

Hva kan forklare de regionale forskjellene i boligprisutvikling i Norge?

En tidsserieanalyse av Oslo, Stavanger, Kristiansand, Bergen og Tromsø
2005-2018

ANDREAS RASMUSSEN
JOHANNES OSELAND

VEILEDER
Anne Wenche Emblem

Universitetet i Agder, 2020
Fakultet for Handelshøyskolen ved UiA
Institutt for Økonomi

Forord

Denne masterutredningen er skrevet ved Universitetet i Agder våren 2020, og markerer avslutningen av masterstudiet i økonomi og administrasjon. Oppgaven er en obligatorisk del av studiet og tilsvarer 30 studiepoeng. Hensikten med oppgaven er å gi oss som studenter muligheten til å fordype seg i et eller flere fagfelt, i tillegg til å lære hvordan man gjennomfører et vitenskapelig arbeid.

Denne oppgaven er i hovedsak skrevet innenfor fagfeltet eiendomsøkonomi. Dette er et fagfelt vi i utgangspunktet hadde begrenset kunnskap om. Arbeidet har derfor vært krevende, men har også ført til en veldig spennende og lærerik prosess. I løpet av denne prosessen har vi tilegnet oss uvurderlig kunnskap som vi vil dra nytte av senere i livet.

Vi vil gjerne rette en stor takk til veileder Anne Wenche Emblem for det gode samarbeidet underveis i oppgaveskrivingen. Hennes konstruktive og ærlige tilbakemeldinger har vært med å heve oppgaven og motivere oss i arbeidet.

Kristiansand, 04.06.2020

Andreas Rasmussen

Johannes Oseland

Sammendrag

I denne masterutredningen vil vi undersøke boligprisutviklingen i det norske boligmarkedet. Vi vil undersøke boligprisutviklingen på et regionalt nivå, og vi vil studere utviklingen i henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø. Utviklingen i boligpriser har variert vidt mellom disse byene, og målet med oppgaven er derfor å undersøke om fundamentale faktorer kan bidra til å forklare disse regionale forskjellene. Vi forsøker å forklare utvikling i boligpris med følgende fundamentale faktorer: inntekt, rente, arbeidsledighet, boligbygging og en forventningsindikator.

I første del av utredningen presenterer vi sentral teori for oppgaven, og videre ser vi nærmere på boligmarkedet i Norge, hvor vi tar for oss utvikling i boligpris og andre sentrale faktorer. Vi gjennomgår deretter vårt datagrunnlag for oppgaven og økonometrisk metode som benyttes i de empiriske analysene. Vi konstruerer boligprismodeller for hver av de respektive regionene med utgangspunkt i empirisk metode samt boligprismodellen Dag Henning Jacobsen og Bjørn E. Naug konstruerer i sin artikkel. Vi konstruerer også en paneldatamodell.

Vi bruker modellene til å foreta økonometriske analyser av de regionale boligmarkedene i perioden første kvartal 2005 til fjerde kvartal 2018. Resultatene viser at 38-58% av variasjonen i boligpriser blir forklart av de fundamentale faktorene vi bruker som forklaringsvariabler i modellene våre. Vi finner at modellen for Tromsø har lavest forklaringskraft (38,5%), mens boligprisene i Bergen og Oslo forklares best av modellen (57-58%). Vi finner at forventning har en langsiktig signifikant påvirkning på boligprisene i et flertall av byene. Vi ender opp med få signifikante koeffisientestimat for forklaringsvariablene. Dog forkaster vi ikke etablert teori om hvilke faktorer som antas å ha en påvirkning på boligpriser. Estimering av samtlige modeller og økonometriske tester er utført i STATA 16. Øvrig databehandling er utført i Excel.

Innholdsfortegnelse

Forord.....	i
Sammendrag	ii
Figuroversikt.....	v
Tabelloversikt.....	vi
1. Innledning	1
1.1 Problemstilling.....	1
1.2 Motivasjon for oppgaven.....	1
1.3 Avgrensning.....	2
1.4 Oppgavens disposisjon.....	2
2. Teori.....	4
2.1 Prisdannelsen i boligmarkedet.....	4
2.1.1 Tilbud	4
2.1.2 Etterspørsel.....	6
2.2 Likevekt på kort sikt.....	9
2.3 Likevekt på mellomlang sikt.....	10
2.4 Likevekt på lang sikt.....	11
2.5 Boligboble.....	12
2.6 Jacobsen og Naugs boligprismodell.....	12
2.7 Generelle karakteristika ved boligmarkedet	13
3. Det norske boligmarkedet.....	14
3.1 Historisk utvikling i boligprisindeksen	14
3.2 Utvikling i nyere tid.....	16
3.2.1 Kvadratmeterpris solgte boliger	16
3.2.2 Disponibel Realinntekt	17
3.2.3 Arbeidsledighet.....	19
3.2.4 Husholdningers gjeldsgrad.....	20
3.2.5 Konsumprisindeksen (KPI)	22
3.2.6 Bokostnadsindeksen for norske boliger	23
3.2.7 Byggekostnadsindeksen	24
3.2.8 Boligbygging i lys av folketilvekst.....	25
3.2.9 Førstegangskjøpere og andel sekundærboliger	26
3.3 Regional utvikling i boligpriser	28
4. Data.....	30
4.1 Kvadratmeterpris.....	30
4.2 Inntekt	32

4.3. Rente	33
4.4 Arbeidsledighet.....	35
4.5 Nettoinnflytting.....	37
4.6 Boligbygging.....	38
4.7 Forventning.....	39
4.8 Utledning av forventningsindikator	41
4.9 Transformasjon av nominelle verdier til reelle verdier	42
4.10 Utelatte variabler.....	44
5. Metode.....	46
5.1 Modellspesifikasjon	46
5.2 Metodisk tilnærming	46
5.3 Regresjonsanalyse og minste kvadraters metode (OLS)	47
5.4 Stasjonæritet og Spuriøsitet	52
5.5 Autokorrelasjon	55
5.5.1 Statistisk Test for autokorrelasjon.....	55
5.5.2 Grafisk test for autokorrelasjon	57
5.6 Heteroskedastisitet.....	59
5.7 Normalitet.....	61
5.7.1 Grafisk test for normalfordeling	61
5.7.2 Statistisk test for normalfordeling.....	62
5.8 Multikollinearitet	63
5.8.1 Test for multikollinearitet.....	63
5.9 Kointegrasjon & Feiljusteringsmodell	64
5.9.1 Kointegrasjon.....	64
5.9.2 Feiljusteringsmodell	65
5.10 Paneldata.....	66
6. Økonometriske tester og analyse.....	69
6.1 Testing for Stasjonæritet.....	69
6.2 Feiljusteringsmodell & test for Autokorrelasjon	70
6.3 Test for Heteroskedastisitet	73
6.4 Test for Normalfordeling.....	74
6.5 Test for Multikollinearitet.....	75
6.6 Regresjonsresultater	77
6.7 Regresjonsresultater fra paneldatamodelle	83
7. Konklusjon	91
Referanseliste	94

Figuroversikt

Figur 2.1: Tilbudskurve for boliger på kort, mellomlang og lang sikt	6
Figur 2.2: Epperspørselskurven.....	9
Figur 2.3: Kortsiktig likevekt.....	10
Figur 2.4: Likevekt på mellomlang sikt	11
Figur 2.5: Langsiktig likevekt.....	11
Figur 3.1: Utvikling i boligprisindeksen for Norge, i forhold til konsumprisindeksen.....	14
Figur 3.2: Gjennomsnittlig nominell kvadratmeterpris i Norge for selveierboliger, 2002-2019	17
Figur 3.3: Indeks for norske husholdningers realdisponible inntekt, 2002K1-2018K4.....	18
Figur 3.4: Boligpriser i forhold til disponibel inntekt per innbygger (15-74 år)	19
Figur 3.5: Arbeidsledighet i Norge, 2002M11-2019M11	20
Figur 3.6: Andelen husholdninger med samlet gjeld større enn 3 ganger samlet inntekt, 2007-2017.....	21
Figur 3.7: Konsumprisindeks, 2015=100, 2005M1-2018M12.....	22
Figur 3.8: Bokostnadsindeks for norske boliger, 2010-2017	23
Figur 4.1: Regionale gjennomsnittlige kvadratmeterpriser, 2005K1-2018K4	31
Figur 4.2: Gjennomsnittlig personlig nominell bruttoinntekt etter region, 2005K1-2018K4. 33	
Figur 4.3: Bankenes nasjonale gjennomsnittlige utlånsrente, 2005K1-2018K4.....	35
Figur 4.4: Arbeidsledighet etter region, 2005K1-2018K4	36
Figur 4.5: Innenlandsk nettoinnflytting etter region, 2005K1-2018K4	38
Figur 4.6: Igangsatte eneboliger etter region, 2005K1-2018K4	39
Figur 4.7: Forventningsbarometeret, 2005K1-2018K4.....	40
Figur 4.8: Justert forventningsindikator etter region, 2005K2-2018K4	42
Figur 4.9: Reelle transformasjoner av nominelle boligpris- og inntektsdata, 2005K1-2018K4	43
Figur 4.10: Reell transformasjon av bankenes nominelle utlånsrente, 2005K1-2018K4.....	44
Figur 5.1: Positiv autokorrelasjon	58
Figur 5.2: Negativ autokorrelasjon	58
Figur 5.3: Ingen autokorrelasjon.....	59
Figur 5.4: Plot av homoskedastisitet og heteroskedastisitet	60
Figur 5.5: Normalfordeling vs. skjevfordeling.....	61
Figur 5.6: Normalfordeling vs. leptokurtisk fordeling	62

Tabelloversikt

Tabell 6.1: Testresultater for utvidet Dickey-Fuller test for stasjonæritet, etter by.....	69
Tabell 6.2: Testresultater for Ljung-Box test, etter by	72
Tabell 6.3: Testresultater for Breusch-Pagan test, etter by	73
Tabell 6.4: Testresultater for Skjevhet-Kurtose test, etter by	74
Tabell 6.5: Regresjonskoeffisienter fra regresjonsmodell med standardfeil i parentes, etter by	77
Tabell 6.6: Regresjonsresultater fra paneldataregresjon, med og uten interaksjoner	85

1. Innledning

I kapittel 1 presenteres problemstillingen vår, og motivasjonen for valget av oppgaven.

Videre gir vi en forklaring av oppgavens avgrensning og oppgavens disposisjon.

1.1 Problemstilling

Boligprisene i Norge har over en lengre periode økt. Forskjellene i boligprisutvikling mellom ulike regioner i Norge har dog vært store, hvor noen regioner har opplevd markant økning i boligpriser, mens boligprisutviklingen i andre regioner har vært relativt lav. I noen regionale områder har til og med boligprisene de siste årene blitt redusert. Det ble derfor interessant for oss å finne ut hva som kan være årsakene til dette. Problemstillingen til oppgaven vår er derfor:

Hva kan forklare de regionale forskjellene i boligprisutvikling?

For å kunne svare på dette vil vi studere utviklingen av en rekke faktorer som i litteraturen antas å påvirke boligpriser, og se hvorvidt disse påvirker ulike regioner i ulik grad. Videre i oppgaven tar vi utgangspunkt i fem regioner, henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø. Regioner definerer vi her som kommuner, og vi bruker kommuneinndelingen per 2019.

1.2 Motivasjon for oppgaven

Motivasjonen for oppgaven var i utgangspunktet vår felles interesse for boligmarkedet. Dette er et marked som de aller fleste, på et eller annet tidspunkt, vil involvere seg i.

Innfallsvinkelen vi valgte på oppgaven ved å se på regionale forskjeller i boligprisutvikling og årsaker bak, hadde sitt utspring i at vi begge var interessert i å forstå sammenhenger og dynamikker som kan bidra til å forklare ulikheter i boligmarkedet. Kjøp av bolig er ofte den største investeringen man gjør i løpet av livet, og en bedre og mer gjennomgående forståelse av boligmarkedet vil gi oss verdifull innsikt.

1.3 Avgrensning

Det er mange momenter som kan forklare problemstillingen vår, og det blir vanskelig å inkludere alle. På grunn av begrenset tid og tilgang på data kan vi ikke inkludere alle relevante faktorer som forklaringsvariabler. Vi har derfor valgt ut fundamentale faktorer som vi i tillegg har tilstrekkelig med data på. Disse variablene er henholdsvis inntekt, rente, arbeidsledighet, boligbygging samt en forventningsindikator, som er en variabel som inkluderer befolkningens forventninger til egen og landets økonomi.

Når vi forsøker å forklare forskjellene i regionale boligpriser begrenser vi oss til tidsperioden første kvartal 2005 til fjerde kvartal 2018. Valget av tidsperiode er først og fremst bestemt av tilgangen på data. Det kan dog argumenteres for at den valgte perioden også er interessant med tanke på hva som inntreffer i denne tidsperioden. Både finanskrisen og oljekrisen inntraff i denne perioden, og hvordan disse krisene kan ha påvirket boligpriser og utvalgte variabler i de ulike regionene er av interesse å studere.

For å få en god bredde i oppgaven valgte vi å ta for oss fem regioner. Oslo, Stavanger, Kristiansand og Bergen utgjør fire av de fem største byene i Norge, og dermed er det sentralt å studere disse storbyene nærmere. Tromsø er ikke helt å regne som en storby sammenlignet med de andre byene, og har også en annen geografisk beliggenhet. I kontrast til de andre byene ligger Tromsø langt nord i Norge, hvor klimaet er annerledes enn i de andre byene. Vi mener det er interessant å inkludere en by som skiller seg litt fra de andre byene. Grunnlaget for inklusjonen av Tromsø er dermed i denne utredningen hovedsakelig for å undersøke om boligprisene i Tromsø drives i lik grad av de samme fundamentale faktorene som i de andre byene, eller om boligprisene i Tromsø i større eller mindre grad bestemmes av andre faktorer.

1.4 Oppgavens disposisjon

Masterutredningen er delt inn i 7 kapitler med tilhørende underkapitler. Innledningsvis har vi gjennomgått problemstillingen og motivasjonen for oppgaven, samt presentert de avgrensningene vi har gjort. I kapittel 2 går vi gjennom grunnleggende teori som forklarer tilbud og etterspørsel i boligmarkedet, samt hvordan likevekt oppstår. Videre presenterer vi Jacobsen og Naugs boligprismodell, som legger grunnlaget for vår tilnærming til problemstillingen. Avslutningsvis i kapittelet presenteres generelle karakteristika til boligmarkedet i Norge og boligmarkedet for øvrig. I kapittel 3 ser vi på boligprisutviklingen i det norske boligmarkedet. Vi presenterer utviklingen i sentrale faktorer sett i lys av kriser og

konjunktorendringer. Dette presenteres med den hensikt å få en bedre forståelse av den historiske boligprisutviklingen og utviklingen i sentrale faktorer som antas å påvirke boligpriser.

I kapittel 4 presenterer vi regionale data for de fundamentale faktorene som vi inkluderer i den endelige boligprismodellen vår. Vi presenterer faktorene hver for seg og gir en forklaring på hvorfor vi inkluderer dem. Avslutningsvis går vi gjennom utelatte variabler. I kapittel 5 går vi gjennom vår metodiske tilnærming. Vi presenterer aktuell teori og metode, som vil fungere som et fundament for de empiriske analysene våre. I kapittel 6 gjennomfører vi økonometriske tester basert på metodene vi presenterte i kapittel 5, og gir en drøfting av resultatene. I kapittel 7 trekker vi våre endelige konklusjoner, og reflekterer rundt funnene vi har gjort. Vi gjennomgår også kritikk av oppgaven, samt forslag til videre forskning.

