

Gir garasjeleilighet merverdi til eneboligen?

Simen Gården Aanjesen

Veileder

Theis Theisen

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

Universitetet i Agder, 2015

Fakultet for økonomi og samfunnsvitenskap

Institutt for økonomi / Handelshøyskolen i Kristiansand

FORORD

Denne oppgaven er skrevet som et ledd i masterutdanningen i økonomi og administrasjon, med fordypning i økonomisk styring og prosjektledelse, ved Universitetet i Agder. Oppgaven utgjør 30 studiepoeng.

Arbeidet med denne oppgaven har vært omfattende og særdeles lærerikt. Jeg valgte å skrive om garasjeleiligheters påvirkning på boligpris, fordi jeg interesserer meg for eiendomsøkonomi, samt at utleie av deler av egen eiendomsmasse er svært utbredt.

Jeg kan ikke få fullrost min veileder, professor Theis Theisen, nok. Han har hjulpet meg med god veiledning, konstruktiv tilbakemelding og ikke minst rask respons.

Simen Gården Aanjesen

Kristiansand, 21 Mai 2015.

INNHALDSFORTEGNELSE

Forord	1
Innholdsfortegnelse	2
Figuroversikt	4
Tabelloversikt	5
Vedleggsoversikt	6
Sammendrag	7
1.0 Innledning	8
2.0 Bakgrunn.....	8
2.1 Regelverk	8
3.0 Teori	10
3.1 Innledning	10
3.2 Prisdannelse i eiermarkedet	11
3.3 Alonso-Muth-mills modellen	13
3.4 Den hedonistiske prisfunksjonen	15
3.5 Garasjebygget	17
3.6 Hypoteser	20
4.0 Økonometrisk modell	21
4.1 Innledning.....	21
4.2 Regresjonsanalysens forutsetninger	22
4.3 Regresjon	23
5.0 Datainnsamling og beskrivelse av datamaterialet	26
5.1 Valg av enheter og datainnsamling	26
5.2 Datarensing	27
6.0 Presentasjon av datamaterialet	S.29

6.1 Deskriptiv statistikk	29
6.2 Utvidet deskriptiv statistikk	30
6.3 Frekvenstabeller for utvalgte variable	32
6.4 Korrelasjon	35
7.0 Estimering og testing av hypoteser.....	37
7.1 Innledning	37
7.2 Multippel lineær regresjon	38
7.3 Dobbel logaritmisk regresjon	44
7.4 Semi-logaritmisk regresjon	46
7.5 Valg av modell	48
7.6 Testing av hypoteser	49
8.0 Nærmere drøfting av hovedproblemstilling	50
8.1 Drøfting	50
8.2 Eksempel på gjennomføring av modell	50
8.3 Svakheter ved analysene	51
9.0 Konklusjon	52
10.0 Litteraturliste	53

FIGUROVERSIKT

Figur 3.1 Samlet tilbud og samlet etterspørsel i markedet for brukte boliger. Kort sikt	12
Figur 3.2 Endringer i prisen ved økning i tilbudet av, og etterspørsel etter boliger	13
Figur 3.3 Husleiegradienten.....	15
Figur 3.4 Hedonistisk prisfunksjon	16
Figur 3.5 Hedonistisk prisfunksjon, to attributter.....	17
Figur 4.1 En dummyvariabels effekt på regresjonslinjen og konstantleddet	25
Figur 6.1 Kurtosis, negativ verdi. Kortere haler og spiss/smål fordeling	32
Figur 6.2 Skewness, negativ verdi. Hale til venstre	32
Figur 7.1 Normalskråplott, test av restledd i multippel lineær regresjon	43
Figur 7.2 Histogram, test av restledd i multippel lineær regresjon	43
Figur 7.3 Histogram med normalfordeling i dobbel log regresjon	45
Figur 7.4 Normalskråplott i dobbel log regresjon	45
Figur 7.5 Histogram med normalfordeling i semi log regresjon	47
Figur 7.6 Normalskråplott i semi log regresjon	47

TABELLOVERSIKT

Tabell 5.1 Variablene	28
Tabell 6.1 Deskriptiv statistikk over tallmaterialet	29
Tabell 6.2 Utvidet statistikk for pris, boa/P-rom og byggeår	31
Tabell 6.3 Frekvenstabell for salgpris	33
Tabell 6.4 Frekvenstabell for byggeår	34
Tabell 6.5 Frekvenstabell for garasjeleiligheter	34
Tabell 6.6 Prosentvis fordeling av enheter fordelt mellom byene	35
Tabell 6.7 Bivariat korrelasjonsanalyse	37
Tabell 7.1 Estimatene ved multippel lineær regresjon	39
Tabell 7.2 VIF-test, lineær regresjon.....	40
Tabell 7.3 Estimatene ved multippel regresjon, ny modell.....	41
Tabell 7.4 VIF-test, lineær regresjon, ny modell.....	42
Tabell 7.5 Breusch-pagan test. Multippel lineær regresjon.....	42
Tabell 7.6 Lineær regresjon, uten postnummer-dummy	43
Tabell 7.7 Estimatene ved dobbel logaritmisk regresjon	44
Tabell 7.8 VIF test, dobbelt logaritmisk regresjon	45
Tabell 7.9 Breusch-Pagan test, dobbel log regresjon	45
Tabell 7.10 Estimer ved semi-logaritmisk regresjon	46
Tabell 7.11 VIF, semi-logaritmisk regresjon	47
Tabell 7.12 Breusch-Pagan test, semi-logaritmisk regresjon.....	47
Tabell 7.13 Semi-log regresjon, endelig modell.....	48

VEDLEGGSOVERSIKT

Vedlegg 1: Koder

Vedlegg 2: Endelig regresjonsform

Vedlegg 3: Tilbudstegning fra Helland Garasjer.

SAMMENDRAG

Denne oppgaven tar for seg en garasjeleilighets påvirkning på prisen til enebolig. Innledningsvis i oppgaven gir jeg litt bakgrunnsinformasjon om regelverket omkring oppføring av garasje.

Videre følger en kort og grunnleggende innføring i relevant teori, presentasjon av Alonso-Muth-Mills modellen, hedonistiske prisfunksjoner og en enkel analyse av kostnad og inntekt tilknyttet garasjeleilighet. Hypoteser blir deretter formulert på bakgrunn av teorien.

Deretter følger en beskrivelse av fremgangsmåten for datainnsamlingen og hvordan data har blitt bearbeidet. Etter bearbeiding av data, stod jeg igjen med 2298 eneboliger fra utvalgte byer i Norge, solgt over en periode på 26 måneder. Det er disse boligene som danner grunnlag for analyse, samt hypotesetestingen i oppgaven.

I analysedelen benytter jeg meg av ulike regresjonsmodeller, hvor valget til slutt falt på semi logaritmisk regresjonsmodell. Resultatene i denne regresjonsmodellen er i tråd med teorien jeg la frem. Hypotesene mine ble testet ved hjelp av regresjonen og problemstillingen ble besvart. Det konkluderes med at garasje med tilhørende leilighet, har positiv innvirkning på prisen.

1. INNLEDNING

De aller fleste av oss her i Norge vil på et tidspunkt kjøpe sin egen bopel. Dette er for mange av oss den største enkeltinvesteringen vi foretar i løpet av livet. Vi binder opp millioner av kroner i fast eiendom og betaler renter og avdrag i flere tiår. Vi pusser opp med tanke på å ha det fint rundt oss, samt at det øker annenhåndsverdien ved ett evt. Salg. Men vi kan også øke markedsverdien, samt redusere de løpende netto utgiftene, ved å leie ut en del av boligen. Dette er svært populært i Norge og spesielt i byer med stor etterspørsel etter leiligheter til leie.

Et særnorsk fenomen er å bygge hybel over garasjen. Det er blitt så populært at man i dag kan bestille bygg-moduler av garasjer med tilhørende hybel. Dette fikk meg til å undres om man får tilbake investeringen i form av merverdi ved salg av eiendommen. Temaet for oppgaven er således markedet for eiendommer med garasjeleilighet. Funnene i denne oppgaven mener jeg er av interesse for boligeiere som i praksis har tomteareal og ressurser til å gjennomføre en oppføring.

Problemstillingen er som følger; «Gir tilhørende garasjeleilighet merverdi til eneboligen ved salg?» Jeg vil også se på hvorvidt en eventuell merverdi, er lik investeringskostnaden.

2. BAKGRUNN

I denne oppgaven ser jeg på garasjer med tilhørende hybel i garasjebygget. Plan og bygningsloven stiller strenge krav for hva som kan regnes som hybel for utleie, samt for garasje i seg selv. I punktet under vil jeg nevne litt om det gjeldende regelverket.

2.1 REGELVERK

For oppføring av garasje er det ulike regler avhengig av garasjens størrelse, samt at man må skille mellom nasjonale og kommunale regler. I utgangspunktet er oppføring søknadspliktig med full byggesøknad, men under gitte forutsetninger er det unntakstilfeller. Det er f.eks mulig å føre opp små boder o.l under 15kvm i størrelse på tomten uten å melde kommunen, men det er ikke relevant for min oppgave og omtales således ikke.

Med dagens regelverk må man uansett søke kommunen for oppføring av garasjebygg. Dette regelverket har sittende regjering tatt høyde for å endre fra og med 1 Juli 2015.¹

Disse endringene vil bl.a omfatte at man kan bygge opptil 50m² garasje, minst 1 m fra naboens tomt. Men det er langt inn begrensninger på garasjens høyde, samt krav om at det ikke kan brukes til beboelse, jf forskrift om endring i byggesaksforskriften § 4.1 a.

Dagens regelverk:

Byggesøknad uten krav om ansvarsrett

Det gjeldende regelverket for garasje med en etasje og loft, forutsatt at loftet er under 190cm fra gulv til hanebjelke og garasjen i sin helhet er under 50m², sier at man kan bygge dersom man har samtykke fra nabo og garasjen er 1 m fra naboens tomt, eller 2 m fra naboens hus samt man har sendt inn søknad til kommunen og fått godkjenning på denne. Link til søknadspapirer ligger vedlagt i appendix.

Garasjer mellom 50m²-70m² har samme søknadsprosess, men det stilles strengere krav til avstand grunnet brannkrav. I dette sjiktet må garasjen plasseres 4 m fra naboens tomtegrense og 8m fra alle eksisterende bygg i nabolaget, inkludert egen bolig.

Full byggesøknad med krav om ansvarsrett

For garasjer større enn 70m² kan man som privatperson ikke søke om å oppføre. For å søke om dette må man ha ansvarsrett. Det er kun foretak med ansvarsrett som kan søke, prosjektere, utføre og kontrollere slike byggetiltak. Denne søknaden sendes til myndighetene. For å ha ansvarsrett må man kunne dokumentere tilstrekkelig kompetanse. Med andre ord må de aller fleste privatpersoner som ønsker å oppføre en garasje med tilhørende hybel på loft eller i kjeller, outsource alle ledd i byggeprosessen. Dette minimerer risiko for boligeier, da foretaket med ansvarsrett påtar seg ansvar ovenfor kommunen for å prosjektere og bygge ihht plan og bygningsloven, samt byggteknisk forskrift, kalt TEK 10. På den annen side er det en svært kostbar affære.

Plan og bygningsloven stiller også strenge krav til hva som kan defineres som hybel. Det stilles bl.a krav til tilfredsstillende lysforhold, isolasjon, ventilering, brannsikring, takhøyde, rømningsveier, bodplass og parkering. Uten å gå nærmere inn på dette, kan jeg slå fast at regelverket er så strengt, at brorparten av hyblene i mitt datagrunnlag ikke vil være regnet som hybel i lovens forstand. Dersom jeg skulle basert datagrunnlaget mitt på gjeldende regelverk, ville utvalget blitt svært begrenset. Av nevnte grunn må jeg legge en subjektiv vurdering til grunn for hva som skal med i mitt utvalg.

Jeg velger derfor å innlemme hybler som ikke er omsøkt, samt hybler som ikke er godkjent for varig opphold. Jeg begrenser kravene til at hybelen må tilhøre garasjebygget, den må ha minst to separate rom, hvorav et av rommene må være baderom med innlagt vann og avløp, samt egen inngang. Dette tilfredsstiller skatteetatens definisjon av hva som kan regnes som en selvstendig hybel.

3. TEORI

3.1 INNLEDNING

Boligmarkedet er et marked så godt som oss alle vil bli kjent med, enten som kjøper eller selger, utleier eller leietaker. Kjennetegn ved boligmarkedet kan oppsummeres i følgende punkter (Fra forelesningsfoilene i BE-409, fritt oversatt til norsk):

-Heterogenitet. Så godt som alle boliger er unike, med ulikt innhold og egenskaper. I tilfeller der boligene ser identiske ut ved første øyekast, kan de være ulikt innredet, men de er uansett ulike da de vil ha ulik plassering i terrenget.

-Immobilie. Boligen er bygd fast til grunnen, har eget bruks og gårdsnummer. Å flytte selve boligen fra et sted til et annet blir ikke gjort i praksis, da dette er tidkrevende og svært kostbart. Det finnes dog unntak. F.eks verneverdige boliger.

-Nødvendig/fundamentalt. Vi ønsker alle å ha tak over hodet, i ly for vær og vind.

-Varig konsumkapital. For de fleste den største enkeltinvesteringen vi foretar oss, vi holder boligen i eie over lengre perioder.

-Uelastisk tilbud på kort sikt. Tilgangen bestemmes av nybygging og avgang. Tilpassinger tar tid, noe som gjør tilbudet uelastisk på kort sikt.

-asymmetrisk informasjon.

-beliggenhet er en del av boligens verdi. Da boligtomt er et heterogent gode, vil prisen på boligen avhenge sterkt av fysisk lokalisering.

-store transaksjonskostnader. Å søke etter ny bolig tar tid og er kostbart. Disse kostnadene ligger på omtrent 8-10% av omsetningsverdien på bolig.

-Befolkningens sparing påvirkes av låneopptak. Fordi bolig er en så stor investering og vi nedbetaler lånet, får vi en vridning fra konsum til sparing. Vi må tilpasse bruken av midler deretter.

For bedre å kunne forstå hvordan prisen på boliger blir fastsatt, skal jeg nå presentere litt teori om emnet prisdannelse i boligmarkedet. Jeg vil starte med en enkel modell for prisdannelse i eiermarkedet. Vi utelater leiemarkedet i denne modellen, fordi vi kan skille mellom eie og leiemarked på kort sikt.

3.2 PRISDANNELSE I EIERMARKEDET

Først og fremst vil jeg nevne at denne modellen legger to antakelser til grunn, for å forenkle modellen og få frem sentrale sammenhenger. Disse antakelsene er:

-Alle boliger er identiske (homogene).

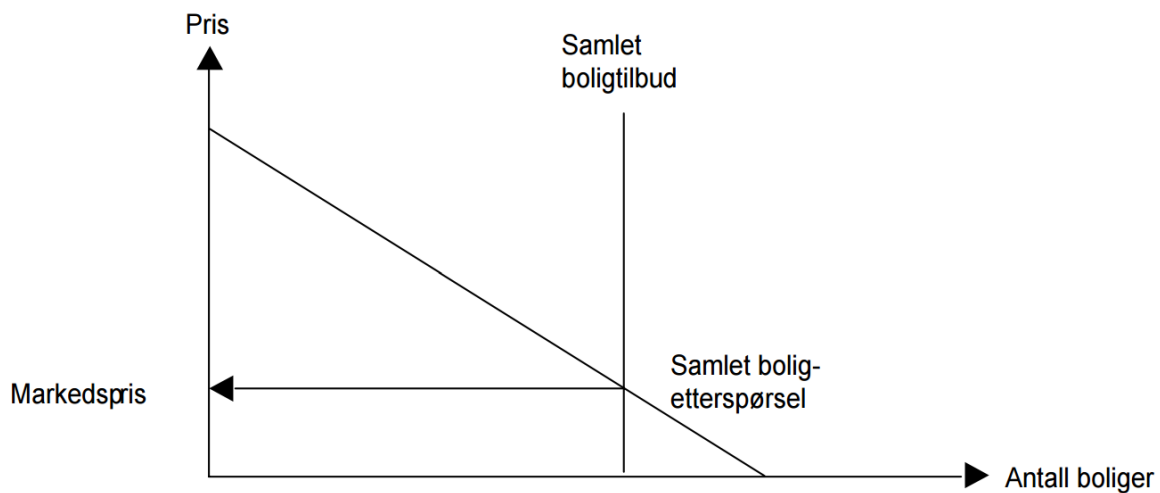
-Alle boliger er eneboliger.

Vi har en tilbudsside, de som ønsker å selge bolig, og en etterspørselsside, de som ønsker å kjøpe bolig. Velger først å se på etterspørselssiden.

Kurven for etterspørsel finner vi ved å sortere alle som ønsker å kjøpe bolig, etter hvor mye de er villige til å betale for boligen. Den som er villig til å betale mest, er førstemann til å få kjøpe bolig. Den som er villig til å betale minst, er sist. Betalingsvilligheten avhenger primært av kjøpers disponible inntekt og formue, der de med høyest inntekt eller formue har høyest betalingsvillighet. Men betalingsvilligheten avhenger også av kjøpers preferanser, hans forventninger, risikovillighet og vektlegging på bolig fremfor annet konsum.

Kurven for tilbud endres ved økt bygging av nye boliger. Nybygging gir en lav årlig endring, tilbudet er således uelastisk på kort sikt. Antar at tilbudet er lavere enn etterspørselen, slik at vi har et press for å komme inn på markedet.

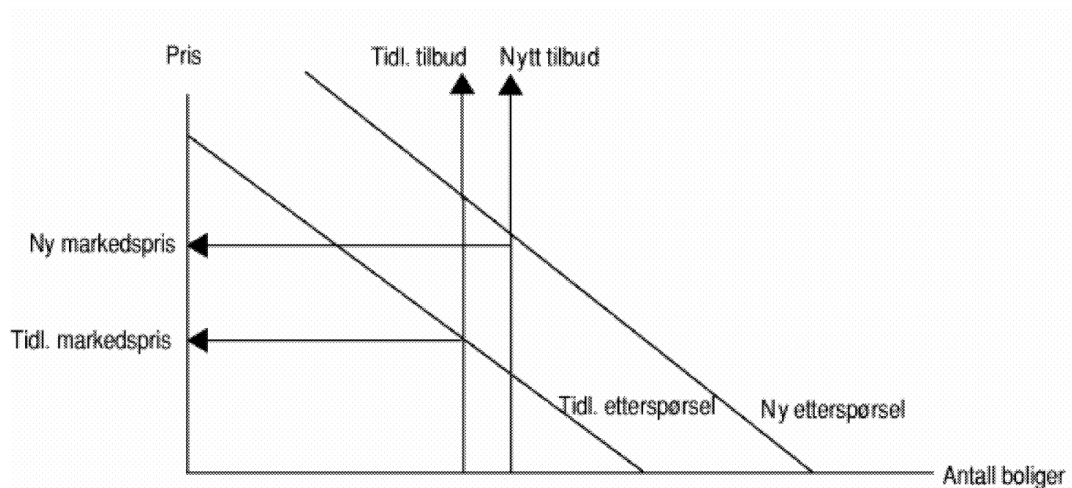
Grafen under illustrerer dette. På y-aksen har vi prisen på bolig og x-aksen viser antallet boliger. Linjen som går fra y-aksen og ned til x-aksen er de som etterspør bolig. De lengst til venstre er de med høyest betalingsvillighet. Den marginale etterspørselen finner vi i skjæringspunktet mellom samlet etterspørsel og tilbud, dette representerer markedspris for bolig. De som ligger til venstre for dette skjæringspunktet har høyere betalingsvillighet enn markedspris, og vil dermed kunne kjøpe bolig. De som ligger til høyre for dette skjæringspunktet har ikke høy nok betalingsvillighet og vil således ikke kunne kjøpe bolig.



Figur 3.1 Samlet tilbud og samlet etterspørsel i markedet for brukte boliger. Kort sikt.²

Vi legger nå til grunn en antakelse om økt nybygging og at andelen nybygg er høyere enn avgangen av boliger. Dette resulterer i en økt boligmasse og samlet boligtilbud vil øke. Dette er illustrert i grafen under, ved at samlet boligtilbud flyttes utover.

På lengre sikt endres også de langsiktige etterspørselsfaktorene som f.eks inntektsvekst. Vi vil da få en økt etterspørsel etter boliger. Dette illustreres i grafen under ved at etterspørselslinjen flyttes utover i grafen. Vi får et «kappløp» mellom nybygging og inntektsvekst, der prisen vil avhenge av hvilken faktor som har størst økning. I grafen under er dette illustrert ved at inntektsveksten har vært større enn veksten i boligtilbud. Som vi ser av grafen resulterer dette i økte boligpriser.



