084	.0172283	3.01	0.003	.0181063	.0857105
år2013	.067115	.0178599	3.76	0.000	.0320737	.1021563
år2014	.0435296	.0178518	2.44	0.015	.0085043	.0785549
år2015	.0776125	.017179	4.52	0.000	.0439071	.111318
år2016	.1306077	.0163469	7.99	0.000	.0985349	.1626805
år2017	.1339921	.0168006	7.98	0.000	.1010291	.1669551
år2018	.1804433	.016976	10.63	0.000	.1471362	.2137503
L1	-.148096	.024725	-5.99	0.000	-.1966067	-.0995853
L2	-.0110031	.0333448	-0.33	0.741	-.076426	.0544198
L3	.069199	.0282229	2.45	0.014	.0138252	.1245727
L4	.0201317	.0133295	1.51	0.131	-.0060209	.0462843
L5	-.0315353	.0242171	-1.30	0.193	-.0790495	.0159789
L6	.0058945	.031656	0.19	0.852	-.056215	.0680039
L7	-.0154801	.0158113	-0.98	0.328	-.046502	.0155418
L8	.0032869	.0189001	0.17	0.862	-.0337954	.0403691
L9	-.0829691	.020274	-4.09	0.000	-.1227469	-.0431913
L11	.2556234	.0247148	10.34	0.000	.2071327	.3041141
L12	.0270074	.0184823	1.46	0.144	-.0092552	.0632699
L13	.0418744	.0264699	1.58	0.114	-.0100599	.0938086
L14	-.1129556	.025091	-4.50	0.000	-.1621844	-.0637268
_cons	11.96628	.07082	168.97	0.000	11.82733	12.10523

Tabell 7.4 Regresjonsmodell for leiligheter med registrert gevinst eller tap

7.2. Sammenlikning av regresjon for to grupper

Problemstillingen for denne oppgaven handler om hvilke attributter som påvirker leilighetens totalpris og om det finnes noen fellestrekk ved de leilighetene som selges med tap.

Jeg ønsker å undersøke om de ulike attributtene har lik eller ulik betydning for leilighetens totalpris for leiligheter solgt med tap og leiligheter solgt med gevinst. For å kunne oppdage forskjellene mellom gruppene, bør ikke for mange variabler inkluderes. Det skal heller ikke inkluderes for mange kategoriske variabler i analysen. I denne analysen har jeg valgt å inkludere de kontinuerlige variablene fra regresjonsmodellen. Det betyr at attributtene leilighetsstørrelse, alder, avstand til sentrum og hvilken etasje leiligheten er i er inkludert. Det

er også inkludert en dummyvariabel som har verdien 1 dersom leiligheten er bygd i løpet av undersøkelsesperioden.

I henhold til Greene (2003) skal følgende regresjon gjennomføres.

$$P_{it} = \alpha + \beta_1 D_i + \beta_2 M + \beta_3 (D_i M) + \beta_4 Y + \beta_5 (D_i Y) + \beta_6 E + \beta_7 (D_i E) + \beta_8 km + \beta_9 (D_i km) + \beta_{10} ny + \beta_{11} (D_i ny) \varepsilon_{it}, t = 0,1. \quad (7.1)$$

Variabelen D_i har verdien 1 dersom leiligheten er solgt med tap. Dersom leiligheten er solgt med gevinst er verdien av variabelen lik 0.

Funksjonen for leiligheter solgt med tap blir da

$$E[P_i | D_i = 1, Z_i] = \alpha + \beta_1 D_i + (\beta_2 + \beta_3 D_i)M + (\beta_4 + \beta_5 D_i)Y + (\beta_6 + \beta_7 D_i)E + (\beta_8 + \beta_9 D_i)km + (\beta_{10} + \beta_{11} D_i)ny + \varepsilon_{it} \quad (7.2)$$

For leiligheter solgt med gevinst vil funksjonen bli som følger

$$E[P_i | D_i = 0, Z_i] = \alpha + \beta_2 M + \beta_4 Y + \beta_6 E + \beta_8 km + \beta_{10} ny + \varepsilon_{it} \quad (7.3)$$

For å gjennomføre regresjon med funksjon 7.2 i Stata må det genereres interaksjonsvariabler. Disse er produktet av dummyvariabelen (D_i) multiplisert med det respektive attributtet.

Resultatet fra regresjonen er presentert i tabellen nedenfor

Linear regression		Number of obs	=	656	
		F(11, 644)	=	83.06	
		Prob > F	=	0.0000	
		R-squared	=	0.7370	
		Root MSE	=	4.9e+05	
P	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]
tap	48770.02	392574.5	0.12	0.901	-722110.6 819650.6
tapM	3058.149	4344.957	0.70	0.482	-5473.845 11590.14
M	28982.44	1800.065	16.10	0.000	25447.74 32517.15
tapY	-6002.41	3886.329	-1.54	0.123	-13633.82 1628.999
Y	-8867.936	1773.4	-5.00	0.000	-12350.28 -5385.591
tapE	69.73864	31515.97	0.00	0.998	-61816.74 61956.21
E	9292.377	12347.77	0.75	0.452	-14954.38 33539.14
tapkm	-20814.69	89145.78	-0.23	0.815	-195866.2 154236.8
km	-171471	28674.77	-5.98	0.000	-227778.3 -115163.7
tapny	-508651.2	283555.4	-1.79	0.073	-1065456 48153.72
ny	527720.7	161984.4	3.26	0.001	209639.3 845802
_cons	1494251	169339.1	8.82	0.000	1161728 1826774

Tabell 7.5 Interaksjonsmodell

Etter en slik interaksjonsmodell er estimert, kan det gjennomføres en Wald test. Resultatet fra Wald-testen er som følger:

$$(1) \text{ tap} = 0$$

$$(2) \text{ tapM} = 0$$

$$(3) \text{ tapY} = 0$$

$$(4) \text{ tapkm} = 0$$

$$(5) \text{ tapE} = 0$$

$$(6) \text{ tapny} = 0$$

$$F(5, 647) = 9.35$$

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

Fra interaksjonsmodellen ser vi at variablene tap, tapM, tapY, tapkm og tapE har relativt høye p-verdier, og er dermed ikke signifikante. Det tyder på at det ikke er betydelige forskjeller i påvirkningen de kontinuerlige variablene har på totalprisen mellom leiligheter solgt med tap og leiligheter solgt med gevinst. Disse resultatene kan settes i sammenheng med histogrammet i figur 4.9 som viste at det ikke var så store forskjeller i leilighetens totalpris mellom de to gruppene. Derimot viser interaksjonsmodellen at variabelen «tapny» er signifikant til et signifikansnivå på 10%.

7.3. T-test om signifikante forskjeller i gjennomsnittsverdier

Histogrammene i kapittel 4 indikerte at det var noen betydelige forskjeller mellom leiligheter solgt med tap og leiligheter solgt med gevinst. I dette kapitlet skal en rekke t-tester (two sample t-test) gjennomføres. T-testen brukes til å undersøke om det er statistisk signifikant forskjell mellom gjennomsnittsverdien til to grupper. I dette kapitlet skal gruppene leiligheter solgt med tap og leiligheter solgt med gevinst undersøkes mot hverandre. Resultatene fra disse testene skal senere brukes til å besvare problemstillingen og de forskningsspørsmålene som er presentert i kapittel 3.

Variabel	Gjennomsnittsverdi leiligheter solgt med tap	Gjennomsnittsverdi leiligheter solgt med gevinst	Signifikant forskjell
Totalpris	2 426 206 kr	2 551 956 kr	Nei
Boligalder	48,71 år	52,24 år	Ja*
Avstand til sentrum	2,45 km	2,57 km	Ja**
Avstand fra UiA	2,065 km	2,100 km	Nei
Størrelse	65,89 m ²	66,27 m ²	Nei
Antall soverom	1,718	1,794	Nei
Avvik pris-prisantydning	- 78 633,06 kr	21 020,55 kr	Ja***
Omsetningshastighet	78,11 dager	33,23 dager	Ja***
Antall år mellom salg	2,94 år	3,33 år	Ja***

Tabell 7.6 Resultat fra T-tester.

*Signifikant til et nivå på 10%

**Signifikant til et nivå på 5%

*** Signifikant til et nivå på 1%

Gruppen leiligheter solgt med tap har en gjennomsnittlig totalpris som er kr 125 750 mindre enn utvalget som består av de leiligheter som er solgt med gevinst.

De leilighetene som selges med tap har lavere gjennomsnittlig boligalder enn utvalget av leiligheter solgt med gevinst. Resultatet fra t-testen viser at det med dette datasettet er en signifikant forskjell mellom de to utvalgene.

Bydelen Lund er en relativt kompakt bydel, og de fleste leilighetene er lokalisert mellom 2 og 3 kilometer fra bysentrum. Gjennomsnittlig avstand til sentrum er 120 meter mindre for gruppen leiligheter som selges med tap. Den gjennomsnittlige avstanden til Universitetet i Agder er relativt lik for de to gruppene leiligheter solgt med tap og leiligheter solgt med gevinst.

Gjennomsnittsverdien for leilighetene som selges med tap er 0,368 m² mindre enn de leilighetene som selges med gevinst. Dette er en veldig liten differanse.

Gjennomsnittsverdiene for antall soverom i leiligheten er omtrent like høy i begge grupper, men leilighetene som selges med tap har en litt mindre gjennomsnittsverdi, og dette kan tyde på at det er flere små, eller eventuelt færre større leiligheter blant de leiligheter som selges med tap.

Gjennomsnittlig avvik mellom salgspris og prisantydning er negativt for leiligheter som selges med tap og positivt for leiligheter som selges med gevinst. Det betyr at for de

leilighetene som selges med tap er gjennomsnittlig prisantydning høyere enn gjennomsnittlig salgspris.

Det er et stort avvik mellom gjennomsnittsverdier i de to gruppene for variabelen omsetningshastighet. Det viser seg at de leilighetene som selges med tap har en gjennomsnittlig omsetningsverdi som er 44,87 dager lenger enn de leilighetene som selges med tap. Det viser seg også at den gjennomsnittlige tiden fra en leilighet er solgt til den selges igjen er litt lavere for de leilighetene som selges med tap.

7.4. Hypotesetesting

Resultatet fra t-testen presentert i tabell 7.6 skal brukes til å besvare de hypotesene som er presentert i kapittel 3.

Forskningsspørsmål 6: Er det de små leilighetene som selges med tap?

H₀: Det er ingen forskjell i leilighetsstørrelse mellom leiligheter solgt med tap og leiligheter solgt med gevinst

H₁: Gruppen med leiligheter som selges med tap har lavere gjennomsnittsstørrelse enn leilighetene som selges med gevinst.

Dette kan uttrykkes som

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 < \mu_2$$

Der μ_1 og μ_2 representerer gjennomsnittlig leilighetsstørrelse for leiligheter solgt med henholdsvis tap og gevinst.

T-testen for gjennomsnittsstørrelse viser at leilighetens gjennomsnittsstørrelse er 0,368 m² mindre for leilighetene som selges med tap enn leilighetene som selges med gevinst.

Resultatet fra t-testen viser at nullhypotesen ikke kan forkastes, det betyr at det ikke er en betydelig forskjell i gjennomsnittsstørrelse for de to gruppene. Dette indikerer at leiligheter i alle størrelser kan selges med tap.

Forskningsspørsmål 7: Er det flest nye leiligheter som selges med tap?

H₀: Det er ingen forskjell i leilighetens alder mellom leiligheter solgt med tap og leiligheter solgt med gevinst

H₁: H₁: De leilighetene som er solgt med tap har lavere gjennomsnittsalder enn de leilighetene som er solgt med gevinst.

Dette kan uttrykkes som

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

Gjennomsnittlig leilighetsalder for de leilighetene som er solgt med tap er 3,57 år lavere enn gjennomsnittsalderen for leiligheter solgt med gevinst. Nullhypotesen som sier at det ikke er noen forskjell i boligalderen mellom gruppene, kan forkastes med 90% sikkerhet. T-testen viser at gruppen med leiligheter solgt med tap har en signifikant lavere gjennomsnittsalder enn de leiligheter solgt med gevinst har.

Forskningsspørsmål 8: Vil leiligheter som omsettes hyppig være mer utsatt for salg med tap?

H₀: Det er ingen sammenheng mellom hvor mange år det er mellom to salg av samme leilighet og om den selges med tap

H₁: Leiligheter som selges med tap omsettes hyppigere enn de som selges med gevinst.

Dette kan uttrykkes som

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 < \mu_2$$

Gjennomsnittlig antall år mellom to salg er 2,9375 år for leiligheter som er solgt med tap og 3,3289 år for de leilighetene som er solgt med gevinst. T-testen indikerer at forskjellen mellom de to gruppene er signifikant forskjellig fra null.

Forskningsspørsmål 9: Øker risikoen for salg med tap desto lenger leiligheten ligger i markedet?