2. Teori

Vi ønsker å identifisere de ulike faktorene som kan forklare de regionale forskjellene i boligprisutvikling i Norge. For å kunne svare på dette blir det naturlig å innledningsvis presentere grunnleggende teori som identifiserer faktorer som antas å ha en påvirkning på boligpriser. Dette vil bidra til å skape et fundament for den videre besvarelsen av problemstillingen.

Først vil det fremstilles en teoretisk gjennomgang av tilbuds- og etterspørselsfunksjonen, og de underliggende faktorene som inngår i fastsettelsen av boligpriser på tilbuds- og etterspørselssiden vil bli gjort rede for. Dette vil gi en bedre forståelse av hva som driver boligprisene. Videre fremstilles det hvordan tilbud og etterspørsel danner likevekt i boligmarkedet i ulike tidshorisonter. Deretter presenteres teori om boligbobler, samt Jacobsen og Naug sin boligprismodell. Avslutningsvis presenteres karakteristika med boligmarkedet generelt, samt i Norge.

2.1 Prisdannelsen i boligmarkedet

Boligmarkedet skiller seg ikke prinsipielt fra andre markeder (NOU, 2004:2). Boligprisen dannes samlet av etterspørselen etter, og tilbudet av boliger. Etterspørselen etter bolig kan deles inn i to segmenter: etterspørsel etter bolig som boformål og etterspørsel etter bolig som investeringsobjekter. Vi vil i dette kapitlet gå nærmere inn på etterspørselen etter bolig som boformål. Tilbudet av boliger deles inn i brukte og nye boliger. Boligprisene på brukte boliger er til enhver tid avhengig av hvilke objekter som er lagt ut til salgs, samt den eksisterende boligmassen i markedet. Tilbudskurven avhenger i stor grad av tidshorisonten, hvor det vil være forskjeller i tilbudet på kort, mellomlang og lang sikt.

2.1.1 Tilbud

På kort sikt antas tilbudet i boligmarkedet å være gitt. Dette fordi det tar tid å bygge nye boliger, samt at nybyggingen er lav sammenlignet med den totale boligmassen. I tillegg vil det oppstå kapasitetsbegrensninger knyttet til arbeidskraft på kort sikt. Dette kalles uelastisk tilbud. Ettersom tilbudet på kort sikt er uelastisk, vil boligprisene i hovedsak endre seg på bakgrunn av endringer i etterspørselen. Over lengre tid vil imidlertid boligmassen tilpasse seg etterspørselen. For å få en forståelse av prisendringer over tid bør en derfor identifisere

faktorer som forklarer utviklingen i boligmassen. Dette vil typisk være bygge- og tomtekostnader samt pris på nye boliger (Jacobsen og Naug, 2004).

Tilbudet av boliger i periode t er et produkt av forrige periodes boligmasse (H_{t-1}) fratrukket depresieringsraten $\delta_t(H_{t-1})$ i tillegg til antall nye boliger (C_t), og kan uttrykkes ved hjelp av følgende uttrykk (Hendry & Wallis, 1984):

$$H_t^T = (1 - \delta_t) H_{t-1} + C_t, \quad (2.1)$$

hvor H_t^T er boligmasse i periode t , δ_t er depresieringsraten, H_{t-1} er forrige periodes boligmasse og C_t er antall nybygg.

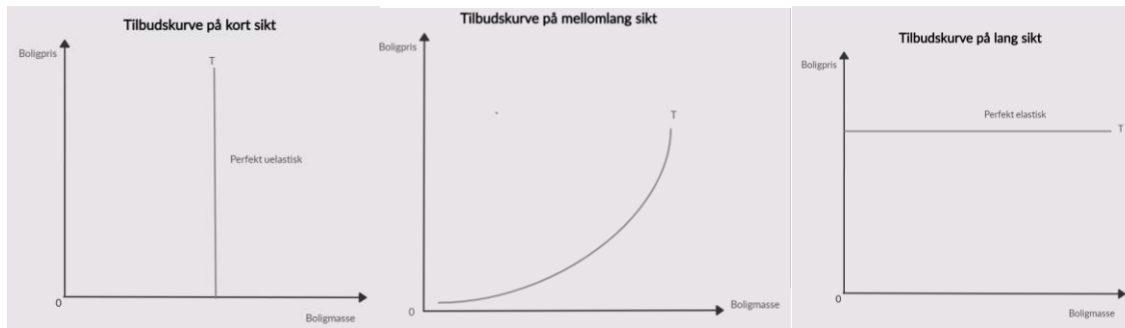
På kort sikt antas tilbudet å være upåvirket av antall nybygg og depresiering, slik at boligtilbudet i periode t vil tilnærmet utgjøre forrige periodes boligtilbud, $H_t^T \approx H_{t-1}$. Boligtilbudet anses å være tilnærmet konstant på kort sikt og tilnærmet perfekt uelastisk. Et sjokk i etterspørselen vil dermed påvirke prisen i stor grad på kort sikt.

På lengre sikt vil tilbudet være mer elastisk, og boligmassen vil øke dersom antall nybygg er høyere enn depresieringen, og under antakelsen at det er lønnsomt å starte nye byggeprosjekter. På lengre sikt vil det likevel kunne eksistere begrensninger i tilgang på arbeidskraft og tomter. Dette gir oss en konveks og stigende tilbudskurve på mellomlang sikt.

På veldig lang sikt vil det være naturlig å anta at tilbudskurven går mot å være perfekt elastisk, under forutsetningen av at det ikke er begrensninger i arbeidskraft og tomter. I et marked med perfekt frikonkurranse vil det være ubegrenset tilgang på arbeidskraft og innsatsfaktorer, og langsiktig likevekt vil oppstå der økt etterspørsel kun medfører økt boligmasse. Dette tilsvarer en perfekt elastisk tilbudskurve. Om det vil være ubegrenset tilgang på tomter kan dog diskuteres. Det kan tenkes at topografien til ulike byer fører til begrensninger i tomtetilgang. Et typisk eksempel på dette er hvis en by er omringet av fjell.

Ifølge Leamer (2002) finnes det ingenting som heter boligmangel i et fritt velfungerende marked. Leamer argumenterer at det kun vil finnes mangel på boliger dersom prismekanismen ikke tillates å fungere slik at tilbud møter etterspørsel. I et velfungerende

boligmarked, slik som i storbyene i Norge, vil et begrenset tilbud av bolig føre til høyere priser. Dette medfører at de som kjøper bolig vil endre sine kjøpspreferanser i form av å endre krav til beliggenhet, og dermed vil likevekten opprettholdes. Dermed vil tilbudet av bolig være perfekt elastisk på veldig lang sikt.



Figur 2.1: Tilbudskurve for boliger på kort, mellomlang og lang sikt

Kilde: Egenutviklede figurer (2020)

2.1.2 Etterspørsel

Etterspørselen av bolig kan som nevnt deles opp i to deler. Etterspørselen etter bolig som boformål og etterspørselen etter bolig som investeringsobjekter. Ifølge Statistisk sentralbyrå (2020) eier 82,1% av den norske befolkning egen bolig, hvor igjen 86% av disse er selveiere, mens de resterende 14% utgjør andels- eller aksjeeiere. Etterspørselen etter bolig som boformål utgjør altså den klart største delen av boligmarkedet, og det blir dermed naturlig å legge vekt på denne delen av etterspørsel videre i oppgaven. Jacobsen og Naug (2004) tar i sin artikkel utgangspunkt i følgende etterspørselsfunksjon:

$$H^D = f\left(\frac{V}{P}, \frac{V}{HL}, Y, X\right), \quad f_1 < 0, f_2 < 0, f_3 > 0, \quad (2.2)$$

hvor:

H^D = Etterspørsel etter boliger

V = Samlet bokostnad for en typisk eier

P = Indeks for prisene for andre varer og tjenester enn bolig

HL = Samlet bokostnad for en typisk leietaker (husleie)

Y = Husholdningens disponible realinntekt

X = En vektor av andre fundamentale faktorer som påvirker boligetterspørselen

f_i = Den deriverte av $f(\cdot)$ med hensyn på argument i

Fra likning 2.2 ser vi at etterspørselen etter eierboliger øker dersom husholdningens disponible realinntekt øker. Vi ser også at etterspørselen vil avta dersom bokostnadene for å eie øker i forhold til kostnadene ved å leie eller prisene på andre varer og tjenester.

Bokostnaden for en selveier måler verdien av godene som selveieren gir avkall på ved å eie og benytte en bolig i en periode. Første ledd i likningen $\frac{V}{P}$ beskriver de reelle bokostnadene.

De reelle bokostnadene kan spesifiseres ytterligere ved hjelp av følgende uttrykk:

$$\frac{V}{P} = \frac{P^H}{P} BK = \frac{P^H}{P} [i(1 - \tau) - E\pi - (E\pi^{PH} - E\pi)], \quad (2.3)$$

hvor

P^H = Pris på en gjennomsnittsbolig (målt i kroner)

BK = Bokostnad per realkrone investert i bolig

i = Nominell rente

τ = Marginalskattesats på kapitalinntekter og -utgifter

$E\pi$ = Forventet inflasjon (den forventede veksten i P og HL , målt som rate)

$E\pi^{PH}$ = Forventet vekst i P^H (målt som rate)

Fra likning 2.3 kan vi hente ut uttrykket for realrenten etter skatt. Uttrykket $[i(1 - \tau) - E\pi]$ måler de reelle rentekostnadene ved boliglån og de reelle renteinntektene en går glipp av ved å ha egenkapital plassert i bolig. Fra likning 2.3 finner vi også uttrykket for den forventede realprisveksten på boligen/boliger som er $[E\pi^{PH} - E\pi]$. Dersom $[E\pi^{PH} - E\pi]$ øker vil også den forventede boligformuen gå opp. En økning av boligformuen betyr igjen at de reelle bokostnadene ved å eie går ned. Dermed blir det relativt billigere å eie kontra å leie. Dette fører til at etterspørselen etter eierboliger går opp. Dette samsvarer med likning 2.2, hvor bokostnadene på førsteordens derivert form har et negativt fortegn.

Likning 2.3 kan videre forenkles til: $\frac{V}{P} = \frac{PH}{P} [i(1 - \tau) - E\pi^{PH}] = \frac{PH}{P} BK$. Variabelen BK angir nå den nominelle renten etter skatt minus den forventede økningen i nominelle boligpriser.

Likningene 2.2 og 2.3 beskriver etterspørselen etter boliger for boformål. Det kan med rimelighet antas at variablene brukt i disse likningene også er med på å forklare etterspørselen etter boliger som investeringsobjekter. Med denne antakelsen kan man forvente at etterspørselen etter boliger som investeringsobjekt øker i takt med realinntektene. På samme måte vil også etterspørselen øke dersom kostnadene ved å leie stiger i forhold til kostnadene ved å eie. En lavere rente eller tilsvarende økning i $E\pi^{PH}$ vil også øke etterspørselen etter bolig som investeringsobjekt ettersom det blir relativt gunstigere å investere i bolig kontra å ha penger i banken.

Det tredje leddet (Y) i likning 2.2 forklarer husholdningens disponible realinntekt. Det kan defineres videre som vist i likning 2.4 nedenfor:

$$Y = \frac{YN}{P^{a_1} HL^{a_2} PH^{a_3}}, a_1 + a_2 + a_3 = 1 \quad (2.4)$$

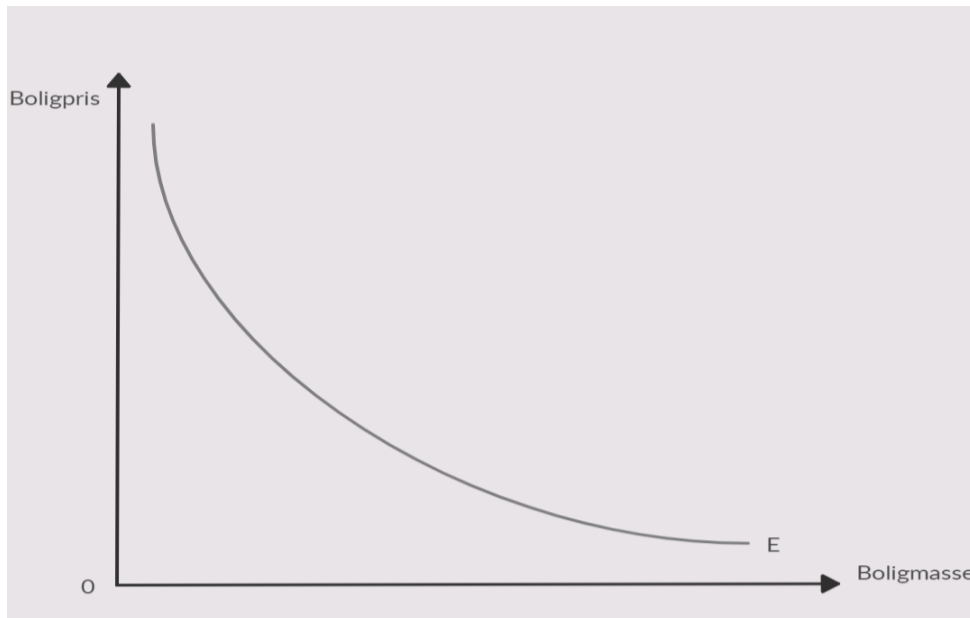
YN angir den nominelle disponible inntekten. Denne likningen tar hensyn til at husholdningens samlede kjøpekraft i boligmarkedet reduseres når boligprisene øker. Når boligprisene øker vil dette, alt annet like, gi redusert kjøpekraft til førstegangskjøpere. Det vil samtidig gi økt kjøpekraft til de som skal tre ut av boligmarkedet. For øvrige husholdninger under ett blir dog virkningen null. Det er viktig å legge merke til at når boligprisene øker vil ikke de som skal ut av boligmarkedet benytte seg av deres økte kjøpekraft til å kjøpe bolig. Dette er grunnen til at den samlede kjøpekraften går ned når boligprisene går opp (Jacobsen og Naug, 2004).

Det fjerde og siste leddet i likning 2.2 er en vektor av andre fundamentale faktorer som påvirker bolig etterspørselen. Dette kan være demografiske forhold, bankenes utlånspolitikk, husholdningenes preferanser, forventninger om fremtidige inntekter og bokostnader.

Det er viktig å merke seg at det er spesielt tre grunner som gjør at fremtidige inntekter og bokostnader er viktige. Det første er at boliger er et varig forbruksgode. Det andre er at boligkjøp for de aller fleste husholdninger utgjør den største investeringen man gjør i løpet av livet. Det tredje er at for de aller fleste husholdninger som er førstegangskjøpere, så finansieres en stor del av kjøpet ved hjelp av lån. At førstegangskjøpere i stor grad finansierer kjøpet ved hjelp av lån er også grunnen til at bankenes utlånspolitikk kan ha potensielt stor

innvirkning på boligprisene. Siden boliger er finansiert hovedsakelig gjennom lån, vil lettere tilgang på kreditt medvirke til økt kjøpekraft og etterspørsel etter boliger. Offentlige reguleringer, bankenes lønnsomhet, kundenes forventede betalingsevne samt panteverdier er faktorer som bidrar til å avgjøre utlånspolitikken (Jacobsen og Naug, 2004).

Vi kan illustrere etterspørselen etter boliger grafisk. Figur 2.2 illustrerer at etterspørselen etter boliger synker når boligprisene øker.