Figur 3.2 Endringer i prisen på boliger ved økning i tilbudet av og etterspørsel etter boliger

Kilde: NOU 2002:2, kap 3. ²

3.3 ALONSO-MUTH-MILLS MODELLEN

I dette del-kapittelet vil jeg redegjøre for Alonso-Muth-Mills modellen. Dette velger jeg å gjøre, fordi jeg vil se på eneboliger i utvalgte byer, av ulik størrelse, i min oppgave. Denne modellen illustrerer godt hvordan prisene på bolig avhenger av boligens beliggenhet i forhold til bysentrum. Benytter meg av forelesningsnotatene i BE-409, samt (Robertsen Karl, Theisen Theis, Boligmarkedet i Kristiansand(2010))

Modellen tar for seg hvordan boligpriser i en by, varierer med avstanden til sentrum. Modellen tar utgangspunkt i en rekke antakelser. Bl.a antas tomteareal å være et fullstendig differensiert gode, tilbudet av areal på et bestemt sted er uelastisk.

Etterspørselen er elastisk og bestemmer prisen. Resultatet av dette er at prisen for en tomt varierer med tomtens beliggenhet.

Følgende forutsetninger legges også til grunn:

-Byen er monosentrisk, med ett bysenter hvor alle innbyggere i byen jobber.

-Tomtearealet, q , per bolig er eksogent gitt.

-Innbyggerne i byen pendler til bysenteret langs ei rett linje, til en kostnad på, k , per kilometer per år. Pendlingsvariabelen er, d .

-Alle husholdningene er identiske. Inntekten deres, y , er brukt på pendling, husleie og annet konsum, x^0 .

- Alle boliger er identiske og husleia er, $R(d)$.
- Husleietjenester produseres ved hjelp av tomtearealet, q , per hus og annen innsats, c .
- Boligene leies ut til leietakeren med høyest betalingsvillighet.
- Forskjellene i husleie korresponderer nøyaktig med forskjellene i transportkostnader.

Videre antas det at annet konsum er likt mellom de ulike husholdningene.

Ut i fra forutsetningene jeg nå har nevnt, ser vi at dette er en svært forenklet modell av virkeligheten. Det er i denne modellen kun beliggenheten til boligen som er avgjørende for prisen. Hvor nærmere bysentrum boligen er lokalisert, hvor høyere er prisen for boligen. Likningen for husleieprisene er som følger:

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

Av denne likningen ser vi at husleien avtar med avstanden fra sentrum og med økningen i transportkostnader. Antar videre at $d=b$ representerer avstanden til bygrensen og at utenfor bygrensen er det kun jordbruk som er aktuelt arealbruk. Avkastningen på jordbruk er, r^a , per mål tomt. Tomteleie ved bygrensen er dermed, $r^a q$. Husleien på bygrensen er dermed summen av tomteleien og byggeleie (c : annuiteten av byggekostnadene): $r^a q + c$.

På bygrensen hvor $d=b$, kan tomteleien skrives som:

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

Alt annet konsum for en husholdning som ligger på bygrensen, blir dermed:

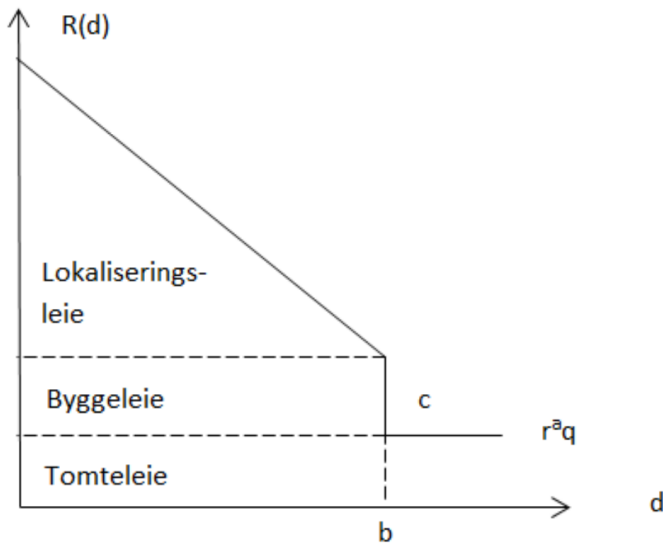
$$X^0 = y - kb - (r^a q + c)$$

Da alle husholdningene er identiske, er likningen ovenfor alt annet konsum for alle husholdningene. Dermed kan husleiegradienten med avstand, d , fra sentrum løses slik:

$$R(d) = y - kd - x^0 = y - kd - y + kb + (r^a q + c)$$

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

Dette illustreres grafisk under.



Figur 3.3 Husleiegradienten.

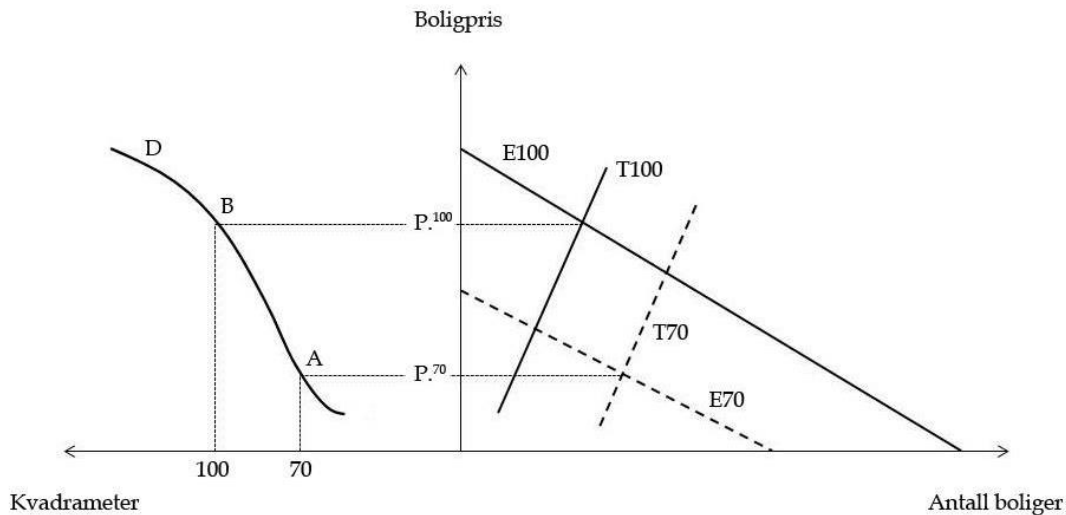
Figuren viser husleien langs Y-aksen, pendleavstanden langs X-aksen og bysenteret i origo. Husleien består av tre komponenter. De to konstante komponentene tomteleie ($r^a q$), byggeleie (c) og lokalisierungsleie ($k(b-d)$), som endres avhengig av avstand til bysenteret.

3.4 DEN HEDONISTISKE PRISFUNKSJONEN

Jeg innledet kapittelet ved å se på noen kjennetegn ved boligmarkedet. Et av disse kjennetegnene var heterogenitet, at boligene innehar ulikt innhold og egenskaper. I det videre arbeidet med oppgaven, vil dette omtales som attributter. Altså faktorer som er med på å bestemme prisen til en bestemt bolig. I dette delkapittelet vil jeg gi en kort og forenklet utledning av den hedonistiske prisfunksjonen. I forrige delkapittel antok vi at avstand til bysenteret var eneste attributt prisen ble bestemt av. I den hedonistiske prisfunksjonen legges det til grunn at prisen bestemmes av flere attributter. Dette kan være attributter for boligmarkedet i en by, som f.eks: nærhet til sentrum, til sjøen, skole, buss-stopp osv. Eller det kan være bygningsattributter, som f.eks: antall soverom, standard på bad osv. Jeg vil benytte meg av (Robertsen Karl, Theisen Theis, Boligmarkedet i Kristiansand(2010)) som kilde.

En sentral antakelse i denne modellen, er at heterogene goder kan oppfattes som en vektor av de ulike attributtene, $z = z_1, z_2, \dots, z_n$. Markedsprisen for et heterogent gode kan sees som en funksjon av godets sammensetning av attributter: $p(z) = p(z_1, z_2, \dots, z_n)$. Det er denne funksjonen som gjerne kalles den hedonistiske prisfunksjonen. Nedenfor følger en enkel modell for

utledning av en hedonistisk prisfunksjon. Modellen tar utgangspunkt i at det bare er ett attributt som bestemmer prisen, i dette tilfellet boligareal.



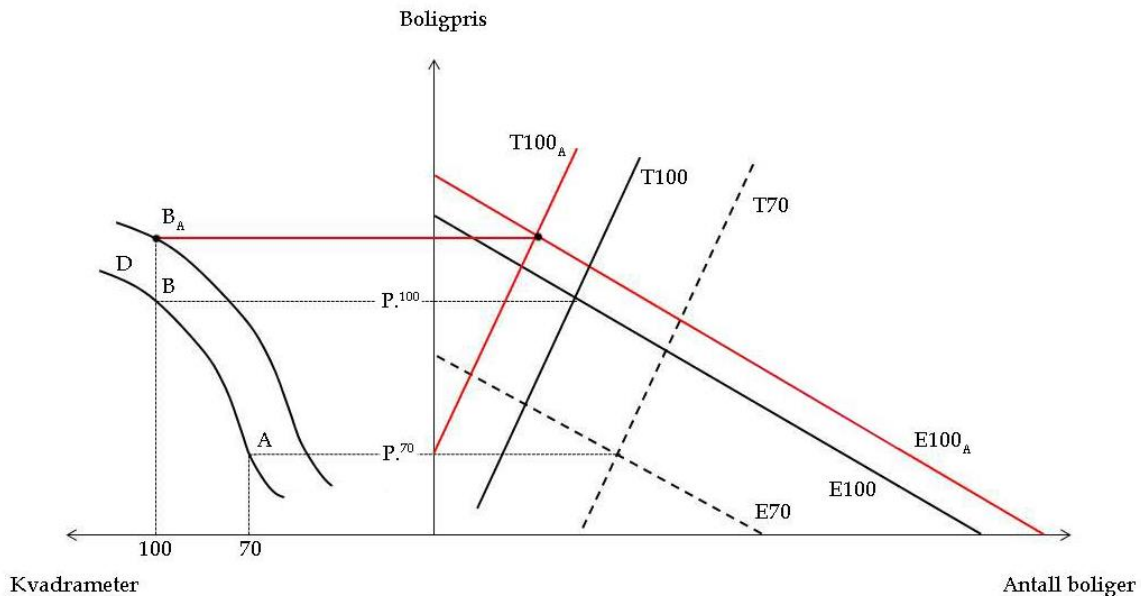
Figur 3.4 Hedonistisk prisfunksjon. (Robertsen Karl, Theisen Theis, Boligmarkedet i Kristiansand(2010))

Til høyre i figuren viser kurven T100, tilbudskurven for en 100kvm stor bolig, mens kurven E100 viser etterspørselskurven. Kurvene T70 og E70 viser også tilbud og etterspørselskurve, men disse gjelder en enebolig på 70kvm. I krysningspunktet mellom T100 og E100, finner vi likevektsprisen, P^{100} . Tilsvarende gjelder krysningspunktet mellom T70 og E70, der vi finner likevektsprisen, P^{70} .

På venstre del av figuren, der boligareal(boa) måles langs den horisontale akse, finner vi punktene A og B, som viser sammenhørende verdier av boligareal og boligpris. Her kan det videre tenkes at vi kunne tegnet inn tilbud og etterspørselskurver for mange andre boligstørrelser og endt opp i punktene, D, osv, på venstre side i figuren. Til sammen utgjør disse punktene den hedonistiske prisfunksjonen, en kontinuerlig stigende kurve på venstre side i figuren. Denne figuren kan utvides til tre dimensjoner, ved å inkludere et attributt utover boa, f.eks et kontinuerlig attributt som tomteareal, målt langs to horisontale akser som står vinkelrett på hverandre.

Jeg vil nå se på et eksempel med to attributter. I dette tilfellet vil jeg benytte meg av garasjeleilighet som attributt. Dvs. at enten så innehar boligen en garasjeleilighet, eller så

innehar den ikke garasjeleilighet. Antar også at alle garasjeleiligheter har en gitt lik størrelse, som ikke er medregnet i boligens boareal. Her er en ny tilbud (T100_A) og etterspørselskurve (E100_A) for bolig på 100kvm tegnet inn med rød skrift. Denne boligen har to attributter, boareal og garasjeleilighet. Nå har den hedonistiske prisfunksjonen gått fra punkt B til punkt B_A.



Figur 3.5 Hedonistisk prisfunksjon, to attributter.

Som vi ser av denne hedonistiske prisfunksjonen, fremgår det at prisen på en bolig vil variere avhengig av mengden attributter som er knyttet til boligen. I eksempelet over er attributtene, boareal og garasjeleilighet de eneste attributtene som skiller boligene. I praksis er det dog en betydelig mengde attributter for heterogene goder. Men modellen illustrerer godt teorien bak prisdifferansen mellom heterogene goder.

3.5 GARASJEBYGGET

I dette avsnittet vil jeg se nærmere på kostnaden ved oppføring av garasje med hybel, samt gjøre en enkel investeringsanalyse basert på kostnad og inntekt fra garasjen.

Da det stilles så strenge krav for oppføring av garasje over 70m², vil som tidligere nevnt de fleste privatpersoner være i behov av en tredjepart med ansvarsrett. Dette er som regel en

entreprenør med byggeteknisk kompetanse. Å finne en markedspris for slike garasjer har vist seg å være svært vanskelig, da det ikke er mange seriøse aktører som leverer disse store garasjene i moduler. Jeg har bl.a. funnet en leverandør, leilighetsgarasjer.no, men velger å ikke bruke denne, da eier av selskapet har en historie med en rekke konkurser, får svært negativ omtale på internett, samt at nåværende selskap står oppført uten inntekt i regnskapene. For å begrense sannsynligheten for uverifiserbar data, ser jeg på store anerkjente nasjonale aktører. I segmentet for garasjer over 70kvm lar dette seg kun gjøre i form av Sveriges største garasjeleverandør, med norsk filial, Mellby-garasje⁶. De har to ferdigarasjer med muligheter for leilighet på toppen. Begge disse garasjene koster ca. 400 000kr inkl. moms. Men det er viktig å påpeke at innvendig innredning som f.eks. dører, gulv, innervegger, kjøkken og bad ikke er inkludert i prisen. Frakt og oppføring er heller ikke inkludert og det forutsetter at du selv står med ansvarsrett, eller skaffer deg en tredjepart som kan gjøre dette.

Jeg har også vært i kontakt med en regional garasjeleverandør på Vestlandet, kalt Koppen industrier AS. De var svært hjelpelige og sendte meg tilbudstegningene til en garasje med leilighet, de nylig har solgt. Tegningene ligger vedlagt. Selgeren jeg pratet med forklarte at denne garasjen kom på rett rundt 600 000kr inkl. moms. Her er også regnestykket uten innvendig innredning, men inkludert frakt og oppføring av bygget på ferdig planert mark. Det er viktig å påpeke at begge eksemplene er for selve garasjen og ikke inkludert planeringsarbeid osv.

For mindre garasjer er Igland-garasjen Norges største leverandør. De leverer ikke garasjer godkjent for leilighet, men leverer garasjer med loft. Dette loftet må kjøper selv i etterkant innrede og klargjøre for leilighet/hybel. Dog må det nevnes at dette alternativet ikke under noen omstendighet er lovlig som utleieenhet. Dette er fordi Igland leverer garasjer i segmentet med unntak fra full byggesøknad. Som tidligere nevnt er dette garasjer der loftet er for lavt for å bli godkjent som beboelsesrom.

Er leieinntekter på garasjeleilighet skattepliktige?

I utgangspunktet er utleie skattepliktig, men i skatteloven § 7.2³ finner vi enkelte unntaksbestemmelser. I punkt 1 a) står det at utleieinntekt fra bolig er skattefritt dersom boligeier benytter minst halvparten av boligen til eget bruk, regnet ut fra utleieverdi.

Men i mange tilfeller er garasjen og boligen to separate bygg og reglene i disse tilfellene er litt mer omfattende. På skatteetatens hjemmesider kan man lese om reglene for utleie dersom det er flere bygninger på boligens tomt. Her er utgangspunktet at hvert bygg skal vurderes hver for

seg, med tanke på om vilkårene for skattefritak av utleieinntekter er oppfylt. Men det finnes unntak. Følgende tekst er kopiert fra skatteetatens sider:

«Frittliggende bygninger skal likevel vurderes sammen med bolighuset dersom bygningene etter sin karakter har en mer eller mindre nødvendig tjenende funksjon til boligen. Dette kan for eksempel gjelde bod, båthus eller garasje»

Om en garasje med leilighet har tjenende funksjon til boligen blir en skjønnsmessig vurdering, da leiligheten i de fleste tilfeller ikke kan anses som tjenende funksjon. Men på den annen side er selve garasjen, altså oppstillingsplassen for bilene, i de fleste tilfeller en tjenende funksjon til boligen. Ut i fra sedvane og rettspraksis kan vi konkludere med at utleie av hybel over garasjen faller innenfor skattefritak i de aller fleste tilfeller, med mindre spesielle omstendigheter tilsier noe annet.

Enkel analyse.

Jeg tar heretter utgangspunkt i garasjen fra Koppen Industrier AS og bruker grove antakelser for hvilke kostnader som påløper. Benytter meg også av en forholdsmessig høy rente.

Garasjen i seg selv kostet 600 000kr, inkludert oppføring. Jeg anslår arbeidet med innredning av boligdelen og planering av grunnen, til ca. 300 000 kr, hvorav 240.000kr tilfaller innredningskostnaden og 60.000 for planering. Investeringskostnaden for ferdigstilt garasje med leilighet blir da 900 000 kr.

La oss anta at disse 900 000kr blir finansiert i sin helhet ved hjelp av refinansiering på eksisterende huslån, med 25års nedbetalingstid, annuitetslån og en rente på 6% p.a. Å regne ut de månedlige kostnadene ved lånet er en enkel sak ved hjelp av boliglånskalkulator på Nordea sine hjemmesider⁴. Jeg ser bort ifra etableringsgebyr og termingebyr, da dette ikke er oppgitt i kalkulatoren. 5849kr er de månedlige utgifter på lånet. Her kan vi igjen trekke fra rentekostnadene på skatten. Da andelen renteutgifter er fallende gjennom lånets periode, velger jeg å se på hele rentekostnaden over perioden og fordele fradraget jevnt over periodene.

Total rentekostnad er 839 483kr. Deler dette på 25 for å få årlig rentekostnad og får 33 579kr. Deler igjen dette på 12 for å få månedlig rentekostnad og får 2798kr. 27% av dette kan trekkes fra de månedlige utgiftene. $2798 \times 0,27 = 756$ kr.

Leieinntekter vil naturlig nok variere fra by til by, avhengig av tilbud og etterspørsel. Brorparten av byene jeg har brukt i min datainnsamling, består av byer med nærhet til universitet eller høyskoler og antas å være velfungerende mht. utleie.

Vi har altså en investeringskostnad på drøye 5094 i måneden. Grunnet skattefritak skal en leieinntekt på 5094kr isolert sett være nok til å dekke både renter (etter rentefradrag) og avdrag på investeringen. Men utleieenheten forbruker vann, samt andre løpende kostnader til vedlikehold av enheten. Allokterer også en del av eiendomsskatt og kommunale avgifter til utleieenheten. La oss sette disse kostnadene til 1500kr månedlig. Strøm holder jeg utenfor regnestykket, da jeg antar leieboer står for å betale strømregningene.

Leieinntekter på ca. 6594kr pr. mnd., vil i mitt tenkte tilfelle være nok til at investeringen betaler seg selv over en 25års periode. Det bør også nevnes at man drar nytte av investeringen fra den er ferdigstilt, i form av garasje plass for bilen(e). Samtidig påtar man seg et ansvar for oppfølging av leieboerforholdet og må regne med å investere litt tid på omstendigheter rundt dette.

Nytten man får ved å ha garasje plass til bilene sine, kan enkelt trekkes ut av regnestykket ved å se på kostnaden av oppføring av garasje, kun tiltenkt biler. Igland garasje sin billigste doble garasje koster 73 100kr for ferdige elementer. Denne prisen er uten frakt og oppføring. Forutsetter at frakt og oppføring gjennomføres av Igland garasjer og et kjapt søk på nett viser at dette kommer på omkring 25.000kr. Planering av tomt antar jeg fortsatt koster 60 000kr. Investeringskostnaden for garasje uten loft blir således 158 100kr. Hvis jeg legger samme lånevilkår og rente til grunn, og benytter meg av samme lånekalkulator, blir den månedlige kostnaden 1069kroner. Rentefradraget trekker jeg i fra på samme måte som ovenfor.