H₀: Det er ingen forskjell i omsetningshastighet mellom de to gruppene

H₁: De leiligheter som selges med tap har høyere gjennomsnittlig omsetningshastighet

Dette kan uttrykkes som

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 > \mu_2$$

Om lag 45% av leilighetene som selges med tap har en omsetningshastighet på mer enn 60 dager. I motsetning har kun 19% av leilighetene som selges med gevinst omsetningshastighet på mer enn 60 dager.

Gjennomsnittlig omsetningshastighet er 44,87 dager lenger for de leilighetene som selges med tap enn for de leilighetene som selges med gevinst. Konklusjonen fra t-testen er at omsetningshastigheten til leiligheter solgt med tap er betydelig høyere enn for de leiligheter som er solgt med gevinst.

Forskningsspørsmål 10: Har leilighetene som er solgt med tap en høyere prisantydning enn salgspris?

H₀: Gjennomsnittlig avvik mellom salgspris og prisantydning er likt for de to gruppene

H₁: Avviket mellom salgspris og prisantydning er mest negativt for de leilighetene som er solgt med tap

Dette kan uttrykkes som

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 < \mu_2$$

Gjennomsnittlig avvik mellom salgspris og prisantydning er negativt for gruppen som selges med tap og positivt for leilighetene som selges med gevinst. Det vil si at gjennomsnittlig salgspris er lavere enn gjennomsnittlig prisantydning for de leilighetene som selges med tap. Som vist i tabell 7.6 er det en betydelig forskjell i gjennomsnittlig avvik for de to gruppene.

T-testen indikerer at nullhypotesen kan forkastes til både et signifikansnivå på 99%. Det gis støtte til alternativhypotesen som sier at differansen mellom de to gruppene er mindre enn 0. Det tilsier at tapet for gruppen leiligheter som selges med tap er mer negativt enn gruppen leiligheter som selges med gevinst.

8. Videre analyse og diskusjon av resultater

I dette kapittelet vil resultatene fra kapittel 6 og 7 brukes til videre analyse. Her vil jeg også knytte resultatene til bakgrunnsinformasjonen og teorien som er presentert i kapittel 2 og 3.

Den dobbellogaritmiske modellen gav signifikante koeffisienter for alle salgsårsdummyer. Alle koeffisientene har positivt fortegn, som vil si at totalprisen var høyere i det respektive året enn i referanse kategorien 2010. Dette tyder på at prisene i hele perioden har vært høyere

enn i 2010. I 2014 er koeffisienten lavere enn i det foregående året, det indiker at prisen økte mindre sammenliknet med referanse kategorien enn hva den gjorde i 2013.

8.1. Leilighetens størrelse

Som antatt og i tråd med resultatene til Osland (2001) viser modellen at desto større leilighet, desto høyere er totalprisen. En 1% økning i leilighetens størrelse gir en økning på 0,762% i leilighetens totalpris.

Dette kan illustreres med et eksempel. Vi antar at det finnes to leiligheter som har samme kombinasjon og mengde av alle attributter, med unntak av størrelse. Den ene leiligheten er 50 m² stor og har en totalpris på kr 2 000 000. Den andre leiligheten har en størrelse på 75 m², og vil ifølge den estimerte modellen ha en totalpris på kr 2 762 000.

For å undersøke om det er noen forskjeller mellom store og små leiligheter gjennomføres regresjoner for mindre grupper av leilighetsstørrelse. I dette kapittelet skal jeg først gjennomføre regresjon for leiligheter større og mindre enn medianen, deretter gjennomføres regresjon for leiligheter større og mindre enn 55 m². Resultatene fra regresjonene skal brukes til å beregne økningen i totalpris ved en økning på 1 kvadratmeter i leilighetens primærrrom. Fullstendige regresjonsresultater finnes i vedlegg 5. Det er undersøkt at forutsetningene for regresjonen er tilstrekkelig oppfylt, slik at estimatene er troverdige.

Leiligheter mindre og større enn medianen

Gjennomsnittsstørrelsen (primærrrom) for de leiligheter inkludert i denne analysen er 68,64 m², og datasettets median er 68 m². Regresjon for leiligheter større og mindre enn 68 m² gjennomføres for å undersøke om det er noen forskjeller mellom gruppene. Til dette brukes den dobbellogaritmiske funksjonsformen.

Modellen for leiligheter mindre enn 68 m² sier at leilighetens totalpris øker med 0,745% ved en økning på 1% i leilighetens størrelse. Mens for de største leilighetene gir en økning i størrelse på 1% en økning på 0,898% i totalprisen. I modellen som omfattet alle størrelser var koeffisienten knyttet til størrelse 0,762.

Koeffisienten for størrelse er forskjellig mellom de to gruppene, men det må også legges til at en økning i størrelse på 1% ikke sier noe om hvor stor økningen er ved en økning på en kvadratmeter. For de små leilighetene tilsvarer en økning på en kvadratmeter et høyere prosent enn for de store leilighetene. Derfor vil jeg beregne økningen på en kvadratmeter for

gjennomsnittleiligheten og medianleiligheten for de to gruppene for å illustrere hvordan totalprisen endres ved en økning i størrelse på 1 m².

	Gj.snitt	Koeffisient	Økning på 1 m ² i prosent	Økning i totalpris
Små leiligheter	52	0,745	1,923%	1,43%
Store leiligheter	88	0,8981	1,136%	1,02%

Tabell 8.1 Endring i pris ved økning på 1 kvm for gjennomsnittleiligheten i hver gruppe

	Median	Koeffisient	Økning på 1 m ² i prosent	Økning i totalpris
Små leiligheter	56	0,7475	1,786%	1,34%
Store leiligheter	79	0,8981	1,266%	1,14%

Tabell 8.2 Endring i pris ved økning på 1 kvm for medianleiligheten i hver gruppe

Av tabell 8.1 og 8.2 er det klart at totalprisen for de små leilighetene øker mer, enn totalprisen for de store leilighetene når størrelsen øker med 1 m². En økning i leilighetens størrelse fra 52 m² til 53 m² gir en økning i totalpris på 1,43%, mens en økning fra 88 m² til 89 m² gir en økning på 1,020% i totalprisen.

Leiligheter mindre og større enn 55 m²

For å undersøke om forskjellen mellom store og små leiligheter endres når definisjonen av store og små leiligheter endres vil jeg også gjennomføre regresjon for leiligheter som er mindre og større enn 55 m². Den prosentvise økningen i totalpris ved en økning på 1 m² beregnes for gjennomsnitt og median i de to gruppene, små leiligheter og store leiligheter.

	Gj.snitt	Koeffisient	Økning på 1 m ² i prosent	Økning i totalpris
Små leiligheter	40	0,7834	2,5%	1,96%
Store leiligheter	77	0,8265	1,299%	1,07%

Tabell 8.3 Endring i pris ved økning på 1 kvm for gjennomsnittleiligheten i hver gruppe

	Median	Koeffisient	Økning på 1 m ² i prosent	Økning i totalpris
Små leiligheter	41	0,7834	2,439%	1,91%
Store leiligheter	69	0,8265	1,449%	1,19%

Tabell 8.4 Endring i pris ved økning på 1 kvm for medianleiligheten i hver gruppe

Tabell 8.3 og 8.4 viser at det er større endring i totalpris når størrelsen øker med 1 m² når små leiligheter defineres som leiligheter som er mindre enn 55 m² enn når små leiligheter defineres som leiligheter mindre enn 68 m². En økning i størrelse fra 40 m² til 41 m² tilsvarer en økning i totalpris på 1,96%. Dette er fordi en kvadratmeter representerer en større prosentandel desto mindre leiligheten er.

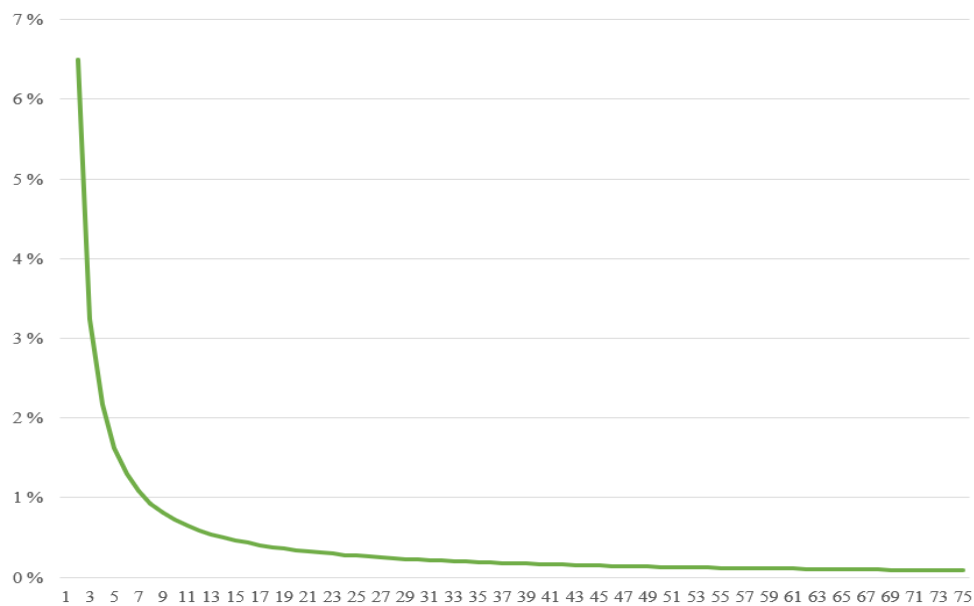
Felles for alle beregninger gjort i dette kapittelet er at en økning på 1 m² i leilighetens størrelse tilsvarer en større økning i totalpris for de små leilighetene enn for de store leilighetene.

8.2. Virkningen av økt alder på leilighet

Med boligens alder øker behovet for vedlikehold, samt at nyere boliger ofte er bedre isolert enn eldre leiligheter. Det vil si at det er knyttet flere kostnader til en eldre leilighet enn til en nyere, og dette kan være blant årsakene til at økt leilighetsalder reduserer leilighetens verdi.

I figur 8.1 er det illustrert hvordan en økning i alder påvirker leilighetens totalpris i henhold til denne modellen. Av figuren er det tydelig at prisen påvirkes mest av en økning i alder de første årene etter leiligheten er ny. Deretter flater kurven ut og stabiliserer seg. Mellom første og andre år synker leilighetens verdi med 6,5%. For en leilighet som kostet kr 2 000 000, innebærer det at totalprisen reduseres med kr 130 000 mellom første og andre år etter den var ny. En økning fra 39 år til 40 år, fører til en reduksjon på kun 0,171%. Det vil si at dersom leiligheten er verdt 2 000 000 i det 39ende året, er verdien etter 40 år kun redusert med kr 3 420.

En mulig årsak til at påvirkningen alder har på leilighetens totalpris flater ut, og nesten ikke har noen betydning ved økt alder kan være at når leiligheten har vært bebodd i en del år, oppgraderes den. Det vil si at selv om leiligheten har høy alder har ikke nødvendigvis et stort oppussingsbehov.



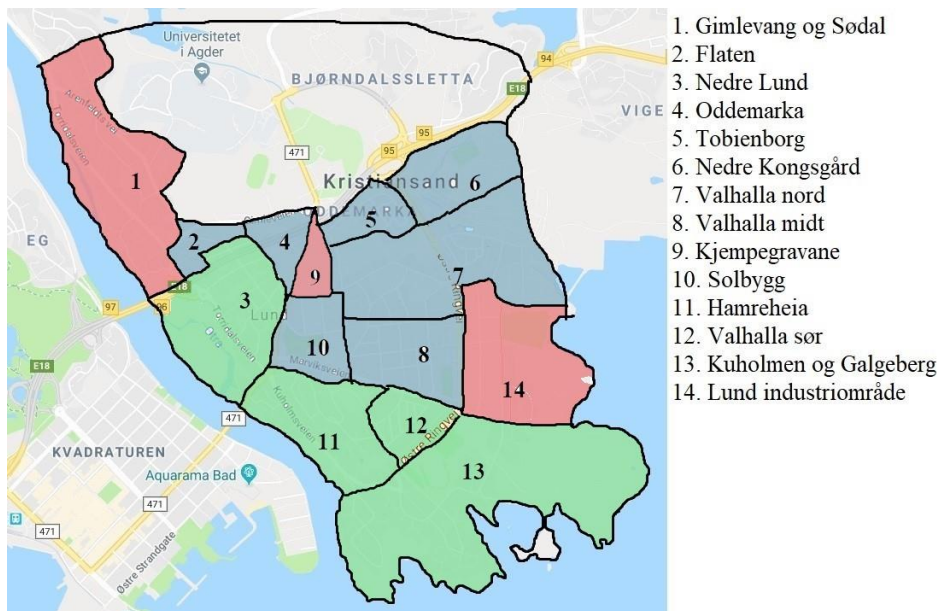
Figur 8.1 Virkningen av økt alder på leilighetens totalpris

8.3. Avstand til sentrum

Økonomisk teori og Alonso-Muth-Mills modellen sier at leilighetens pris reduseres desto lenger fra sentrum leiligheten er lokalisert. Denne modellen viser at for hver prosent lenger fra bykjernen en leilighet er lokalisert, reduseres leilighetens totalpris med 0,0829%.