Figur 2.2: Etterspørselskurven

Kilde: Egenutviklet figur (2020)

2.2 Likevekt på kort sikt

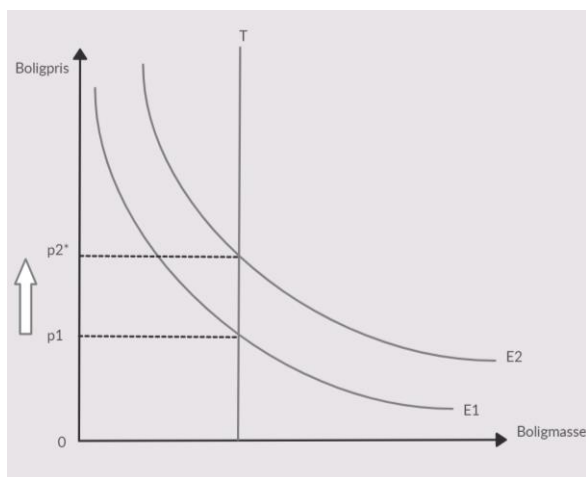
Likevektstilpasningen dannes i punktet hvor tilbøyeligheten til kjøper og selger er lik i de forskjellige tidshorisontene. Jacobsen og Naug (2004) argumenterer for at tilbudet av boliger på kort sikt vil være gitt. Dette på grunn av lang byggetid samt at nybyggingen er lav i forhold til den totale boligmassen. Vi tar derfor utgangspunkt i at tilbudskurven i boligmarkedet på kort sikt er tilnærmet konstant og perfekt uelastisk, mens etterspørselskurven som tidligere beskrevet har en fallende form og avtar med høyere pris. I henhold til Jacobsen og Naug (2004) kan likevekt på kort sikt dermed uttrykkes ved:

$$H^D = f\left(\frac{V}{P}, \frac{V}{HL}, Y, X\right) = H^S \quad (2.5)$$

hvor H^S = tilbudet av boliger

Grafisk vil likevektsprisen i boligmarkedet være gitt ved skjæringspunktet mellom etterspørselskurven og den uelastiske tilbudskurven. Boligprisene vil på kort sikt variere med endringer i etterspørselen, og en økning i etterspørsel vil føre til en økning i pris. Med andre ord vil etterspørselen bestemme prisene.

Figur 2.3 illustrerer konsekvensene av et positivt etterspørselssjokk fra E1 til E2. Tilbudet er som sagt tilnærmet konstant på kort sikt, noe som resulterer i en økning i boligpris fra p_1 til p_2^* , og boligtilbudet forblir uendret.

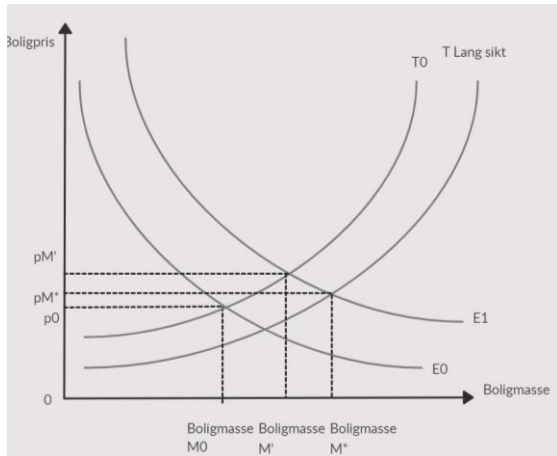


Figur 2.3: Kortsiktig likevekt

Kilde: Egenutviklet figur (2020)

2.3 Likevekt på mellomlang sikt

På mellomlang sikt forventes det at endringer i etterspørselssiden i boligmarkedet vil påvirke boligprisene. Figur 2.4 illustrerer likevektstilpasningen på mellomlang sikt. Med utgangspunkt i tilpasningen i skjæringspunktet mellom E0 og T0, vil en økning i etterspørsel fra E0 til E1 i begynnelsen øke boligprisen til pM' og tilbudet til Boligmasse M' . På mellomlang sikt antas det at økt etterspørsel vil gjøre byggeaktivitet mer attraktivt, og nye aktører vil komme på markedet. Dermed vil tilbudskurven få et positivt skift. Det positive skiftet i tilbudskurven fører til at boligprisen synker, og vi får en ny likevektstilpasning i pM^* og Boligmasse M^* .

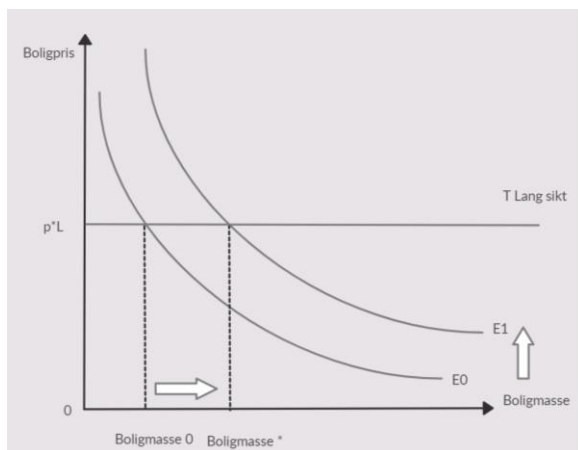


Figur 2.4: Likevekt på mellomlang sikt

Kilde: Egenutviklet figur (2020)

2.4 Likevekt på lang sikt

Som nevnt i delkapittel 2.1.1 vil boligtilbudet på veldig lang sikt antas å være nærmest perfekt elastisk, og dermed vil et positivt etterspørselssjokk medføre en økning i boligtilbudet, mens prisen forblir uforandret. Dette oppstår fordi det på lang sikt antas å ikke eksistere noen begrensninger i form av arbeidskraft eller tilgjengelige tomter, slik at en endring i etterspørselen vil bli møtt av en tilsvarende endring i boligtilbudet. Figur 2.5 illustrerer hvordan boligetterpørselen påvirker likevektstilpasningen på lang sikt. Initielt vil tilpasningen finnes i skjæringspunktet mellom E0 og T Lang sikt. En økning i etterspørsel fra E0 til E1 medfører at prisene først vil øke på kort sikt, men prisen vil falle tilbake til opprinnelig verdi p^*L etter hvert som etterspørselen blir møtt av tilbudet, som øker fra Boligmasse 0 til Boligmasse*.



Figur 2.5: Langsiktig likevekt

Kilde: Egenutviklet figur (2020)

2.5 Boligboble

Når vi skal analysere boligprisene og årsakene bak peker vi ofte på det vi kaller fundamentale faktorer. Disse ble definert i likning 2.1 og 2.2 og sentrale variabler er blant annet antall nybygg, samlede bokostnader, husholdningens disponible realinntekt og renteendringer. Variabel X i likning 2.2 er også viktig i denne sammenhengen, som er en vektor av andre fundamentale forhold som for eksempel demografiske forhold, bankenes utlånspolitikk, husholdningenes preferanser og forventninger om fremtidige inntekter. Dette er grunnleggende faktorer som er med på å forklare tilbud og etterspørsel og er med på å gi et helhetlig bilde på hvorfor boligprisene er som de er.

Men hva om boligprisene er kunstig høye og fundamentale faktorer ikke kan forklare hvorfor prisene er som de er? Vi snakker ofte da om det vi kaller for en boligboble. Stiglitz (1990) forklarer begrepet slik: *“If the reason that the price is high today is only because investors believe that the selling price will be high tomorrow—when “fundamental” factors do not seem to justify such a price—then a bubble exists.”* Når vi snakker om en boligboble snakker vi som regel om en boligboble på et nasjonalt nivå. Målet med denne oppgaven er ikke å utforske om det eksisterer en boligboble i Norge, men prinsippene bak en boligboble er interessant å ta med seg når vi skal utforske regionale forskjeller i boligprisene. Det kan for eksempel tenkes at noen regioner har boligpriser som ikke like enkelt kan forklares ved hjelp av fundamentale faktorer.

2.6 Jacobsen og Naugs boligprismodell

Dag Henning Jacobsen og Bjørn E. Naug publiserte i desember 2004 artikkelen *Hva driver boligprisene?* i tidsskriftet *Penger og Kreditt* ved Norges Bank. I denne artikkelen konstruerer forfatterne en regresjonsmodell med hensikt å forklare hva de viktigste fundamentale forklaringsfaktorene for boligprisene det siste tiåret var. Boligprismodellen de konstruerer inneholder fundamentale forklaringsvariabler som historisk sett har vist å ha den mest signifikante effekten på boligprisene. Boligprismodellen de estimerer viser den samlede effekten en endring i lønnsinntekter, arbeidsledighetsraten, boligmassen, bankenes utlånsrente etter skatt og husholdningers forventning til egen og landets økonomi har på boligprisene. Forfatterne identifiserer i denne artikkelen de viktigste drivkreftene bak boligpriser til å være inntekt, arbeidsledighet, rente og boligmasse.

Langtidssammenhengen for modellen de estimerer ser ut som følger:

$$\text{Boligpris} = -0.12[\text{boligpris}_{t-1} + 4.47 * \text{rente}(1 - \tau)_{t-1} + 0.45 * \text{arbeidsledighet} \\ - 1.66(\text{inntekt} - \text{boligmasse})_{t-1}]$$

2.7 Generelle karakteristika ved boligmarkedet

Pirounakis (2013) argumenterer at boligmarkedet kjennetegnes av å være i aller størst grad differensiert av beliggenhet. Beliggenhet er et unikt kjennetegn ved hver enkelt bolig, hvor to boliger aldri kan ha identisk beliggenhet. Beliggenheten i tillegg til andre egenskaper med boligen gjør at ingen boliger er identiske. Dette gjør boligmarkedet til et heterogent marked, både når det gjelder boligtype og regionale variasjoner (NOU, 2004:2). Boliger er immobile, og det forekommer ved en sjeldenhet at boliger fysisk flyttes på i Norge, grunnet de høye transaksjonskostnadene knyttet til dette. Boliger har typisk en veldig lang fysisk levetid.

Heterogeniteten, substituerbarhet mellom eiendommer, ufullstendig informasjon og et stort antall kjøpere og selgere gjør at boligmarkedet er preget av monopolistisk konkurranse (Pirounakis, 2013). Det vil si at tilbyderne produserer differensierte produkter som skiller seg fra hverandre, men de konkurrerer om de samme etterspørerne.

I tillegg bærer det norske boligmarkedet preg av særegne kjennetegn. Kjennetegn for det norske boligmarkedet er (NOU, 2004:2):

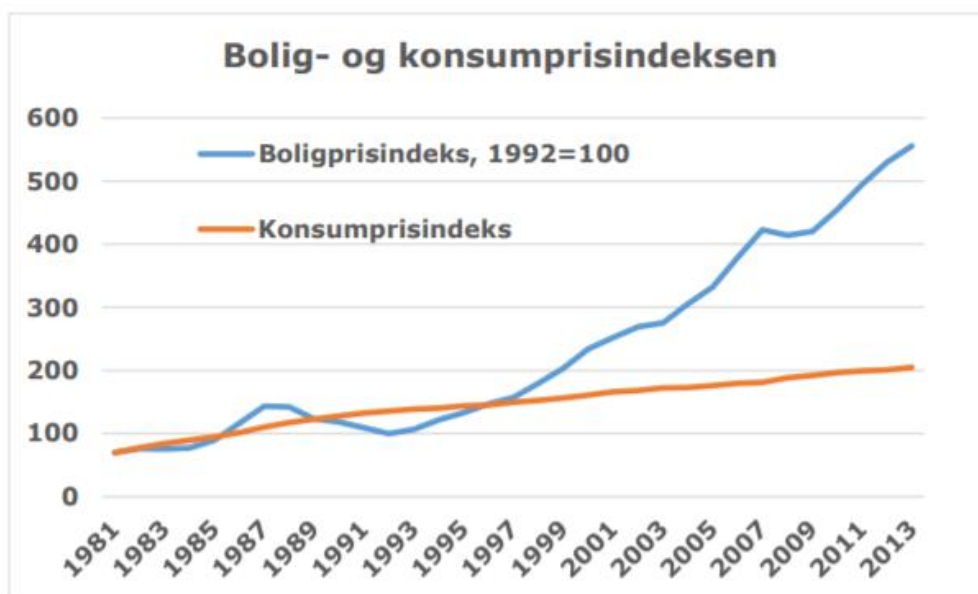
- Boligen er et “nødvendighetsgode”, i den forstand at alle behøver et sted å bo.
- Majoriteten av nordmenn eier sin bolig, i kontrast til utlandet hvor det er mer vanlig å leie boliger. Dette gjør boligen til den normalt sett største enkeltinvesteringen et norsk hushold gjør i løpet av livet.
- Innbyggerne har normalt en boligkarriere, hvor et boligskifte avhenger av alder, arbeidssituasjon, familie og eventuell migrasjon.
- Boligprisene varierer kraftig regionalt. I pressområder er boligprisen høy, som følge av høye tomtepriser, begrenset tilbud av og høy etterspørsel etter boliger. Det motsatte er tilfellet i fraflyttingsområder, der etterspørselen og dermed prisene er lave.
- De høye tomteprisene i pressområder, sammen med byggekostnadene, gjør nybygging av boliger kostbart.

3. Det norske boligmarkedet

Boligprisene i Norge har på et nasjonalt nivå opplevd en markant økning de siste tiårene. I det følgende kapittelet tar vi innledningsvis for oss den historiske boligprisutviklingen i Norge. Deretter presenteres utviklingen blant annet i de ulike fundamentale faktorene som vi identifiserte i forrige kapittel. Vi vil blant annet kaste lys over utviklingen i gjennomsnittlig kvadratmeterpris og disponibel realinntekt, samt utviklingen i sentrale indekser for blant annet konsumpris, bokostnader og byggekostnader. Vi presenterer også forholdstall mellom ulike faktorer, med hensikten å gi et perspektiv på boligmarkedet i Norge. Avslutningsvis ser vi på utviklingen i bruktboligpriser i ulike regioner i Norge.

3.1 Historisk utvikling i boligprisindeksen

For å få et bedre perspektiv på dagens boligpriser er det av interesse å studere hvordan boligprisene i Norge har utviklet seg gjennom de siste tiårene. For å få et mer nøyaktig bilde av boligprisutviklingen, sidestilles boligprisindeksen med konsumprisindeksen (KPI). Dette vil gi et mer reelt bilde på utviklingen i boligpriser, fordi man ser utviklingen i boligpriser i lys av forbrukernes kjøpekraft. For en grundigere gjennomgang av konsumprisindeksen viser vi til delkapittel 3.2.5.



Figur 3.1: Utvikling i boligprisindeksen for Norge, i forhold til konsumprisindeksen

Kilde: Finans Norge (2014)

Indeksen for reelle boligpriser i Norge har tilnærmet tredoblet seg siden starten av 1980-tallet. Med unntak av periodevis nedgang på slutten av 1980-tallet og 2000-tallet, har boligprisutviklingen vært preget av sterk vekst, og vi ser en generell positiv trend fra 1993 og årene etter. Reinhart og Rogoff (2009) finner at utviklingen i realboligprisene er den fremste og viktigste indikatoren når en skal vurdere risikoen for en bankkrise, og vi vil videre drøfte de underliggende krisene bak boligprisindeksen vi presenterte ovenfor.

Den første perioden som skiller seg ut, er den høye prisveksten i forkant av bankkrisen på slutten av 1980-tallet. Ved inngangen til 1984 opphevet regjeringen de kvantitative reguleringene for bankutlån (Aamo, 2019). Denne dereguleringen av finansmarkedene i 1984-1985 førte til at både husholdninger og foretak lettere fikk tilgang på lån.