Total rentekostnad er 147.349kr. Deler dette på 25 for å få årlig rentekostnad og får 5894kr. Deler igjen dette på 12 for å få månedlig rentekostnad og får 491kr. 27% av dette kan trekkes fra de månedlige utgiftene. $491 \times 0,27 = 133 \text{ kr}$

Veldig forenklet kan man si at en hybel over garasjen genererer en merkostnad på $6594 - 1069 - 133 = 5392 \text{ kr}$ i måneden, kontra en garasje uten hybel.

3.6 HYPOTESER

I dette avsnittet vil jeg formulere hypoteser på bakgrunn av oppgavens problemstilling og teorien gjennomgått i dette kapittelet. Problemstillingen er kort og konsis og går ut på hvorvidt en garasje med leilighet gir økt verdi til eneboligen. Jeg vil sette opp hypotesesett, der hvert sett

vil bestå av en nullhypotese og en alternativ hypotese. Alternativhypotesen er påstanden jeg fremlegger, mens nullhypotesen er det motsatte. Dersom jeg finner støtte i de empiriske resultatene, vil nullhypotesen forkastes, mens den beholdes hvis alternativhypotesen ikke støttes av resultatene.

Hypotesesett1/Hovedhypotese: /

H0: Garasjeleilighet har ingen betydning for eneboligens salgspris

H1: Garasjeleilighet har betydning for boligens salgspris

Jeg vil også legge til to kontrollhypoteser, da det er flere faktorer som er med på å bestemme boligpriser.

Hypotesesett2/kontrollhypotese:

H0: Boareal har ingen betydning for boligens salgspris

H1: Boareal har betydning for boligens salgspris

Hypotesesett3/kontrollhypotese:

H0: Boligens alder har ingen betydning for salgsprisen

H1: Boligens alder har betydning for salgsprisen

4. ØKONOMETRISK MODELL

4.1 INNLEDNING

Relasjonen som skal estimeres er en garasjeleilighets påvirkning på salgsprisen til enebolig. Jeg velger salgspris som avhengig variabel og kontrollerer videre for uavhengige variabler som kan ha innvirkning på prisen. Disse er boligareal/P-rom og boligens opprinnelses-år. Videre har jeg flere dummyvariabler for postnummer, enebolig med og uten garasjeleilighet, samt årstall for salg. Dummyvariabelen for postnummer er tatt med for å fange opp avstand til sentrum og det faktum at byene har ulik størrelse, jf. teorien i Alonso-Muth-Mills modellen. Postnummer vil også fange opp avstanden til andre attributter. Dummyen for årstall for salg, er tatt med for å fange opp en eventuell pris-stigning mellom 2013 og 2014(to måneder i 2015 tas med her).

Videre i dette kapitlet, vil jeg si litt om regresjonsanalysens forutsetninger, så se på de aktuelle regresjonsformene og til sist si litt om et par relevante statistiske estimater.

4.2 REGRESJONSANALYSENS FORUTSETNINGER

En lineær regresjonsanalyse er basert på følgende forutsetninger:

1. Lineær sammenheng mellom variablene
2. Fravær av multikollinearitet
3. Restleddvariasjonene er: a) Homoskedastiske b) Normalfordelt c) Uavhengige av hverandre
4. Uavhengige variabler og restleddet er ukorrelert med hverandre.

Lineær sammenheng mellom variablene

Dersom det ikke foreligger lineær sammenheng mellom variablene, risikerer vi feil estimat på koeffisientene, feil i standardfeil og feil i signifikanstest.

Fravær av multikollinearitet

I estimeringen og tolkningen av resultatene i regresjonen, kan det oppstå et problem kalt multikollinearitet. Dette går ut på at de uavhengige variablene kan ha en relasjon seg imellom og korrelere. En korrelasjonsmatrise vil kunne si oss noe om multikollinearitet, men i denne matrisen avsløres bare bivariat multikollinearitet. Jeg vil likevel si litt om korrelasjon i kapittel 6. Det mest hensiktsmessige er å foreta en analyse kalt varianseinflasjonsfaktor. Denne analysen går ut på å kjøre regresjon på hver enkelt variabel mot de andre variablene i likningen. Den matematiske formelen er som følger:

$$VIF = 1 / (1 - R_i^2)$$

R_i^2 står for den samlede korrelasjonskoeffisienten mellom en variabel «i» og de andre uavhengige variablene opphøyd i andre. Denne testen vil jeg kjøre i kapittel 7. Dersom VIF er under 10 og samlet gjennomsnittlig VIF for variablene ikke ligger mye over 1, kan vi med sikkerhet si at stor grad av multikollinearitet ikke foreligger.

Homoskedastiske

Forutsetningen om at restleddvariasjonene skal være homoskedastiske vil si at variasjonene rundt regresjonslinjen er like store overalt, altså for alle verdier av uavhengige variabler. Da jeg har et utvalg som er så stort som 2298 enheter, antar jeg at eventuell heteroskedansitet (det motsatte av homoskedansitet) ikke vil være et problem, men vil likevel kontrollere for dette i kapittel 7.

Normalfordelt restledd

Forutsetningen om normalfordelt restledd er viktig i mindre utvalg, for å kontrollere for at feilmarginer og signifikansnivåer er korrekte. I større utvalg er dette av mindre viktighet, men jeg vil likevel kjøre test for dette i stata.

Restledd uavhengige av hverandre

Jeg vil i kapittel 7 teste om restledd er uavhengige av hverandre, ved å bruke en kommando i stata, kalt «clustered».

Uavhengige variabler og restledd ukorrelert med hverandre.

Denne forutsetningen byr på større problemer. Forutsetningen innebærer at det ikke må foreligge bakenforliggende årsaksfaktorer til Y som også er korrelert med årsaksfaktoren til X.

4.3 REGRESJON

Regresjonsanalyse er en statistisk metode brukt for å bestemme hvordan en eller flere uavhengige variabler avhenger eller forklares av avhengig variabel. I lineær regresjon skiller vi mellom enkel og multippel regresjon.

Enkel regresjon

Ved enkel regresjon forsøkes det å forklare en uavhengig variabels innvirkning på den avhengige variabelen. Den matematiske formelen for enkel regresjon er som følger:

$$Y = \alpha + \beta X + \varepsilon$$

Hvor Y står for avhengig variabel, α står for konstantledd, β står for regresjonskoeffisienten, X står for uavhengig variabel og ε står for feilledd. Teoretisk gir β den eksakte effekten av uavhengig variabelen, X, sin påvirkning på avhengig variabel, Y.

For mitt vedkommende, som ser på boligpris, vil prisen avhenge av flere uavhengige variabler, fordi boliger er et heterogent gode. Jeg vil derfor se nærmere på multippel lineær regresjon, som tar hensyn til flere uavhengige variabler.

Multippel lineær regresjon

Jeg starter med den matematiske formelen og forklarer ut i fra denne.

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

Her står Y for avhengig variabel, α står for konstantledd, β er helningskoeffisienten, X står for den uavhengige/forklaringsvariabelen, k står for antallet uavhengige variabler og ε står for feilledd. Regresjonskoeffisientene β_1 , β_2 opp til β_k og feilleddet ε , er ukjente størrelser vi ønsker å estimere. Når dette er gjort, må det gjennomføres statistiske tester. Her er det mest prekärt å undersøke om sammenhengen mellom avhengig og uavhengig variabel er signifikant ulik fra null.

Dobbel logaritmisk regresjon

Denne funksjonsformen er ikke lineær i forholdet mellom variablene Y og X, men den er lineær i koeffisientene. Den matematiske formelen skrives som følger:

$$\ln Y_i = \alpha + \beta_1 \ln X_i + \varepsilon_i$$
, der vi benytter oss av naturlig logaritme for variablene.

Dobbelt logaritmisk form benyttes ofte når egenskapen ved funksjonen er sånn at elastisiteten er konstant, men ikke forholdet mellom variablene. Dette er i kontrast til lineær form, som jeg nevnte antar et konstant forhold mellom variablene. I dobbelt logaritmisk form kan vi tolke regresjonskoeffisienten som elastisitet. Begrunnelsen for dette utleder jeg matematisk.

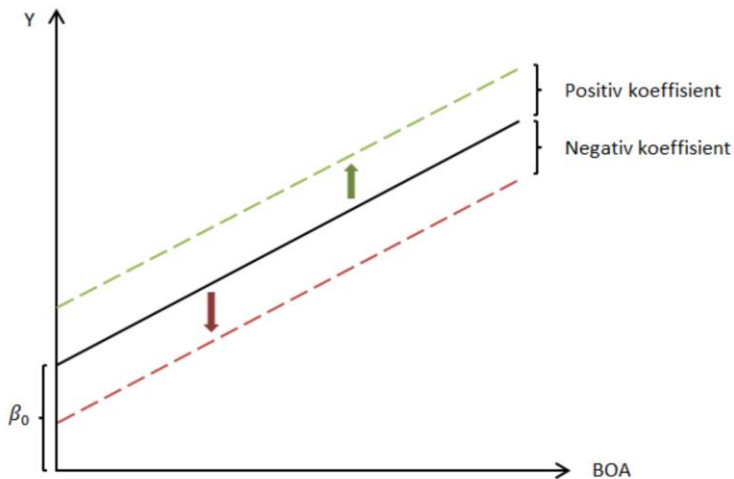
$$\beta_k = \Delta(\ln Y) / \Delta(\ln X) = (\Delta Y / Y) / (\Delta X_k / X_k) = \text{Elastisitet } Y, X_k$$

Da regresjonskoeffisientene er konstante er forutsetningen om at modellen har konstant elastisitet oppfylt av dobbelt logaritmisk funksjon sin likning. β_k kan tolkes som at dersom X_k har 1% økning og de andre X holdes konstante, vil Y_k endres med β_k %.

Dummyvariabel

En dummyvariabel er en variabel som beskriver nærvær eller fravær av en effekt eller egenskap. Variabelen innehar verdien 1 eller 0, og gir et tillegg eller fratrukk i konstanten α . Denne endringen vi vises som et skift i regresjonslinjen, oppover for positiv verdi av dummyvariabelens koeffisient, nedover ved negativ verdi. Dette illustreres grafisk nedenfor. I

den grafiske fremstillingen, er konstantleddet kalt β_0 . Når et sett med dummyvariabler brukes i regresjonen må en «kategori» holdes utenfor regresjonen, for å forhindre at det oppstår multikollinearitet. Dersom alle kategorier inkluderes i regresjonen, vil det danne lineær avhengighet i parameterne, da hver enkelt kategori kan forklares som en funksjon av de andre kategoriene. Dersom vi har X antall kategorier, må X-1 kategorier hentes inn i regresjonen.



Figur 4.1 En dummyvariabels effekt på regresjonslinjen og konstantleddet.

Jeg vil nå forklare litt kort om noen statistiske estimater som benyttes for å vurdere kvaliteten av regresjoner.

R^2

I regresjonsanalysen gjenspeiler R^2 hvor mange prosent av variasjonen i avhengig variabel som kan forklares av uavhengig variabel. Den matematiske formelen skrives som følger:

$$R^2 = (1 - (SSE/SST))$$

SSE gjenspeiler regresjonsfeil (sum of squares error) og SST gjenspeiler alle kvadrater (sum of squares total).

Justert R^2

Måler det samme som R^2 , justert for frihetsgrader. Den tar altså hensyn til hvor mange uavhengige variabler som inngår i regresjonsmodellen. Dersom frihetsgradene går ned, vil også estimatenes pålitelighet gå ned. Justert R^2 øker bare dersom variabel som legges til,

forholdsmessig øker forklaringskraften mer enn tapet av frihetsgrader som brukes til estimering av koeffisienten til ny variabel.

5. DATAINNSAMLING OG BESKRIVELSE AV DATAMATERIALET

5.1 VALG AV ENHETER OG DATAINNSAMLING

Det neste steget i undersøkelsen er å velge undersøkelsesdesign, enheter, innsamling av data, koding, analysering og konklusjon, i nevnte rekkefølge. I dette kapitlet velger jeg å beskrive mitt valg av undersøkelsesdesign, samt utvalg av enheter.

Problemstillingen i oppgaven er som tidligere nevnt; «Gir tilhørende garasjeleilighet merverdi til eneboligen ved salg?» Ønsket mitt med denne problemstillingen er å kunne forklare hvorvidt en garasje med hybel gir utslag i pris. Gjennom utvalget mitt håper jeg å kunne si noe om hvorvidt en garasje med hybelloft genererer merverdi også til tilsvarende objekter utenfor min undersøkelse. Antallet enheter burde derfor være relativt stort. Faktorer som tid og tilgjengelighet vil naturlig nok være med og begrense mitt antall enheter. I min oppgave har jeg valgt en kvantitativ tilnærming for innsamling og analysering av data. Variablene som tas med i innsamlingsprosessen er valgt i samråd med veileder, på bakgrunn av tilgjengelig informasjon. Jeg har benyttet meg av eiendomsverdi.no for innsamlingen.

Den teoretiske populasjonen er eneboliger i Norge. Den faktiske populasjonen er eneboliger solgt i utvalgte byer i Norge, i perioden mellom 1.1.2013 og 20.2.2015. Jeg har ved hjelp av eiendomsverdi.no gått gjennom eneboliger solgt i Kristiansand, Lillesand, Grimstad, Arendal, Skien, Sandefjord, Tønsberg og Drammen. Grunnet begrensninger på eiendomsverdi.no, har jeg ikke kunnet gjennomgå boliger solgt uten megler. Dette er fordi eiendomsverdi registrerer alle boliger som er solgt i Norge, men for å fysisk kunne se boligene og om de har garasje med hybel, må jeg trykke på eiendomsverdi sine linker til eiendomsmeglerens prospekt på finn.no. Boliger som er solgt uten eiendomsmegler, eller overdratt gjennom arv, er på eiendomsverdi.no kun oppført med salgsdato og salgspris, uten bilder. Da tidsbruk er en vesentlig begrensning, har jeg gått gjennom byene på følgende måte;

Kristiansand og Grimstad; alle eneboliger solgt av eiendomsmegler i perioden 1.1.2013-20.2.2015.

Arendal, Skien, Sandefjord og Tønsberg; Alle eneboliger solgt av eiendomsmegler i perioden 1.1.2013-20.2.2015, med byggeår etter 1980.

Lillesand og Drammen; Alle eneboliger solgt av eiendomsmegler i perioden 1.1.2014-20.2.2015, med byggeår etter 1980.

Jeg måtte gjøre det på følgende vis, da populasjonen ble veldig stor og innsamlingen tidkrevende. Årsaken til at jeg etter hvert begrenset søket ytterligere, til boliger med byggedato etter 1980, var fordi jeg bet meg merke i at brorparten av eneboligene med garasjeleilighet i Kristiansand og Grimstad, var bygget etter 1980. Årsaken til at jeg kun gikk igjennom perioden 1.1.2014-20.2.2015 i Lillesand og Drammen, var fordi utvalget av garasjeleiligheter her var svært begrenset. (1 i hver by). Jeg har også gått gjennom alle eneboliger solgt i Trondheim, Bergen og Stavanger i perioden mellom 1.1.2014-20.2.2015, uten å ta dette med i populasjonen. Dette valgte jeg å gjøre, da jeg ikke fant en eneste enebolig med tilhørende garasjeleilighet i disse byene.

Grunnet min metode for innsamling av data, kan det tenkes at jeg har mistet noen eneboliger som har garasjeleilighet, men jeg måtte gjøre en avveining mhp. tidsbruk. I teorien hadde den beste metoden for datainnsamling, vært å se på boliger som relativt nylig er solgt uten garasjeleilighet, deretter har garasjeleilighet blitt oppført, for så å se på hva boligene er solgt for etter oppføringen av garasjeleiligheten. I praksis er dette en vanskelig, trolig umulig oppgave. Da utvalget mitt er såpass stort, antar jeg at resultatene i analysen gjenspeiler populasjonen i de byene jeg har undersøkt.

5.2 DATARENSING

Datagrunnlaget mitt består av en populasjon på 2702 enheter. Jeg har benyttet meg av excel i innsamlingen. Fra eiendomsverdi har jeg fått oppgitt variabler som salgsdato, adresse, salgpris, eieform, bygningsdato, megler, boa/p-rom. De 2702 observasjonene ble så lastet opp i stata, via et tekst-dokument jeg opprettet i excel. I stata benyttet jeg meg av en rekke kommandoer i en såkalt do-file. Innholdet i denne do-filen ligger vedlagt i vedleggene, da det er 29 sider med koding. På denne måten fikk jeg sortert og rensset dataene mine. Jeg har valgt å benytte meg av salgpris som avhengig variabel og boa/p-rom og bygningsdato som uavhengige variabler. Jeg har også lagt inn postnummer og dato som dummy-variabel. Jeg ser bort i fra

eiendom fordi alle enhetene i utvalget er selveierboliger. Jeg måtte fjerne 339 enheter som ikke var omsatt via eiendomsmegler, 11 enheter som manglet informasjon om når eneboligen var ferdigstilt, 30 enheter der P-rom ikke var oppgitt, samt 1 enhet der postnummer ikke var oppgitt. Etter å ha rensset data for eventuelle feil og mangler, bestod utvalget i 2298 enheter, alle fullstendige mhp. Variabler. Det er verdt å merke seg at jeg ikke har sjekket for eventuelle dobbeltføringer i utvalget. Dette er en forholdsvis tidkrevende prosess. Jeg antar dessuten at andelen dobbeltføringer er uvesentlig, da boligen må være solgt to ganger i løpet av en 26 måneders periode for at en slik dobbeltføring skal finne sted. At populasjonen bare består av eneboliger taler også for at mengden dobbeltføringer er svært liten, da jeg antar at eneboliger holdes lengre før et videresalg.

46 av enhetene har garasjeleilighet. I henhold til Roscoe (1975) sin tommelfinger-regel vil et utvalg mellom 30-500 være passende i de aller fleste undersøkelser. På den andre siden er et større utvalg å foretrekke, fremfor et lite. Et endelig utvalg bestående av 46 garasjeleiligheter må kunne sies å være smalt. Jeg er inneforstått med at dette øker risikoen for feilmarginer, kontra et større utvalg. Dette er beklagelig, men tidsbegrensninger må også hensyn tas. I positiv retning trekker tilgjengelighet av verdier for variablene som ønskes belyst. Her kan blant annet salgspris, P-rom og postnummer hensyn tas med stor grad av sikkerhet. Tabell over variablene, med kort forklaring følger under.

Tabell 5.1 Variablene

Variabel	Beskrivelse
Pris	Salgspris målt i NOK
Boa/P-rom	Areal av gulvflate i P-rom.
Bygningsdato	Dato for når bygget var ferdigstilt.
Salgsår	1 hvis boligen er solgt etter 31.12.2013, =0 hvis solgt før
Garasje	1 hvis garasjeleilighet, 0 hvis ikke.
Postnummer	Har et sett av dummyvariabler her. Gått gjennom hvert postnummer og gitt de verdier 1 og 0.
By	Sett bestående av 9 dummy-variable, en for hver by.

Jeg må også påpeke at jeg ikke har rensset dataene ytterligere for såkalte ekstreme verdier. I utvalget mitt er det en håndfull eneboliger som er solgt for over 10.000.000kr. Siden jeg har gått gjennom byer på Sørlandet og Sør-Østlandet, kan man nok argumentere for at denne prisklassen er en ekstrem verdi. Variablen for postnummer håper jeg vil fange opp dette, da boligens lokalisering inkluderer attributter som driver prisen. F.eks. nærhet til sjøen, sentrum,

handlesenter m.fl. Postnummervariabelen bød på litt problemer, da enkelte boliger i utvalget mitt var eneste, eller en av svært få, boliger med et gitt postnummer. For å løse dette valgte jeg å slå sammen noen postnumre. Dette er ikke ideelt, men jeg antar at tilnærmet like postnummer gir tilnærmet like attributter for disse eneboligene. Totalt hadde jeg 159 postnumre, som etter sammenslåing ble 112 postnumre.

6. PRESENTASJON AV DATAMATERIALET

I dette kapittelet vil jeg presentere datamaterialet mitt. Datamaterialet er stort og inneholder en mengde informasjon. For at denne informasjonen skal være forståelig er det viktig å ha en enkel, hensiktsmessig og forståelig fremstilling av dataene.

6.1 DESKRIPTIV STATISTIKK

Jeg vil starte med en hovedtabell over datamaterialet, med deskriptiv statistikk. Tabellen sees nedenfor.

Tabell 6.1 Deskriptiv statistikk over tallmaterialet.