Totalprisen til en leilighet som er lokalisert 3 km fra sentrum vil være 4,15% lavere enn for en leilighet som er lokalisert 2 km fra sentrum, dersom sammensetningen av de øvrige attributtene er identiske.

I områdekartet under er de estimerte prisnivåene fra regresjonsmodellen inkludert. De røde områdene er de områdene som ifølge regresjonsmodellen hadde lavere prisnivå enn referanse-kategorien Solbygg. De blåskraverte områdene er de områdene med middels prisnivå, de som ikke skiller seg betydelig fra Solbygg. De grønne områdene er de områdene som ifølge modellen har høyest prisnivå.



Figur 8.2 Områdekart med estimerte prisnivåer

Flertallet av de leilighetene som har mindre enn 2 km til sentrum er lokalisert i områdene Nedre Lund og Hamreheia, som også er blant de områdene med høyest gjennomsnittlig leilighetspris, dette er i tråd med Alonso-Muth-Mills modellen.

Kartet viser at områdene med høyest prisnivå ligger i nærhet til sjøen. Området Kuholmen og Galgeberg har lenger vei til sentrum enn det mange av de andre leilighetene i bydelene har, men dette området ligger nær sjøen som kan være en faktor som fører til at prisen er høy i dette området. Det finnes også en del nyere leiligheter i flere av de områdene med høyest

prisnivå. I disse områdene finner vi også relativt nye toppleiligheter med sjøutsikt i 10-millionersklassen.

Midt mellom områder med middels og høy prisklasser finner vi Kjempegravane som har lavere prisnivå. Dette området har omtrent samme avstand til sentrum som Oddemarka og Solbygg, men prisnivået i området er lavere. Hva som er årsaken til at prisnivået i Kjempegravane er lavere enn nærliggende områder er vanskelig å bestemme med denne modellen. Gjennomsnittlig kvadratmeterpris er lavere for området Kjempegravane enn den er for bydelen sett sammen. Regresjonsmodell av bare det gitte området viser at en økning i størrelse på 1% kun fører til en økning på 0,64% i totalpris, sammenliknet med 0,76% økning som modellen for hele bydelen viste.

De to resterende områdene med lavest prisnivå, Lund Industriområde og Gimlevang, er blant områdene med lengst avstand til sentrum, og at prisnivået er lavest i disse områdene er dermed i tråd med økonomisk teori.

8.4. Betydningen av parkeringsplass og heis

Om leiligheten hadde egen parkeringsplass har i denne modellen ingen påvirkning på totalprisen. I bydelen Lund er gateparkering tillatt i mange gater, og dette kan føre til at parkeringsplass tilhørende leiligheten ikke har like stor betydning som det vil ha i sentrumsgater med gateparkering forbudt, eller i storbyer med flere biler enn tillatte parkeringsplasser. Som nevnt ligger Universitetet i Agder i bydelen, og mange studenter bor på Lund. Mange studenter har ikke bil, og dette kan også være en faktor som fører til at parkeringsplass ikke øker totalprisen betraktelig i bydelen. Fra et sentralt punkt i bydelen tar det omtrent 10 minutter å sykle til Kristiansand sentrum, og bussforbindelsene både til sentrum og til Sørlandsparken, som kan regnes som et sub-arbeidssentrum, er relativt gode. Det betyr at det er mulig å bo i bydelen uten å ha bil. Det er naturlig å tenke at parkeringsplassen verdsettes mer i et område eller bydel hvor gateparkering er vanskeligere, samt at behovet for bil er større enn det er i bydel Lund.

Modellen indikerer at leiligheter som befinner seg i en blokk hvor det finnes heis har høyere verdi enn leiligheter hvor det ikke er tilgang på heis. Heis er et gode som er særlig viktig for småbarnsforeldre med barnevogn og eldre som ikke bor i første etasje. I kapittel 2 ble det kjent at det i bydelen bor en betydelig større andel eldre sammenliknet med kommunen sett under ett. Dette kan være blant faktorene som gjør at heis øker leilighetens verdi. Andelen små barn er lavere enn for kommunen sett under ett, men høyere enn for Kvadraturen. En stor

del av innbyggerne i bydelen er mellom 20 og 29 år, og i etableringsfasen. I denne aldersgruppen kan man ønske heis fordi man planlegger barn i løpet av et par år.

8.5. Leiligheter som er solgt med tap i perioden 2010 – 2018.

Andelen tap per år varierer, og det er viktig å ta med i betraktningen at alle leiligheter inkludert i denne analysen er solgt mellom 2010 og 2018. Dermed viser tabell 7.2 bare andelen leiligheter i undersøkelsesperioden som er solgt med tap og det kan være flere leiligheter enn hva som kommer frem av mine resultater som er solgt med tap. Av tabell 7.2 ser vi at andelen tap er ganske stabil på mellom 20% og 30% av leiligheter solgt per år, med unntak av 2017 og 2018 hvor andelen tap var lavere. I årene 2013 og 2014 er det klart størst andel leiligheter solgt med tap. Det vil si leiligheter som maksimalt er eid i tre eller fire år.

Tabell 7.3 viser at av de leilighetene som er solgt etter å bare ha blitt eid i ett år er 29,70% er blitt solgt med tap. Andelen leiligheter solgt med tap er nest størst når det er tre år mellom salgene. Gjennomsnittsverdiene til gruppene tap og gevinst indikerer at antall år mellom salgene er lavere for leiligheter solgt med tap enn for leiligheter solgt med gevinst. T-testen viser at det er en signifikant forskjell mellom gjennomsnittsverdiene til leiligheter solgt med tap og leiligheter solgt med gevinst. Dette er i tråd med hva Schultz og Kaspersen (2016) presenterte – at de leilighetene som omsettes hyppigst er mest utsatt for salg med tap.

Resultatet fra regresjonen viste at områdene Hamreheia, Kuholmen, Valhalla sør og Nedre Lund er de områdene i bydelen med høyest prisnivå. Valhalla sør er blant områdene med færrest tap registrert, det er heller ikke så stor andel tap i områdene Hamreheia og Kuholmen. Derimot har Nedre Lund i perioden har hatt en relativt høy andel leiligheter solgt med tap. Halvparten av de leilighetene som er solgt med tap i området Nedre Lund har en boligalder på mer enn 80 år. Den nyeste leiligheten som er solgt med tap har en alder på 11 år. Det er en adresse hvor det i løpet av perioden er tre tap, ellers er det få likheter i lokalisasjonen til de leiligheter som er solgt med tap i området Nedre Lund.

Områdene Solbygg, Valhalla nord og Kjempegravane er de områdene i bydelen Lund med lavest gjennomsnittlig totalpris. Av de leilighetene som er omsatt mer enn en gang er 19,7% av leilighetene i området Solbygg solgt med tap. I Valhalla nord og Kjempegravane er henholdsvis 16,35% og 23,53% solgt med tap. For bydelen generelt er 19,48% solgt med tap, så tallene jeg ser fra de områdene med lavest gjennomsnittspris avviker ikke mye fra andelen i bydelen totalt.

8.5.1. Området Tobienborg

Området Tobienborg skiller seg ut ved at det i perioden er her det har vært klart størst andel leiligheter solgt med tap. Av de leilighetene som er omsatt mer enn en gang i området Tobienborg er 40% av leilighetene solgt med tap, og det største tapet er på kr 425 000. Den gjennomsnittlige endringen i totalpris mellom salgene er også negativ. Det er i dette området Tobienborg hage med 70 leiligheter har stått ferdig i slutten av 2018 og begynnelsen av 2019. Per 15.april 2019 er om lag 28% av disse leilighetene usolgt (Eiendomsverdi AS, 2019a).

Av de leilighetene som er solgt mer enn en gang i området Tobienborg er alle yngre enn 10 år gamle, det vil si at det i alle tilfeller er snakk om relativt nye leiligheter. I området er det 6 leiligheter som er solgt med tap, og alle disse har en alder på 7, og er bygd i 2012. Fire av seks leiligheter er lokalisert i samme blokk og de resterende to er lokalisert i samme blokk et annet sted i området. I blokken hvor to leiligheter er solgt med tap, finnes det også seks leiligheter som er solgt med gevinst. Leilighetene som er solgt med tap er lokalisert i både første og andre etasje. En av leilighetene har ett soverom, mens de fem andre har to. Alle leilighetene som er solgt med tap er solgt under prisantydning. Alle leilighetene har sør eller vestvendt balkong og parkeringsplass, samt at det er heis i begge bygg. Utfra den informasjonen jeg besitter er det vanskelig å konkludere på hvorfor leiligheter i akkurat disse to blokkene er utsatt for salg med tap. Det er naturlig å anta at det er relativt lave kostnader både i forbindelse med vedlikehold og oppvarming forbundet med en ny leilighet.

Boligannonser fra Finn.no viser at de månedlige felleskostnadene ligger på rundt kr 35 per kvadratmeter i den blokken hvor fire av fire leiligheter solgt i perioden er solgt med tap.

Det er vanskelig å finne en klar grunn til hvorfor fire av fire solgte leiligheter i samme blokk er solgt med tap. Disse leilighetene er heller ikke solgt i samme år. Jeg tenkte derfor det kunne være interessant å se på hvordan disse leilighetene er markedsført i boligannonsene. Den leiligheten med klart lengst omsetningshastighet, på 195 dager, er markedsført som «nyere og meget lekker 2-roms leilighet». Meget lekker er et uttrykk som gjerne forbindes med noe luksuriøst og noe som er ekstra fint som skaper høye forventninger hos potensielle kjøpere. Når bildene av leiligheten viser en helt ordinær, nyere leilighet med hvit veggpanel i alle rom, hvit kjøkkeninnredning og flislagt bad med standardinnredning kan interessen svekkes.

Da blokkene er relativt nye, kan det tenkes at leilighetens første eier har vært villig til å betale relativt mye for et nybygg. Når leiligheten etter et par år er solgt er ikke neste eier villig til å

betale like mye for en brukt leilighet, dermed selges leilighetene med tap fordi prisen i utgangspunktet var høy da leiligheten var ny.

Det viser seg at 27% av de leilighetene i bydelen som er bygd etter 2009 er solgt med tap. Dette kan støtte oppom min teori om at den første eieren har vært villig til å betale mer for ett nybygg enn hva neste eier er villig til å betale for en brukt leilighet. Tobienborg er som nevnt et område hvor det i løpet av de siste årene er bygd mange nye leiligheter. Dette kan forårsake mer støy enn først antatt, samt at området har blitt tettere befolket.

En takstingeniør hevder at boligutvikling i nærområdet kan øke verdien på bolig med mer enn kr 200 000. Men det ligger også mye verdi i en usjenert og solrik uteplass. I noen av områdene bygges det tett, og nær vei eller fortau. Dette fører til at mange har verandaer med fullt innsyn, som antas å redusere boligens verdi. Han sier også at støy og trafikk kan redusere boligens verdi med mer enn kr 100 000 (Lindvoll, 2018). I området Tobienborg er det i løpet av de siste årene bygd mange nye leiligheter, og området virker innbydende og moderne. Men med økt leilighetsmasse øker også trafikken og støyen i området. Ifølge takstingeniørens utsagn drar disse kreftene totalprisen i hver sin retning. Resultatet viste at 40% av leilighetene som er solgt mer enn en gang i løpet av perioden i området Tobienborg er solgt med tap og som nevnt er leilighetene lokalisert i to nærliggende blokker. Disse blokkene er blant de nærmeste naboene til Tobienborg hage, og har i løpet av de siste årene fått dårlige utsikt, samt at støynivået antakelig har økt. Dette kan være blant årsakene til at dette området er særlig utsatt for hyppige salg og tap i denne perioden. Fædrelandsvennen skriver 10.mai 2019 om en sak i en annen bydel, Vågsbygd, der det er planlagt leiligheter tett på eksisterende boliger. Naboene viser misnøye, og mener at det ikke bare handler om tap av utsikt, men også om hvordan man har det i eget hjem. Det naboene er mest bekymret for er den økende trafikken blokkene vil føre med seg (Damsgaard, 2019). Dette er et eksempel på at selv om flere ønsker å bo i blokk, ønsker man fortsatt privatliv og følelsen av å være alene i egen stue, samt at trafikksituasjonen i en bydel i utvikling bekymrer innbyggerne.

Området Tobienborg er et område med mange nye blokker, og innbyggere i enkelte av disse nye blokkene opplever rystelser i forbindelse med trafikk i en nærliggende tunnel. Det har vært en pågående sak der leilighetseiere i 39 leiligheter har saksøkt byggherren på grunn av at de opplever det som beskrives som jevnlig «jordskjelv» på grunn av Oddernestunnelen som ligger rett ved (Damsgaard, 2014). Ingen av de 39 aktuelle leilighetene er solgt med tap i denne perioden, men blokker i samme område er hardt rammet av andelen tap i

undersøkellesperioden. Alle leilighetene som er solgt med tap i området er solgt etter at saken hadde blitt omtalt. Selv om ingen av leilighetene i de blokkene som har saksøkt byggherren er rammet av tap, kan sakens mediedekning likevel være en medvirkende faktor til at området Tobienborg har hatt en stor andel tap i løpet av perioden.