Dereguleringen falt sammen med høykonjunkturen på midten av 1980-tallet, og bankenes utlån steg i denne perioden med omtrent 20 prosent i året (Torsvik, 1999). Efterspørselen var allerede mer enn stor nok, mye på grunn av at rentene reelt sett var negative tatt hensyn til rentefradrag og inflasjon (Aamo, 2019). 1980-tallet var altså preget av sterk utlånsvekst, og når eiendomsprisene i 1987 falt kraftig, resulterte dette i at både husholdninger og foretak fikk problemer med å innfri sine lån. Samtidig førte en generell nedgang på børsene til at bankene tapte betydelige beløp i verdipapirer. Den samlede konsekvensen av dette var at flere banker tapte hele sin egenkapital (Torsvik, 1999).

Med fall i oljeprisen, inflasjonspress og renteøkning, ble høykonjunkturen i første del av 1980-tallet etterfulgt av lavkonjunktur i årene 1987-1990. Denne lavkonjunkturen i Norge sammenfalt med en generell internasjonal konjunkturedgang, noe som forsterket de finansielle problemene til Norge. Krisen var på sitt sterkeste rundt 1991, da staten måtte gripe inn for å redde bankene (Torsvik, 1999).

I etterkant av bankkrisen lå fokuset i å gjenopprette vekst i økonomien gjennom reguleringer i penge- og kredittpolitikk som styringsmiddel. Norges Bank la gradvis økt vekt på å påvirke utviklingen i prisene som forutsetning for mer stabil kronekurs over tid. Regjeringen la i mars 2001 frem nye retningslinjer for den økonomiske politikken. Finansdepartementet innførte et inflasjonsmål for pengepolitikken, med et mål om en årsvekst i konsumprisene over tid på nær 2,5 prosent (Gjedrem, 2008).

Høsten 2008 ble Norge igjen rammet av en alvorlig finansiell krise, da kollapsen i den store amerikanske investeringsbanken Lehman Brothers spredte seg internasjonalt. Store tap på boliglån skapte omfattende tap i amerikanske banker, som også spredte seg til internasjonale banker som hadde kjøpt deler av lånene. Krisen utløste svikt i tillit mellom banker, og verdensøkonomien gikk inn i en kraftig nedgangskonjunktur (Gjedrem, 2009).

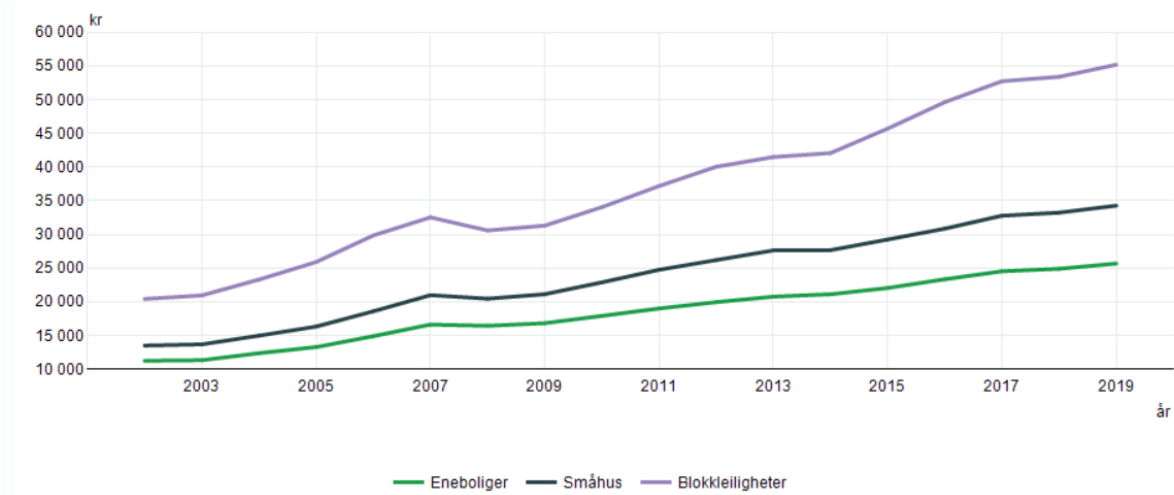
Rentepåslagene på lån mellom banker steg markert, og økte risikopremier på rentene i markedene i USA og Europa smittet dermed raskt over på norske renter, ettersom norske banker lånte i internasjonale markeder. Bankene strammet inn sin kredittpraksis, i tillegg til at aksjekursene falt svært mye (Norges Bank, 2008). Svakere etterspørsel etter norske eksportvarer og stor usikkerhet om den økonomiske utviklingen påvirket Norge. Det norske finansmarkedet ble som et resultat av dette nærmest helt inntørket. Norges Bank var raskt ute og tilførte svært mye likviditet i markedene, samtidig som de lempet kravene om sikkerhetsstillelse for lån. I tillegg til dette ble styringsrenten betydelig nedjustert ofte i forhold til utsiktene for inflasjon og vekst ellers (Gjedrem, 2009). Solid oppfølging og tiltak fra myndighetene sammen med en stabil oljeindustri gjorde at Norge i mindre grad enn andre land ble påvirket av finanskrisen, og muliggjorde at Norge var blant de tidligste ute med å gjenopprette økonomien.

3.2 Utvikling i nyere tid

I dette delkapittelet vil vi se nærmere på den nasjonale utviklingen i ulike faktorer og indekser i nyere tid. Vi vil se på utviklingen i fundamentale faktorer og drøfte betydningen disse har for boligpriser.

3.2.1 Kvadratmeterpris solgte boliger

Statistisk sentralbyrå publiserer statistikk i tabell 06035 i Statistikkbanken for gjennomsnittlige kvadratmeterpriser nasjonalt for selveierboliger etter boligtype og år målt i kroner. SSB sin datakilde er elektroniske data innhentet fra matrikkelen, Norges offisielle eiendomsregister, som inneholder opplysninger om eiendommer, eiendomsgrenser, adresser og bygninger i Norge. Disse dataene publiseres offentlig i Statistikkbanken. Tallene for gjennomsnittlig kvadratmeterpris oppgis med nominell verdi, og oppgis med årlig frekvens fra 2002 til 2019. Figur 3.2 nedenfor illustrerer disse tallene hentet fra SSB.



Figur 3.2: Gjennomsnittlig nominell kvadratmeterpris i Norge for selveierboliger, 2002-2019

Kilde: Statistisk sentralbyrå, Statistikkbanken

Økningen i gjennomsnittlig kvadratmeterpris for eneboliger beregnes til å være 128% i denne perioden, mens kvadratmeterpris for blokkleiligheter beregnes å ha økt med hele 170% i samme periode. Tallene fra SSB viser at gjennomsnittlig kvadratmeterpris har negativ samvariasjon med boligstørrelse, der de største type boliger i gjennomsnitt har lavere kvadratmeterpris enn de mindre boligene. Dette samsvarer med mikroøkonomisk teori om avtakende grensenytte, hvor nytten en får av en ekstra enhet av et gode avtar med økende mengde, etter hvert som behovet i økende grad blir tilfredsstilt.

3.2.2 Disponibel Realinntekt

En sentral variabel å studere for å få et bedre bilde av veksten i boligpriser er den disponible realinntekten til husholdninger. Den disponible inntekten en husholdning har defineres til å være summen av alle medlemmene i husholdningens inntektskilder etter fratrukk av skatt. Denne nominelle inntekten sier derimot lite om kjøpekraften inntekten gir. Det er dermed viktig å skille mellom nominell inntekt og realinntekt. Realdisponibel inntekt definerte vi i kapittel 2.1.2, og er kort forklart den nominelle inntekten i kroner korrigert for endringer i KPI. Realdisponibel inntekt gir dermed et mer realistisk bilde på kjøpekraften en husholdning har i sin inntekt i forhold til den nominelle disponible inntekten.

Melsom (2019) har under Statistisk sentralbyrå utarbeidet en indeks som illustrerer veksten i husholdningenes realdisponible inntekt, presentert i figur 3.3. Tallene er hentet ut fra

nasjonalregnskapets kvartalsvise inntekts- og kapitalregnskap. Tallene er sesongjustert, hvor indeksen tar utgangspunkt i en verdi på 100 i 2009.

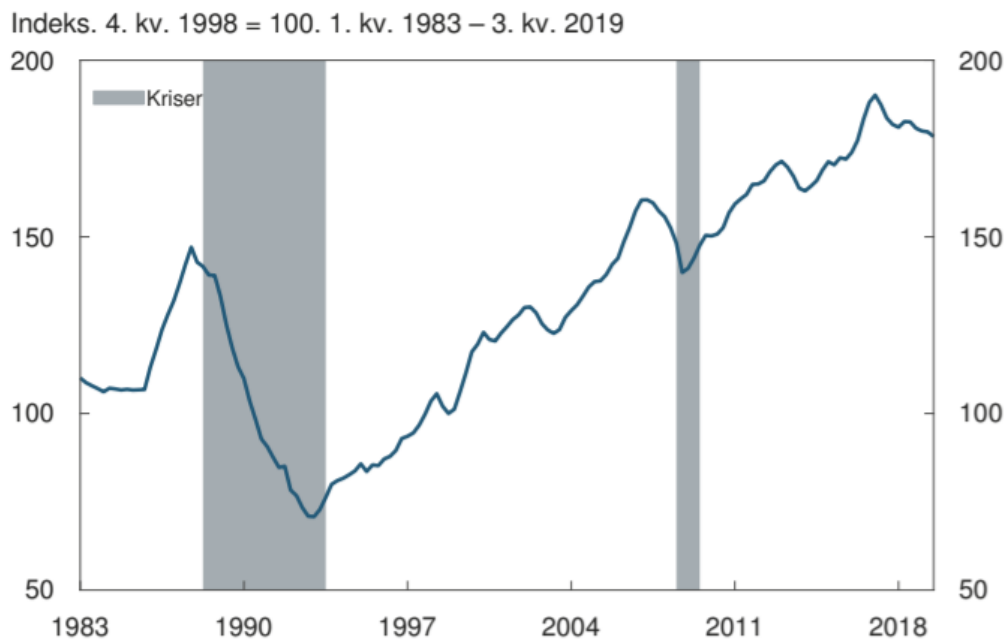


Figur 3.3: Indeks for norske husholdningers realdisponible inntekt, 2002K1-2018K4

Kilde: Melsom (2019), Nasjonalregnskap, inntekts- og kapitalregnskapet, Statistisk sentralbyrå

Fra 2009 til fjerde kvartal 2018 har husholdningenes realdisponible inntekt økt med 28,4%. Med unntak av svingninger opp og ned fra periode til periode har den realdisponible inntekten helhetlig sett økt stabilt. Økningen i realdisponibel inntekt er en viktig forklaringsfaktor på økte boligpriser, da økt realdisponibel inntekt styrker husholdningenes kjøpekraft og betalingsvillighet.

Norges Bank (2019) publiserte i sin pengepolitiske rapport for desember en indeks for utviklingen i nasjonale boligpriser relativt til husholdningenes disponible inntekt, gjengitt i figur 3.4 nedenfor. Det er av interesse å studere dette forholdstallet, fordi det fungerer som en boligprisindikator på hvorvidt boligprisene er høye eller lave.



Figur 3.4: Boligpriser i forhold til disponibel inntekt per innbygger (15-74 år)

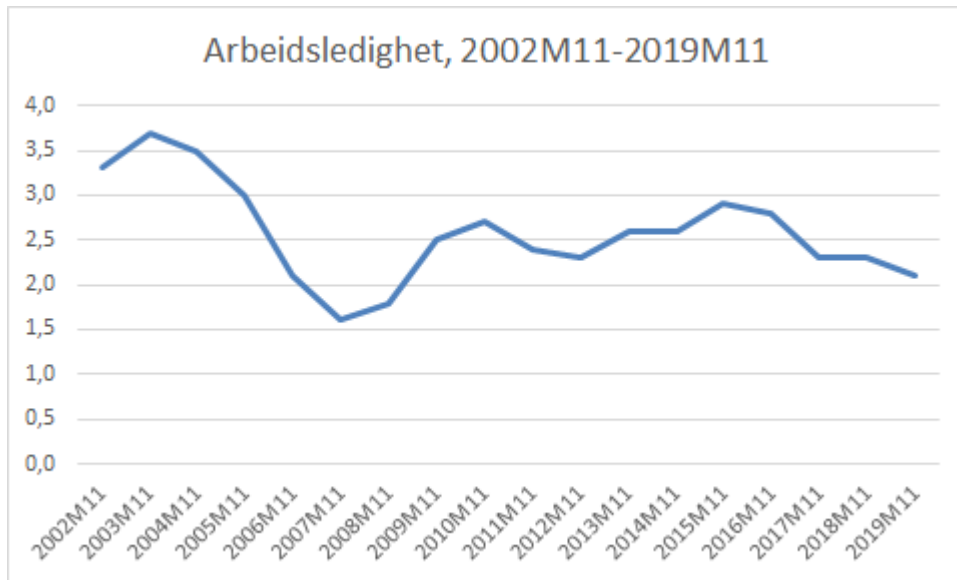
Kilde: Norges Bank (2019)

Boligprisene i Norge har steget kraftig over lengre tid, og utgjorde i 2018 en stor andel i forhold til disponibel inntekt. Det kommer frem av figuren at veksten i boligpriser helt tilbake til 1993 frem til 2018 har hovedsakelig vært større enn veksten i disponibel inntekt. Over de siste 20 årene, fra 1998 til 2018, har boligprisene blitt nærmest det dobbelte av den disponible inntekten. Indeksen preges av kraftige fall ved inngangen til 1990-tallet, og i 2008/2009, noe som godt forklares av finanskrisene i disse periodene. Boligprisene har i denne perioden nådd historisk høye nivåer, men de siste årene bærer også preg av en lavere vekst i boligpriser relativt til disponibel inntekt, sammenlignet med tidligere år. Denne avdempingen bidrar til å redusere risikoen for en bråere nedgang, og dermed også sårbarheten i boligmarkedet (Norges Bank, 2019).

3.2.3 Arbeidsledighet

Arbeidsledighet er annen faktor som i litteraturen antas å påvirke boligpriser, og blir blant annet ansett som en fundamental faktor av Jacobsen og Naug (2004). Økt arbeidsledighet vil føre til bortfall av inntekt og dermed redusere betalingsevne til etterspørere i boligmarkedet. Jacobsen og Naug (2004) argumenterer også for at arbeidsmarkedet påvirker hvordan husholdninger vurderer egne og andres fremtidige inntekter. De presiserer at økt arbeidsledighet vil gi forventninger om lavere lønnsvekst og økt usikkerhet om fremtidig

inntekt og betalingssevne, som igjen vil gi redusert betalingsvillighet. SSB publiserer statistikk for registrerte helt arbeidsledige i tabell 10540 i Statistikkbanken. Figur 3.5 gjengir disse dataene og viser arbeidsledigheten på landsbasis i månedlig frekvens fra november 2002 til november 2019.