Variable	Obs	Mean	Std. Dev.	Min	Max
Boaprom	2298	163.4617	51.67664	50	480
Bygningsr	2298	1981.718	29.93542	1691	2015
Salgsaar	2298	.5809399	.4935126	0	1
Pris	2298	3533376	1453196	590000	1.55e+07
Garasje	2298	.0200174	.1400901	0	1
postnr43	2298	.0021758	.0466049	0	1
Postnr104	2298	.0369887	.1887751	0	1

Tabellen viser antall variabler, observasjoner, gjennomsnittsverdi, standardavvik, minimumsverdi og maksimumsverdi i nevnte rekkefølge. Minimum og maksimumsverdiene forteller oss hvor mye variablene varierer og er nyttige for å oppdage og eventuelt fjerne ekstremverdier og feil. Gjennomsnittet(mean) er summen av dataverdiene for hver variabel, delt på antall observasjoner. «Standardavviket måler hvor mye en serie verdier avviker fra

seriens gjennomsnitt» (Wenstøp (2006)). Ut i fra denne forklaringen er det tydelig at et høyt standardavvik indikerer en stor spredning fra gjennomsnittet.

Kort kommentar til tallene i tabellen

Variablene boa/P-rom og bygningsår er informative og gir en klar mening. Variabelen salgsår gir ved første øyekast liten mening. Som tidligere nevnt er denne laget som en dummy, der hus solgt før 1.1.2014 har fått verdien 0, resten har verdien 1. Gjennomsnittet er her ca 58%, noe som betyr at mer enn halvparten av utvalget mitt er solgt etter 1.1.2014, mens litt under halvparten er solgt i 2013. Dummyvariabelen «garasje» betyr at jeg har gitt eneboliger med garasjeleilighet verdien 1 og eneboliger uten garasjeleilighet verdien 0. Gjennomsnittet her er på ca 2%, noe som indikerer at en veldig liten del av dataene mine har garasjeleilighet. Dummyvariablene post43 og post104 er de to variablene, i ett sett bestående av 112 dummyvariabler, med lavest og størst antall observasjoner for hvert postnummer. Post43 er postnummer 3241 i Sandefjord, her har jeg 5 observasjoner. Post104 er postnummer 4870 på Fevik i Grimstad, her har jeg 85 observasjoner.

6.2 UTVIDET DESKRIPTIV STATISTIKK

Nedenfor følger mer detaljert deskriptiv statistikk for de kontinuerlige variablene pris, boa/P-rom og bygningsår.

Tabell 6.2 Utvidet statistikk for pris, boa/P-rom og byggeår.

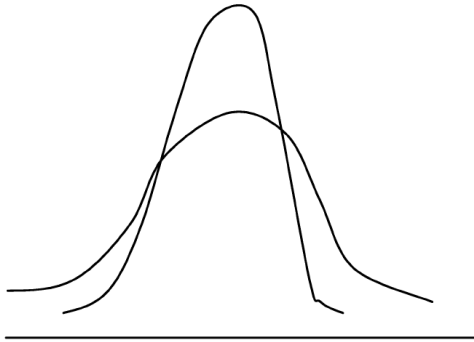
```
. summarize pris boaprom bygningsår, detail
```

pris				
Percentiles	Smallest			
1%	1390000	590000		
5%	1900000	775000		
10%	2180000	890000	Obs	2298
25%	2580000	1020000	Sum of Wgt.	2298
50%	3250000		Mean	3533376
		Largest	Std. Dev.	1453196
75%	4100000	1.30e+07		
90%	5300000	1.42e+07	Variance	2.11e+12
95%	6100000	1.48e+07	Skewness	2.191908
99%	8750000	1.55e+07	Kurtosis	12.39736

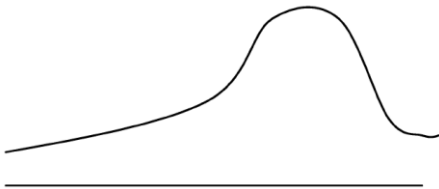
boa/P-rom				
Percentiles	Smallest			
1%	63	50		
5%	94	51		
10%	106	51	Obs	2298
25%	128	51	Sum of Wgt.	2298
50%	157		Mean	163.4617
		Largest	Std. Dev.	51.67664
75%	192	412		
90%	230	418	Variance	2670.475
95%	252	448	Skewness	.9575066
99%	327	480	Kurtosis	5.209389

byggingår				
Percentiles	Smallest			
1%	1860	1691		
5%	1932	1750		
10%	1956	1764	Obs	2298
25%	1972	1764	Sum of Wgt.	2298
50%	1987		Mean	1981.718
		Largest	Std. Dev.	29.93542
75%	2000	2014		
90%	2009	2014	Variance	896.1294
95%	2012	2014	Skewness	-3.054195
99%	2013	2015	Kurtosis	18.05083

I denne utvidede analysen er skewness og kurtosis for de kontinuerlige variablene av interesse. Disse målene er matnyttige for mine kontinuerlige variable, men bærer ingen hensikt for variable med nominale verdier, slik som dummyvariablene. Skewness og kurtosis brukes som indikator på hvorvidt en variabel er normalfordelt. Skewness måler skjevheten i datagrunnlaget til variabelen. Ved høy positiv verdi på skewness har fordelingen til variabelen en lang «hale» til høyre. Ved høy negativ verdi er halen til venstre. Om verdien er null, er variabelen fullstendig normalfordelt. Kurtosis er et mål på hvordan fordelingen av data er spredt mellom laveste og høyeste verdi i variabelen. Dersom kurtosis har høy positiv verdi, er variabelens haler flatere og lengre enn ved normalfordeling. Dersom kurtosis har høy negativ verdi vil halene være kortere og indikere smalere/mer spiss fordeling enn ved normalfordeling. Dette er illustrert i figurene nedenfor.



Figur 6.1 Kurtosis, negativ verdi. Kortere haler og spiss/smål fordeling



Figur 6.2 Skewness, negativ verdi. Hale til venstre.

I prisvariabelen ser vi at skewness er 2,19 og kurtosis er 12,39. At kurtosis er såpass høy indikerer at en stor mengde av observasjonene i variabelen har samme verdi som gjennomsnittet. I variabelen boa/P-rom ser dataene ok ut. I variabelen bygningsår er skewness negativ og kurtosis så høy som 18. Dette ser ved første øyekast abnormalt ut, men dette skyldes min «manipulering» av innsamlet data. Som nevnt tidligere, begynte jeg etter hvert å se på boliger bygd etter 1980. Dette resulterer i at skewness blir som i figur 6.2, til stor forskjell fra normalfordeling. Kurtosis er høyt positiv av samme grunn. Gjennomsnittlig byggealder i variabelen er 1987. Men andelen hus bygd før 1980 er betydelig lavere enn andelen hus bygd etter nevnte dato.

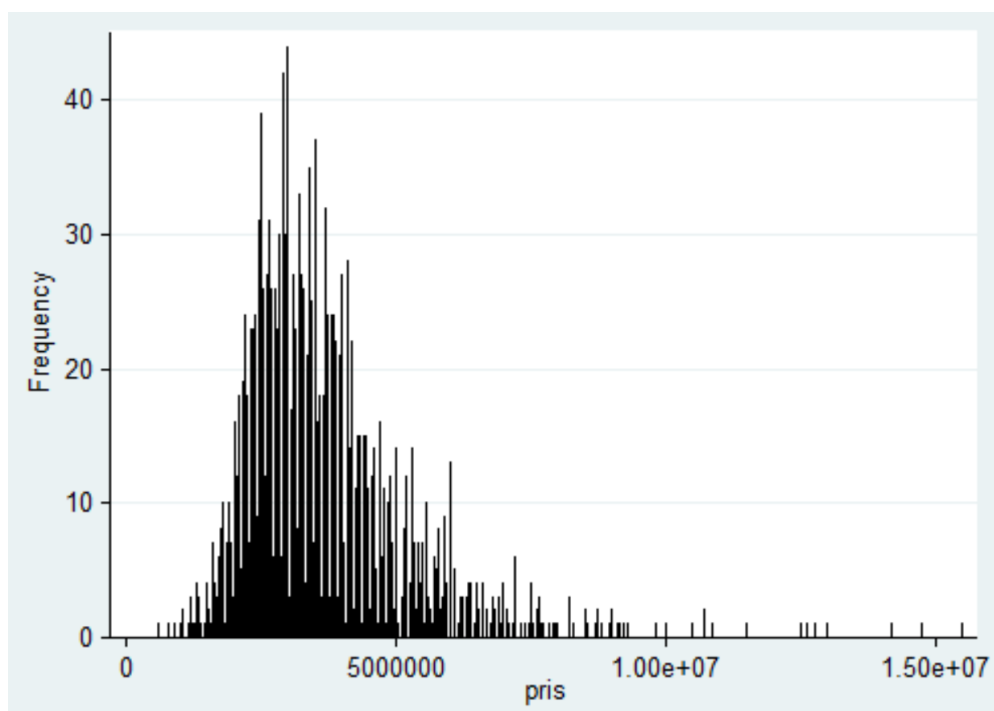
6.3 FREKVENSTABELLER FOR UTVALGTE VARIABLER

Her vil jeg skille mellom ulike typer variabler. Velger å dele opp i kontinuerlige variabler og diskrete variabler. De kontinuerlige variablene i mitt tilfelle, er pris, boa/P-rom, byggedato og postnummer. De diskrete variablene er dummyvariablene. Jeg vil her vise frekvenstabeller for de kontinuerlige variablene, pris og bygningsår, samt dummyvariabelen garasje. Jeg velger å presentere variablene i histogrammer, da jeg føler det gir best oversikt. Histogrammer for alle variable ligger vedlagt.

Salgspris

Nedenfor følger et histogram av variabelen salgspris. Jeg har delt opp i 2298 søyler, en for hver enebolig i variabelen. Her ser vi pris per bolig på X-aksen og frekvensen på Y-aksen.

Tabell 6.3 Frekvenstabell for salgspris.

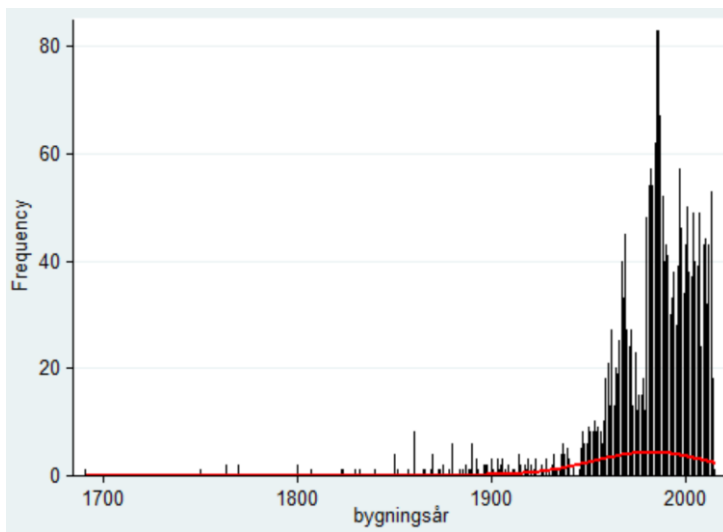


Som vi ser av tabellen, ligger de fleste observasjonene samlet, mens enkelte ekstremverdier ligger langt til høyre i histogrammet. Disse eneboligene er de boligene jeg omtalte som ekstremverdier i kapittel 5.2

Byggedato

Nedenfor følger et histogram over variabelen byggeår. I histogrammet er det også tegnet en normalfordelingskurve, for å få en grafisk fremstilling av teorien jeg gjennomgikk i kap. 6.2. Normalfordelingskurven er tegnet rød. Y-aksen representerer frekvens, altså hvor mange eneboliger som er bygd i det enkelte år. Fra fremstillingen kan vi se at størstedelen av observasjonene er plassert helt til høyre på X-aksen, bygningsår. Også her har vi noen «ekstremverdier», dvs. ett fåtall boliger bygd før 1800tallet.

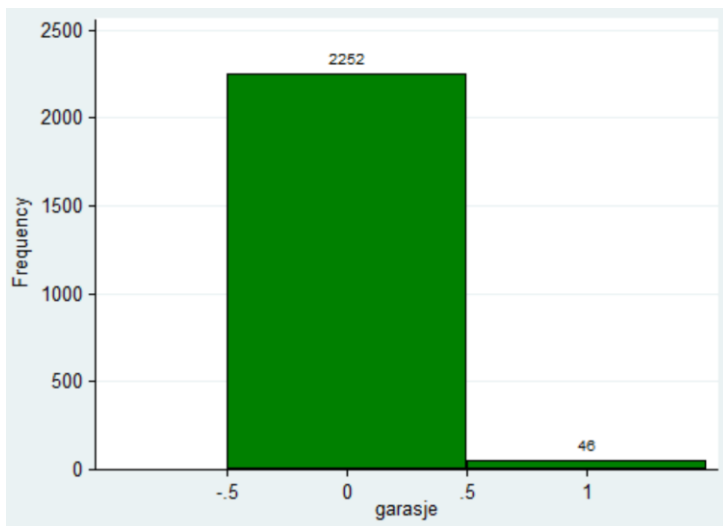
Tabell 6.4 Frekvenstabell for byggeår



Garasje.

Nedenfor følger et histogram av den uavhengige variabelen garasje. Her ser man tydelig at andelen boliger med garasjeleilighet er betydelig lavere enn andelen boliger uten garasjeleilighet. Vi har 2252 observasjoner uten garasjeleilighet, 46 observasjoner med.

Tabell 6.5 Frekvenstabell for garasjeleiligheter



Byer.

Jeg velger å legge ved en fordeling av byene i utvalget mitt. Denne kan sees under. Som vi ser av tabellen er fordelingen av enheter i hver by nokså ujevn. Dette skyldes faktorer som antallet eneboliger solgt i en gitt by, men i hovedsak skyldes dette min utvelgelsesprosess som er gjort rede for tidligere i kapittelet. I Kristiansand gjennomgikk jeg alle eneboliger solgt i en periode på 26 måneder, mens i Lillesand gikk jeg gjennom eneboligene solgt i en periode på 14 måneder. Årsaken til at totalsummen i kolonnen for prosent er 0,17% for stor, er avrundinger.

Tabell 6.6 Prosentvis fordeling av enheter fordelt mellom byene

By	Antall	%	Garasjeleilighet
Kristiansand	886	38,71	15
Grimstad	375	16,32	10
Arendal	221	9,62	8
Drammen	113	4,92	1
Lillesand	39	1,7	1
Sandefjord	224	9,75	5
Skien	207	9,01	2
Tønsberg	233	10,14	4
Total:	2298	100,17	46

6.4 KORRELASJON

Jeg uttalte meg kort om korrelasjon i kapittel 4 under punktet multikollinearitet. I dette delkapittelet vil jeg gå nærmere inn på korrelasjon. Korrelasjon måler samvariasjon mellom to variabler. Samvariasjon er et mål på styrken og retningen på den lineære avhengigheten mellom to variabler. Det finnes ulike korrelasjonskoeffisienter, men jeg vil benytte meg av pearsons r. Pearsons r har den egenskapen at den gir mål på samvariasjon i intervallet mellom -1 og 1, hvor ytterpunktet 1 representerer perfekt positiv relasjon mellom de to variablene. Dette vil si at dersom en av variablene øker eller reduseres i verdi, vil den andre variabelen opptre identisk. Dersom pearsons r er -1 er vi i en situasjon med perfekt negativ korrelasjon. Variablene har da et inverst forhold, slik at dersom en variabel øker vil den andre variabelen reduseres tilsvarende. Hvis r er 0, er variablene totalt uavhengige av hverandre. I min oppgave vil samvariasjon mellom den avhengige og de uavhengige variablene være en nødvendig, men ikke tilstrekkelig, forutsetning for å avdekke kausalitet(sammenheng) i alternativ-hypotesene. Samvariasjon og sammenheng må ikke forveksles, da samvariasjon bare måler en del av sammenhengen. For å kunne påberope sammenheng mellom variablene, må tre krav være oppfylt. Samvariasjon er

allerede nevnt, men årsak må også være før virkningen i tid. I så måte kan det argumenteres for at jeg i arbeidet med innsamlingen av data, burde sett på eneboliger som nylig er omsatt uten garasjeleilighet, for så å få oppført en garasje med leilighet og deretter solgt med tilhørende leilighet. Det sier seg selv at dette er en tidkrevende og omstendelig prosess, som nok ikke lar seg gjennomføre i praksis. Det siste kravet er det vanskeligste kravet å følge opp. Her må vi ha kontroll over alle relevante variabler (Jacobsen(2005)). Å finne alle attributter som er med på å styre eiendomsprisen er svært krevende og det å ha kontroll over alle de relevante variablene blir således en vanskelig oppgave.

I neste kapittel vil jeg estimere modeller. En av modellene vil være multippel lineær regresjon. Å bestemme variablene som skal tas med i denne analysen er enklere etter at jeg har utført korrelasjonsmatriser. Dersom jeg i matrisen finner ut at enkelte uavhengige variabler har verdi på 0, sett i sammenheng med den avhengige variabelen, er det ikke hensiktsmessig å ha disse med i den lineære regresjonen. Dette er fordi de ikke er med å «driver» den avhengige variabelen. Man må også være oppmerksom på dersom uavhengige variabler har verdier seg i mellom som nærmer seg ytterpunktene 1 og minus 1. I disse tilfellene kan det ha oppstått multikollinearitet, se kapittel 4.2 om multikollinearitet. Dersom dette er tilfelle, bør kanskje den ene av de uavhengige variablene tas ut av regresjonen.

Med tanke på modellvalget, multippel lineær regresjon, velger jeg å gå gjennom en bivariat korrelasjonsanalyse. Denne analysen måler direkte samvariasjon mellom de ulike uavhengige variablene på den avhengige variabelen, samt at den måler samvariasjonen mellom de uavhengige variablene. Nedenfor følger en tabell over korrelasjonsmatrisen. Jeg utelater postnummervariablene, da det er 112 stykker av disse

Tabell 6.7 Bivariat korrelasjonsanalyse

```
. correlate pris boaprom bygninger garasje salgsaar Kristiansand Grimstad Arendal Drammen Lillesand Sandefjord Skien Tønsberg
(obs=2298)
```

	pris	boaprom	bygning-r	garasje	salgsaar	Kristi-d	Grimstad	Arendal	Drammen	Lilles-d	Sandef-d	Skien	Tønsberg
pris	1.0000												
boaprom	0.5974	1.0000											
bygninger	0.1257	0.0622	1.0000										
garasje	0.0918	0.0843	0.0094	1.0000									
salgsaar	0.0743	0.0269	0.0184	-0.0171	1.0000								
Kristiansand	0.1022	0.0668	-0.3169	-0.0111	-0.0502	1.0000							
Grimstad	-0.1677	-0.0875	-0.1378	0.0210	-0.0545	-0.3498	1.0000						
Arendal	-0.1181	-0.0491	0.1280	0.0271	-0.0490	-0.2584	-0.1440	1.0000					
Drammen	0.1155	0.0315	0.1021	-0.0181	0.1931	-0.1801	-0.1004	-0.0742	1.0000				
Lillesand	-0.0115	-0.0301	0.0550	0.0053	0.1116	-0.1041	-0.0580	-0.0429	-0.0299	1.0000			
Sandefjord	0.0751	0.0207	0.1711	0.0054	0.0115	-0.2603	-0.1451	-0.1072	-0.0747	-0.0432	1.0000		
Skien	-0.1183	0.0278	0.1398	-0.0233	-0.0316	-0.2492	-0.1389	-0.1026	-0.0716	-0.0413	-0.1034	1.0000	
Tønsberg	0.1164	-0.0089	0.1572	-0.0068	0.0282	-0.2661	-0.1483	-0.1096	-0.0764	-0.0441	-0.1104	-0.1057	1.0000

Velger i det videre å kommentere resultatene for boa/P-rom, bygningsår og garasje, mot pris, da det er disse variablene som er av interesse for mine hypoteser. Variablene har alle positiv korrelasjon og er signifikant på et 0,05 nivå. Variabelen garasje korrelerer positiv med pris, men ikke på et signifikant nivå. Dette kan bety at garasjeleilighet ikke gir særlig utslag på prisen, men det kan også ha sammenheng med begrensninger i datagrunnlaget.

7. ESTIMERING OG TESTING AV HYPOTESER

7.1 INNLEDNING

I forrige kapittel presenterte jeg hovedmengden av datamaterialet mitt. I dette kapittelet vil jeg gå videre med å forsøke å besvare problemstillingen i oppgaven, ved å estimere regresjoner og teste hypotesene mine. Jeg presenterte analyseteknikken korrelasjon i slutten av forrige kapittel og vil derfor starte dette kapittelet med å se på momenter i (Hair m.fl.(2006)) sin 6 steps prosess for å estimere en regresjonsmodell, for så å så sette opp en multippel regresjon og vurdere hvorvidt det er hensiktsmessig å teste andre regresjonsformer.