8.5.2. Leiligheter som er yngre enn 10 år

Som nevnt tidligere viste det seg at 27% av de leilighetene som er yngre enn 10 år, og som er solgt mer enn en gang i perioden er solgt med tap. T-testen viste også at gjennomsnittlig alder på de leilighetene som selges med tap er betydelig lavere enn alderen på de leilighetene som selges med gevinst. Av histogrammene i figur 4.10 ser vi at det er flere leiligheter som er yngre leiligheter som er solgt med tap enn det er som er solgt med gevinst.

Resultatet fra interaksjonsmodellen som er presentert i tabell 7.4 gir en negativ, signifikant koeffisient for variabelen «tapny». Variabelen «ny» har positiv, signifikant verdi. Dette indikerer at om leiligheten er ny i utgangspunktet øker totalprisen. Derimot tyder det på at dersom leiligheten er solgt med tap er det at leiligheten er ny en faktor som trekker ned leilighetens totalpris. Dette gir incentiv til å undersøke de leilighetene som er solgt for første gang i perioden ytterligere.

Av de leilighetene som har vært helt nye, og omsatt mer enn en gang i løpet av perioden er omtrent 25% solgt med tap mellom første og andre eier av leiligheten. Det viser seg at 14 av 16 registrerte tap på leiligheter yngre enn 10 år er ved overdragelse fra første til andre eier. Av disse 14 leilighetene er ingen solgt en gang til i perioden, det gjør at det ikke er mulig å si noe mer om prisutviklingen på disse leilighetene.

Dette underbygger også hva som kom frem for området Tobienborg, at nye leiligheter som selges etter kort tid er særlig utsatt for salg med tap. At mange kjøper nye leiligheter før de er ferdigstilt kan være en av årsakene til at disse leilighetene er ekstra utsatt for salg med tap.

Når en leilighet kjøpes før den er ferdigstilt, er kjøpers beslutning gjort på bakgrunn av prospekt og hva utbygger eller selger har lovet. Da har kjøper ikke mulighet til å kjenne på og kvalitetssikre de materialer som er brukt før man har kjøpt leiligheten. Et sentralt spørsmål i forhold til dette er om vi som kjøpere er rasjonelle. Det kan være vanskelig å klare å se for seg hvordan livet i leiligheten blir når leiligheten er kjøpt før den er ferdigbygd. Det er vanskelig å vurdere i hvilken grad leiligheten blir plaget av innsyn og støy, samt at utsikten kan bli noe annet enn hva som var forespeilet. Dette er faktorer som kan føre til at den første kjøperen

verdsetter leiligheten mer enn det den neste kjøperen gjør, fordi den neste kjøperen har mulighet til å se og kjenne på materialer og egenskaper ved leiligheten. Et annet viktig moment er at dersom en leilighet selges relativt raskt etter den første eieren har kjøpt den, uten en god forklaring om for eksempel endring i livssituasjon, kan det gi de signaler om at noe er galt med leiligheten. Dette kan da føre til at ny eier ikke er villig til å betale like mye for leiligheten som det den forrige eieren var.

8.5.3. Leiligheter i første etasje

En leilighet som ligger i første etasje er som regel mer utsatt for innsyn enn det en leilighet i høyere etasjer er. På Lund bygges det tett, og mange av blokkene ligger nær en vei. I mange av blokkene er det store stuevinduer, og graden av innsyn er høy. En leilighet i første etasje vil også kunne være mer plaget av støy. Regresjonsresultatene viste at en økning på 1% i leilighetens etasjeplassering økte prisen med 0,036%. Det vil si at prisforskjellen mellom en leilighet i første og andre etasje, med lik sammensetning av de øvrige attributtene er 3,6%. Mellom andre og tredje etasje skiller 1,8% i leilighetens totalpris ifølge denne modellen.

En av fire leiligheter som ligger i første etasje er solgt med tap i perioden. Det er en betydelig større andel enn i andre og tredje etasje. Likevel finner jeg at det er mellom sjette og tiende størst andel leiligheter er solgt med tap. Det er ikke så mange blokker i bydelen med mellom seks og ti etasjer, dermed er det følgelig relativt få leiligheter i disse etasjene som er solgt mer enn en gang i perioden. At leiligheter i de øverste etasjene selges med tap kan skyldes andre faktorer enn leilighetens etasjebelighet.

8.5.4. Omsetningshastighet og prisantydning

De leilighetene som selges med tap har betydelig høyere gjennomsnittlig omsetningshastighet enn de leilighetene som selges med gevinst. Dette kan tyde på at disse leilighetene ikke fremstår som like attraktive som de leilighetene som selges med gevinst. Men akkurat hvilke faktorer det er som gjør at leilighetene selges med tap er vanskelig å bestemme med de variablene jeg har hatt tilgang til. Knight (2002) peker på at en av årsakene til høy omsetningshastighet kan være at prisantydningen er for høy, da kan man miste potensielle kjøpere fordi prisantydningen overstiger deres budsjett. På Finn.no, hvor boligannonser for de fleste leiligheter ligger er det mulig å avgrense søket slik at søkerresultatene for eksempel bare viser de leilighetene som er innenfor kjøpers budsjett. Dersom prisantydningen er for høy, kan dermed leiligheten falle utenfor søkerresultatene til potensielle kjøpere. Faktorer jeg ikke har

undersøkt er blant annet om leiligheten har balkong og solforholdene. Dette kan selvsagt også påvirke salgsprisen og om leiligheten selges med gevinst eller tap. En annen årsak til at det er høy omsetningshastighet på leilighetene som selges med tap kan være at selger blir mer villig til å senke prisen, desto lengere tid det tar å selge leiligheten.

Dataen fra perioden viser at også en stor andel av de leilighetene som selges med gevinst selges til en pris som er lavere enn prisantydningen. I følge tall fra Eiendom Norge selges de fleste boliger i Norge til prisantydningen (Lindvold, 2018). At omtrent 40% av leilighetene i datasettet er solgt under prisantydning kan tyde på at selger og megler har høyere forventninger til salgsprisen enn hva kjøpere er villige til å betale i dagens marked.

Det viser seg at omtrent 20% av leilighetene som er omsatt to eller flere ganger i undersøkelsesperioden har hatt en lavere prisantydning enn ved forrige omsetningspris. For leiligheter som er yngre enn 10 år, er andelen litt høyere – 22,5%. Dette understøtter også teorien om at det å kjøpe ny leilighet før den er ferdig bygd medfører en viss risiko for at man betaler en høyere pris enn hva man ville gjort dersom man hadde sett leiligheten ferdigstilt før man kjøpte den.

8.6. Tilbud og etterspørsel

Grunnleggende teori om tilbud og etterspørsel sier at dersom tilbudet overstiger etterspørselen vil prisene presses ned. I bydelen Lund har tilbudet av leiligheter økt i perioden. Noen eneboliger har blitt revet for å gjøre plass til nye blokker, men mange av blokkene er også bygget der det tidligere har vært butikker, industribygg eller bare en åpen plass. Dette gjør at det er litt usikkert hvor mye det totale boligtilbudet har økt, men det har med sikkerhet økt i perioden. Gjennomsnittlig leilighetspris har også økt i perioden, det tyder på at det foreløpig ikke er noe problem med at tilbudet overstiger etterspørselen. Det har også i perioden vært en jevn befolkningsøkning på rundt 1% årlig. Det som er interessant er at befolkningsveksten, justert for antall studenter som bodde i Kristiansand fra før var lavere enn vanlig. Mellom 2016 og 2017 økte antallet innbyggere i bydelen Lund med bare 24 personer. Dette var det klart laveste tallet gjennom hele perioden. Hvordan dette påvirker boligprisene er litt usikkert, da det er så nært i tid, og effektene av dette kan komme senere. Men dersom årlig befolkningsvekst er lavere enn estimert over en lengere periode, kan det bli en situasjon hvor tilbudet overstiger etterspørselen og prisen presses ned.

Gjennomsnittlig leilighetspris har i perioden økt, men det kan også være et resultat av at de nye leilighetene har en høyere pris, samt at det finnes flere store toppleiligheter enn tidligere

som trekker opp gjennomsnittsprisen. Tabell 8.5 viser gjennomsnittlig totalpris for leiligheter som er mindre enn 100 m². Av tabellen ser vi at gjennomsnittlig totalpris sank mellom 2011 og 2012, samt mellom 2015 og 2016. Når alle leiligheter er inkludert er gjennomsnittlig totalpris stigende i alle perioder. Her ser vi virkningen av avvikene presentert i kapittel 4, det finnes noen leiligheter som trekker gjennomsnittsprisen opp.

2010	2011	2012	2013	2014	2015	2016	2017	2018
2 153 888	2 315 235	2 248 228	2 306 318	2 367 648	2 417 945	2 411 672	2 590 565	2 626 647

Tabell 8.5 Gjennomsnittlig totalpris for leiligheter som har primærrom mindre enn 100 m²

Det er vanskelig å finne noen tydelige sammenhenger mellom antall leiligheter solgt med tap og mengden nybygg. Men av tabell 7.2 ser vi at andelen tap var størst i 2014, tabell 2.3 viser at dette er også året hvor det er bygd nest flest nye leiligheter. I 2015 var andelen tap stor, men antall nye leiligheter var lavt. I 2014 ble Norge rammet av et oljeprisfall, dette førte til at arbeidsledigheten økte. Kristiansand ble relativt hardt rammet, da mange jobbet i oljeindustrien. Dette gav ringvirkninger, og andre næringer ble også rammet. Dette kan ha ført til at leiligheter har blitt solgt til en lavere pris enn det de ble solgt for i årene før oljeprisfallet.

Andelen sysselsatte i Kristiansand kommune var lavest i 2016 og 2017, og den var i 2018 lavere enn hva den var i årene før oljeprisfallet. Av tabell 8.5 kommer det frem at gjennomsnittlig leilighetspris for de leiligheter som er mindre enn 100 m² var lavere i 2016 enn i 2015. Andelen leiligheter solgt med tap var stor i 2015 da var sysselsetningsgraden omtrent 3% lavere enn året før. Det er uvisst om det faktisk er den reduserte sysselsetningen som har skyld i en stor andel leiligheter solgt med tap eller om det er andre faktorer, eller en kombinasjon.

Det er 264 observasjoner med omsetningshyppighet større enn eller lik tre. Av disse observasjonene er 52 leiligheter solgt med tap. Av disse 52 leilighetene er 42 også solgt med gevinst ved en anledning i løpet av perioden. Det vil si at 80% av de leilighetene som er solgt tre eller flere ganger og solgt med tap, også er solgt med gevinst i samme periode. Dette gjør arbeidet med å konkludere med hvorfor noen leiligheter er solgt med tap vanskeligere. Av disse leilighetene er klart størst andel solgt med tap i årene 2014 og 2015. Dette kan ha en sammenheng med oljeprisfallet og at arbeidsledigheten i Kristiansand økte.

8.7. Relevansen og viktigheten rundt dette temaet

Bolig er av NOU 2002:2 (2002) et kapitalobjekt, og å kjøpe bolig blir sett på som en god investering. Det er stort fokus på boligprisveksten i de største byene, men det er svært få som uttaler seg om at boliger også kan selges med tap. Derfor synes jeg det er viktig å sette fokus på at det er en risiko for å selge bolig med tap.

Dagens boligmarked er karakterisert av høye boligpriser i de største byene. Mange, særlig unge mennesker i etableringsfasen, ender opp med et stort lån fordi boligprisene er høye (Myhre, Bugge & Nordstrøm, 2013). Bankene setter i dag krav til at låntaker skal ha en egenkapital på minimum 15%, samt at totalt lån ikke skal overstige fem ganger egen inntekt. Men selv om låntakeren tilfredsstiller bankens krav kan lånesummen bli høy. Dersom planen er å kjøpe en leilighet for å komme inn på boligmarkedet og bare eie noen år for å spare til større leilighet eller hus er det viktig å tenke seg om.

Denne analysen viser at sjansen for å selge med tap er størst dersom en bare eier leiligheten noen år. Smarte penger (2018) har utviklet en kalkulator som kan hjelpe til å bestemme om det lønner seg å leie eller eie bolig. Jeg har brukt denne kalkulatoren til å lage et enkelt eksempel. Å leie leilighet på rundt 60 m² Lund koster rundt kr 11 000 i måneden (Finn.no, 2019). Å kjøpe en tilsvarende leilighet koster i gjennomsnitt kr 2 227 077 i henhold til mitt datasett. En rentekostnad på 2,85% legges til grunn, og det antas 2% årlig verdistigning på leiligheten. Det antas felleskostnader på kr 3 000 i måneden. I kalkulatoren er også kostnader forbundet med salget medregnet. Resultatet fra kalkulatoren er som følger:

Antall år boligen eies	1 år	2 år	3 år	4 år	5 år
Lønnsomhet ved å eie	-98 000	-65 000	-34 000	-4 000	22 000

Tabell 8.6 Lønnsomhet ved å eie leilighet (Smarte penger, 2018)

Fra tabell 8.6 ser vi at det ifølge kalkulatoren utarbeidet av Smarte penger (2018) må man eie leiligheten på Lund i minst 5 år for at det skal være lønnsomt å eie fremfor å leie. I denne tabellen er det forutsatt at årlig verdistigning på Lund er 2%, men likevel lønner det seg ikke å eie leiligheten i en kort periode.