Figur 3.5: Arbeidsledighet i Norge, 2002M11-2019M11

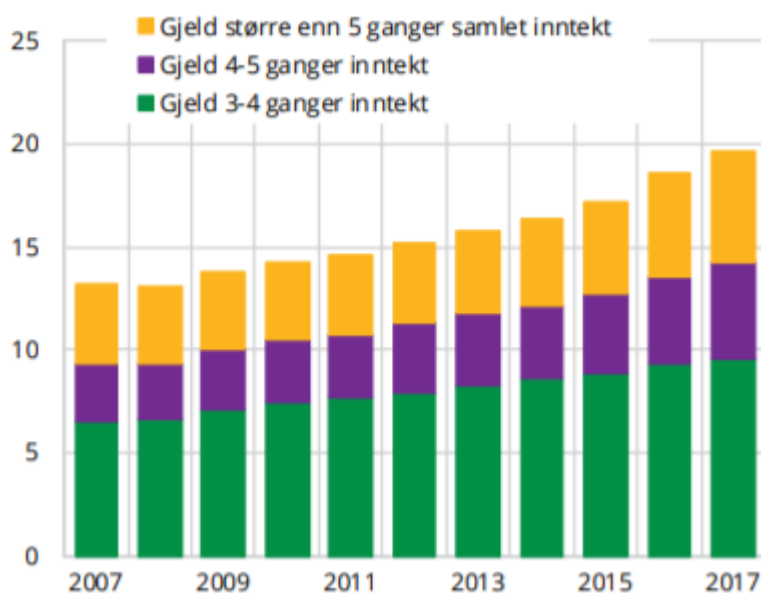
Kilde: Statistisk sentralbyrå, Statistikkbanken

Arbeidsledigheten har variert en del fra år til år, men aldri vært over 4 prosent i denne perioden. Arbeidsledigheten nådde sitt laveste i november 2007 og var på dette tidspunktet nede i 1,6 prosent. Arbeidsledigheten steg deretter frem til 2010, noe som sterkt knyttes opp mot finanskrisen som inntraff i 2008.

3.2.4 Husholdningers gjeldsgrad

Kjøp av bolig er typisk lånefinansiert. Dersom vi ser på aldersgruppen 20 til 29 år var 92% av boligkjøp helt eller delvis finansiert ved lån i perioden 2010 til 2015 (Revold, Sandvik & With, 2018, s.41). I tillegg utgjør lån med pant i boliger vel halvparten av bankers utlån (Emblem, Theisen & Aamo, 2017, s. 17). Tilgang til og pris på kreditt er derfor viktig. Anundsen og Jansen (2013) fant i sin utredning at det er gjensidig avhengighet mellom gjeld og boligpriser over lang tid. Dette finner vi også dersom vi ser på utviklingen i gjeldsgrad og boligpriser i Norge over tid. Over en periode på ti år har stadig flere norske husholdninger opparbeidet seg en relativt høy gjeld. Ifølge Halvorsen (2019) er mye av grunnen til dette økte boligpriser, og den økte gjelden blir møtt av en tilsvarende økning i boligformue.

Andelen husholdninger som hadde samlet gjeld som utgjorde mer enn tre ganger samlet inntekt, har økt fra 13,2 prosent i 2007 til 19,6 prosent i 2017.



Figur 3.6: Andelen husholdninger med samlet gjeld større enn 3 ganger samlet inntekt, 2007-2017

Kilde: Statistisk sentralbyrå (2019a, s.90)

Den økte gjeldsgraden, altså graden av gjeld i forhold til inntekt, må sees i sammenheng med utviklingen i utlånsrenta. Utlånsrenta har vært stabil lav over lengre tid og siden 2009 har den ikke overstegyet fem prosent. En lav rente fører til lavere rentekostnader, som igjen betyr at husholdninger har bæreevne til å låne mer. Dette gjør det mulig å låne større beløp og konsekvensen av dette blir en gjennomsnittlig høyere gjeldsgrad. Den økte gjeldsgraden må også ses i sammenheng med endringer i når man entrer boligmarkedet. I en artikkel fra e24 (2017), som tar utgangspunkt i tall fra SSB, kommer det frem at andelen unge mellom 25 og 30 år som eier egen bolig, har steget med 25 prosent fra 2004 til 2016. Denne aldersgruppen er i en tidlig fase av yrkeskarrieren og vil typisk være preget av lavere inntekt i tillegg til studielån, som igjen vil føre til en høyere gjeldsgrad.

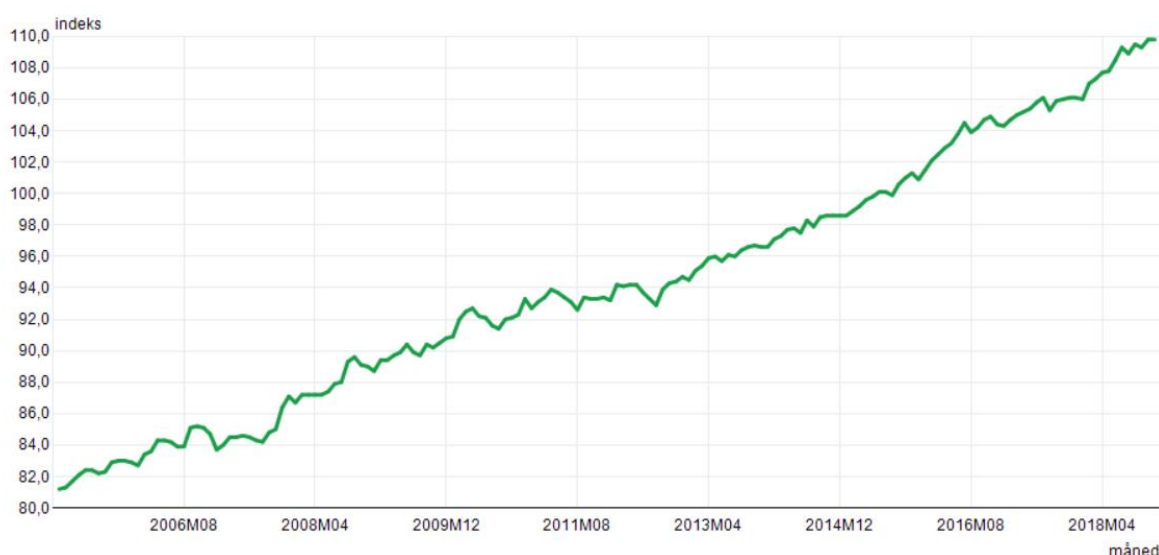
En høyere gjeldsgrad gjør også at husholdninger er mer sårbare for fremtidige renteøkninger. Dette gjelder spesielt i Norge hvor de aller fleste låner med flytende rente. I utgangen av juni 2019 var andelen husholdninger som hadde nedbetalingslån med pant i bolig med fast rente, nede i 8,1 prosent (Øksnes, 2019). Dette i sammenheng med at gjeldsgraden er relativt høy

gjør at selv moderate endringer i renten kan få store konsekvenser for økonomien til husholdninger.

3.2.5 Konsumprisindeksen (KPI)

En sterk og vedvarende vekst i det generelle prisnivået i økonomien betegnes som inflasjon. Et mål på prisstigning er Konsumprisindeksen (KPI). KPI er en levekostnadsindeks som skal besvare hvilken inntektskompensasjon som er nødvendig for at en gjennomsnittlig husholdning skal kunne opprettholde sin levestandard når prisene på konsumgoder som varer og tjenester endres (Johannessen, 2014). En levekostnadsindeks skal i prinsippet kunne følge utviklingen i levekostnadene knyttet til å beholde et gitt nyttenivå for et individ eller husholdning (Bye og Hægeland, 2014).

KPI-veksten måler prisutviklingen i et utvalg representative varer og tjenester som etterspørres av norske husholdninger (Johannessen, 2014). Konsumprisindeksen kommer til anvendelse i den forstand at den ofte brukes som en deflator for ulike mål og indekser i økonomien. At en deflaterer en indeks mot KPI gjør at indeksen justeres for veksten i det generelle prisnivået på varer og tjenester, og transformerer nominelle tall til reelle tall. Dette er relevant for å skape et reelt bilde på den faktiske utviklingen i eksempelvis boligpris-, disponibel inntekts- og renteutvikling, uavhengig av generell prisvekst. SSB publiserer statistikk for konsumprisindeksen i tabell 08183 i Statistikkbanken, som gjengis i figur 3.7. Figuren løper fra første måned 2005 til tolvte måned 2018.

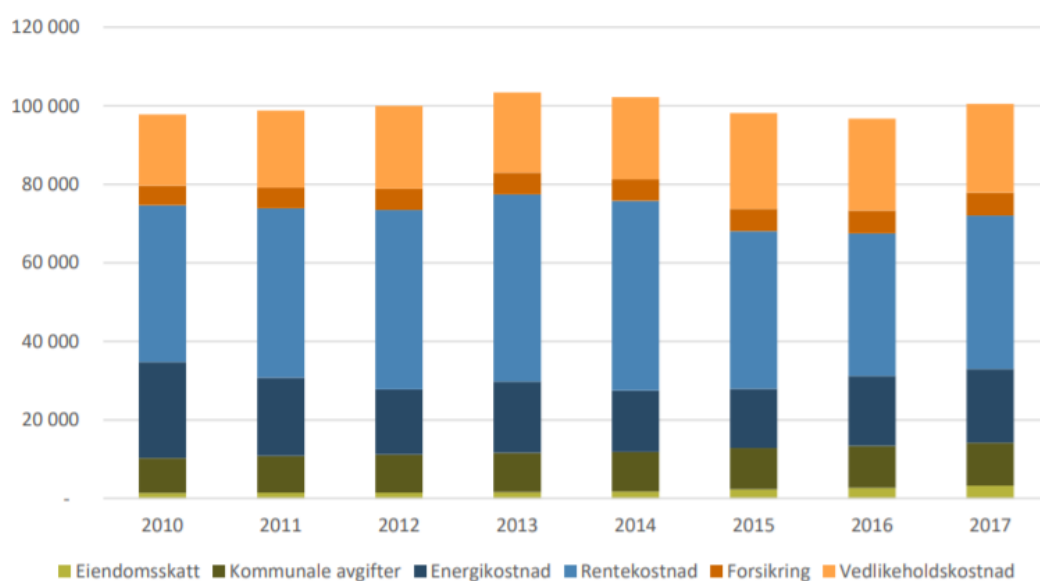


Figur 3.7: Konsumprisindeks, 2015=100, 2005M1-2018M12

Kilde: Statistisk sentralbyrå, Statistikkbanken

3.2.6 Bokostnadsindeksen for norske boliger

I kapittel 2 introduserte vi hva som påvirker tilbud og etterspørsel i boligmarkedet. Etterspørselsfunksjonen ble gitt i likning 2.1 og bokostnadene for en typisk eier er en av faktorene. Det er derfor interessant å studere utviklingen i bokostnadene over tid og hva som påvirker disse. I september 2018 utga Samfunnsøkonomisk analyse AS en rapport hvor de beregner en bokostnadsindeks der hensikten er å vise en detaljert utvikling i bokostnader for en gjennomsnittlig husholdning. Bokostnader definerte vi i delkapittel 2.1.2 som husholdningens utgifter knyttet til å bo i egen bolig. I rapporten til Samfunnsøkonomisk analyse (2018) blir disse utgiftene avgrenset til seks ulike kostnadselementer. Disse er eiendomsskatt, kommunale gebyrer, energikostnader, rentekostnader, forsikring og vedlikehold av egen bolig. Disse kostnadselementene gjengis i figur 3.8.



Figur 3.8: Bokostnadsindeks for norske boliger, 2010-2017

Kilde: Samfunnsøkonomisk Analyse (2018)

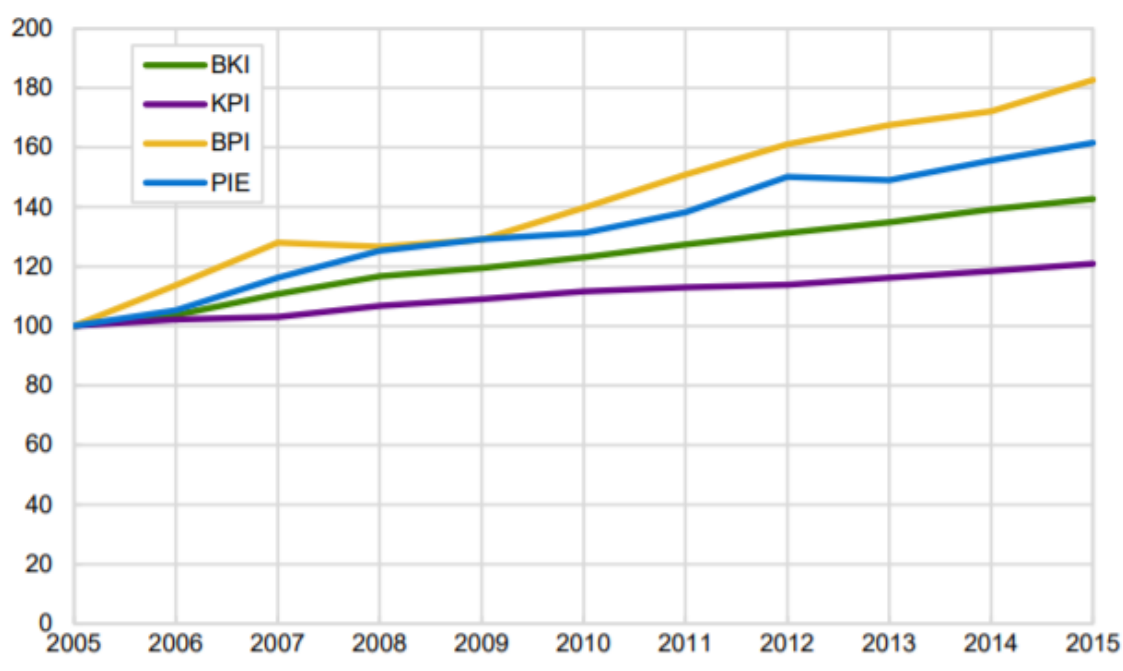
Dersom vi ser på hele perioden fra 2010 til 2017, har bokostnadene økt med 3 prosent. Det har imidlertid vært variasjoner innad i perioden og spesielt rentekostnader har vært utsatt for endringer. I perioden 2010 til 2013 økte bokostnadene med 6 prosent, og dette var i hovedsak grunnet økte rentekostnader som var en konsekvens av en generell økning i gjeld. Fra 2013 til 2016 falt bokostnadene med 6 prosent, noe som var drevet av en nedgang i rentene (Samfunnsøkonomisk analyse, 2018). Lavere bokostnader, som følge av for eksempel lavere

realrente eller forventning om økt appresiering av boligformue, trekker i retning av økt etterspørsel etter boliger (Emblem et al., 2017, s.15).

Som nevnt i kapittel 3.2.4 er det stadig flere husholdninger i Norge som opparbeider seg relativt høy gjeld. Dette gjør at flere er utsatt for endringer i rentenivået. Små endringer i renten i fremtiden kan dermed føre til store økninger i rentekostnadene og dermed også bokostnader sett under ett.

3.2.7 Byggekostnadsindeksen

Byggekostnader er en annen faktor som i litteraturen anses å påvirke boligprisene. Når byggekostnadene ved en bolig øker, vil en konsekvens av dette ofte være at prisen av boligen øker tilsvarende. Høiby (2017) presenterer i sin artikkel en byggekostnadsindeks som sammenstilles med andre indekser, blant annet KPI. Artikkelen tar utgangspunkt i data fra Statistisk sentralbyrå. Byggekostnadsindeksen her måler prisutviklingen på innsatsfaktorene som inngår i boligproduksjonen mot basisåret 2005, og alt annet enn prisene holdes konstant. Utviklingen til alle innsatsfaktorene vektet sammen til totalindeksen som presenteres i figur 3.9 nedenfor.