Det første steget i modellen er å se på hva man ønsker å gjøre og hvor man vil ende opp, bestemme målene for analysen. Her inngår det å identifisere variabler man ønsker å ha med i undersøkelsen. Målet for analysen er konkretisert i problemstillingen i oppgaven(kapittel1) og ytterligere operasjonalisert gjennom testbare hypoteser (kapittel3). Ved å bruke multippel regresjon ønsker jeg å lage en prisfunksjon som kan besvare hypotesene. Den avhengige variabelen er pris. Ved valg av forklaringsvariabler, tok jeg utgangspunkt i deres forventede relasjon til prisen. Her inngår de kontinuerlige variablene for byggeår og boa/P-rom, samt en

dummy for garasjeleilighet og for salgsår. I fra faget eiendomsøkonomi og Alonso-muth-mills modellen i teorikapittelet, fremkommer det at den viktigste driveren til boligpris er lokasjon. Jeg har derfor lagt inn postnummer som et sett av dummyvariabler, da postnummer fanger opp beliggenhet.

Det andre steget i modellen er å velge forskningsdesign. Valg av antall variabler, samt utvalgsstørrelse står sentralt her. Utvalgsstørrelsen påvirker statistisk signifikans, fordi antallet observasjoner inngår i beregningens forklaringsgrad, R^2 . Mitt utvalg er på 2298 enheter, dette er nokså stort og kan føre til at vi ser en tendens hvor alle variabler blir regnet som signifikante. Det tredje steget i modellen går ut på å teste variablene. Man tester om forutsetningene for gjennomføring av regresjon ligger til grunn, for den enkelte variabel. Forutsetningene som bør være oppfylt for hver variabel, nevnt i kapittel 4.2, er bl. a normalfordeling. Tidligere i oppgaven, bl. a ved å se på skjevhet og spissitet er det klare indikatorer på at dette ikke er tilfelle. Dette kan løses ved å transformere de aktuelle variablene. Dette kan eksempelvis gjøres ved å ta logaritmen av variabelen. Ved å gjøre dette, vil en ikke-lineær variabel bli gjort lineær. Dette vil jeg se på senere i kapittelet.

7.2 MULTIPPEL LINEÆR REGRESJON

Vi husker fra kapittel 4 at den matematiske utledningen for multippel lineær regresjon er som følger;

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k + \varepsilon$$

Jeg skriver likningen inn i stata og tar i første omgang med alle relevante variabler, men viser ikke data for de fleste dummyvariable for postnummer. Jeg har to sett av dummyvariabler. I det ene settet har vi byene i utvalget, i det andre har vi postnumre. Som nevnt i kapittel 4 må en dummy i fra hvert sett holdes utenfor regresjonen. Jeg holder variabelen Drammen og variabelen Postnr 112 utenfor. I den grafiske fremstillingen utelar jeg postnummerdummyene, for å få en kort og oversiktlig tabell.

Hele tabellen kan sees vedlagt. Estimatene følger i tabellen under. I Tabellen sees avhengig variabel pris, med de uavhengige variablene loddrett under. Vannrett på samme linje som pris, sees ulike statistiske observasjoner, disse vil forklares i sammenheng med estimatene, under tabellen.

Tabell 7.1 Estimatenes ved multippel lineær regresjon

Source	SS	df	MS	
Model	2.8848e+15	116	2.4869e+13	Number of obs = 2298
Residual	1.9660e+15	2181	9.0141e+11	F(116, 2181) = 27.59
Total	4.8508e+15	2297	2.1118e+12	Prob > F = 0.0000
				R-squared = 0.5947
				Adj R-squared = 0.5732
				Root MSE = 9.5e+05

pris	Coef.	Std. Err.	t	P> t
boaprom	15138.11	405.7387	37.31	0.000
byggningsr	8137.207	847.1632	9.61	0.000
garasje	594232.5	145477.2	4.08	0.000
salgsaar	86530.94	42346.97	2.04	0.041
Kristiansand	-1805606	492976	-3.66	0.000
Grimstad	-1771693	403462.6	-4.39	0.000
Arendal	-2267873	656177.3	-3.46	0.001
Lillesand	-6492.174	555994.9	-0.01	0.991
Sandefjord	-1379919	556411.4	-2.48	0.013
Skien	-2085425	507844.6	-4.11	0.000
Tønsberg	-580386.1	508668.8	-1.14	0.254
_cons	-1.40e+07	1723496	-8.13	0.000

Jeg vil i det videre forklare de relevante statistiske uttrykkene og estimatene deres.

R-squared: Forklart i kapittel 4. 59,38% av variansen i avhengig variabel, pris, er forklart gjennom de uavhengige forklaringsvariablene.

Coef: Angir hver enkelt variabls regresjonskoeffisient (β_i). I lineær regresjon er disse partielle. De angir hvilken påvirkningskraft endring i en enhet i dens tilhørende uavhengige variabel, får på avhengig variabel. F. eks vil en økning i boa på $1m^2$ gi prisøkning på drøyt 15.150 kr.

Helt nederst i kolonnen for variablene, finner vi `_cons`, som står for konstantleddet, α .

Vi ser at den kontinuerlige variabelen, byggningsår, har en koeffisient på 8137. Hver årlige økning i byggningsår, vil da øke prisen med byggeåret*8137. For eksempel får en bolig fra 1850, $1850*8137= 15.053.450$ kr i utslag på prisen. Dette ser ved første øyekast voldsomt ut, men vær obs på at konstantleddet har en verdi på minus 14.000.000 i denne regresjonsformen.

Std.Err: Angir hver enkelt regresjonskoeffisient sin standardfeil.

t: Tallene i denne kolonnen står for statistikken en variabel har fått ved en tosidig T-test. I en T-test formuleres en nullhypotese, der β ikke er signifikant fra null. Denne hypotesen forkastes dersom observert t-verdi er innenfor intervallet på $100(1-a)$ prosent. a representerer feilmargen, som i dette tilfellet er satt til 0.05. Intervallet blir da på 95%. Ytterpunktene for hvor

nullhypotesen er gyldig ligger fra -1,96 til 1,96. Det vil si at dersom tallet i kolonne t er større enn 1.96 eller mindre enn minus 1.96 kan nullhypotesen forkastes. Jeg går kolonne t og ser på t-verdien for garasje, boa/P-rom og Byggeår. Disse har t-verdier på 4,1 og 37,41 og 9,64. Med andre ord taler dette for at vi kan forkaste alle nullhypotesene.

Tester nå VIF for å se hvorvidt de uavhengige variablene korrelerer med hverandre. I den grafiske fremstillingen tar jeg bare med observasjoner med VIF over 10.

Tabell 7.2 VIF test, lineær regresjon

Variable	VIF	1/VIF
Kristiansand	146.77	0.006813
Arendal	95.41	0.010481
Skien	53.89	0.018556
Grimstad	52.59	0.019017
Sandefjord	36.25	0.027590
Tønsberg	34.04	0.029379
Lillesand	13.15	0.076059
Mean VIF	7.07	

Her ser vi at alle dummyvariablene for by ligger langt over grenseverdien på 10, noe som indikerer multikolliaritet.

Jeg velger nå å trekke by-dummyene helt ut av regresjonen og gjør som tidligere og holder postnummerdummy, postnr112 utenfor regresjonen.

Multipel regresjon, ny modell

Tabell 7.3 Estimaterne ved multipel regresjon, ny modell

Source	SS	df	MS	
Model	2.8841e+15	115	2.5079e+13	Number of obs = 2298
Residual	1.9666e+15	2182	9.0130e+11	F(115, 2182) = 27.83
Total	4.8508e+15	2297	2.1118e+12	Prob > F = 0.0000
				R-squared = 0.5946
				Adj R-squared = 0.5732
				Root MSE = 9.5e+05

pris	Coef.	Std. Err.	t	P> t
boaprom	15148.2	405.5438	37.35	0.000
bygningssr	8126.037	847.0122	9.59	0.000
garasje	593883.8	145467.8	4.08	0.000
salgsaar	86210.91	42342.77	2.04	0.042
_cons	-1.58e+07	1693397	-9.34	0.000

De statistiske uttrykkene er i dette tilfellet omtrent de samme som jeg gjorde rede for tidligere. R^2 har falt minimalt, til 0,5946. Dette indikerer som tidligere nevnt at de uavhengige variablene forklarer 59,46% av variasjonen i pris. Den prosentandelen det ikke er gjort rede for, $100 - 59,46 = 40,54\%$, forklares av restleddet ϵ . Forklaringskraften har altså falt med 0.01 i forhold til regresjonen hvor by-dummyene var inkludert. Regresjonsmodellen har et konstantledd på $-1.58e+07$. Denne størrelsen indikerer prisen på en teoretisk bolig, dersom de andre variablene settes til verdien null. De kontinuerlige variablene boaprom og byggedato har betastørrelser på hhv. 15 148 og 8126 kr. Dersom disse variablene økes med én enhet, vil prisen økes med disse summene. La oss f.eks si at boaprom øker med 1kvm. Da vil prisen øke med $15\ 148 \times 1 = 15\ 148$ kr. Øker boaprom med 2kvm, øker prisen med $15\ 148 \times 2 = 30\ 296$ kr. Da har vi med andre ord ingen marginalt avtakende økning, noe som er rimelig å anta ut fra økonomisk teori. Konsumentene vil på et eller annet tidspunkt bli mett og økningen vil flate ut. Dette gjenspeiles ikke i lineær regresjon, men hensyntas i logaritmisk regresjon. Koeffisientene i en slik modell tolkes nemlig som priselastisiteter og ikke konstante marginale økninger, slik som i lineær regresjon.

En god del av dummyvariablene for postnummer ligger ikke innenfor et signifikansnivå på 0.05, men jeg vil uansett beholde alle dummyene i regresjonsanalysene, fordi hensikten med disse dummyene er å ivareta en boligs beliggenhet. De er også en del av et sett og dersom jeg fjerner noen utvalgte dummyer, kan dette gi utslag for de gjenværende dummyene.

Jeg foretar i det følgende en ny VIF-test. Legger bare ved de høyeste VIF og utelar dermed størstedelen av variablene i den grafiske fremstillingen:

Tabell 7.4 VIF-test, lineær regresjon, ny modell

Variable	VIF	1/VIF
postnr104	1.63	0.614745
bygninger	1.63	0.615174
postnr108	1.56	0.642568
postnr80	1.55	0.644664
postnr71	1.48	0.675509
postnr85	1.47	0.681496
postnr105	1.40	0.714184
postnr110	1.39	0.718637
postnr74	1.39	0.718974
postnr87	1.37	0.728495
Mean VIF	1.17	

Vi ser nå at alle VIF ligger under 10 og er således på et akseptabelt nivå. Gjennomsnittlig VIF på 1,17 er også godt innenfor akseptabelt nivå.

Jeg vil nå teste restleddet for å se om det foreligger homoskedasitet. Dette velger jeg å gjøre ved hjelp av Breusch-Pagan testen. Denne testen konstruerer en null-hypotese som sier at det foreligger konstant varians, altså homoskedasitet, og tester denne. Som vi ser av tabellen under, kan nullhypotesen forkastes for residualene. Vi kan konkludere med at det foreligger heteroskedasitet.

Tabell 7.5 Breusch-pagan test. Multippel lineær regresjon.

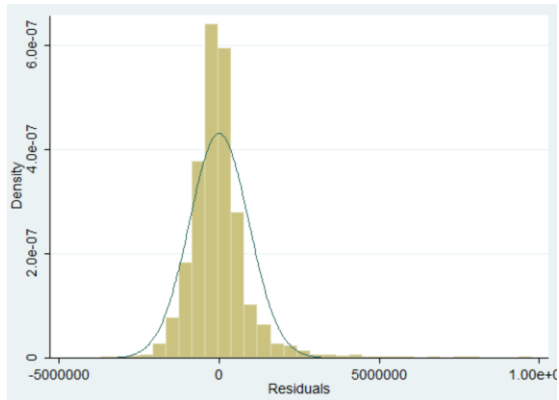
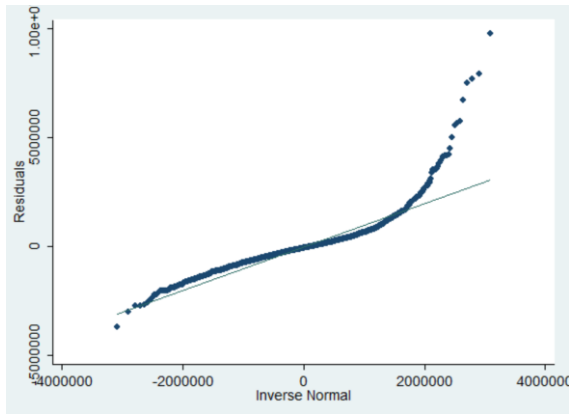
```
. predict residual
(option xb assumed; fitted values)

. hettest residual

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

chi2(1)      = 1219.11
Prob > chi2  = 0.0000
```

Den siste testen jeg vil gjøre, er å sjekke om restleddet er tilnærmet normalfordelt eller ei. Kommandoene jeg har skrevet inn i stata, er som tidligere nevnt vedlagt. Valg av grafisk modell for å sjekke for normalfordeling, følger under. Jeg velger å bruke både histogram med normalfordeling og normalskråplott.



Figur 7.1 Normalskråplott, test av restledd **Figur 7.2** Histogram, test av restledd

Som vi ser av den første grafen, ligger den tykke linjen omtrent på linje med den smale linjen gjennom det meste av grafen, før den skrår oppover. Dette henger sammen med ekstremverdiene jeg omtalte i kapittel 5.2. I histogrammet ser det ut til at restleddet er rimelig normalfordelt. Perfekt «fit» er høyst usannsynlig i undersøkelser. Min subjektive mening er at restleddet tilsynelatende ser nok så normalfordelt ut.

Jeg har også kjørt en lineær regresjon der jeg utelot alle postnummer-variablene. Dette gjorde jeg for å se om by-dummyene alene kunne «matche» postnummer-dummyene sin forklaringskraft. Den grafiske fremstillingen sees nedenfor.

Tabell 7.6 Lineær regresjon, uten postnummer-dummy

Source	SS	df	MS	
Model	2.0892e+15	11	1.8993e+14	
Residual	2.7615e+15	2286	1.2080e+12	
Total	4.8508e+15	2297	2.1118e+12	

Number of obs =	2298
F(11, 2286) =	157.22
Prob > F =	0.0000
R-squared =	0.4307
Adj R-squared =	0.4280
Root MSE =	1.1e+06

	pris	Coef.	Std. Err.	t	P> t
	boaprom	16076.37	450.0252	35.72	0.000
	byggningsr	4677.51	844.0911	5.54	0.000
	garasje	487306.7	164545.7	2.96	0.003
	salgsaar	86690.05	47847.81	1.81	0.070
Kristiansand		-351859.2	114005.6	-3.09	0.002
Grimstad		-870416.4	121947.2	-7.14	0.000
Arendal		-984017.2	129495.4	-7.60	0.000
Lillesand		-563687.9	204328.5	-2.76	0.006
Sandefjord		-324880.5	128299.9	-2.53	0.011
Skien		-1201824	130495.8	-9.21	0.000
Tønsberg		-71409.64	127338.7	-0.56	0.575
_cons		-7895051	1685290	-4.68	0.000

Av tabellen ser vi at R^2 har falt markant uten postnummer-dummyene. Da by-dummyene ikke gir betydelig økt forklaringskraft og det foreligger multikollinearitet mellom by og postnummer-

dummyer, velger jeg å ikke innlemme by-dummyer i resten av regresjonene jeg kjører. Jeg vil i det videre arbeidet undersøke om log-lineær funksjon er en bedre regresjonsform for min oppgave.

7.4 DOBBEL LOGARITMISK REGRESJON

Det er flere muligheter for valg av logaritmisk eller logaritmisk-lineær funksjonsform, jeg vil i første omgang se på dobbelt logaritmisk funksjonsform og evt. se videre på en annen funksjonsform dersom det syns å være behov for det. Den matematiske formelen for dobbelt logaritmisk regresjon er oppgitt i kapittel 4 og gjentas under, med små endringer.

$$\ln Y_i = \alpha + \beta_1 \ln X_i + \beta_2 D_i + \varepsilon_i$$

Vi skriver de kontinuerlige variablene (her uttrykt som Y_i og X_i) på logaritmeform. For denne oppgaven blir dette pris, boa/P-rom og byggeår. Dummyvariablene har jeg i formelen ovenfor gitt uttrykket (D_i) og man tar ikke logaritmen av disse. Denne modellen skiller seg fra multippel lineær regresjon, ved at vi ser på hvor mye avhengig kontinuerlig variabel, pris, endres med en prosent økning i de kontinuerlige uavhengige variablene. Dummyvariablene angir prosentvis endring i prisen, hvis hver enkelt av disse variablene inntreffer. Nedenfor følger resultatene fra denne regresjonen. Her, som i fremstillingen for lineær regresjon, tar jeg ikke med postnummervariablene i selve tabellen.

Tabell 7.7 Estimaten ved dobbel logaritmisk regresjon

Source	SS	df	MS	Number of obs = 2298
Model	204.668591	115	1.77972688	F(115, 2182) = 38.78
Residual	100.132776	2182	.045890365	Prob > F = 0.0000
Total	304.801367	2297	.132695415	R-squared = 0.6715
				Adj R-squared = 0.6542
				Root MSE = .21422

lnpris	Coef.	Std. Err.	t	P> t
lnboaprom	.6673305	.0149981	44.49	0.000
lnbyggningsaar	4.158042	.3686837	11.28	0.000
garasje	.1643416	.0327999	5.01	0.000
salgsaar	.022611	.0095547	2.37	0.018
_cons	-20.15538	2.795643	-7.21	0.000

Hvis vi sammenlikner med tabell 7.3, ser vi at R^2 har økt, forklaringsgraden har dermed økt i denne regresjonsformen. Boa/P-rom har her en koeffisient på 0,667. Det betyr at dersom

boligarealet øker en prosent, vil prisen øke med 0,667 prosent. Jeg sjekker om det foreligger multikollinearitet, og foretar derfor en VIF-test.

Tabell 7.8 VIF test, dobbelt logaritisk regresjon

Variable	VIF	1/VIF
postnr104	3.70	0.270086
Mean VIF	1.66	

Nå som by-dummyene er tatt ut av regresjonen, ligger høyeste VIF på 3,7 mens gjennomsnittlig VIF er 1,66. Dette er langt innenfor referanseområdet.

Kjører så en test for om det foreligger heteroskedastisitet, samme breusch-pagan test som jeg bruke i multipel lineær regresjonsform.

Tabell 7.9 Breusch-Pagan test, dobbel log regresjon

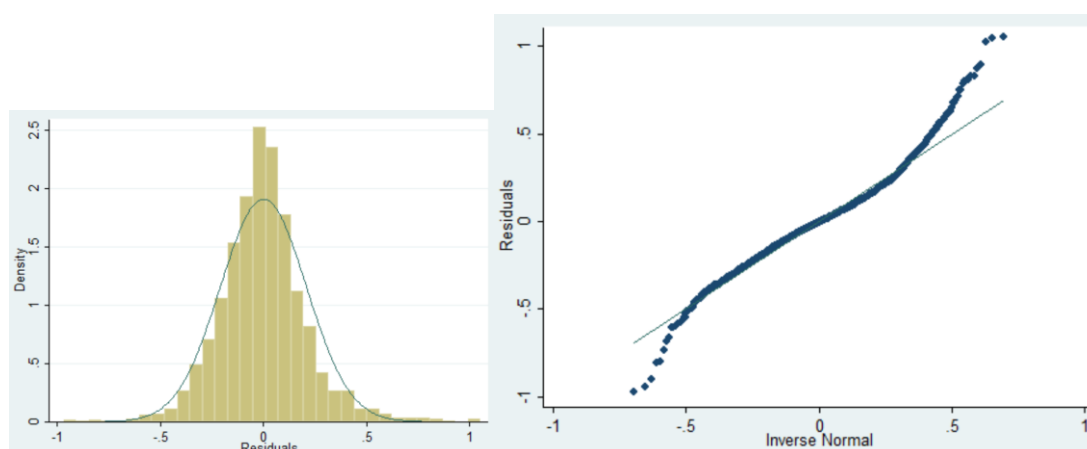
```

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

chi2(1)      =    35.67
Prob > chi2  =    0.0000
    
```

I dette tilfellet er $\chi^2 > \text{prob}$ for residualene. Det foreligger altså heteroskedastisitet.

Sjekker til sist om restleddet er normalfordelt eller ei. Benytter igjen samme metode som under multipel lineær regresjon.



Figur 7.3 Histogram med normalfordeling i dobbel log regresjon. **Figur 7.4** Normalskråplott i dobbel log regresjon.