I Norge er vi veldig opptatt av å eie bolig, og rundt 80% eier den boligen de bor i. Men er det alltid mest lønnsomt å eie fremfor å leie, hvis man ikke har planlagt til å bo i leiligheten over en lengere periode? Det er selvsagt mange faktorer som spiller inn når dette spørsmålet skal evalueres, men det er viktig å ha med i beregningen at en av fem leiligheter i bydelen Lund

ble solgt med tap i perioden 2010-2018, samt at så mye som en av tre leiligheter i en analyse av boliger i Norge i perioden 2012 -2016, gjennomført for Sparebank 1 ble solgt med tap. Og i begge analyser viser det seg at desto færre år det er mellom salgene av boligen, desto større er sannsynligheten for at den selges med tap (Schultz & Kaspersen, 2016). Ved å leie leilighet får man bare konsumgodedelen av boligen, altså at man har et sted å bo. Fordelen med å leie, når man ikke vet hvor lenge man skal bli boende, er at man ikke sitter på et høyt boliglån når man ønsker å flytte.

8.8. Analysens svakheter

På grunn av at arbeidet med masteroppgaven går over en begrenset periode, har jeg vært nødt til å begrense datamengden da dataen måtte samles inn manuelt og systematiseres. Derfor er oppgavens omfang begrenset til bydelen Lund. Dette er et veldig begrenset område, og det er ikke gitt at resultatene i denne analysen gjelder i andre deler av Kristiansand kommune, i andre deler av landet, eller i andre tidsperioder enn den angitte perioden. Jeg har også avgrenset oppgaven til bare leiligheter, dermed er det ikke slik at resultatene fra denne analysen gjelder for eneboliger og rekkehus.

I denne analysen er kun leiligheter som er omsatt i perioden 2010 til 2018 inkludert. Dette medfører at dersom en leilighet er solgt i 2009 og igjen i 2016 er bare salget fra 2016 inkludert i denne analysen. Dermed er bare et begrenset utvalg av leilighetene i bydelen inkludert. Det ideelle hadde vært å gjennomføre denne analysen over en lengre tidsperiode, men tidsbegrensning samt at dataen måtte samles inn manuelt førte til at jeg måtte gjøre et valg i forhold til undersøkelsesperiode.

Det er også litt usikkerhet rundt hvor stor andel av leilighetsmarkedet som er inkludert i analysen. Estimater fra tabell 2.3 er at total leilighetsmasse i 2018 var 2 662. Hver observasjon har fått sitt unike identifikasjonsnummer, og det er litt mer enn 1 200 ulike leiligheter inkludert i analysen. Det betyr at mindre enn halvparten av leilighetene i bydelen er inkludert. En analyse med en større andel av leilighetsmassen inkludert kan gi andre resultater enn det jeg har fått i denne analysen.

Dataen er samlet fra Eiendomsverdi, og det finnes noen mangler i denne databasen. Som vist i kapittel 4 måtte en del observasjoner utelates fordi sentral informasjon manglet. Det hadde vært mulig å finne informasjon til en del av observasjonene ved å bruke prospekt eller kontakte meglere, men dette ble vurdert som både tidkrevende og unøyaktig. Dataen kunne også blitt feilregistrert ved å bruke prospekt eller snakke med meglere, da det ikke alltid er

mulig å skille leiligheter fra hverandre. Det er vanskelig å si om de utelatte leilighetene ville utgjort en betydelig forskjell i resultatene.

For å forklare prisen ved hjelp av den hedonistiske metoden har jeg i denne analysen valgt å inkludere attributtene primærrrom, boligens alder, avstand til sentrum, etasjen leiligheten er plassert i, om leiligheten har tilhørende parkeringsplass og om det er heis i blokken. I tillegg er dummyvariabler for om leiligheten ligger i blokk, eierform, salgsår, samt området leiligheten ligger i inkludert. Det antas at flere egenskaper ved boligen er med å påvirker prisen, blant annet tomtens verdi, bygningskostnader, hvor utsatt leiligheten er for støy, om det er balkong og avstand til kollektivtransport eller Europaveier. Disse attributtene er ikke inkludert i denne analysen.

Jeg har heller ikke tatt hensyn til makroøkonomiske forhold som befolkning og arbeidsledighet i analysen. Disse faktorene er kun forklart innledende, og deretter brukt til å foreslå årsaker i diskusjonsdelen. Men noen årssaksammenheng kan ikke bekreftes eller avkreftes.

8.9. Forslag til videreføring av oppgaven

Til slutt ønsker jeg å komme med noen forslag til videreføring av oppgaven. Dette er basert på spørsmål jeg selv skulle ønske jeg kunne funnet svar på i løpet av prosessen, men som jeg ikke har hatt anledning til å besvare på nåværende tidspunkt.

Det kunne vært interessant å undersøke hvilken påvirkning valg av ord, særlig adjektiv, i boligannonser overskrift har på potensielle kjøpere og prisen på boligen. Det kunne for eksempel vært gjort en undersøkelse av boligannonser til leilighetene som er solgt med både gevinst og tap i en periode og sett på om det var noen forskjell i ordbruk og knytte dette opp mot pris, tap og omsetningshastighet.

Det hadde også vært interessant å gjennomføre tilsvarende analyse for andre bydeler i Kristiansand eller andre byer i Norge. Det hadde vært særlig interessant å undersøke hvordan boligprisen i Kvadraturen har utviklet seg i denne perioden. Kvadraturen er også en bydel med høy byggeaktivitet, hvor det er mange nye leiligheter.

I løpet av de neste årene er det ventet at flere nye, store leilighetsprosjekt skal stå ferdig. Dette betyr at tilbudet av leiligheter øker. Befolkningsveksten i bydelen har vært laber de siste årene. Hvordan vil resultatet av denne analysen se ut om et par år? Og vil det bli en situasjon der leiligheter står tomme fordi det er bygd flere leiligheter enn det etterspørres?

En videreføring som kunne gitt interessante og nye opplysninger er å se på forholdet mellom å eie og leie i bydelen over for eksempel en treårsperiode. Deretter undersøke innbyggernes gjeld og rentekostnader, og undersøke lønnsomheten ved å eie fremfor å leie. Disse opplysningene kunne blitt knyttet til antall tap, og størrelsen på disse, og et helhetlig bilde av hvor lønnsomt det er å eie leilighet kan tegnes.

9. Konklusjon

I denne analysen er boligmarkedet i bydelen Lund i Kristiansand studert. Analysen som bygger på data fra Eindomsverdi.no har vist at leilighetens størrelse, leilighetens etasje, om det er heis i blokken og om leiligheten er selveierleilighet påvirker leilighetens totalpris i positiv retning. Ved økende alder og økende avstand til sentrum reduseres leilighetens totalpris. Modellen har også vist at det er prisforskjeller mellom de ulike områdene i bydelen Lund.

Av de variablene som er inkludert og de testene som er gjennomført er det vanskelig å trekke noen konklusjoner om hvilke leiligheter som selges med tap. Det viser seg imidlertid at det er noen adresser hvor flere leiligheter er solgt med tap i undersøkelsesperioden. I perioden 2010 til 2018 er nesten 20% av de leilighetene som er solgt mer enn en gang solgt med tap. Denne analysen indikerer at det er noen forskjeller mellom leiligheter som selges med gevinst og leiligheter med tap, og det er betydelige forskjeller mellom de to gruppene for attributtet avstand til sentrum. Det viser seg også at avviket mellom salgpris og prisantydning er negativt for gruppen leiligheter som er solgt med tap, mens det er positivt for de leilighetene som er solgt med gevinst. Omsetningshastigheten er betydelig høyere for leiligheter solgt med tap. Gjennomsnittlig omsetningshastighet for leiligheter solgt med tap er 78,11 dager, det er mer enn dobbelt så mye som for leiligheter som er solgt med gevinst. Antall år mellom salg av samme leilighet er betydelig lavere for de leilighetene som er solgt med tap enn for de som er solgt med gevinst.

Tallene i denne analysen viser det er en signifikant forskjell mellom gjennomsnittsalderen for leiligheter solgt med tap og leiligheter solgt med gevinst. Leiligheter solgt med tap har lavere gjennomsnittsalder og dette gav incentiv til å undersøke nye leiligheter. Det viste seg at 1 av 4 leiligheter som er yngre enn 10 år, som både er kjøpt av første eier og overdratt fra første til andre eier i løpet av perioden er solgt med tap. Dette kan tyde på at den første eieren verdsetter nybygg eller den informasjon som er gitt i prospektet, mer enn hva neste kjøper er

villig til å betale for en relativt ny, men brukt leilighet hvor han har mulighet til å gå på en visning i motsetning til første eier som kanskje kjøpte leiligheten før den sto ferdig.

Oppgavens omfang er begrenset både i tidsperiode og område. Da denne analysen som nevnt kun er basert om omsetninger foretatt mellom 1.1.2010 og 31.12.2018, er det en mulighet for at resultatene hadde vært litt annerledes dersom en lengere tidsperiode hadde blitt benyttet.

Analysen sier heller ingenting om hvor stor andel leiligheter som selges med tap i Kristiansand generelt eller på landsbasis. Men selv om analysens tidsperiode og sted er begrenset er den relevant i tiden da en nyere tidsperiode er benyttet. Konklusjonen om at 1 av 5 leiligheter selges med tap, og at andelen er større for de leilighetene som er yngre enn 10 år kan derfor anses å være korrekt i dagens boligmarked på Lund. Dette er en viktig å relevant analyse sett i lys av at det i løpet av de siste årene er bygget mange nye leiligheter i bydelen og at Kristiansand kommune legger til rette for ytterligere utbygging i området.

10. Referanser

- Aktiv Eiendomsmegling. (2019). Fellesgjeld. Hentet fra <https://aktiv.no/boligtips/Fellesgjeld>
- Anderssen, H. B. (2019a). Primærrrom. I *Store Norske Leksikon*. Hentet fra <https://snl.no/prim%C3%A6rrrom>
- Bolstad, E. (2019). Postnummer i Kristiansand Kommune. Hentet fra <https://www.erikbolstad.no/postnummer-koordinatar/kommune.php?kommunennummer=1001>
- Brooks, C. (2008). *Introductory econometrics for finance* (2. utgave. utg.). New York: Cambridge university press.
- Dahlum, S. (2018). Kvalitativ analyse. I *Store Norske Leksikon*. Hentet fra https://snl.no/kvantitativ_analyse
- Damsgaard, V. (2014). Det er jo den lille mann mot den store. *Fædrelandsvennen*, s. 8.
- Damsgaard, V. (2019). En horisont byttes med en vegg. *Fædrelandsvennen*. Hentet fra <https://www.fvn.no/nyheter/lokalt/i/wPxeRd/-En-horisont-byttes-med-en-vegg>
- DiPasquale, D. & Wheaton, W. C. (1996). *Urban economics and real estate markets* Prentice Hall Englewood Cliffs, NJ.
- Eiendomsverdi AS. (2019). Eiendomsinformasjon satt i system. Hentet fra <https://eiendomsverdi.no/>
- Eiendomsverdi AS. (2019a). Nybygg [Statistikk]. Hentet fra <https://eiendomsverdi.no/app/appNewDwelling.aspx>
- Eiendomsverdi AS. (2019b). Område/utvalg [Statistikk]. Hentet fra <https://eiendomsverdi.no/app/appAreaSelection.aspx>
- Emblem, A. W. (2018). *Forelesningsnotater og Powerpoints fra emnet BE-409: Real Estate Economics*.
- Eretveit, S. & Theisen, T. (2016). Efficiency and justice in the market for cooperative dwellings. *International Real Estate Review*, 19(3), 297-326.
- Finn.no. (2019). Bolig til leie. Hentet fra <https://www.finn.no/realestate/lettings/search.html?location=0.20011&location=1.20011.20179&location=2.20011.20179.20536>
- Google Maps. (2019). Kristiansand [kart]. I. Hentet fra <https://www.google.com/maps/place/Kristiansand/@58.1504318,8.0169354,14.25z/data=!4m5!3m4!1s0x46380258d5607a5b:0xdf0c0d6fc81c58a4!8m2!3d58.1599119!4d8.0182064>
- Greene, W. H. (2003). *Econometric analysis* Pearson Education India.
- Gripsrud, G., Olsson, U. H. & Silkoset, R. (2017). *Metode og dataanalyse* (3. utgave. utg.). Oslo: Cappelen Damm Akademisk.
- Hellevik, O. (2011). *Forskningsmetode i sosiologi og statsvitenskap* (7. utg.). Oslo: Universitetsforlaget.
- Jungeilges, J. (2018). *Forelesningsnotater og power points fra emnet SE-414: Econometrics for finance*.
- Kjetil Reite. (2018, 27.april). Omstridt ordning ga rekordstor folkevekst. *Fædrelandsvennen*, s. 11.
- Knight, J. R. (2002). Listing price, time on market, and ultimate selling price: Causes and effects of listing price changes. *Real Estate Economics*, 30(2), 213-237.
- Kristiansand Kommune. (2017). Bolig, boligmasse, boligbygging. Hentet fra <https://www.kristiansand.kommune.no/politikk-og-administrasjon/om-kristiansand/Statistikkportalen/bearbeidet-statistikk/bolig-boligbygging>