Figur 3.9: Sammenligning av boligprisindeks (BPI), prisindeks for nye eneboliger (PIE), byggekostnadsindeks (BKI) og konsumprisindeks (KPI), 2005-2015, (100=2005)

Kilde: Høiby (2017, s. 16), Statistisk sentralbyrå

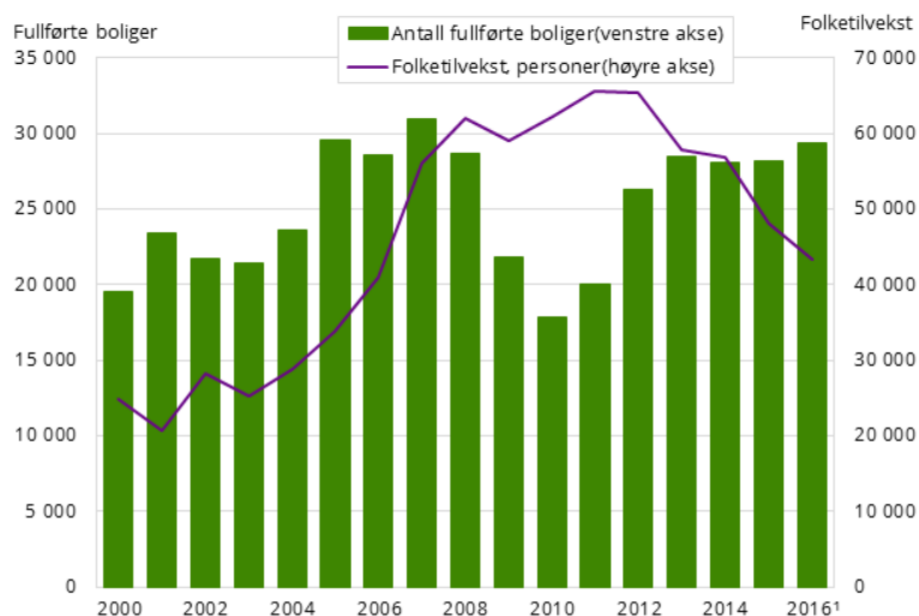
Fra figuren ser vi at byggekostnadene øker utover den generelle prisveksten i økonomien. I korte trekk betyr dette at de reelle byggekostnadene har økt i løpet av perioden 2005 til 2015. Høyere byggekostnader trekker også opp prisen på brukte boliger (Emblem et al., 2017, s.16).

I tillegg fremgår det av figuren til Høiby (2017) at boligprisindeksen og prisindeksen for nye boliger også har økt i enda større grad enn byggekostnadsindeksen. Prisen på nye boliger reflekterer til en viss grad kostnadene ved å bygge nytt, men indeksen indikerer også at kostnadene knyttet til andre forhold enn det som måles gjennom innsatsfaktorene i byggeprosessen også har økt. Eksempelvis er fortjeneste og produktivitet faktorer som inngår i prisindeksen på nye boliger, og som byggekostnadsindeksen ikke tar høyde for. Prisutvikling på tomter er også en faktor som kan bidra til boligprisvekst (Høiby, 2017). Høye priser i boligmarkedet muliggjør at en kan selge nye boliger til en høyere pris, og dette residualet mellom pris og byggekostnader tilfaller eieren av tomten.

3.2.8 Boligbygging i lys av folketilvekst

En relevant faktor vi presenterer for å skape et bilde av boligmarkedet i Norge er befolkningsveksten. En befolkning i vekst impliserer flere etterspørrere etter boliger, som isolert sett bidrar til høyere boligpriser. Dermed er det aktuelt å betrakte befolkningsveksten i lys av boligbygging for å skape et bedre bilde av tilbudet og etterspørselen, og om markedet er i likevekt.

I Statistisk sentralbyrå (2017) sin rapport omhandlende byggeareal presenteres det statistikk som setter antall fullførte boliger i lys av folketilveksten, gjengitt i figur 3.10 nedenfor. De siste årene er det registrert et gjennomsnitt på omtrent 2,15 personer per husholdning (Statistisk sentralbyrå, 2019b). Dette tallet gir en indikasjon på forholdet mellom behovet for og tilgangen på nye boliger. Med andre ord betyr dette omtrentlig at for hver 2. nye person behøves det en ny bolig, og følgelig kan dette anvendes som et vagt uttrykk for likevekt. Avstanden mellom kurven for befolkningsvekst og søylene for antall fullførte boliger gir i figuren et bilde på dette forholdet. Hvis boligbyggingen ikke holder tritt med befolkningsveksten vil dette bidra til et økt boligpress i boligmarkedet, og konsekvent vil boligprisene stige.



¹ Beregnet folketilvekst
Kilde: Statistisk sentralbyrå.

Figur 3.10: Antall fullførte boliger og folketilvekst målt i antall personer, 2000-2016

Kilde: Statistisk sentralbyrå (2017)

Antall fullførte boliger har i løpet av perioden 2000 til 2016 for det meste holdt tritt med befolkningsveksten. Nye boliger økte i stigende grad frem til finanskrisen i 2008, hvor nybygging av boliger ble kraftig redusert i årene som fulgte frem til 2012.

Befolkningsveksten var økende i denne perioden, noe som skapte et ubalansert forhold og økt press i boligmarkedet. Boligbyggingen var i 2016 tilbake på samme nivå som før finanskrisen, men det ble imidlertid bygget flere boliger enn det kom til nye husholdninger.

3.2.9 Førstegangskjøpere og andel sekundærboliger

I 2020 utarbeidet Samfunnsøkonomisk analyse i samarbeid med Norges

Eiendomsmeglerforbund og Ambita en rapport som dokumenterer utviklingen i antall førstegangskjøpere fra 2010 til og med 4. kvartal 2019. I rapporten viser de også statistikk for sekundærboliger i hele Norge med detaljert beskrivelse av både geografisk fordeling og fordeling på ulike boligtyper.

Tall fra rapporten viser at antallet førstegangskjøpere har holdt seg stabilt fra 2010 og frem til i dag, dog med variasjoner fra år til år. 2016 var preget av nedgang i førstegangskjøpere.

Det kan skyldes den sterke boligprisveksten som oppstod samme år. Fra 2016 og utover har

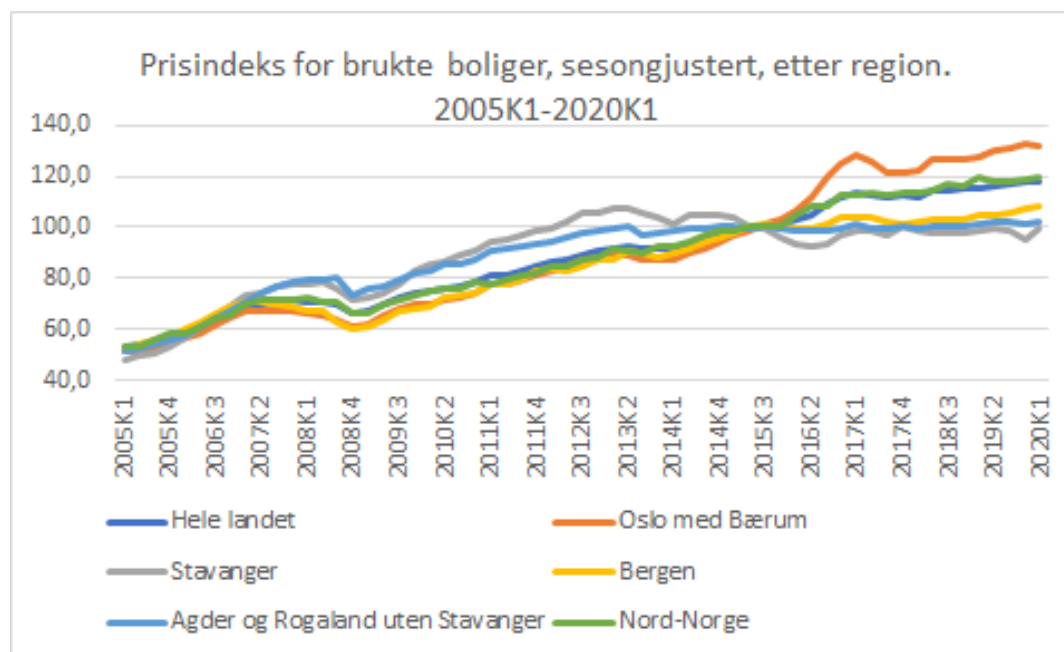
imidlertid andelen førstegangskjøpere steget jevnt og var i 2019 oppunder 50 000 (Samfunnsøkonomisk analyse, 2020, s.6).

Førstegangskjøpere er en relativt prissensitiv gruppe. Dette er en gruppe som er preget av unge mennesker med relativt lav lønn og lav egenkapital. Dette kan føre til at de tidlig vil bli begrenset av høye priser. I 2017 kom også boliglånsforskriften, som stilte større krav til egenkapital og inntekt og dermed bidro til å begrense førstegangskjøpernes handlingsrom i første omgang. Det argumenteres dog i rapporten for at boliglånsforskriften i dag bidrar til å dempe boligprisene og dermed isolert sett gjør det enklere å etablere seg i boligmarkedet, selv om kravene til egenkapital og inntekt fører til begrenset handlingsrom for en del (Samfunnsøkonomisk analyse, 2020, s.6).

Ifølge rapporten fra Samfunnsøkonomisk analyse var det i 2019 fjerde kvartal registrert 399 855 sekundærboliger i Norge. Dette utgjør omtrent 15% prosent av den totale boligmassen i Norge. Den geografiske fordelingen av sekundærboliger varierer stort og distriktskommunene er ofte preget av en høy andel sekundærboliger, men også enkelte bydeler i de store byene har et høyt antall sekundærboliger. Dette er to svært forskjellige markeder hvor sekundærboligene i byene stort sett er investeringsobjekter mens de i distriktene trolig i stor grad er preget av fraflyttede boliger (Samfunnsøkonomisk analyse, 2020, s.18).

3.3 Regional utvikling i boligpriser

Så langt har vi sett på utviklingen i nasjonale gjennomsnittstall i boligmarkedet i Norge. Bak disse tallene skjuler det seg til dels store variasjoner mellom regioner og byer. Figuren under gir et bilde av variasjonen i utviklingen av boligpriser mellom regioner, som ikke fanges opp av den nasjonale prisindeksen vi diskuterte tidligere. Tallene er hentet fra prisindeksen for brukte boliger i tabell 07221 i Statistikkbanken til SSB.



Figur 3.11: Prisindeks for brukte boliger, sesongjustert, etter region. 2015=100. 2005K1-2020K1.

Kilde: Statistisk sentralbyrå, Statistikkbanken

Regionene vi ser på her er ikke de endelige regionene vi vil studere videre i oppgaven. De presenteres med den hensikt å belyse at ulike regioner har opplevd ulik vekst målt ved denne indeksen. Den blå linjen markerer prisindeksen for hele landet, og fungerer her som et gjennomsnittstall for utviklingen i boligpriser. Fra figuren ser vi at prisutviklingen på bruktboliger i samtlige regioner ble negativt påvirket av finanskrisen i 2008-2009. I etterkant av finanskrisen ser vi at regionene Stavanger og Agder og Rogaland uten Stavanger preges av en høyere vekst i boligpriser utover det nasjonale gjennomsnittet. Veksten i bruktboligpriser i Stavanger, som hadde 15,4 prosent sysselsatte i petroleumsnæringen i sin kommune per 2012 (Ekeland, 2014), ble kraftig dempet etter oljekrisen inntraff i 2014. Andelen sysselsatte i petroleumsnæringen var på denne tiden blant den høyeste på nasjonal basis, noe som følgelig gjorde at Stavanger i større grad ble påvirket av oljekrisen enn de andre regionene. Fra

basisåret i indeksen som i dette tilfellet er 2015 og frem til 2020 har Stavanger hatt den laveste gjennomsnittsvæksten i pris på brukte boliger. I tillegg ser vi at veksten i Agder og Rogaland uten Stavanger dempes over tid, og har hatt en vekst under gjennomsnittet siden 2015.

I motsetning til dette, har regionen Oslo med Bærum opplevd markant vekst i priser fra 2015 og utover, og innehar den høyeste veksten i bruktboligpriser per første kvartal 2020. Veksten i bruktboligpriser i Bergen har lenge utviklet seg omtrent gjennomsnittlig, men vi ser at veksten ligger under det nasjonale gjennomsnittet de siste fem årene. Boligprisveksten i Nord-Norge holdt tritt med gjennomsnittsvæksten, og av regionene som presenteres innehar Nord-Norge den nest høyeste veksten i bruktboligpriser per første kvartal 2020.

Bak variasjonene i regionale boligprisutvikling skjuler det seg også regionale variasjoner i utviklingen av fundamentale faktorer. I neste kapittel vil vi følgelig gå inn i detalj på utviklingen av faktorer som, i tråd med litteraturen vi har presentert i tidligere kapitler, anses å påvirke boligpriser. Vi vil videre studere utviklingen i disse faktorene på et regionalt nivå.

4. Data

I dette kapittelet vil vi gjennomgå de data vi legger til grunn i de empiriske analysene i oppgaven. All data er innhentet i perioden første kvartal 2005 til fjerde kvartal 2018, og vi henter inn data for henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø. Vi tar hovedsakelig utgangspunkt i databasen til Statistisk sentralbyrå samt Finans Norge og Kantar AS som kilde til de data vi anvender. Definisjonen på regionene som ligger til grunn er kommuneinndelingen per 2019. Vi refererer til kommunene som byer/regioner videre i oppgaven. Data for kvadratmeterpris, inntekt, arbeidsledighet, nettoinnflytting, boligbygging og forventning er regionale, mens data for renten er nasjonal. Observasjonsperioden er på 14 år, noe som defineres å være en relativt kort observasjonsperiode. Dette er en konsekvens av begrensninger i tilgjengeligheten på data.

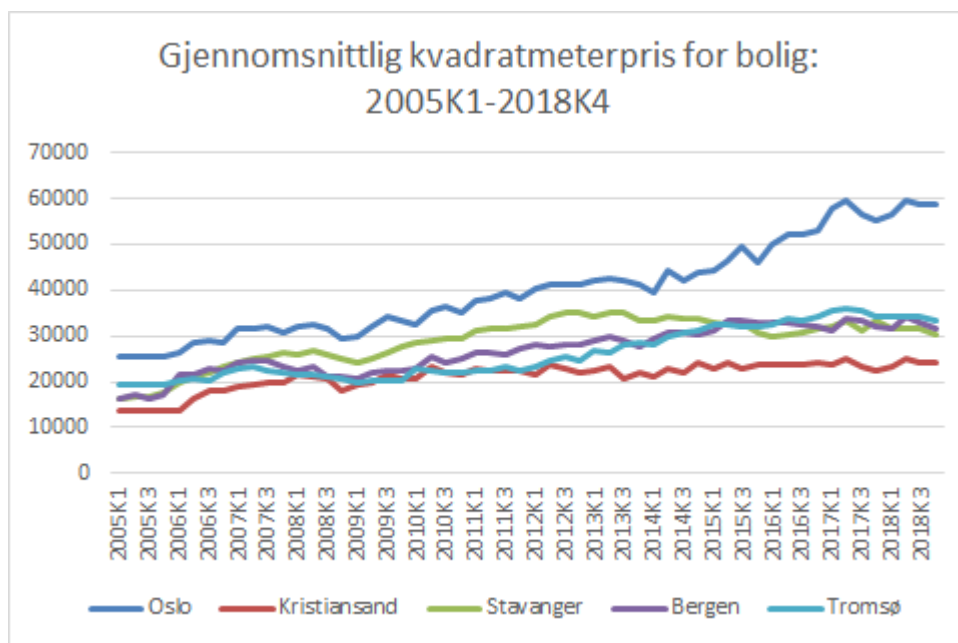
Noen av variablene blir omgjort til reelle verdier, mens andre beholder sin nominelle form. Denne transformasjonen blir redegjort for og utført etter gjennomgåelsen av data. Til slutt gjennomgås det hvilke variabler som vi må utelate fra modellen vår, og hvorfor. I tillegg vil et utvalg variabler transformeres til logaritmisk form, og en begrunnelse for dette vil gis i det følgende.