Hvis jeg sammenlikner histogrammet, med histogrammet i Figur 7.2, ser denne versjonen mer normalfordelt ut. I normalskråplottet ser den tykke linjen nærmest symmetrisk ut, da den ligger oppå den tynne linjen i intervallet mellom -0,5 og 0,5, mens den fraviker fra linjen på nærmest identisk vis for verdier lavere enn -0,5 og verdier over 0,5.

7.5 SEMI-LOGARITMISK REGRESJONSANALYSE

Denne regresjonsformen var opprinnelig ikke tenkt gjennomgått, og er således ikke forklart i kapittel 4. I denne regresjonsformen bruker jeg ln for den avhengige variabelen, pris, mens resten av forklaringsvariablene står på lik linje som i lineær regresjon. Dersom de kontinuerlige variablene i denne modellen øker med en enhet, vil tilhørende koeffisient ganger hundre, si hvor mange prosent prisen vil øke eller reduseres med. For dummyvariablene vil prisen øke med den prosentsatsen koeffisienten har, dersom kjennetegnet til dummyen er til stede. I tabellen utelar jeg igjen postnummerdummyene.

Tabell 7.10 Estimerer ved semi-logaritmisk regresjon

Source	SS	df	MS	
Model	199.421958	115	1.73410398	
Residual	105.37941	2182	.048294871	
Total	304.801367	2297	.132695415	

Number of obs =	2298
F(115, 2182) =	35.91
Prob > F =	0.0000
R-squared =	0.6543
Adj R-squared =	0.6360
Root MSE =	.21976

lnpris	Coef.	Std. Err.	t	P> t
boaprom	.0039442	.0000939	42.01	0.000
bygninger	.0023611	.0001961	12.04	0.000
garasje	.1562439	.0336731	4.64	0.000
salgsaar	.0232835	.0098016	2.38	0.018
_cons	9.446492	.3919901	24.10	0.000

Som vi ser er alle variable signifikante på et 0.05 nivå. Dette gjelder dog ikke alle postnummer-dummyer, men som tidligere nevnt tillegger jeg ikke dette vekt, da jeg antar at det skyldes ulik mengde enheter i de forskjellige postnummer-dummyene, samt at det ikke bærer noen hensikt å sortere ut de som ikke er innenfor 0.05 nivå. Videre kan vi se at forklaringsgraden er gått minimalt ned. Vil i det videre sjekke restleddet for homoskedastitet, om vi har multikollinearitet

og normalfordeling. Begynner med VIF-test for multikollinearitet, tar bare inn høyeste verdi for VIF og gjennomsnittlig VIF.

Tabell 7.11 VIF, semi-logaritmisk regresjon.

Variable	VIF	1/VIF
postnr104	3.70	0.270086
Mean VIF	1.66	

Høyeste VIF er på 3,7 og gjennomsnittlig VIF er på 1,66. Dette er langt innenfor godkjent nivå. Legg også merke til at ingen av de kontinuerlige variablene er å finne i tabellen, da disse lå langt nede med en VIF på omtrent 1.

Kjører så en test for om det foreligger heteroskedastisitet, samme breusch-pagan test som jeg bruke i multippel lineær regresjonsform og dobbel logaritmisk regresjonsform.

Tabell 7.12 Breusch-Pagan test, semi-logaritmisk regresjon

```

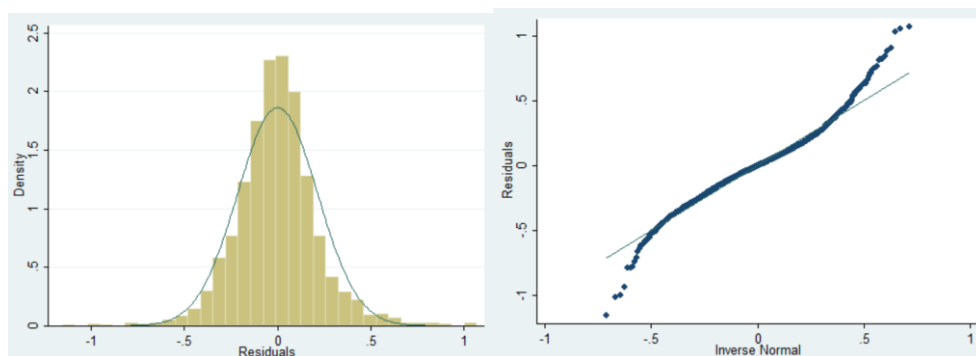
Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
Ho: Constant variance
Variables: residual

chi2(1)      =    40.69
Prob > chi2  =    0.0000

```

Igjen indikeres det heteroskedastisitet.

Til slutt vil jeg se om restleddet er normalfordelt eller ei. Benytter igjen samme metode som under multippel lineær regresjon.



Figur 7.5 Histogram med normalfordeling i semi log regresjon. **Figur 7.6** Normalskråplott i semi log regresjon.

Kommenterer ikke dataene ytterligere, da de er relativt lik dataene for dobbelt-log regresjon.

7.6 VALG AV MODELL

Før jeg starter med hypotesetestingen, må jeg bestemme hvilken funksjonsform jeg vil benytte meg av. Kommentarer til funksjonene er gitt underveis i kapitlet og på bakgrunn av mine funn og begrunnelser gitt i kommentarer, velger jeg å benytte meg av semi logaritmisk regresjonsmodell. Kunne like gjerne valgt dobbelt log regresjonsmodell, da disse modellene gav relativt like data, men jeg velger å benytte meg av semi-log modell. Både i denne modellen, i multipel lineær regresjonsmodell og dobbelt logaritmisk modell, har jeg sterke indikatorer på at det foreligger heteroskedastisitet. Dette medfører ikke feil i parameterestimatet, men kan gi gal standardfeil og følgelig feil i signifikanstest. (Tuftes, Per Arne(2000)⁵). Jeg vil derfor opprette en variabel kalt «Z». For denne variabelen vil jeg bruke kommandoen cluster, for å teste om restleddene er uavhengige av hverandre og for å rette opp eventuelle gale standardfeil, som følge av heteroskedastisitet. Jeg velger også å redefinere variabelen bygningsdato til «boligalder». Dette gjør jeg fordi jeg ser av andre oppgaver at det er vanlig å bruke alder, fremfor dato. Derfor tenker jeg at å redefinere denne variabelen, vil gjøre mine resultater mer forståelig for den uinnvidde leser, samt at resultatene enklere lar seg overføre til tilsvarende oppgaver og analyser.

Semi-logaritmisk regresjonsanalyse, endelig modell:

Tabell 7.13 Semi-log regresjon, endelig modell

```
Number of obs = 2298
F( 3, 111) = .
Prob > F = .
R-squared = 0.6543
Root MSE = .21976

(Std. Err. adjusted for 112 clusters in Z)
```

lnpris	Coef.	Robust Std. Err.	t	P> t
boaprom	.0039442	.0001575	25.04	0.000
boligalder	-.0023617	.0005328	-4.43	0.000
salgsaar	.0232863	.0094866	2.45	0.016
garasje	.1562441	.0331826	4.71	0.000
_cons	14.20421	.0255301	556.37	0.000

Alle variable er fortsatt signifikante. Variabelen boligalder har nå negativt fortegn i koeffisienten, men tallmaterialet er fortsatt det samme. Dette skyldes naturlig nok fordi jeg nå

tar utgangspunkt i alder. Koeffisienten for byggeår var positiv tidligere, fordi jeg tok utgangspunkt i byggedato.

7.7 TESTING AV HYPOTESER

I kapittel 3 lagde jeg tre hypotesesett. Disse vil det følgende delkapittelet testes, ved hjelp av den semi logaritmiske regresjonsanalysen.

Hypotesesett1/Hovedhypotese:

H0: Garasjeleilighet har ingen betydning for eneboligens salgspris

H1: Garasjeleilighet har betydning for boligens salgspris

Koeffisienten til variabelen garasje har en p-verdi på 0,000. Dette tilsier at den er signifikant på et 5 prosent nivå. Jeg kan dermed med 95% sannsynlighet fastslå at garasjeleilighet har betydning på prisen. Den samme koeffisienten viser at betydningen er på 15,62 prosent. Dvs at dersom fenomenet garasjeleilighet er til stede, vil den øke prisen til boligen med 15,62 prosent. Nullhypotesen kan dermed forkastes og alternativhypotesen beholdes.

Hypotesesett2/kontrollhypotese:

H0: Boareal har ingen betydning for boligens salgspris

H1: Boareal har betydning for boligens salgspris

Koeffisienten til denne variabelen (boaprom) har en p-verdi på 0,000. Dette tilsier at den er signifikant på et 5 prosent nivå. Jeg kan dermed med 95 prosent sannsynlighet fastslå at boareal har betydning for boligens salgspris. Boa-koeffisienten har en høy positiv verdi, noe som indikerer at boligareal er et svært priselastisk attributt. Nullhypotesen kan dermed forkastes og alternativhypotesen beholdes.

Hypotesesett3/kontrollhypotese:

H0: Boligens alder har ingen betydning for salgsprisen

H1: Boligens alder har betydning for salgsprisen

Koeffisienten til variabelen boligalder har en p-verdi på 0,000, som indikerer at den er signifikant på et 5 prosent nivå. Kan da med 95 prosent sannsynlighet fastslå at bygningsåret

har betydning for boligens salgspris. Koeffisienten har negativt fortegn. Boligprisen faller med økning i alder. Nullhypotesen forkastes dermed og alternativhypotesen beholdes.

8. NÆRMERE DRØFTING AV

HOVEDPROBLEMSTILLING

8.1 DRØFTING

I dette kapittelet vil jeg se nærmere på resultatene fra analysen. Den endelige analyseformen jeg falt ned på fremgår av tabell 7.13. I slutten av kapittel 6 testet jeg hypotesene mine, på bakgrunn av funn i nevnte analyse. Funnene støttet opp om problemstillingen i oppgaven og var i tråd med teorien fra kapittel 3.

Hovedhypotesen om at garasjeleilighet har betydning for boligens pris, har empirisk støtte i alle de tre regresjonsmodellene jeg har satt opp og koeffisienten er ikke nevneverdig forskjellig mellom de ulike modellene. Henholdsvis rett i underkant av 600.000kr i lineær modell, mens i dobbelt logaritmisk modell forklares tilstedeværelse av garasjeleilighet en økning i prisen på 16,43 prosent. I semi-logaritmisk modell forklarer tilstedeværelse av garasjeleilighet en prisøkning på 15,62 prosent.

Begge kontrollhypotesene har også empirisk støtte i regresjonsmodellene.

8.2 EKSEMPEL PÅ GJENNOMFØRING AV MODELL

Jeg har altså valgt å benytte meg av semi-logaritmisk regresjonsmodell. I dette kapittelet vil jeg demonstrere hvordan modellen fungerer. Lager i den forbindelse en teoretisk bolig(basisboligen), med ulike kvaliteter og setter deretter inn verdier i regresjonsmodellen, for å illustrere hvordan den fungerer. La oss si at boligen har et boa på 200kvm og at den er bygget i år 2000. Setter først opp likningen:

$$\ln Y_i = \alpha + \beta_1 X_i + \beta_2 D_i + \epsilon_i$$

Denne likningen kan skrives om til;

$$Y = e^{\alpha + \beta_1 X_i + \beta_2 D_i + \epsilon_i}$$

For den teoretiske boligen vil likningen se ut som følger;

$$\text{Pris} = e^{14.20421 + (0,0039442*200) - (0,0023617*15)} = 3.133.085 \text{ kr.}$$

La oss nå anta at boligen ligger på Fevik i Grimstad (postnr-variabel 104);

$$\text{Pris} = e^{14.20421 + (0,0039442*200) - (0,0023617*15) + 0,0600537} = 3.327.322 \text{ kr.}$$

La oss i tillegg anta at eneboligen har garasjeleilighet;

$$\text{Pris} = e^{14.20421 + (0,0039442*200) - (0,0023611*15) + 0,0600537 + 0,1562439} = 3.890.044 \text{ kr.}$$

Hvis vi ser på investeringskostnaden jeg kom frem til i kapittel 3, antok jeg at en investering i garasjeleilighet, kom på omkring 900.000kr. Hvis jeg nå tar utgangspunkt i gjennomsnittlig pris på eneboligene i utvalget mitt, som var på rett i overkant av 3.500.000kr og gir denne boligen 15,62 prosent påslag for garasjeleilighet, blir påslaget 546.700kr. Hvis vi trekker fra «nytt» i form garasje plass for egen bil, blir summen $900.000 - 158.100 = 741.900$ kr. Denne summen er fortsatt høyere enn påslaget for garasjeleilighet for gjennomsnittsboligen. Men det er verdt å merkes at mine forutsetninger i analysen i kapittel 3, er nok så konservative.

8.3 SVAKHETER VED ANALYSENE

En opplagt svakhet ved de ulike regresjonsformene, er som tidligere nevnt, problemene knyttet til å få med alle attributtene som driver prisen og at jeg således ikke har fått belyst alle årsaksfaktorene relatert til bolig. Jeg har f.eks ikke sett på tomteareal, antall soverom osv. En annen svakhet er behovet for å ta gitte forutsetninger i enkelte av variablene. Diskusjonen i kapittel 3 dekker min begrunnelse for bruk av skjønn i definisjonen av garasjeleilighet. Et annet problem som kan oppstå som følge av dette, er at garasjeleilighet er inkludert i boa for noen eneboliger, men kanskje ikke for alle eneboligene i utvalget. Dette er fordi de garasjeleilighetene som av ulike grunner ikke er lovlig for varig opphold, høyst trolig ikke blir tatt med i boa-beregningene fra eiendomsmeglers side. Dette vil resultere i at effekten av garasjeleilighet bare delvis fanges opp av boareal-variabelen. Dette kan illustreres ved regning. La oss ta utgangspunkt i basisboligen på 200kvm på Fevik, fra år 2000, med garasjeleilighet. Antar at denne garasjeleiligheten ikke er lovlig for varig opphold og således ikke regnes med i boa. Denne boligen vil ha en pris på 3.890.044 kr. Men hva skjer om vi antar at garasjeleiligheten er lovlig for varig opphold? La oss si at garasjeleiligheten er på 70kvm. Disse 70kvm må nå legges til boa-variabelen og regnestykket blir som følger:

$$\text{Pris} = e^{14.20421 + (0,0039442*270) - (0,0023611*15) + 0,0600537 + 0,1562439} = 5.126.968 \text{ kr.}$$

Vi ser av eksempelet at vi får stor differanse i prisen, avhengig av om garasjeleiligheten innlemmes i boa, eller ei.

Jeg vil også si at inndelingen av dummyvariablene for postnummer kan være en svakhet, da jeg måtte slå sammen enkelte postnumre og la til grunn at tilnærmet like postnumre innebar tilnærmet like kvaliteter, som nærhet til by, sjø, handlesenter og andre attributter. Til sist vil jeg si at mangelen på by-variabel gjør så jeg ikke får hensyntatt det ulike prisnivået i de forskjellige byene i så stor grad som ønsket. Dette ble en avveining mellom å beholde by-variablene og godta multikollinearitet, eller å utelate dem. Valget falt på sistnevnte, da jeg på bakgrunn av resultatene i kapittel 6 antar at postnummervariablen vil fange opp prisdifferansene.

9. KONKLUSJON

Problemstillingen i oppgaven har vært å se på om garasjeleilighet gir merverdi til boligen. Ut i fra teori har jeg konstruert hypoteser, og testet disse ved hjelp av regresjon. Resultatene i regresjonen har støttet opp om alternativhypotesene og de kan således ikke forkastes.

For eventuelle videre studier, hadde det vært interessant å se på et større utvalg av garasjeleiligheter og se på flere attributter, for å ytterligere konkretisere. En naturlig videreføring av dette arbeidet, vil være å ta hensyn til begrensningene i utvalget, samt å på hvilke momenter som mangler. Et eksempel på dette, kan være å måle boa for hver enkelt garasjeleilighet. Det er interessant å se at den relative andelen garasjeleiligheter er større for Sørlandet, enn byene lengre østover i utvalget, jf tabell 6.6 i kapittel 6. Som tidligere nevnt fant jeg ingen garasjeleiligheter i Bergen, Stavanger og Trondheim. Jeg har også ringt en bekjent som er eiendomsmegler i Oslo. Han forklarte at han aldri hadde solgt en enebolig i Oslo, med garasjeleilighet. Dette begrunnet han med at kravene for total mønehøyde (se kapittel 3 om garasjeleilighet) var svært strenge i Oslo. Hadde jeg hatt tid og anledning, ville jeg utformet en hypotese om at tomtearealet hadde betydning for hvorvidt en enebolig har garasjeleilighet, eller ei.

10. LITTERATURLISTE/KILDER

- Hair, J., Black, B. Babin, B., Anderson, R. and Tatham, R. (2006). *Multivariate Data Analysis* (6th edition). Upper Saddle River, New Jersey: Prentice-Hall
- Roscoe, J.T. (1975), *Fundamental Research Statistics for the Behavioral Sciences*, 2nd ed., New York: Holt, Rinehart and Winston, Inc.
- Robertsen Karl, Theisen Theis(2010), *Boligmarkedet i Kristiansand*. Publisert i Sødal, Sigbjørn og Jon P. Knudsen (red.): *Økonomi og tid*. Fagbokforlaget.
- Wenstøp, Fred (2006), *statistikk og dataanalyse*, (9 utgave). Oslo: Universitetsforlaget.
- Jacobsen, Dag Ingvar(2005), *Hvordan gjennomføre undersøkelser?* (2.utgave) *Innføring samfunnsvitenskapelig metode*. Cappelen Damm Høgskoleforlaget.
- Emblem, A.W (2014). *Forelesningsnotater BE-409, Eiendomsøkonomi*. Kristiansand: Universitetet i Agder
- Andersen, O. (2014). *Forelesningsnotater ME-407, Research methods*. Kristiansand: Universitetet i Agder
- Fredriksen, K.M (2013). *Bybanens innvirkning på boligpriser I Bergen*. Masteroppgave I økonomi og administrasjon. Kristiansand: Universitetet I Agder.
- Pedersen, A.G (2005): *Hvilken betydning har distanse til sjøen for prisene på fritidseiendommer? En studie av fritidseiendommer for salg langs kystkommunene i Agderfylkene 2005*. Masteroppgave i økonomi og administrasjon. Kristiansand: Høgskolen i Agder.
- Isaksen, H (2007): *Ungdom på boligmarkedet. Hvilke attributer etterspør unge boligkjøpere I Kristiansand?* Masteroppgave I økonomi og administrasjon. Kristiansand: Høgskolen I Agder.
- Byremo,K (2006): *Bomringens effekt på boligprisene I Kristiansand*. ? Masteroppgave I økonomi og administrasjon. Kristiansand: Høgskolen I Agder.