- Kristiansand Kommune. (2018a). *Boligprogram 2019 - 2022*. Hentet fra <https://www.kristiansand.kommune.no/contentassets/556119e9149745189974f0e3961f11e7/boligprogram-2019-2022.pdf>
- Kristiansand Kommune. (2018b). Folkemengde. Hentet fra <http://stat.kristiansand.kommune.no/portal/?locale=no#LandingPlace:r=2742700655167912710>
- Kristiansand Kommune. (2018c). Befolkning Hentet fra <https://www.kristiansand.kommune.no/politikk-og-administrasjon/om-kristiansand/Statistikkportalen/bearbeidet-statistikk/befolkning/>
- Krogsveen. (2018). Å kjøpe bolig i "kjøpers marked" - hvordan går man frem. Hentet fra <https://krogsveen.no/Boligtips/Trude-Larsen/Kjoepe-bolig-i-kjoepers-marked>
- Larsen, T. & Nordheim, S. (2018). *Studentenes preferanser for boligområder i Kristiansand* (Mastergradsavhandling). Universitetet i Agder, Kristiansand.
- Lindvold, E. (2018). Bolig under prisantydning - ikke fortell meklerin hva som er smertegrensen din. Hentet fra <https://www.dinside.no/bolig/ikke-fortell-mekleren-hva-som-er-smertegrensen-din/69578266>
- Lindvoll, E. (2018). Dette kan heve eller senke boligens verdi med opp mot én million kroner. Hentet fra <https://www.dinside.no/bolig/dette-kan-heve-eller-senke-boligens-verdi-med-opp-mot-n-million-kroner/69860761>
- MittBjørndalen. (2019). Salgstrinn. Hentet fra <https://www.mittbjoerndalen.no/nav/salgstrinn>
- Myhre, A., Bugge, S. & Nordstrøm, J. (2013, 23.04.2019). Mange unge har mer enn 4 mill. i gjeld. *E24*. Hentet fra <https://e24.no/privat/penger/mange-unge-har-mer-enn-4-mill-i-gjeld/10084373>
- NOU 2002:2. (2002). *Boligmarkedene og boligpolitikken*. Hentet fra <https://www.regjeringen.no/contentassets/80899d9e55ef499c86359694e816207f/no/pdfa/nou200220020002000dddpdfa.pdf>
- Osland, L. (2001). Den hedonistiske metoden og estimering av attributtpriser. *Norsk økonomisk tidsskrift*, 115, 22.
- Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of political economy*, 82(1), 34-55.
- Rudberg, K. (2018, 4.12.2018). Flere flytter til byen - færre bor spredt. *Fædrelandsvennen* s. 7.
- Schnare, A. B. & Struyk, R. (1976). Segmentation in urban housing markets. *Journal of Urban Economics*, 3(2), 146-166.
- Schultz, J. & Kaspersen, L. (2016, 11.4.2019). 1 av 3 selger boligen med tap. *Dagens Næringsliv*. Hentet fra <https://www.dn.no/boligmarkedet/eiendom/sparebank-1-gruppen/boligpriser/1-av-3-selger-boligen-med-tap/1-1-5767150>
- Sekaran, U. & Bougie, R. (2016). *Research Methods for Business* (7.utgave. utg.). Chichester Wiley.
- Smarte penger. (2018). Bør du eie eller leie bolig. Hentet fra <https://www.smartepenger.no/kalkulatorer/901-eie-eller-leie-kalkulator>
- Statistisk Sentralbyrå. (2018a). *Boliger*. Hentet fra <https://www.ssb.no/bygg-bolig-og-eiendom/statistikker/boligstat>
- Statistisk Sentralbyrå. (2018b). *Befolkning*. Hentet fra <https://www.ssb.no/befolkning/statistikker/folkemengde>
- Statistisk Sentralbyrå. (2019a). *Fakta om Bolig*. Hentet fra <https://www.ssb.no/bygg-bolig-og-eiendom/faktaside/bolig>
- Statistisk Sentralbyrå. (2019d). *Arbeid og lønn* [statistikk]. Hentet fra <https://www.ssb.no/arbeid-og-lonn>

- Theisen, T. (2018). *Forelesningsnotater og Powerpoints fra emnet BE-409: Real Estate Economics*.
- Tobienborg hage. (2019). Leilighetene. Hentet fra <http://tobienborghage.no/?menuid=3>
- Trading Economics. (2019). Homeownership rate | Europe. Hentet fra <https://tradingeconomics.com/country-list/home-ownership-rate?continent=europe>
- Verbeek, M. (2008). *A guide to modern econometrics* John Wiley & Sons.

Vedlegg

Vedlegg 1: Refleksjonsnotat

Dette refleksjonsnotatet er skrevet i forbindelse med min mastergradsavhandling i økonomi og administrasjon, med spesialisering i finansiell økonomi ved Universitetet i Agder. Som en avsluttende del av masteroppgaven skal jeg reflektere rundt begrepene internasjonalisering, innovasjon og samfunnsansvar og knytte dette til kunnskapen jeg har tilegnet meg gjennom masterstudiet og tema på masteroppgaven.

Jeg har skrevet masteroppgaven alene, og jeg ser både positive og negative sider ved å skrive oppgaven alene. På den positive siden sto jeg helt fritt til å velge tema, dette førte til at jeg kunne velge det temaet jeg var aller mest interessert i. Det å ikke måtte forholde seg til noen, og jobbe når det passet meg best opplevde jeg også som positivt. Det negative ved å skrive alene er at jeg ikke har noen å diskutere valg i oppgaven med. Dette førte blant annet til at jeg brukte veldig lang tid på å formulere den rette problemstillingen, selv om oppgaven var definert og avgrenset.

Hovedtema for min masteroppgave har vært prisutviklingen i bydelen Lund. Jeg synes dette var et veldig interessant tema fordi jeg i løpet av studietiden min har sett hvordan den ene blokken etter den andre har blitt ferdigstilt på Lund. I tillegg hadde jeg før jeg begynte å skrive oppgaven kunnskap om at det finnes leiligheter som selges med tap i bydelen. Dette trigget meg til å finne ut av hvordan prisutviklingen i bydelen var og om det var noen fellestrekk ved de leiligheter som selges med tap. Jeg har alltid vært interessert i eiendom, men det var emnet BE-409 Real Estate Economics som inspirerte meg til å skrive oppgave innenfor temaet eiendomsøkonomi.

I tillegg til at jeg synes eiendomsøkonomi er spennende, ønsket jeg å bruke analyseteknikker lært i emnene SE-414 Econometrics for Finance, SE-506 Advanced Econometrics for Finance og BE-506 Computational Finance and Portfolio Management, og da passet det bra med en kvantitativ analyse. Disse tre emnene har gitt meg mye kunnskap om dataanalyser og grunnleggende ferdigheter i Stata og R som jeg nå tar med meg inn i en analytikerstilling i bank.

Resultatene viser at en økning i leilighetens størrelse fører til en økning i leilighetens totalpris, og dette er i tråd med antakelser og konsumentteori. En større leilighet har plass til flere mennesker, det er følgelig naturlig å anta at en større leilighet er dyrere enn en mindre.

Leilighetens totalpris reduseres ved en økning i avstand til sentrum, og når leilighetens alder øker. Analysen avdekket også at eierformen selveier øker leilighetsprisen, den samme effekten ser vi av at det er heis i blokken.

Om lag en av fem leiligheter i bydelen Lund er solgt med tap i undersøkelsesperioden. Dersom vi bare ser på leiligheter som er nyere enn ti år er 1 av fire leiligheter solgt med tap. T-testen indikerte også at gjennomsnittlig leilighetsalder var lavere for leiligheter som selges med tap. Leiligheter som selges med tap har også mye høyere gjennomsnittlig omsetningshastighet enn leiligheter som selges med gevinst, samt at avviket mellom pris og prisantydning er negativt for de leiligheter som er solgt med tap.

Videre i dette refleksjonsnotatet vil jeg nå reflektere rundt de tre begrepene internasjonalisering, innovasjon og samfunnsansvar og knytte dette til tema på min masteroppgave og masterprogrammet som en helhet.

Internasjonalisering

Min oppgave omhandler boligmarkedet i en bydel i Kristiansand. Boligprisene i Kristiansand påvirkes av internasjonale krefter.

Da Lehman Brothers gikk konkurs og finanskrisen var et faktum påvirket dette styringsrenten i Norge. Styringsrenten er sentralbankens virkemiddel for å nå inflasjonsmålet. Når styringsrenten settes ned vil renten på boliglånet reduseres, og det vil bli billigere å låne penger. Dermed kan følgene av konkurser i utlandet være at det blir gunstigere å ta opp boliglån i Norge og etterspørselen etter bolig kan øke.

I mars 2019 hevet Norges Bank styringsrenten fra 0,75% til 1%. I makroøkonomi lærte jeg at dersom Norges Bank øker styringsrenten vil det bli dyrere for bankene å låne penger, da det fører til at bankene øker sine utlånsrenter. Det vil føre til at renten på boliglånet øker, og det blir dyrere å låne penger. En av følgene kan være at færre tar opp boliglån og etterspørselen etter bolig kan avta. I denne analysen er ikke den siste renteøkningen fanget opp, da undersøkelsesperioden stopper før renten økte til 1%. Det er heller ikke undersøkt hvordan endringen i styringsrenten fra 0,5% til 1% i 2018 påvirket boligmarkedet.

I 2014 sank oljeprisene betraktelig, det førte til at arbeidsledigheten i Norge økte fordi mange var ansatt i oljesektoren. Kristiansand ble relativt hardt rammet av dette. I resultatene ser vi at andelen leiligheter solgt med tap er størst i 2014 og 2015. Dette kan være en ringvirkning av at oljeprisen sank og arbeidsledigheten økte i Kristiansand. Årsakene til at oljeprisen sank

hevdes å være at det produksjonen i Libya økte, samt at store økonomier som Storbritannia og Kina nedjusterte forventningene om økonomisk vekst i 2015.

Innovasjon

Jeg ønsket å skrive en oppgave med et tema som var litt utenfor boksen. Det er mange analyser og masteroppgaver om hvilke attributter som påvirker boligprisen, men det er svært få som uttaler seg om at det finnes en risiko for at boliger selges med tap. Ved å se på de boligene som er solgt med tap, og undersøke forskjeller mellom boliger som er solgt med tap og gevinst ytterligere føler jeg at jeg har tatt inn et nyskapende element i min oppgave.

Også i boligmarkedet søker man å være nyskapende og komme opp med løsninger som gir mer nytte enn tidligere. Et eksempel på innovasjon i boligmarkedet er at det finnes nye boligblokker, for eksempel Tobienborg hage, som har solcellepaneler på taket som skal dekke strømkostnadene i fellesarealene.

Samfunnsansvar

Samfunnsansvar forbindes med å ta sosiale og miljømessige hensyn, over det som er bestemt i loven. Kjøp av bolig er en stor investering og det er derfor viktig at alle aktører i boligmarkedet opptrer på en etisk riktig måte. Viktig informasjon om boligen bør ikke holdes tilbake, og når man kjøper bolig basert på prospekt er det viktig at utbygger tilfredsstillende egenskaper som er beskrevet i prospektet.

For meg har det vært viktig å prøve å anonymisere de blokkene som er mest utsatt for tap, da det er en mulighet for at potensielle kjøpere i fremtiden kan bruke slike analyser til egen vinning. Slike analyser kan også påvirke boligsalg i fremtiden, og derfor er det viktig å være bevisst på at det nødvendigvis ikke skyldes egenskaper ved selve blokken, men heller at den første salgsprisen kan ha vært for høy i forhold til markedet. Selv om jeg har delt bydelen inn i områder, gis det i oppgaven ingen presis definisjon av områdene, bare et kart som er omtrentlig tegnet. Jeg har bevisst heller ikke utlevert noen adresser i oppgaven, og datamaterialet blir fjernet fra min pc når masteroppgaven leveres.