4.1 Kvadratmeterpris

Vi velger å anvende den gjennomsnittlige kvadratmeterprisen for eneboliger som avhengig variabel i vår modell. Utgangspunktet for å måle forskjeller legges i prisutviklingen i eneboliger da enebolig er den dominerende boligtypen i Norge, der 55,8% av befolkningen er bosatt i denne typen bolig per 2019 (Statistisk sentralbyrå, 2020). SSB publiserer statistikk i tabell 05963 i Statistikkbanken som måler gjennomsnittlig kvadratmeterpris etter region fra første kvartal 2006 til fjerde kvartal 2018. Data for Stavanger og Bergen fra første til fjerde kvartal 2005 er hentet fra tabell 03637 i Statistikkbanken. Data for Oslo, Kristiansand og Tromsø for år 2005 er hentet fra prisindeksen for brukte boliger som publiseres i tabell 06035 i Statistikkbanken. Dataen her angår også gjennomsnittlig kvadratmeterpris for enebolig, men oppgis med årlig frekvens.

Dette begrensede datautvalget fører til at gjennomsnittlig kvadratmeterpris for Oslo, Kristiansand og Tromsø fra første til fjerde kvartal 2005 er lik årsgjennomsnittlig

kvadratmeterpris dette året. Følgelig vil alle kvartalene i 2005 ha samme verdi, og vil dermed ikke variere innad i dette året. Det må nevnes at dette ikke er optimalt for modellen, men dette er en konsekvens av det begrensede datautvalget vi har til disposisjon. Alle tallene oppgis nominelt i kroner, og er verken sesong- eller kalenderjusterte. Vi vil transformere tidsserien for kvadratmeterpriser om til reelle priser før vi estimerer boligprismodellen, noe vi redegjør for i delkapittel 4.9. Ettersom boligprisdata oppgis i nominelle kroner og inneholder høye verdier, vil vi transformere variabelen om til logaritmisk form når vi estimerer modellene. Dette er ikke uvanlig å gjøre for denne typen data, ettersom det gir oss mulighet til å tolke estimatene av koeffisientene som elastisiteter, samt at det er med på å senke effekten av eventuelle ekstremverdier i datasettet. Figur 4.1 presenterer de sammensatte data for boligpris vi benytter i vår modell.



Figur 4.1: Regionale gjennomsnittlige kvadratmeterpriser, 2005K1-2018K4

Kilde: Statistisk sentralbyrå, Statistikkbanken

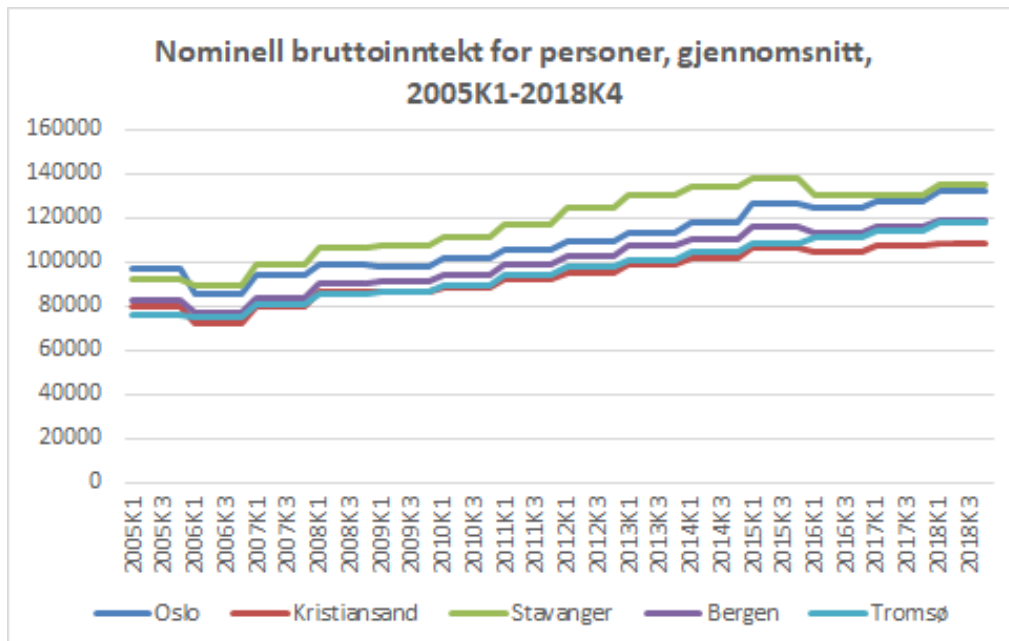
Fra figuren ser vi at enkelte regionale områder opplever kraftigere vekst enn andre. Enkelte områder har i løpet av de siste årene opplevd kontinuerlig vekst i boligpris, mens boligprisen i andre regionale områder har vært preget av nedgang. Fra figuren fremkommer det at Oslo innehar den høyeste gjennomsnittlige kvadratmeterprisen, og har også opplevd brattest vekst, særlig fra 2014 og utover. Stavanger har opplevd en stabil vekst frem mot 2014, men årene etter oljekrisen preges av nedgang i kvadratmeterpris. Kvadratmeterprisene i Tromsø og

Bergen har vokst stabilt i tidsperioden, mens kvadratmeterprisen i Kristiansand har holdt seg på noenlunde samme nivå fra 2009 til 2018.

4.2 Inntekt

Inntektsnivå er en sentral faktor som bidrar til å forklare etterspørselen etter boliger. Dette kom frem i likning 2.1 hvor vi definerte etterspørselen etter bolig, hvor det tredje leddet Y , beskriver husholdningens disponible realinntekt. På et regionalt nivå vil det være av interesse å studere inntektsnivået, ettersom inntektsnivå i teorien påvirker kjøpekraft og betalingsvillighet blant husholdninger. Dermed er et varierende inntektsnivå en mulig forklaringsfaktor til varierende boligpriser i ulike områder, og dermed en sentral variabel å inkludere.

I vår modell vil vi benytte oss av gjennomsnittlig nominell bruttoinntekt som datagrunnlaget for inntekt. Dataene vi benytter er hentet fra tabell 03068 i Statistikkbanken til SSB, og omhandler gjennomsnittlig nominell bruttoinntekt for bosatte personer 17 år og eldre. Dataene oppgis med årlig frekvens, og er verken sesong- eller kalenderjustert. Årlig data sammenfaller ikke med frekvensen på data vi benytter i vår modell, og derfor konstruerer vi kvartalstall for bruttoinntekt ved å dele årsgjennomsnittet til hvert år på fire. Dette fører til at kvartalsgjennomsnittene innenfor samme år ikke varierer, men kun varierer mellom år. Variabelens egenskaper er dermed ikke optimale for analysen, ettersom lite variasjon i inntekt mellom kvartalene i et år sannsynligvis vil svekke påvirkningskraften inntekt har på boligpris i modellen. Vi vil transformere nominell bruttoinntekt om til realinntekt før vi estimerer boligprismodellen, noe vi nærmere redegjør for i delkapittel 4.9. Som tilfellet med kvadratmeterpris vil vi også transformere denne serien om til logaritmisk form i estimeringen av våre modeller. Figuren under presenterer de deskriptive inntektsdata vi benytter i vår modell.



Figur 4.2: Gjennomsnittlig personlig nominell bruttoinntekt etter region, 2005K1-2018K4.

Kilde: Statistisk sentralbyrå, Statistikkbanken

Fra figuren fremgår det at Stavanger er regionen med høyest vekst fra 2006 og utover, frem til oljekrisen i 2014, med deretter et lite fall i inntektsnivå rundt og etter oljekrisen inntraff. Dette henger sammen med Stavangers høye andel sysselsatte i petroleumsnæringen, som i 2012 var på 15,4 prosent (Ekeland, 2014). Den høye andelen gjør at Stavanger er mer utsatt for svingninger i petroleumsnæringen og reduksjonen i inntekt vi ser i 2014 var et resultat av blant annet dette.

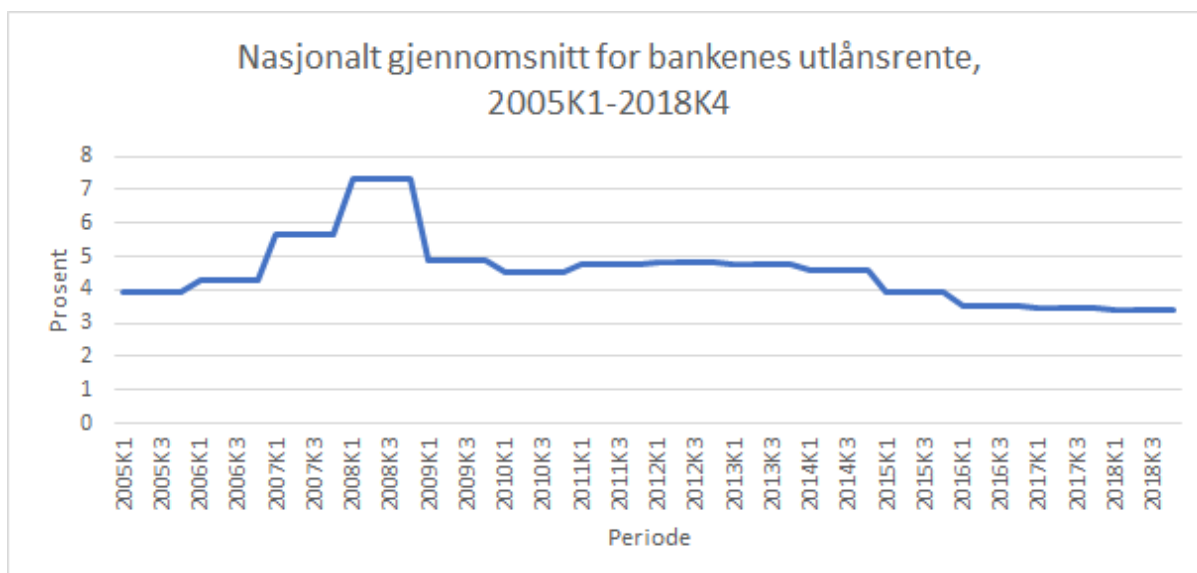
4.3. Rente

Rente blir i litteraturen ansett som en av de mest sentrale faktorene som driver boligpriser, og regnes også som en såkalt fundamental faktor. Vi velger å anvende bankenes utlånsrente som målet vårt på rentevariabelen. Grunnen til at vi benytter utlånsrenten er fordi det er denne renten konsumenten påvirkes av når de skal ta opp boliglån, og følgelig denne renten som egner seg best som vurderingsgrunnlag i avgjørelsen om konsumenten har råd til å kjøpe bolig eller ikke. Jacobsen og Naug (2004) fant også i sin artikkel sterk signifikant effekt av bankenes utlånsrente på boligpris, mens markedrenten derimot viste seg å være insignifikant.

En økning i utlånsrenten gjør at boligutgiftene stiger. Dette fører igjen til at etterspørselen etter bolig avtar og som en konsekvens synker prisene. Vi får en motsatt effekt dersom utlånsrenten reduseres. Da vil boligutgiftene reduseres, forbrukerne vil få lettere og billigere tilgang på kreditt, og dermed stiger også etterspørselen etter bolig. Det er derfor interessant å studere utviklingen i boliglånsrenten i sammenheng med utviklingen i boligpriser.

Vi velger å benytte bankenes nominelle utlånsrente oppgitt med årlige tall fra 2005 til 2018. Dataene er innhentet fra tabell 08175 i Statistikkbanken til SSB. Tallene som oppgis er gjennomsnittsrenten for hvert år, og tallene oppgis i prosent. Vi velger å beholde dataene på prosentform. Ettersom vi benytter oss av kvartalsvise data i modellen, velger vi å sette de kvartalsvise verdiene innad i et år lik årsgjennomsnittet i det respektive året. I likhet med inntektsvariabelen vil rentevariabelen ikke variere innad i årene, kun mellom årene. Dette vil sannsynligvis også svekke påvirkningskraften rente har på boligpris i modellen. Siden tallene oppgir den nasjonale utlånsrenten vil de naturligvis ikke variere regionalt, men dette er ikke et problem ettersom rente på boliglån i Norge ikke er betinget av hvor i Norge en kjøper bolig. Jacobsen og Naug (2004) tar høyde for marginalsattesatsen på kapitalinntekter og kapitalutgifter i sin modell, noe som må tas hensyn til i tolkningen av koeffisientestimatene for rente i den langsiktige sammenhengen.

Vi vil transformere den nominelle utlånsrenten om til reelle verdier før vi estimerer boligprismodellen, noe vi nærmere redegjør for i kapittel 4.9. Figuren under gir en presentasjon av dataene vi benytter for utlånsrenten.



Figur 4.3: Bankenes nasjonale gjennomsnittlige utlånsrente, 2005K1-2018K4

Kilde: Statistisk sentralbyrå, Statistikkbanken

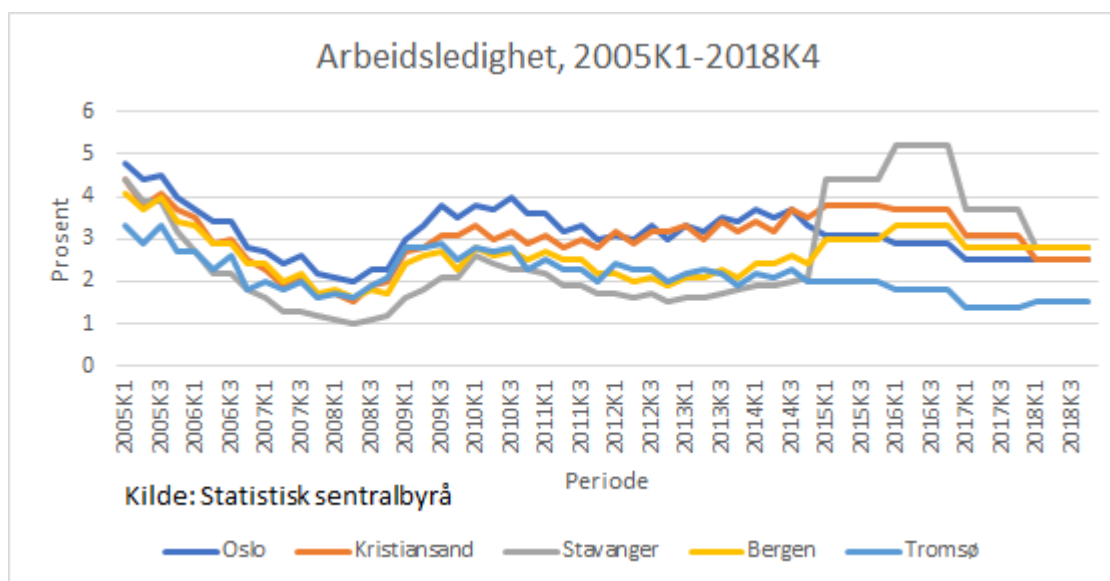
4.4 Arbeidsledighet

Arbeidsledighet er en annen fundamental faktor som også antas å påvirke boligpriser på et regionalt nivå. Arbeidsledighet oppstår når det er et høyere tilbud på arbeidskraft enn etterspørselen etter det i et arbeidsmarked. Bortfall av arbeidsinntekt vil trekke i retning av lavere etterspørsel etter bolig i aktuell region (Emblem et al., 2017, s. 20). Økt arbeidsledighet vil føre til økt usikkerhet og redusert betalingsevne, og på bakgrunn av dette vil etterspørselen etter boliger reduseres. I kontrast vil lav arbeidsledighet medføre høyere etterspørsel etter boliger. Endringer i arbeidsledighet er dermed en mulig forklaringsfaktor på endringer i boligprisene i de regionale områdene.