Lovdata(2015):

<https://lovdata.no/dokument/LTI/forskrift/2015-03-08-206> ¹

NOU(2002:2):

<https://www.regjeringen.no/contentassets/80899d9e55ef499c86359694e816207f/no/pdfa/nou200220020002000dddpdfa.pdf> ²

Lovdata(2014):

https://lovdata.no/dokument/NL/lov/1999-03-26-14/KAPITTEL_8#KAPITTEL_8 ³

Andre internettkilder:

Nordea.no:

http://www.nordea.no/Privat/L%C3%A5n/L%C3%A5n+til+hus+leilighet+og+hytte/Boligl%C3%A5nskalkulator/1642682.html?lnkID=top-user-goals_privat_boliglanskalkulator_24-03-2015 ⁴

Sifo.no:

http://www.sifo.no/files/file48351_arbeidsnotat08-2000web.pdf ⁵

Mellbygarasje.no:

<http://www.mellbygarasje.no/no/privat/garasjer/villagarasjer/garasje-med-leilighet/index1,210.htm> ⁶

Dibk.no:

<http://www.dibk.no/no/BYGGEREGLER/Nyheter-byggeregler/na-kan-du-bygge-mer-uten-soke/>

Huseierne.no:

<http://www.huseierne.no/hus-bolig/tema/juss/utleie-av-bolig/streng-krav-til-hybler/>

Garasjer.no:

<http://www.garasjer.no/hvor-stor-garasje-kan-du-bygge/>

<http://www.garasjer.no/wp-content/uploads/2014/08/5153-S%C3%B8knad-uten-ansvarsrett-A.pdf>

<http://www.garasjer.no/wp-content/uploads/2014/08/5155-Opplysninger-gitt-i-nabovarsel-A.pdf>

<http://www.garasjer.no/wp-content/uploads/2014/08/5156-Kvittering-for-Nabovarsel-A.pdf>

<http://www.garasjer.no/wp-content/uploads/2014/08/5154-Nabovarsel-A.pdf>

<http://www.garasjer.no/wp-content/uploads/2014/08/5167-S%C3%B8knad-om-ferdigattest-A.pdf>

VEDLEGG 1:

Kopierer programfilen jeg benyttet meg av i stata:

```
insheet using C:\Users\Simen\Desktop\Datagrunnlag.txt
```

```
*deskriptiv statistikk
```

```
generate salgsaar=0
```

```
generate cas=_n
```

```
generate Kristiansand=0
```

```
replace Kristiansand=1 if cas>=0 & cas<=924
```

```
generate Grimstad=0
```

```
replace Grimstad=1 if cas>=925 & cas<=1421
```

```
generate Arendal=0
```

```
replace Arendal=1 if cas>=1422 & cas<=1688
```

```
generate Drammen=0
```

```
replace Drammen=1 if cas>=1689 & cas<=1832
```

```
generate Lillesand=0
```

```
replace Lillesand=1 if cas>=1833 & cas<=1880
```

```
generate Sandefjord=0
```

```
replace Sandefjord=1 if cas>=1881 & cas<=2163
```

```
generate Skien=0
```

```
replace Skien=1 if cas>=2164 & cas<=2423
```

```
generate Tønsberg=0
```

```
replace Tønsberg=1 if cas>=2424 & cas<=2700
```

```
replace salgsaar=1 if cas>=431& cas<=1176
```

```
replace salgsaar=1 if cas>=1422&cas<=1552
```

```
replace salgsaar=1 if cas>=1689&cas<=2040
```

```
replace salgsaar=1 if cas>=2164&cas<=2295
```

```
replace salgsaar=1 if cas>=2424&cas<=2589
```

```
generate pris=real(salgspris)
```

```
summarize pris
```

```

drop if pris==.
drop cas
*sorterer variablen megler.
generate casenr=_n
drop if casenr <=339
drop v3
drop salgspris
split adresse
generate p1=real(adresse3)
generate p2=real(adresse4)
generate p3=real(adresse5)
generate post1=0
replace post1=p1 if p1>=1000
replace post1=0 if p1==.
generate post2=0
replace post2=p2 if p2>=1000
replace post2=0 if p2==.
generate post3=0
replace post3=p3 if p3>=1000
replace post3=post1 if post3==.
replace post3=post2 if post3==0
generate p4=real(adresse6)
replace post3=p4 if post3==0
drop p1
drop p2
drop p3
drop post1
drop post2
drop p4
drop adresse1

```

```

drop adresse2
drop adresse3
drop adresse4
drop adresse5
drop adresse6
drop adresse7
drop adresse8
drop if bygningsr==.
drop if casenr==.
drop if post3==.
drop if boaprom==.
tabulate post3, gen(p)
generate garasje=0
replace garasje=1 if garasjehybel=="ja"
drop meglar
drop cas
drop eieform
drop garasjehybel
*slår sammen postnummer-variabler
generate postnr1=0
replace postnr1=p1 if p1>=0.1
replace postnr1=p2 if p2>=0.1
replace postnr1=p3 if p3>=0.1
replace postnr1=p4 if p4>=0.1
replace postnr1=p5 if p5>=0.1
replace postnr1=p6 if p6>=0.1
generate postnr2=0
replace postnr2=p7 if p7>=0.1
generate postnr3=0
replace postnr3=p8 if p8>=0.1

```

```
generate postnr4=0
replace postnr4=p9 if p9>=0.1
generate postnr5=0
replace postnr5=p10 if p10>=0.1
generate postnr6=0
replace postnr6=p11 if p11>=0.1
replace postnr6=p12 if p12>=0.1
replace postnr6=p13 if p13>=0.1
drop p1
drop p2
drop p3
drop p4
drop p5
drop p6
drop p7
drop p8
drop p9
drop p10
drop p11
drop p12
drop p13
generate postnr7=0
replace postnr7=p14 if p14>=0.1
generate postnr8=0
replace postnr8=p15 if p15>=0.1
drop p14
drop p15
generate postnr9=0
replace postnr9=p16 if p16>=0.1
replace postnr9=p17 if p17>=0.1
```

replace postnr9=p18 if p18>=0.1
drop p16
drop p17
drop p18
generate postnr10=0
replace postnr10=p19 if p19>=0.1
replace postnr10=p20 if p20>=0.1
replace postnr10=p21 if p21>=0.1
drop p19
drop p20
drop p21
generate postnr11=0
replace postnr11=p22 if p22>=0.1
replace postnr11=p23 if p23>=0.1
drop p22
drop p23
generate postnr12=0
replace postnr12=p24 if p24>=0.1
drop p24
generate postnr13=0
replace postnr13=p25 if p25>=0.1
replace postnr13=p26 if p26>=0.1
drop p25
drop p26
generate postnr14=0
replace postnr14=p27 if p27>=0.1
replace postnr14=p28 if p28>=0.1
drop p27
drop p28
generate postnr15=0

replace postnr15=p29 if p29>=0.1
drop p29
generate postnr16=0
replace postnr16=p30 if p30>=0.1
replace postnr16=p31 if p31>=0.1
replace postnr16=p32 if p32>=0.1
drop p30
drop p31
drop p32
generate postnr17=0
replace postnr17=p33 if p33>=0.1
drop p33
generate postnr18=0
replace postnr18=p34 if p34>=0.1
drop p34
generate postnr19=0
replace postnr19=p35 if p35>=0.1
drop p35
generate postnr20=0
replace postnr20=p36 if p36>=0.1
drop p36
generate postnr21=0
replace postnr21=p37 if p37>=0.1
drop p37
generate postnr22=0
replace postnr22=p38 if p38>=0.1
drop p38
generate postnr23=0
replace postnr23=p39 if p39>=0.1
drop p39

```
generate postnr24=0
replace postnr24=p40 if p40>=0.1
drop p40
generate postnr25=0
replace postnr25=p41 if p41>=0.1
drop p41
generate postnr26=0
replace postnr26=p42 if p42>=0.1
drop p42
generate postnr27=0
replace postnr27=p43 if p43>=0.1
replace postnr27=p44 if p44>=0.1
drop p43
drop p44
generate postnr28=0
replace postnr28=p45 if p45>=0.1
drop p45
generate postnr29=0
replace postnr29=p46 if p46>=0.1
replace postnr29=p47 if p47>=0.1
drop p46
drop p47
generate postnr30=0
replace postnr30=p48 if p48>=0.1
drop p48
generate postnr31=0
replace postnr31=p49 if p49>=0.1
drop p49
generate postnr32=0
replace postnr32=p50 if p50>=0.1
```

replace postnr32=p51 if p51>=0.1
drop p50
drop p51
generate postnr33=0
replace postnr33=p52 if p52>=0.1
replace postnr33=p53 if p53>=0.1
drop p52
drop p53
generate postnr34=0
replace postnr34=p54 if p54>=0.1
replace postnr34=p55 if p55>=0.1
drop p54
drop p55
generate postnr35=0
replace postnr35=p56 if p56>=0.1
replace postnr35=p57 if p57>=0.1
drop p56
drop p57
generate postnr36=0
replace postnr36=p58 if p58>=0.1
drop p58
generate postnr37=0
replace postnr37=p59 if p59>=0.1
drop p59
generate postnr38=0
replace postnr38=p60 if p60>=0.1
drop p60
generate postnr39=0
replace postnr39=p61 if p61>=0.1
drop p61

generate postnr40=0
replace postnr40=p62 if p62>=0.1
replace postnr40=p63 if p63>=0.1
drop p62
drop p63
generate postnr41=0
replace postnr41=p64 if p64>=0.1
replace postnr41=p65 if p65>=0.1
replace postnr41=p66 if p66>=0.1
drop p64
drop p65
drop p66
generate postnr42=0
replace postnr42=p67 if p67>=0.1
drop p67
generate postnr43=0
replace postnr43=p68 if p68>=0.1
drop p68
generate postnr44=0
replace postnr44=p69 if p69>=0.1
drop p69
generate postnr45=0
replace postnr45=p70 if p70>=0.1
replace postnr45=p71 if p71>=0.1
drop p70
drop p71
generate postnr46=0
replace postnr46=p72 if p72>=0.1
drop p72
generate postnr47=0

replace postnr47=p73 if p73>=0.1
drop p73
generate postnr48=0
replace postnr48=p74 if p74>=0.1
drop p74
generate postnr49=0
replace postnr49=p75 if p75>=0.1
drop p75
generate postnr50=0
replace postnr50=p76 if p76>=0.1
drop p76
generate postnr51=0
replace postnr51=p77 if p77>=0.1
drop p77
generate postnr52=0
replace postnr52=p78 if p78>=0.1
drop p78
generate postnr53=0
replace postnr53=p79 if p79>=0.1
replace postnr53=p80 if p80>=0.1
replace postnr53=p81 if p81>=0.1
drop p79
drop p80
drop p81
generate postnr54=0
replace postnr54=p82 if p82>=0.1
drop p82
generate postnr55=0
replace postnr55=p83 if p83>=0.1
drop p83

```
generate postnr56=0
replace postnr56=p84 if p84>=0.1
replace postnr56=p85 if p85>=0.1
replace postnr56=p86 if p86>=0.1
drop p84
drop p85
drop p86
generate postnr57=0
replace postnr57=p87 if p87>=0.1
replace postnr57=p88 if p88>=0.1
replace postnr57=p89 if p89>=0.1
drop p87
drop p88
drop p89
generate postnr58=0
replace postnr58=p90 if p90>=0.1
replace postnr58=p91 if p91>=0.1
drop p90
drop p91
replace postnr58=p92 if p92>=0.1
drop p92
generate postnr59=0
replace postnr59=p93 if p93>=0.1
replace postnr59=p94 if p94>=0.1
drop p93
drop p94
generate postnr60=0
replace postnr60=p95 if p95>=0.1
drop p95
generate postnr61=0
```

replace postnr61=p96 if p96>=0.1
drop p96
generate postnr62=0
replace postnr62=p97 if p97>=0.1
replace postnr62=p98 if p98>=0.1
drop p97
drop p98
generate postnr63=0
replace postnr63=p99 if p99>=0.1
drop p99
generate postnr64=0
replace postnr64=p100 if p100>=0.1
drop p100
generate postnr65=0
replace postnr65=p101 if p101>=0.1
replace postnr65=p102 if p102>=0.1
drop p101
drop p102
generate postnr66=0
replace postnr66=p103 if p103>=0.1
replace postnr66=p104 if p104>=0.1
drop p103
drop p104
generate postnr67=0
replace postnr67=p105 if p105>=0.1
drop p105
generate postnr68=0
replace postnr68=p106 if p106>=0.1
drop p106
generate postnr69=0

replace postnr69=p107 if p107>=0.1
drop p107
generate postnr70=0
replace postnr70=p108 if p108>=0.1
drop p108
generate postnr71=0
replace postnr71=p109 if p109>=0.1
drop p109
generate postnr72=0
replace postnr72=p110 if p110>=0.1
drop p110
generate postnr73=0
replace postnr73=p111 if p111>=0.1
generate postnr74=0
replace postnr74=p112 if p112>=0.1
drop p111
drop p112
generate postnr75=0
replace postnr75=p113 if p113>=0.1
generate postnr76=0
replace postnr76=p114 if p114>=0.1
drop p113
drop p114
generate postnr77=0
replace postnr77=p115 if p115>=0.1
generate postnr78=0
replace postnr78=p116 if p116>=0.1
drop p115
drop p116
generate postnr79=0

replace postnr79=p117 if p117>=0.1
generate postnr80=0
replace postnr80=p118 if p118>=0.1
drop p117
drop p118
generate postnr81=0
replace postnr81=p119 if p119>=0.1
generate postnr82=0
replace postnr82=p120 if p120>=0.1
drop p119
drop p120
generate postnr83=0
replace postnr83=p121 if p121>=0.1
generate postnr84=0
replace postnr84=p122 if p122>=0.1
drop p121
drop p122
generate postnr85=0
replace postnr85=p123 if p123>=0.1
generate postnr86=0
replace postnr86=p124 if p124>=0.1
generate postnr87=0
replace postnr87=p125 if p125>=0.1
drop p123
drop p124
drop p125
generate postnr88=0
replace postnr88=p126 if p126>=0.1
replace postnr88=p127 if p127>=0.1
drop p126

drop p127
generate postnr89=0
replace postnr89=p128 if p128>=0.1
drop p128
generate postnr90=0
replace postnr90=p129 if p129>=0.1
drop p129
generate postnr91=0
replace postnr91=p130 if p130>=0.1
drop p130
generate postnr92=0
replace postnr92=p131 if p131>=0.1
drop p131
generate postnr93=0
replace postnr93=p132 if p132>=0.1
replace postnr93=p133 if p133>=0.1
drop p132
drop p133
generate postnr94=0
replace postnr94=p134 if p134>=0.1
drop p134
generate postnr95=0
replace postnr95=p135 if p135>=0.1
drop p135
generate postnr96=0
replace postnr96=p136 if p136>=0.1
drop p136
generate postnr97=0
replace postnr97=p137 if p137>=0.1
drop p137

```
generate postnr98=0
replace postnr98=p138 if p138>=0.1
drop p138
generate postnr99=0
replace postnr99=p139 if p139>=0.1
replace postnr99=p140 if p140>=0.1
replace postnr99=p141 if p141>=0.1
replace postnr99=p142 if p142>=0.1
drop p139
drop p140
drop p141
drop p142
generate postnr100=0
replace postnr100=p143 if p143>=0.1
replace postnr100=p144 if p144>=0.1
drop p143
drop p144
generate postnr101=0
replace postnr101=p145 if p145>=0.1
generate postnr102=0
replace postnr102=p146 if p146>=0.1
drop p145
drop p146
generate postnr103=0
replace postnr103=p147 if p147>=0.1
replace postnr103=p148 if p148>=0.1
drop p147
drop p148
generate postnr104=0
replace postnr104=p149 if p149>=0.1
```



```

drop p149
generate postnr105=0
replace postnr105=p150 if p150>=0.1
drop p150
generate postnr106=0
replace postnr106=p151 if p151>=0.1
drop p151
generate postnr107=0
replace postnr107=p152 if p152>=0.1
drop p152
generate postnr108=0
replace postnr108=p153 if p153>=0.1
drop p153
generate postnr109=0
replace postnr109=p154 if p154>=0.1
generate postnr110=0
replace postnr110=p155 if p155>=0.1
drop p154
drop p155
generate postnr111=0
replace postnr111=p156 if p156>=0.1
generate postnr112=0
replace postnr112=p157 if p157>=0.1
drop p156
drop p157
replace postnr112=p158 if p158>=0.1
replace postnr112=p159 if p159>=0.1
drop p158
drop p159
recode
pris(0/1000000=1)(1000001/1500000=2)(1500001/2000000=3)(2000001/2500000=4)(25000

```

01/3000000=5)(3000001/3500000=6)(3500001/4000000=7)(4000001/4500000=8)(4500001/5000000=9)(5000001/5500000=10)(5500001/6000000=11)(6000001/6500000=12)(6500001/7000000=13)(7000001/7500000=14)(7500001/8000000=15)(8000001/8500000=16)(8500001/9000000=17)(9000001/9500000=18)(9500001/10000000=19)(10000001/max=20), gen(p)

*Multippel lineær regresjon

```
reg pris boaprom bygningsr garasje salgsaar postnr1 postnr2 postnr3 postnr4 postnr5 postnr6
postnr7 postnr8 postnr9 postnr10 postnr11 postnr12 postnr13 postnr14 postnr15 postnr16
postnr17 postnr18 postnr19 postnr20 postnr21 postnr22 postnr23 postnr24 postnr25 postnr26
postnr27 postnr28 postnr29 postnr30 postnr31 postnr32 postnr33 postnr34 postnr35 postnr36
postnr37 postnr38 postnr39 postnr40 postnr41 postnr42 postnr43 postnr44 postnr45 postnr46
postnr47 postnr48 postnr49 postnr50 postnr51 postnr52 postnr53 postnr54 postnr55 postnr56
postnr57 postnr58 postnr59 postnr60 postnr61 postnr62 postnr63 postnr64 postnr65 postnr66
postnr67 postnr68 postnr69 postnr70 postnr71 postnr72 postnr73 postnr74 postnr75 postnr76
postnr77 postnr78 postnr79 postnr80 postnr81 postnr82 postnr83 postnr84 postnr85 postnr86
postnr87 postnr88 postnr89 postnr90 postnr91 postnr92 postnr93 postnr94 postnr95 postnr96
postnr97 postnr98 postnr99 postnr100 postnr101 postnr102 postnr103 postnr104 postnr105
postnr106 postnr107 postnr108 postnr109 postnr110 postnr111
```

vif

predict residual

hettest residual

predict res1,residuals

histogram res1, normal

qnorm res1

generate lnpris=ln(pris)

generate lnboaprom=ln(boaprom)

generate lnboligalder=ln(boligalder)

*dobbel logaritmisk regresjon

```
reg lnpris lnboaprom lnbygningsr garasje salgsaar postnr1 postnr2 postnr3 postnr4 postnr5
postnr6 postnr7 postnr8 postnr9 postnr10 postnr11 postnr12 postnr13 postnr14 postnr15
postnr16 postnr17 postnr18 postnr19 postnr20 postnr21 postnr22 postnr23 postnr24 postnr25
postnr26 postnr27 postnr28 postnr29 postnr30 postnr31 postnr32 postnr33 postnr34 postnr35
postnr36 postnr37 postnr38 postnr39 postnr40 postnr41 postnr42 postnr43 postnr44 postnr45
postnr46 postnr47 postnr48 postnr49 postnr50 postnr51 postnr52 postnr53 postnr54 postnr55
postnr56 postnr57 postnr58 postnr59 postnr60 postnr61 postnr62 postnr63 postnr64 postnr65
postnr66 postnr67 postnr68 postnr69 postnr70 postnr71 postnr72 postnr73 postnr74 postnr75
postnr76 postnr77 postnr78 postnr79 postnr80 postnr81 postnr82 postnr83 postnr84 postnr85
postnr86 postnr87 postnr88 postnr89 postnr90 postnr91 postnr92 postnr93 postnr94 postnr95
postnr96 postnr97 postnr98 postnr99 postnr100 postnr101 postnr102 postnr103 postnr104
postnr105 postnr106 postnr107 postnr108 postnr109 postnr110 postnr111
```

vif

hettest residual

predict res2, residuals

histogram res2, normal

qnorm res2

*Semi logaritmisk regresjon

```
reg lnpris boaprom bygningsr garasje salgsaar postnr1 postnr2 postnr3 postnr4 postnr5
postnr6 postnr7 postnr8 postnr9 postnr10 postnr11 postnr12 postnr13 postnr14 postnr15
postnr16 postnr17 postnr18 postnr19 postnr20 postnr21 postnr22 postnr23 postnr24 postnr25
postnr26 postnr27 postnr28 postnr29 postnr30 postnr31 postnr32 postnr33 postnr34 postnr35
postnr36 postnr37 postnr38 postnr39 postnr40 postnr41 postnr42 postnr43 postnr44 postnr45
postnr46 postnr47 postnr48 postnr49 postnr50 postnr51 postnr52 postnr53 postnr54 postnr55
postnr56 postnr57 postnr58 postnr59 postnr60 postnr61 postnr62 postnr63 postnr64 postnr65
postnr66 postnr67 postnr68 postnr69 postnr70 postnr71 postnr72 postnr73 postnr74 postnr75
postnr76 postnr77 postnr78 postnr79 postnr80 postnr81 postnr82 postnr83 postnr84 postnr85
postnr86 postnr87 postnr88 postnr89 postnr90 postnr91 postnr92 postnr93 postnr94 postnr95
postnr96 postnr97 postnr98 postnr99 postnr100 postnr101 postnr102 postnr103 postnr104
postnr105 postnr106 postnr107 postnr108 postnr109 postnr110 postnr111
```

vif

hettest residual

predict res3, residuals

histogram res3, normal

qnorm res3

*redefingering av kontinuerlig variabel

generate boligalder=0

replace boligalder=324 if bygningsr==1691

replace boligalder=265 if bygningsr==1750

replace boligalder=251 if bygningsr==1764

replace boligalder= 245 if bygningsr==1770

replace boligalder=215 if bygningsr==1800

replace boligalder=208 if bygningsr==1807

replace boligalder= 192 if bygningsr==1823

replace boligalder=191 if bygningsr==1824

replace boligalder=185 if bygningsr==1830

replace boligalder=182 if bygningsr==1832
replace boligalder=175 if bygningsr==1840
replace boligalder=165 if bygningsr==1850
replace boligalder=163 if bygningsr==1852
replace boligalder=158 if bygningsr==1857
replace boligalder=155 if bygningsr==1860
replace boligalder=150 if bygningsr==1865
replace boligalder=149 if bygningsr==1866
replace boligalder=146 if bygningsr==1869
replace boligalder=145 if bygningsr==1870
replace boligalder=142 if bygningsr==1873
replace boligalder=141 if bygningsr==1874
replace boligalder=140 if bygningsr==1875
replace boligalder=137 if bygningsr==1878
replace boligalder=135 if bygningsr==1880
replace boligalder=131 if bygningsr==1884
replace boligalder=130 if bygningsr==1885
replace boligalder=128 if bygningsr==1887
replace boligalder=127 if bygningsr==1888
replace boligalder=126 if bygningsr==1889
replace boligalder=125 if bygningsr==1890
replace boligalder=123 if bygningsr==1892
replace boligalder=122 if bygningsr==1893
replace boligalder=119 if bygningsr==1896
replace boligalder=118 if bygningsr==1897
replace boligalder=117 if bygningsr==1898
replace boligalder=115 if bygningsr==1900
replace boligalder=114 if bygningsr==1901
replace boligalder=112 if bygningsr==1903
replace boligalder=111 if bygningsr==1904