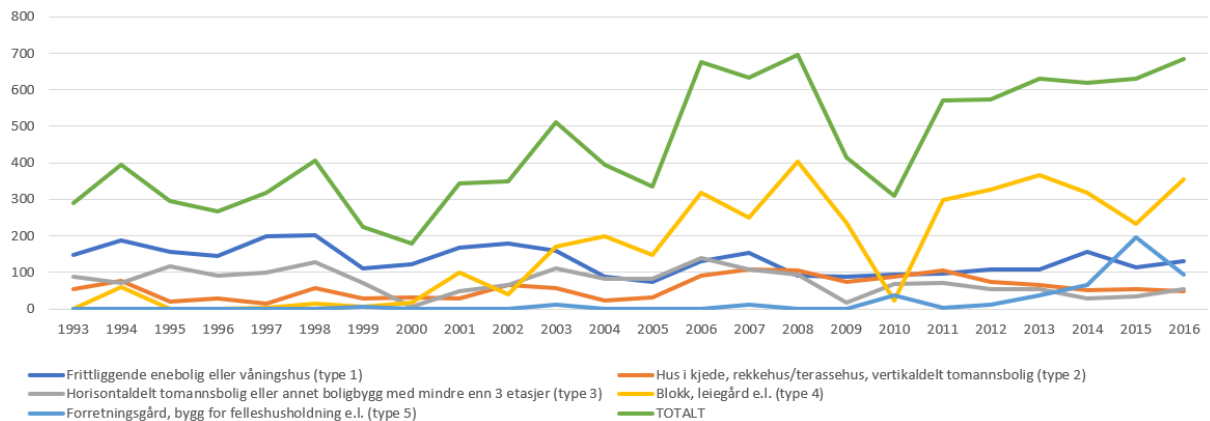
Vedlegg 2: Adresser på 4633 som tilhører bydelen Lund

Bjørndalsheia

Vollebakken

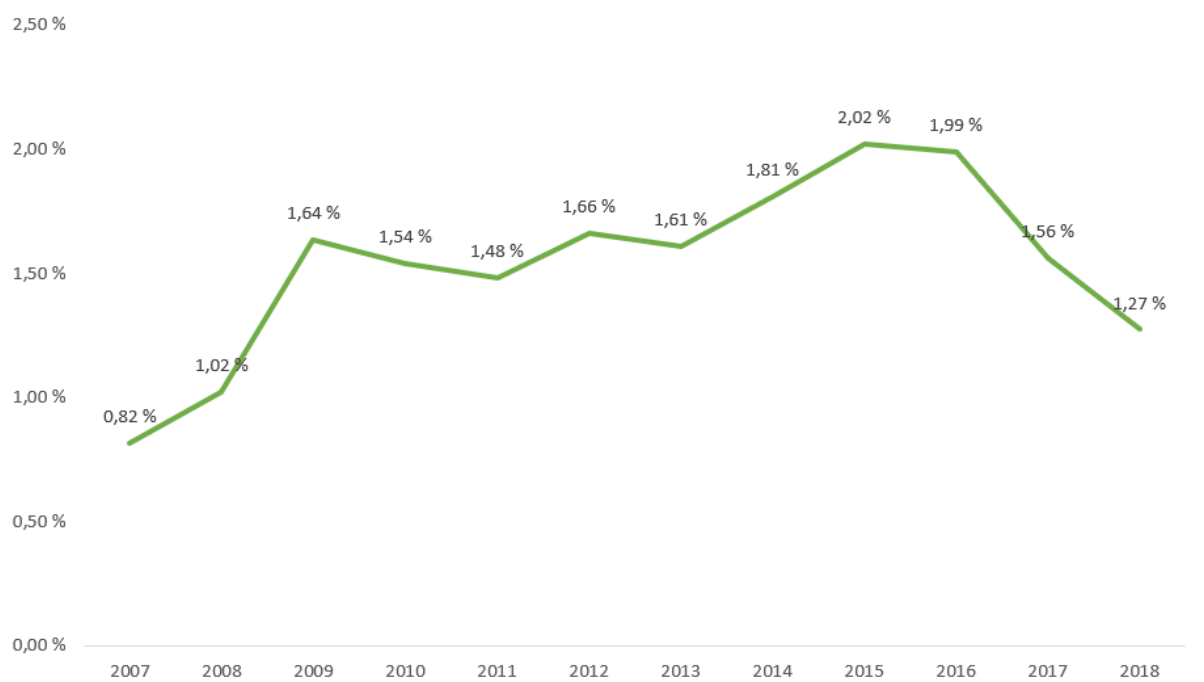
Østerveien

Vedlegg 3: Utvikling i boligbygging i Kristiansand, sortert etter boligtype.



Økningen i «forretningsgård, bygg for felleshusholdning» i 2015 skyldes at det ble bygd 182 nye studentboliger.

Vedlegg 4: Arbeidsledighet



Vedlegg 5: Regresjonsmodeller for små og store leiligheter

Leiligheter mindre og større enn medianen

Source	SS	df	MS	Number of obs	=	900
Model	62.7153777	29	2.16259923	F(29, 870)	=	150.17
Residual	12.5286851	870	.014400787	Prob > F	=	0.0000
				R-squared	=	0.8335
				Adj R-squared	=	0.8279
Total	75.2440628	899	.083697511	Root MSE	=	.12

lnP	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
lnM	.7474665	.013856	53.95	0.000	.7202714 .7746616
lnY	-.0922145	.0098173	-9.39	0.000	-.1114829 -.0729462
lnkm	.0056114	.0260233	0.22	0.829	-.0454644 .0566872
lnE	.0230225	.0081665	2.82	0.005	.0069942 .0390508
G	.0165249	.0171131	0.97	0.334	-.0170628 .0501126
H	.0292897	.0232015	1.26	0.207	-.0162478 .0748272
BD	-.0590564	.0177414	-3.33	0.001	-.0938774 -.0242353
SD	.0384546	.013543	2.84	0.005	.0118739 .0650352
år2011	.0870596	.0171903	5.06	0.000	.0533202 .120799
år2012	.0637922	.0171128	3.73	0.000	.0302049 .0973795
år2013	.0745483	.0172861	4.31	0.000	.0406208 .1084757
år2014	.0529103	.0176211	3.00	0.003	.0183256 .0874951
år2015	.1080001	.0175871	6.14	0.000	.0734819 .1425183
år2016	.134685	.016307	8.26	0.000	.1026793 .1666908
år2017	.1627336	.0176528	9.22	0.000	.1280866 .1973807
år2018	.1085755	.016875	6.43	0.000	.075455 .141696
L1	-.1234148	.035801	-3.45	0.001	-.1936812 -.0531484
L2	-.015292	.0311116	-0.49	0.623	-.0763546 .0457706
L3	-.0190511	.0277708	-0.69	0.493	-.0735567 .0354545
L4	.0137883	.0193368	0.71	0.476	-.024164 .0517406
L5	.0003053	.0351587	0.01	0.993	-.0687005 .0693111
L6	.0090673	.0413087	0.22	0.826	-.072009 .0901437
L7	-.0241853	.0194669	-1.24	0.214	-.0623929 .0140223
L8	.0151105	.0213314	0.71	0.479	-.0267565 .0569774
L9	.0665994	.0323446	2.06	0.040	.003117 .1300819
L11	.142598	.0239005	5.97	0.000	.0956886 .1895074
L12	.0428881	.0208315	2.06	0.040	.0020022 .0837739
L13	.0072642	.0300273	0.24	0.809	-.0516703 .0661986
L14	-.0987259	.0314756	-3.14	0.002	-.1605028 -.0369489
_cons	11.80927	.0728675	162.06	0.000	11.66625 11.95228

Linear regression						Number of obs	=	748
						F(29, 718)	=	120.36
						Prob > F	=	0.0000
						R-squared	=	0.8356
						Root MSE	=	.15308
lnP	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]			
lnM	.8980946	.0340262	26.39	0.000	.8312919	.9648973		
lnY	-.0935025	.011111	-8.42	0.000	-.1153145	-.0716905		
lnkm	-.1618215	.0417421	-3.88	0.000	-.2437726	-.0798705		
lnE	.0503829	.01167	4.32	0.000	.0274715	.0732942		
G	-.0142205	.0176481	-0.81	0.421	-.0488686	.0204276		
H	.0653041	.0211578	3.09	0.002	.0237655	.1068427		
BD	.0402481	.0215429	1.87	0.062	-.0020466	.0825427		
SD	.0481908	.018305	2.63	0.009	.012253	.0841286		
år2011	.000702	.0257647	0.03	0.978	-.0498812	.0512852		
år2012	.0232398	.0261943	0.89	0.375	-.0281867	.0746663		
år2013	.0223064	.0264991	0.84	0.400	-.0297187	.0743314		
år2014	.0244632	.0272219	0.90	0.369	-.0289808	.0779072		
år2015	.0527175	.027608	1.91	0.057	-.0014846	.1069197		
år2016	.1195836	.0267036	4.48	0.000	.0671572	.1720101		
år2017	.0958698	.0260846	3.68	0.000	.0446585	.1470811		
år2018	.1795928	.0257227	6.98	0.000	.1290921	.2300935		
L1	-.068729	.033891	-2.03	0.043	-.1352664	-.0021916		
L2	-.0161635	.0447552	-0.36	0.718	-.1040301	.0717031		
L3	.1983933	.0294813	6.73	0.000	.1405134	.2562731		
L4	.0103672	.0271163	0.38	0.702	-.0428695	.0636038		
L5	.1997148	.0455476	4.38	0.000	.1102923	.2891372		
L6	.0976535	.0465817	2.10	0.036	.0062009	.1891061		
L7	.0345108	.0277961	1.24	0.215	-.0200605	.0890821		
L8	-.018392	.0274071	-0.67	0.502	-.0721996	.0354156		
L9	-.1013084	.0265351	-3.82	0.000	-.153404	-.0492129		
L11	.3006081	.0257083	11.69	0.000	.2501356	.3510806		
L12	.0644072	.0338324	1.90	0.057	-.0020151	.1308295		
L13	.1740924	.0357785	4.87	0.000	.1038495	.2443353		
L14	.0394479	.0536982	0.73	0.463	-.0659764	.1448722		
_cons	11.17816	.1647225	67.86	0.000	10.85477	11.50156		

Leiligheter mindre og større enn 55 kvm

Source	SS	df	MS	Number of obs = 420		
Model	32.085611	28	1.14591468	F(28, 391) = 87.83		
Residual	5.10120193	391	.013046552	Prob > F = 0.0000		
				R-squared = 0.8628		
				Adj R-squared = 0.8530		
Total	37.1868129	419	.088751343	Root MSE = .11422		

lnP	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
lnM	.7834553	.0260007	30.13	0.000	.7323367	.8345739
lnY	-.0908975	.0175479	-5.18	0.000	-.1253975	-.0563975
lnkm	.1003321	.0523764	1.92	0.056	-.0026425	.2033067
lnE	.0412593	.0118103	3.49	0.001	.0180397	.0644788
G	-.0010915	.0270303	-0.04	0.968	-.0542345	.0520514
H	-.0107007	.0361926	-0.30	0.768	-.0818571	.0604556
BD	-.0323635	.0425268	-0.76	0.447	-.1159733	.0512463
SD	.0892664	.0279393	3.20	0.002	.0343364	.1441965
år2011	.0640437	.0237592	2.70	0.007	.017332	.1107554
år2012	.0514811	.0249487	2.06	0.040	.0024308	.1005314
år2013	.0484607	.0243842	1.99	0.048	.0005201	.0964012
år2014	.0659249	.0259392	2.54	0.011	.0149271	.1169227
år2015	.1325354	.0269706	4.91	0.000	.0795099	.1855609
år2016	.1194982	.0224681	5.32	0.000	.0753249	.1636716
år2017	.1615041	.0244188	6.61	0.000	.1134955	.2095127
år2018	.0342432	.023799	1.44	0.151	-.0125468	.0810332
L2	.0224144	.0482626	0.46	0.643	-.0724722	.117301
L3	-.0508069	.0511272	-0.99	0.321	-.1513256	.0497118
L4	-.027231	.0501209	-0.54	0.587	-.1257712	.0713091
L5	.0098542	.0514003	0.19	0.848	-.0912013	.1109097
L6	.031652	.0553748	0.57	0.568	-.0772175	.1405216
L7	-.0213018	.0388255	-0.55	0.584	-.0976347	.0550312
L8	.0516785	.0385639	1.34	0.181	-.0241399	.127497
L9	.0183116	.0403481	0.45	0.650	-.0610148	.097638
L11	.1280547	.0358401	3.57	0.000	.0575913	.1985181
L12	.0303112	.0342995	0.88	0.377	-.0371234	.0977458
L13	.0050758	.0394516	0.13	0.898	-.0724881	.0826397
L14	-.1631165	.0529175	-3.08	0.002	-.267155	-.059078
_cons	11.56579	.1262022	91.64	0.000	11.31767	11.81391

*Gimlevang/sødal er utelatt fordi det ikke finnes noen leiligheter som er mindre enn 55 kvm i området.