replace boligalder=110 if bygningsr==1905
replace boligalder=109 if bygningsr==1906
replace boligalder=108 if bygningsr==1907
replace boligalder=106 if bygningsr==1909
replace boligalder=104 if bygningsr==1911
replace boligalder=103 if bygningsr==1912
replace boligalder=101 if bygningsr==1914
replace boligalder=100 if bygningsr==1915
replace boligalder=98 if bygningsr==1917
replace boligalder=97 if bygningsr==1918
replace boligalder=96 if bygningsr==1919
replace boligalder=95 if bygningsr==1920
replace boligalder=93 if bygningsr==1922
replace boligalder=92 if bygningsr==1923
replace boligalder=89 if bygningsr==1926
replace boligalder=88 if bygningsr==1927
replace boligalder=87 if bygningsr==1928
replace boligalder=85 if bygningsr==1930
replace boligalder=84 if bygningsr==1931
replace boligalder=83 if bygningsr==1932
replace boligalder=82 if bygningsr==1933
replace boligalder=81 if bygningsr==1934
replace boligalder=80 if bygningsr==1935
replace boligalder=79 if bygningsr==1936
replace boligalder=78 if bygningsr==1937
replace boligalder=77 if bygningsr==1938
replace boligalder=76 if bygningsr==1939
replace boligalder=75 if bygningsr==1940
replace boligalder=73 if bygningsr==1942
replace boligalder=70 if bygningsr==1945

replace boligalder=69 if bygningsr==1946
replace boligalder=68 if bygningsr==1947
replace boligalder=67 if bygningsr==1948
replace boligalder=66 if bygningsr==1949
replace boligalder=65 if bygningsr==1950
replace boligalder=64 if bygningsr==1951
replace boligalder=63 if bygningsr==1952
replace boligalder=62 if bygningsr==1953
replace boligalder=61 if bygningsr==1954
replace boligalder=60 if bygningsr==1955
replace boligalder=59 if bygningsr==1956
replace boligalder=58 if bygningsr==1957
replace boligalder=57 if bygningsr==1958
replace boligalder=56 if bygningsr==1959
replace boligalder=55 if bygningsr==1960
replace boligalder=54 if bygningsr==1961
replace boligalder=53 if bygningsr==1962
replace boligalder=52 if bygningsr==1963
replace boligalder=51 if bygningsr==1964
replace boligalder=50 if bygningsr==1965
replace boligalder=49 if bygningsr==1966
replace boligalder=48 if bygningsr==1967
replace boligalder=47 if bygningsr==1968
replace boligalder=46 if bygningsr==1969
replace boligalder=45 if bygningsr==1970
replace boligalder=44 if bygningsr==1971
replace boligalder=43 if bygningsr==1972
replace boligalder=42 if bygningsr==1973
replace boligalder=41 if bygningsr==1974
replace boligalder=40 if bygningsr==1975

replace boligalder=39 if bygningsr==1976
replace boligalder=38 if bygningsr==1977
replace boligalder=37 if bygningsr==1978
replace boligalder=36 if bygningsr==1979
replace boligalder=35 if bygningsr==1980
replace boligalder=34 if bygningsr==1981
replace boligalder=33 if bygningsr==1982
replace boligalder=32 if bygningsr==1983
replace boligalder=31 if bygningsr==1984
replace boligalder=30 if bygningsr==1985
replace boligalder=29 if bygningsr==1986
replace boligalder=28 if bygningsr==1987
replace boligalder=27 if bygningsr==1988
replace boligalder=26 if bygningsr==1989
replace boligalder=25 if bygningsr==1990
replace boligalder=24 if bygningsr==1991
replace boligalder=23 if bygningsr==1992
replace boligalder=22 if bygningsr==1993
replace boligalder=21 if bygningsr==1994
replace boligalder=20 if bygningsr==1995
replace boligalder=19 if bygningsr==1996
replace boligalder=18 if bygningsr==1997
replace boligalder=17 if bygningsr==1998
replace boligalder=16 if bygningsr==1999
replace boligalder=15 if bygningsr==2000
replace boligalder=14 if bygningsr==2001
replace boligalder=13 if bygningsr==2002
replace boligalder=12 if bygningsr==2003
replace boligalder=11 if bygningsr==2004
replace boligalder=10 if bygningsr==2005

```
replace boligalder=9 if bygningsr==2006
replace boligalder=8 if bygningsr==2007
replace boligalder=7 if bygningsr==2008
replace boligalder=6 if bygningsr==2009
replace boligalder=5 if bygningsr==2010
replace boligalder=4 if bygningsr==2011
replace boligalder=3 if bygningsr==2012
replace boligalder=2 if bygningsr==2013
replace boligalder=1 if bygningsr==2014
replace boligalder=0 if bygningsr==2015
*opretter cluster
generate Z=0
replace Z=1 if postnr1==1
replace Z=2 if postnr2==1
replace Z=3 if postnr3==1
replace Z=4 if postnr4==1
replace Z=5 if postnr5==1
replace Z=6 if postnr6==1
replace Z=7 if postnr7==1
replace Z=8 if postnr8==1
replace Z=9 if postnr9==1
replace Z=10 if postnr10==1
replace Z=11 if postnr11==1
replace Z=12 if postnr12==1
replace Z=13 if postnr13==1
replace Z=14 if postnr14==1
replace Z=15 if postnr15==1
replace Z=16 if postnr16==1
replace Z=17 if postnr17==1
replace Z=18 if postnr18==1
```


replace Z=19 if postnr19==1
replace Z=20 if postnr20==1
replace Z=21 if postnr21==1
replace Z=22 if postnr22==1
replace Z=23 if postnr23==1
replace Z=24 if postnr24==1
replace Z=25 if postnr25==1
replace Z=26 if postnr26==1
replace Z=27 if postnr27==1
replace Z=28 if postnr28==1
replace Z=29 if postnr29==1
replace Z=30 if postnr30==1
replace Z=31 if postnr31==1
replace Z=32 if postnr32==1
replace Z=33 if postnr33==1
replace Z=34 if postnr34==1
replace Z=35 if postnr35==1
replace Z=36 if postnr36==1
replace Z=37 if postnr37==1
replace Z=38 if postnr38==1
replace Z=39 if postnr39==1
replace Z=40 if postnr40==1
replace Z=41 if postnr41==1
replace Z=42 if postnr42==1
replace Z=43 if postnr43==1
replace Z=44 if postnr44==1
replace Z=45 if postnr45==1
replace Z=46 if postnr46==1
replace Z=47 if postnr47==1
replace Z=48 if postnr48==1

replace Z=49 if postnr49==1
replace Z=50 if postnr50==1
replace Z=51 if postnr51==1
replace Z=52 if postnr52==1
replace Z=53 if postnr53==1
replace Z=54 if postnr54==1
replace Z=55 if postnr55==1
replace Z=56 if postnr56==1
replace Z=57 if postnr57==1
replace Z=58 if postnr58==1
replace Z=59 if postnr59==1
replace Z=60 if postnr60==1
replace Z=61 if postnr61==1
replace Z=62 if postnr62==1
replace Z=63 if postnr63==1
replace Z=64 if postnr64==1
replace Z=65 if postnr65==1
replace Z=66 if postnr66==1
replace Z=67 if postnr67==1
replace Z=68 if postnr68==1
replace Z=69 if postnr69==1
replace Z=70 if postnr70==1
replace Z=71 if postnr71==1
replace Z=72 if postnr72==1
replace Z=73 if postnr73==1
replace Z=74 if postnr74==1
replace Z=75 if postnr75==1
replace Z=76 if postnr76==1
replace Z=77 if postnr77==1
replace Z=78 if postnr78==1

replace Z=79 if postnr79==1
replace Z=80 if postnr80==1
replace Z=81 if postnr81==1
replace Z=82 if postnr82==1
replace Z=83 if postnr83==1
replace Z=84 if postnr84==1
replace Z=85 if postnr85==1
replace Z=86 if postnr86==1
replace Z=87 if postnr87==1
replace Z=88 if postnr88==1
replace Z=89 if postnr89==1
replace Z=90 if postnr90==1
replace Z=91 if postnr91==1
replace Z=92 if postnr92==1
replace Z=93 if postnr93==1
replace Z=94 if postnr94==1
replace Z=95 if postnr95==1
replace Z=96 if postnr96==1
replace Z=97 if postnr97==1
replace Z=98 if postnr98==1
replace Z=99 if postnr99==1
replace Z=100 if postnr100==1
replace Z=101 if postnr101==1
replace Z=102 if postnr102==1
replace Z=103 if postnr103==1
replace Z=104 if postnr104==1
replace Z=105 if postnr105==1
replace Z=106 if postnr106==1
replace Z=107 if postnr107==1
replace Z=108 if postnr108==1

replace Z=109 if postnr109==1

replace Z=110 if postnr110==1

replace Z=111 if postnr111==1

replace Z=112 if postnr112==1

*Endelig valg av regresjonsform

```
reg lnpris boaprom bygningsr garasje salgsaar postnr1 postnr2 postnr3 postnr4 postnr5
postnr6 postnr7 postnr8 postnr9 postnr10 postnr11 postnr12 postnr13 postnr14 postnr15
postnr16 postnr17 postnr18 postnr19 postnr20 postnr21 postnr22 postnr23 postnr24 postnr25
postnr26 postnr27 postnr28 postnr29 postnr30 postnr31 postnr32 postnr33 postnr34 postnr35
postnr36 postnr37 postnr38 postnr39 postnr40 postnr41 postnr42 postnr43 postnr44 postnr45
postnr46 postnr47 postnr48 postnr49 postnr50 postnr51 postnr52 postnr53 postnr54 postnr55
postnr56 postnr57 postnr58 postnr59 postnr60 postnr61 postnr62 postnr63 postnr64 postnr65
postnr66 postnr67 postnr68 postnr69 postnr70 postnr71 postnr72 postnr73 postnr74 postnr75
postnr76 postnr77 postnr78 postnr79 postnr80 postnr81 postnr82 postnr83 postnr84 postnr85
postnr86 postnr87 postnr88 postnr89 postnr90 postnr91 postnr92 postnr93 postnr94 postnr95
postnr96 postnr97 postnr98 postnr99 postnr100 postnr101 postnr102 postnr103 postnr104
postnr105 postnr106 postnr107 postnr108 postnr109 postnr110 postnr111,cluster(Z)
```

VEDLEGG 2:

Semi-logaritmisk regresjon, endelig modell:

Number of obs = 2298
 F(3, 111) = .
 Prob > F = .
 R-squared = 0.6543
 Root MSE = .21976

(Std. Err. adjusted for 112 clusters in Z)

lnpris	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
boaprom	.0039442	.0001575	25.04	0.000	.0036321	.0042563
boligalder	-.0023617	.0005328	-4.43	0.000	-.0034175	-.0013059
salgsaar	.0232863	.0094866	2.45	0.016	.004488	.0420847
garasje	.1562441	.0331826	4.71	0.000	.0904906	.2219975
postnr1	.4762986	.0076771	62.04	0.000	.4610858	.4915114
postnr2	.3629006	.005016	72.35	0.000	.352961	.3728402
postnr3	.4981517	.0050796	98.07	0.000	.4880861	.5082172
postnr4	.3278866	.0047296	69.33	0.000	.3185147	.3372586
postnr5	.2519231	.0066058	38.14	0.000	.2388333	.2650128
postnr6	.4665571	.0070383	66.29	0.000	.4526103	.4805039
postnr7	.4458949	.00496	89.90	0.000	.4360664	.4557235
postnr8	.21859	.006505	33.60	0.000	.2056999	.2314801
postnr9	.2442702	.0079558	30.70	0.000	.2285053	.2600351
postnr10	.546402	.0059725	91.49	0.000	.534567	.5582369
postnr11	.3910783	.0037183	105.18	0.000	.3837102	.3984464
postnr12	.3526156	.0037677	93.59	0.000	.3451497	.3600815
postnr13	.4257619	.0053104	80.18	0.000	.415239	.4362848

postnr14	.177488	.0024078	73.71	0.000	.1727168	.1822591
postnr15	.5350514	.0085311	62.72	0.000	.5181463	.5519564
postnr16	.6013728	.0025359	237.14	0.000	.5963477	.6063979
postnr17	.4810515	.0068655	70.07	0.000	.4674471	.494656
postnr18	.2669672	.0039032	68.40	0.000	.2592328	.2747015
postnr19	.2621346	.0054835	47.80	0.000	.2512686	.2730005
postnr20	.3319485	.0048575	68.34	0.000	.3223231	.341574
postnr21	.3416076	.0014404	237.15	0.000	.3387533	.344462
postnr22	.1472845	.0065874	22.36	0.000	.1342312	.1603379
postnr23	.1060923	.0040033	26.50	0.000	.0981596	.114025
postnr24	.1820271	.0039556	46.02	0.000	.1741887	.1898654
postnr25	.606605	.0040195	150.92	0.000	.5986401	.6145699
postnr26	.349018	.0061034	57.18	0.000	.3369237	.3611124
postnr27	.3028761	.0067755	44.70	0.000	.28945	.3163023
postnr28	.1001262	.0061486	16.28	0.000	.0879424	.11231
postnr29	.4050835	.0036913	109.74	0.000	.3977689	.4123981
postnr30	.2414213	.0055867	43.21	0.000	.2303509	.2524917
postnr31	.3107861	.0063591	48.87	0.000	.2981851	.3233871
postnr32	.092284	.0062209	14.83	0.000	.0799568	.1046112
postnr33	.2719714	.0072639	37.44	0.000	.2575776	.2863652
postnr34	.2359725	.0061909	38.12	0.000	.2237048	.2482402
postnr35	.191947	.0033006	58.16	0.000	.1854067	.1984873
postnr36	.4384527	.0070946	61.80	0.000	.4243943	.4525111
postnr37	.2305232	.0025149	91.66	0.000	.2255397	.2355066
postnr38	.2831994	.0056574	50.06	0.000	.2719888	.2944099
postnr39	.4603531	.0024282	189.58	0.000	.4555414	.4651648
postnr40	.0697683	.0049759	14.02	0.000	.0599082	.0796285
postnr41	.1401985	.001919	73.06	0.000	.1363959	.1440011
postnr42	.0597134	.0018521	32.24	0.000	.0560434	.0633834
postnr43	.1955123	.0027501	71.09	0.000	.1900628	.2009617
postnr44	.2332127	.0032556	71.63	0.000	.2267615	.2396638

postnr45	.0633016	.0025983	24.36	0.000	.058153	.0684502
postnr46	.1464619	.002588	56.59	0.000	.1413337	.1515901
postnr47	.0572534	.004357	13.14	0.000	.0486197	.0658872
postnr48	-.0096493	.0048597	-1.99	0.050	-.0192791	-.0000195
postnr49	.1188814	.0102407	11.61	0.000	.0985888	.1391739
postnr50	-.019767	.0056799	-3.48	0.001	-.0310222	-.0085119
postnr51	-.2081143	.0039838	-52.24	0.000	-.2160084	-.2002202
postnr52	.2262517	.0050459	44.84	0.000	.2162529	.2362504
postnr53	.0816523	.0052679	15.50	0.000	.0712137	.092091
postnr54	-.0466098	.0061111	-7.63	0.000	-.0587192	-.0345003
postnr55	.1170823	.0063523	18.43	0.000	.1044948	.1296698
postnr56	-.0344481	.0025041	-13.76	0.000	-.0394101	-.0294862
postnr57	.0200103	.005195	3.85	0.000	.009716	.0303045
postnr58	-.0566223	.0028624	-19.78	0.000	-.0622942	-.0509503
postnr59	-.0086193	.0029717	-2.90	0.004	-.0145079	-.0027307
postnr60	.0309357	.0048671	6.36	0.000	.0212912	.0405802
postnr61	.1082682	.0037921	28.55	0.000	.1007538	.1157826
postnr62	.636738	.0324133	19.64	0.000	.5725089	.7009672
postnr63	-.0107625	.0206044	-0.52	0.602	-.0515915	.0300664
postnr64	.7946458	.0595049	13.35	0.000	.6767328	.9125588
postnr65	.4432646	.0307366	14.42	0.000	.3823579	.5041713
postnr66	-.0095326	.0100572	-0.95	0.345	-.0294616	.0103965
postnr67	-.0884151	.0100904	-8.76	0.000	-.10841	-.0684203
postnr68	.272805	.0100229	27.22	0.000	.252944	.2926659
postnr69	.2551317	.0251659	10.14	0.000	.2052637	.3049997
postnr70	.2166784	.0143373	15.11	0.000	.1882681	.2450887
postnr71	.3803563	.0056209	67.67	0.000	.369218	.3914945
postnr72	.1077422	.0097556	11.04	0.000	.0884108	.1270736
postnr73	.3053038	.0158249	19.29	0.000	.2739456	.3366619
postnr74	.0793222	.0120191	6.60	0.000	.0555056	.1031389
postnr75	.1263038	.0060552	20.86	0.000	.1143051	.1383026

postnr76	.1542273	.014292	10.79	0.000	.1259068	.1825478
postnr77	.5099187	.0278439	18.31	0.000	.4547441	.5650932
postnr78	.6267639	.0228872	27.38	0.000	.5814115	.6721164
postnr79	.8027548	.0153486	52.30	0.000	.7723405	.8331692
postnr80	.4454484	.0105183	42.35	0.000	.4246057	.4662911
postnr81	.1760346	.0035088	50.17	0.000	.1690818	.1829875
postnr82	.2985239	.0051755	57.68	0.000	.2882683	.3087795
postnr83	.3518959	.0123495	28.49	0.000	.3274245	.3763673
postnr84	.295209	.0067504	43.73	0.000	.2818326	.3085854
postnr85	.4229223	.0032925	128.45	0.000	.4163981	.4294465
postnr86	.3714603	.0081198	45.75	0.000	.3553704	.3875502
postnr87	.0239011	.018985	1.26	0.211	-.0137189	.061521
postnr88	.4847435	.005252	92.30	0.000	.4743364	.4951507
postnr89	.1659483	.004733	35.06	0.000	.1565697	.175327
postnr90	.0152231	.006002	2.54	0.013	.0033298	.0271164
postnr91	.0701626	.0027282	25.72	0.000	.0647565	.0755686
postnr92	-.1036029	.0034719	-29.84	0.000	-.1104827	-.096723
postnr93	.2347015	.0024367	96.32	0.000	.2298731	.23953
postnr94	.1899825	.0032927	57.70	0.000	.1834579	.1965072
postnr95	-.0994007	.0040499	-24.54	0.000	-.1074258	-.0913756
postnr96	.0809012	.0014788	54.71	0.000	.0779709	.0838315
postnr97	.0746685	.0025728	29.02	0.000	.0695703	.0797667
postnr98	.23672	.0068325	34.65	0.000	.2231809	.250259
postnr99	.2260718	.0041242	54.82	0.000	.2178995	.2342441
postnr100	.0031439	.0051101	0.62	0.540	-.006982	.0132698
postnr101	-.1288383	.0090816	-14.19	0.000	-.1468341	-.1108424
postnr102	.0322909	.0031644	10.20	0.000	.0260204	.0385614
postnr103	.0068553	.0041235	1.66	0.099	-.0013158	.0150263
postnr104	.0600608	.0065586	9.16	0.000	.0470646	.073057
postnr105	.3880819	.0366879	10.58	0.000	.3153824	.4607813
postnr106	.0896484	.013775	6.51	0.000	.0623523	.1169445
postnr107	.2017477	.0220045	9.17	0.000	.1581444	.2453511
postnr108	.1615631	.0063654	25.38	0.000	.1489497	.1741765
postnr109	.0562389	.0081744	6.88	0.000	.0400408	.0724371
postnr110	.1325812	.008046	16.48	0.000	.1166375	.1485249
postnr111	.0791131	.0104073	7.60	0.000	.0584903	.0997359
_cons	14.20421	.0255301	556.37	0.000	14.15363	14.2548

VEDLEGG 3:

7,8 x 8,4 meter garasje med leilighet. 23,5° tak.