Linear regression		Number of obs		=		1,436	
		F(29, 1406)		=		164.47	
		Prob > F		=		0.0000	
		R-squared		=		0.8243	
		Root MSE		=		.13972	
lnP	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]		
lnM	.8265383	.0261399	31.62	0.000	.7752609	.8778156	
lnY	-.096109	.0090427	-10.63	0.000	-.1138476	-.0783705	
lnkm	-.15862	.0253024	-6.27	0.000	-.2082544	-.1089855	
lnE	.040676	.0076999	5.28	0.000	.0255715	.0557806	
G	-.0093956	.0136398	-0.69	0.491	-.0361522	.0173609	
H	.0494558	.0165617	2.99	0.003	.0169675	.0819441	
BD	.016164	.0142	1.14	0.255	-.0116914	.0440194	
SD	.018214	.0102766	1.77	0.077	-.001945	.038373	
år2011	.0411422	.0173515	2.37	0.018	.0071045	.0751798	
år2012	.0449034	.0159921	2.81	0.005	.0135325	.0762743	
år2013	.0521932	.0166354	3.14	0.002	.0195604	.0848261	
år2014	.0416711	.0164638	2.53	0.011	.0093748	.0739674	
år2015	.0809521	.0170594	4.75	0.000	.0474875	.1144167	
år2016	.1295352	.0166168	7.80	0.000	.0969389	.1621316	
år2017	.1220878	.0166072	7.35	0.000	.0895102	.1546654	
år2018	.1832469	.0162636	11.27	0.000	.1513432	.2151505	
L1	-.0663729	.0238441	-2.78	0.005	-.1131468	-.0195991	
L2	.0042211	.0350852	0.12	0.904	-.0646038	.073046	
L3	.1854119	.025451	7.29	0.000	.135486	.2353379	
L4	.039614	.0130854	3.03	0.003	.013945	.0652829	
L5	.0852296	.0318832	2.67	0.008	.0226858	.1477735	
L6	.0899396	.0384067	2.34	0.019	.0145989	.1652802	
L7	.0204899	.0161089	1.27	0.204	-.0111102	.0520899	
L8	.0275856	.0192382	1.43	0.152	-.0101531	.0653243	
L9	-.0996382	.020024	-4.98	0.000	-.1389182	-.0603581	
L11	.2979026	.0219278	13.59	0.000	.2548879	.3409172	
L12	.0572149	.0198031	2.89	0.004	.0183681	.0960617	
L13	.1593844	.0292177	5.46	0.000	.1020694	.2166995	
L14	.0203823	.026422	0.77	0.441	-.0314484	.0722131	
_cons	11.54439	.1200167	96.19	0.000	11.30896	11.77982	

Vedlegg 6: Stata-kommandoer

Datsett med alle observasjoner inkludert

```
//Klargjøring av datsett
destring Omsetningshastighet, generate(omsetningshastighet_num) ignore('»d»', illegal)
tabulate Adresse
drop if Prom == 0
drop if Pris == . //fjerner missing values for Pris
drop if omsetningshastighet_num == .
drop BA
drop if byggeår==0
//endre navn på variabler og lables
rename m2prom kvadratmeterpris
label variable kvadratmeterpris «Kvadratmeterpris»
rename kvadratmeterpris kvmp
rename omsetningshastighet_num Ohast
label variable Pris «Omsetningspris»
rename Prisantydning prisant
label variable Prom «Primærrom»
label variable Fellesgjeld «Fellesgjeld»
rename boligalder alder
label variable alder «Boligalder»
rename Prom M
label variable M «Primærrom»
rename Pris P
rename Fellesgjeld F
rename Soverom R
rename Etasje E
rename omsetningshyppighet20102018 Z
label variable Z «Antall ganger omsatt i perioden»
rename Parkering G
rename Heis H
```

```

rename Blokk1 Blokk
label variable Blokk «Blokk_dummy»
rename Blokk BlokkD
label variable Y «Boligalder»
label variable nrnybygg «Antall nybygg»
rename nrnybygg N
drop Y
rename Boligalder Y
label variable SelveierD «Selveier_dummy»
rename Storeleil StoreleilD
label variable StoreleilD «Storeleil_dummy»
replace Z = 1 in 1258
rename Salgsår S
rename SelveierD SD
rename StoreleilD SL
//omkodingoding
gen boligalder = 2019- byggeår //genererer en variabel som angir boligens alder i dag
gen Y = S- Byggeår // boligens alder på kjøpstidspunkt
encode Eierform, generate(Eierform_num) // borettslag=1, selveier =2
gen selveierD = Eierform_num==2 // dummy for selveier
gen borettslagD = Eierform_num==1 //dummy for borettslag
replace Fellesgjeld = 0 if missing(Fellesgjeld) //erstatter missing values med 0
gen totalpris = Pris + Fellesgjeld

//Datarensing

summarize Prom
summarize omsetningshastighet_num
summarize Y
summarize P
// Sjekk for outliers
graph box Y
graph box P
graph box M

```

```

graph box km
//Lage dummyvariabler for hvert salgsår
gen år2010 = S==2010
gen år2011 = S==2011
gen år2012 = S==2012
gen år2013 = S==2013
gen år2014 = S==2014
gen år2015 = S==2015
gen år2016 = S==2016
gen år2017 = S==2017
gen år2018 = S==2018
//Deskriptiv statistikk
summarize P M F Y R E N Z A G H SD BD
histogram P, width(300000) percent addlabel normal
histogram M , width(5) percent normal
histogram Z, width(1) percent normal
//splitte adresse for å bruke til georoute
split Adresse, p(,)
split Adresse2, p( )
drop Adresse21
split Adresse2, p(« «)
drop Adresse23
rename Adresse1 gatenr
drop Adresse2
rename Adresse21 zip
rename Adresse22 krs
gen markens = «Markens gate 19» //sentrum
gen zipM = «4611»
gen land = «Norway»
//Bytte ut bokstaver georoute ikke kan lese
gen gatenr1 = usubinstr(gatenr, «Ø»,»O»,.)
gen gatenr2 = usubinstr(gatenr1,»é»,»e»,.)
gen hus = usubinstr(gatenr2, «å», «a»,.)

```

```

drop gatenr
drop gatenr1
drop gatenr2
//Beregne avstand til sentrum

georoute, hereid(*****) herecode(*****) startad(hus zip krs land) endad(markens zipM krs
land) km di(dist) ti(time)

*hereid og herecode er skjult da dette er private koder

//Valg av alternative variabler og estimering av regresjonsmodell

regress P M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017 år2018
L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

predict resL,r

regress P R Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017 år2018
L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress lnP R Y km E G H BD SD SL1 år2011 år2012 år2013 år2014 år2015 år2016 år2017
år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress lnP M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017
år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

predict resS,r

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

predict resD, r

regress lnP lnR lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress P M Y km E G H BD SD S L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress P R Y km E G H BD SD S L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress lnP M Y km E G H BD SD S L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

regress lnP R Y km E G H BD SD S L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

```

```
regress lnP lnM lnY lnkm lnE G H BD SD lnS L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13  
L14
```

```
regress lnP lnR lnY lnkm lnE G H BD SD lnS L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13  
L14
```

```
//test av mindre utvalg
```

```
gen n =_n
```

```
keep if !mod(_n,6)
```

```
regress P M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017 år2018
```

```
regress lnP M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017  
år2018
```

```
regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016  
år2017 år2018
```

```
//Test av forutsetninger hele populasjon
```

```
regress P M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017 år2018  
L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14
```

```
estat hettest
```

```
regress P M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017 år2018  
L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14, robust
```

```
predict resL1, r
```

```
pnorm resL1
```

```
regress lnP M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017  
år2018 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14
```

```
estat hettest
```

```
vif
```

```
regress lnP M Y km E G H BD SD år2011 år2012 år2013 år2014 år2015 år2016 år2017  
år2018 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14, robust
```

```
predict resS1,r
```

```
vif
```

```
pnorm resS1
```

sktest resS1

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14

estat hettest

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L10 L11 L12 L13 L14, robust

predict yest,xb

predict resD2, r

vif

pnorm resD2

sktest resD2

//store og små leiligheter

summarize M, d

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M<55

estat hettest

predict res1,r

pnorm res1

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M>55

estat hettest

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M>55, robust

predict res2,r

pnorm res2

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M<68

estat hettest

predict res3,r


```
pnorm res3
```

```
regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016  
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M>68
```

```
estat hettest
```

```
regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016  
år2017 år2018 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14 if M>68, robust
```

```
predict res4,r
```

```
pnorm res4
```

```
summarize M if M<55,d
```

```
summarize M if M>55,d
```

```
summarize M if M<68,d
```

```
summarize M if M>68,d
```

```
//Beregne gjennomsnittlig pris per år
```

```
summarize P if år2010==1
```

```
summarize P if år2011==1
```

```
summarize P if år2012==1
```

```
summarize P if år2013==1
```

```
summarize P if år2014==1
```

```
summarize P if år2015==1
```

```
summarize P if år2016==1
```

```
summarize P if år2017==1
```

```
summarize P if år2018==1
```

```
//interaksjonsmodell
```

```
gen dummy_S1 = Z==1
```

```
replace dummy_S1 = 0 if Z>1
```

```
gen S1km = dummy_S1*km
```

```
gen S1M = dummy_S1*km
```

```
gen S1E = dummy_S1*E
```

```
gen S1Y = dummy_S1*Y
```

```
regress P dummy_S1 M S1M km S1km Y S1Y E S1E, robust
```

Datsett med bare leiligheter omsatt mer enn en gang inkludert

```
//Klargjøring
rename EndringP C
drop lnP
rename gevinststapnull gtn
rename antallårmellomsalgene t
gen SL1 = 1 if M>69 // dummy leiligheter større enn gjennomsnittet
replace SL1 = 0
replace SL1 = 1 if M>69
drop nyeleiligheter
gen nyeleiligheter = 1 if Y<10
replace nyeleiligheter = 0 if Y>10
replace nyeleiligheter = 1 if Y==10
// deskriptiv statistikk
tabulate S gtn
tabulate SL1 gtn
tabulate nyeleiligheter gtn
destring Omsetningshastighet, generate(Omsetningshastighet_num) ignore('»d»', illegal)
replace L1=0 if (L1>=.)
replace L2=0 if (L2>=.)
replace L3=0 if (L3>=.)
replace L4=0 if (L4>=.)
replace L5=0 if (L5>=.)
replace L6=0 if (L6>=.)
replace L7=0 if (L7>=.)
replace L8=0 if (L8>=.)
replace L9=0 if (L9>=.)
replace L10=0 if (L10>=.)
replace L11=0 if (L11>=.)
replace L12=0 if (L12>=.)
```

```

replace L13=0 if (L13>=.)
replace L14=0 if (L14>=.)
gen avvik = p-Prisantydning
label variable avvik «Avvik»
rename avvik d
summarize P,d
summarize M,d
summarize C,d
summarize H
summarize G
summarize Y,d
summarize R,d
summarize E,d
summarize tap
summarize km,d
summarize SD
summarize BD
summarize Z,d
summarize Omsetningshastighet_num,d
//frekvenstabell, tap per år
tabulate gtn, subpop(år2011)
tabulate gtn, subpop(år2012)
tabulate gtn, subpop(år2013)
tabulate gtn, subpop(år2014)
tabulate gtn, subpop(år2015)
tabulate gtn, subpop(år2016)
tabulate gtn, subpop(år2017)
tabulate gtn, subpop(år2018)
//figurer
histogram P if gtn==1, width(500000) percent
histogram P if gtn==2, width(500000) percent
histogram Y1 if gtn==1, width(5) percent
histogram Y if gtn==2, width(5) percent

```

```

histogram M if gtn==1, width(10) percent
histogram M if gtn==2, width(10) percent
histogram d if gtn==1, width(50000) percent
histogram d if gtn==2, width(50000) percent
histogram R if gtn==1, percent width(1)
histogram R if gtn==2, percent width(1)
histogram t if gtn==1, percent width(1)
histogram t if gtn==2, percent width(1)
tabulate Adresse gtn
//analyse
gen lnP = ln(P)
gen lnkm=ln(km)

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 if gtn==2

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 if gtn==1

ttest P, by (gtn)
ttest Y, by (gtn)
ttest km, by (gtn)
ttest M, by (gtn)
ttest R, by (gtn)
ttest Omsetningshastighet_num, by (gtn)
ttest d, by(gtn)
// Bergene avstand til UiA
georoute, hereid(*****) herecode(*****) startad(hus zip krs land) endad(uiia zipuia krs land)
km di(dist) ti(time)

// Regresjon for leiligheter solgt mer enn en gang
regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14

estat hettest

```

```

regress lnP lnM lnY lnkm lnE G H BD SD år2011 år2012 år2013 år2014 år2015 år2016
år2017 år2018 L1 L2 L3 L4 L5 L6 L7 L8 L9 L11 L12 L13 L14, robust

predict res1,r

pnorm res1

//interaksjonsmodell

gen tap =1 if gtn==1

replace tap=0 if gtn==2

gen tapM = tap*M

gen tapY = tap*Y

gen tapkm = tap*km

gen tapE = tap*E

gen ny = Byggeår>=2010

gen tapny = tap*ny

regress P tap tapM M tapY Y tapkm km tapE E tapny ny,robust

```