SSB publiserer statistikk i tabell 10540 i Statistikkbanken for registrerte helt arbeidsledige etter region og måned, gjengitt i figur 4.4 nedenfor. Utviklingen av registrerte arbeidsledige måles prosentvis. Det er viktig å legge merke til at tallene vi bruker er for registrerte helt arbeidsledige. Dette defineres som personer som aktivt går til NAV og registreres som arbeidsledige. Dette er en noe streng definisjon av arbeidsledighet, men tilgangen på data gjorde at vi valgte å bruke dette målet på arbeidsledighet i vår analyse. Utviklingen i arbeidsledighet registreres med månedlig frekvens frem til 2014. Fra 2015 frem til 2018 oppdateres tabellen kun en gang i året i statistikkmåned november. Gjennom analysen vår bruker vi konsekvent kvartalsvise data. Vi har derfor regnet ut kvartalsgjennomsnitt for

perioden 2005K1 til 2018K4 der første kvartal er et gjennomsnitt av januar, februar og mars og de påfølgende kvartalsverdiene beregnes ut fra de tilhørende månedene.

Fra 2015 til 2018 har vi kun data fra statistikk måned november og de kvartalsvise tallene blir dermed like for hele året. Dette er en svakhet med modellen vår som er verdt å legge merke til, og konsekvensen blir at utviklingen i arbeidsledighet fra 2015 og utover blir mer stykkevis. SSB melder i tillegg til dette om et brudd i tidsserien fra 2018, slik at tallene for dette året ikke er sammenlignbare med tidligere år. Vi må derfor ta i kritikk av oppgaven at data for arbeidsledighet for år 2018 er noe unøyaktige, noe som kan slå ut negativt i regresjonsresultatene og medføre svake resultater.



Figur 4.4: Arbeidsledighet etter region, 2005K1-2018K4

Kilde: Statistisk sentralbyrå, Statistikkbanken

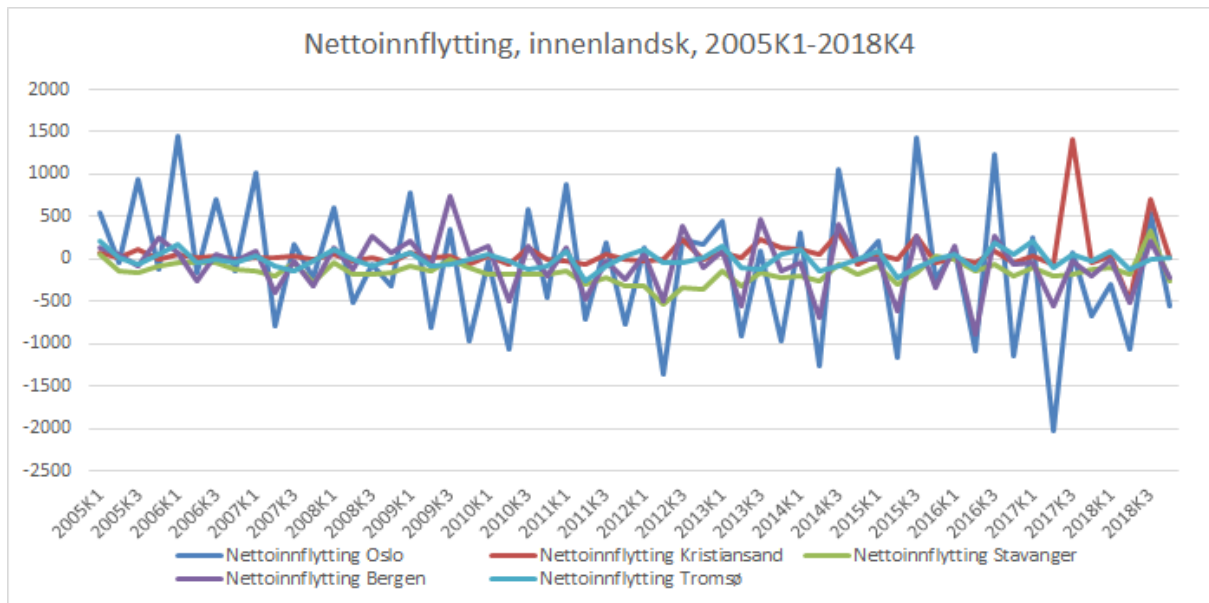
For samtlige byer er arbeidsledigheten på sitt laveste i forkant av finanskrisen. Vi ser deretter en økning i arbeidsledighet for alle byer etter finanskrisen inntreffer, noe som sammenfaller med at dette var en nasjonal krise, og dermed økte også arbeidsledigheten på et nasjonalt nivå. Regionen som ellers skiller seg ut er Stavanger, hvor vi ser at arbeidsledigheten steg kraftig i perioden rundt og etter 2014. I likhet med utviklingen i inntekt for Stavanger henger dette sammen med oljekrisen som oppstod i 2014, og det faktum at Stavanger har en høy andel sysselsatte i petroleumsnæringen.

4.5 Nettoinnflytting

Det er en klar urbaniseringstrend i Norge, så vel som i resten av verden. Selv om byene i Norge er små i en internasjonal sammenheng, er byene viktige for samfunnsutviklingen. For 150 år siden bodde bare 20 prosent av Norges befolkning i byer og tettsteder, mens i dag er denne andelen 80 prosent (NOU 2015:1). At befolkningen i økende grad flytter til de største byene i Norge, kan bidra til å øke etterspørselen etter boliger i byene, og det kan skape et økt boligpress i byene. Det kan tenkes at folk i dagens samfunn flytter til de større byene i Norge i økende grad på bakgrunn av blant annet bedre karrieremuligheter og utdanning. Jacobsen og Naug (2004) fant at befolkningsvekst ikke hadde en signifikant påvirkning på boligpris i sin analyse, men vi velger å inkludere et mål på befolkningsvekst til tross for dette.

Befolkningsvekst regnes ofte ikke som en av de fundamentale faktorene som påvirker boligpris, men vi er fremdeles interessert i å undersøke om det har en signifikant påvirkning i perioden vi undersøker.

SSB har produsert en rekke statistikker som omhandler befolkningsutvikling og -vekst. Blant disse er statistikk som viser innenlandsk innflytting og utflytting i antall personer etter region/kommune. Dataene finner vi i tabell 01222 i Statistikkbanken til SSB. Dette er dataene vi velger å anvende for å måle befolkningsvekst på et regionalt nivå. Dataene strekker seg fra 2005 til 2018, og oppgis med kvartalsvis frekvens. Dataene er verken sesong- eller kalenderjusterte. Vi bruker nettoinnflytting som målet på denne dataen, som beregnes ut ifra innflytting i hvert kvartal fratrukket utflyttingen i samme kvartal for hver respektive region. Vi velger å bruke nettoinnflytting for å måle befolkningsvekst i kommunene, fordi vi tror at dette i større grad fanger opp urbanisering enn ved å bruke den generelle folketilveksten for hver region. Hvis vi hadde brukt folketilveksten som mål, ville dette inkludert faktorer som barnefødsel og inn- og utvandring. Dette mener vi ville blitt et mer upresist mål, ettersom dette er grupper som nødvendigvis ikke umiddelbart deltar i boligmarkedet. Figuren nedenfor gir en presentasjon av de data for nettoinnflytting vi benytter i vår modell.

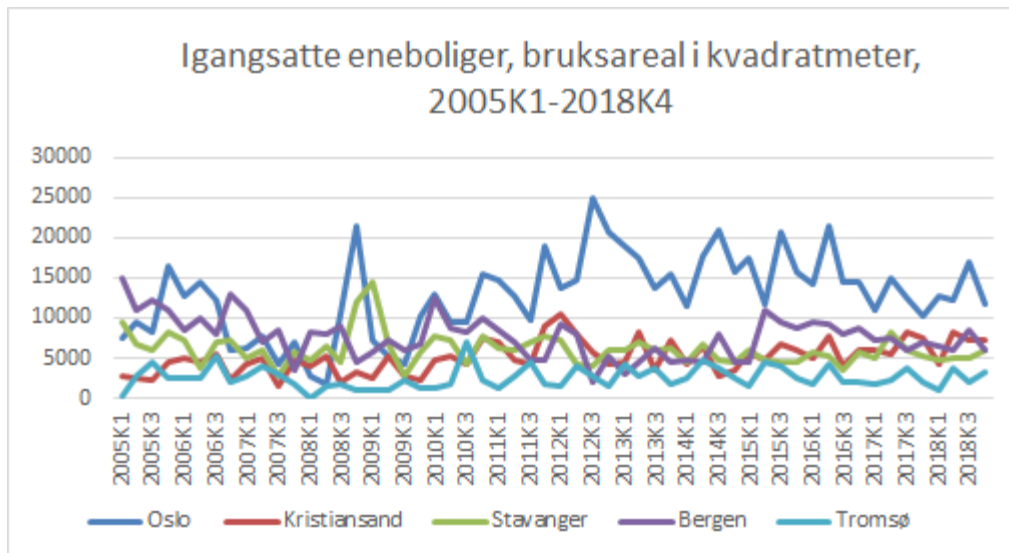


Figur 4.5: Innenlandsk nettoinnflytting etter region, 2005K1-2018K4

Kilde: Statistisk sentralbyrå, Statistikkbanken

4.6 Boligbygging

En annen fundamental faktor som påvirker boligprisene, er boligbygging. Igangsatte boliger målt i kvadratmeter er en potent indikator på tilbudssiden i boligmarkedet, da igangsatte boliger til slutt vil utgjøre ferdigstilte boliger, som summert med eksisterende boligmasse utgjør tilbudet i markedet. Igangsatte boliger blir da målet på nytt tilbud som tilføres i markedet. I vår modell benytter vi data for antall igangsatte eneboliger målt i kvadratmeter, hentet fra tabell 05889 i Statistikkbanken til SSB. Dataene oppgis med kvartalsvis frekvens. Vanligvis ville en brukt boligbygging målt i kroner som det mest presise målet, men ettersom vi studerer data for ulike byer er datautvalget begrenset. Likevel mener vi at boligbygging målt på denne måten vil være et nyttig mål for boligtilbudet på lang sikt. Som for kvadratmeterpris og inntekt vil vi transformere data for boligbygging om til logaritmisk form i modellen vi estimerer. Figuren under gir en presentasjon av de data vi benytter for boligbygging i vår modell.



Figur 4.6: Igangsatte eneboliger etter region, 2005K1-2018K4

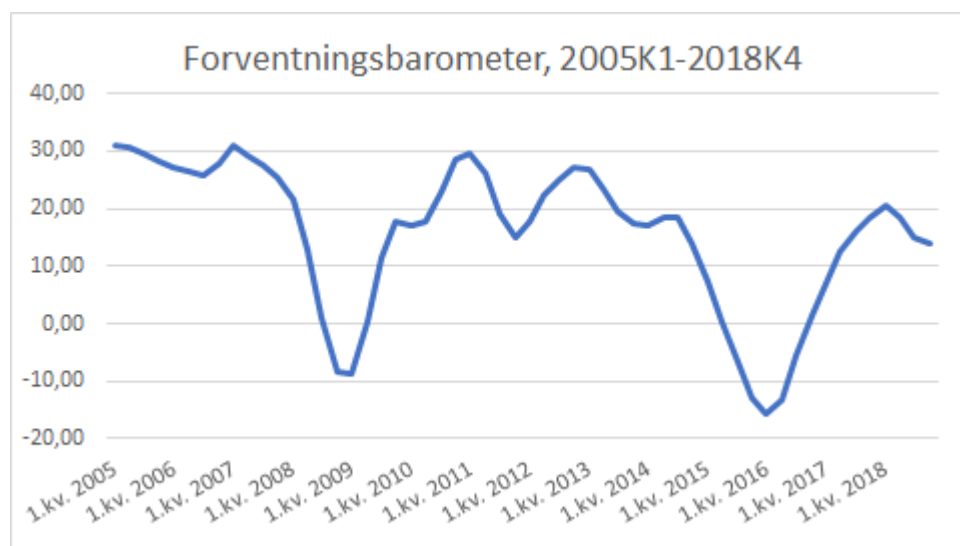
Kilde: Statistisk sentralbyrå, Statistikkbanken

4.7 Forventning

Konsumenters forventninger om egen økonomi samt den økonomiske utviklingen på et nasjonalt nivå har potensiale til å påvirke etterspørselen etter bolig i stor grad og dermed også boligprisene. Jacobsen og Naug (2004) argumenterer i sin artikkel at utviklingen i arbeidsmarkedet er viktig for husholdningenes vurdering av egne og andres fremtidige inntekter. De presiserer at økt arbeidsledighet vil gi forventninger om lavere lønnsvekst og økt usikkerhet om fremtidig inntekt og betalingssevne, som igjen vil gi redusert betalingsvillighet. Jacobsen og Naug (2004) presiserer også at en reduksjon i rente vil kunne gi forventninger om at realboligprisen skal øke i større grad enn før. I tillegg kan forventninger om økte boligpriser føre til en økning i boligpriser, og motsatt kan forventninger om lavere boligpriser føre til nettopp dette.

Disse forventningene regnes i litteraturen å være en fundamental faktor, og dermed også en relevant variabel som vi velger å inkludere i modellen vår. Kantar TNS produserer i samarbeid med Finans Norge en trendindikator som vi velger å bruke som basis i konstruksjonen av vår egen forventningsindikator. Denne trendindikatoren produseres kvartalsvis og måler antatt fremtidig etterspørsel fra forbrukernes side (Finans Norge, u.å.). Trendindikatoren fra Kantar TNS/Finans Norge konstrueres ved hjelp av en spørreundersøkelse hvor et tilfeldig utvalg av ca 1000 personer får fem ulike spørsmål knyttet til egen og Norges økonomi. Svarene blir deretter tillagt en tallverdi basert på responsen, der

positive verdier er knyttet til positive forventninger til egen og landets økonomi. Jo høyere verdier jo høyere er forventningene. Figur 4.7 viser utviklingen i forventningsbarometeret fra 1. kvartal 2005 frem til 4. kvartal 2018.



Figur 4.7: Forventningsbarometeret, 2005K1-2018K4

Kilde: Finans Norge/Kantar AS

Fra figur 4.7 ser vi at perioden fra 2005 til 2018 er preget av både oppgang og nedgang i forventninger til egen og landets økonomi. Perioden preges av to store nedgangsperioder, henholdsvis i 2008-2009 og 2014-2016. Den første nedgangen i 2008-2009 er en refleksjon av finanskrisen som inntraff i 2008, og var et nasjonalt sjokk. Etter denne krisen ser vi at forventningene til egen og landets økonomi øker igjen.

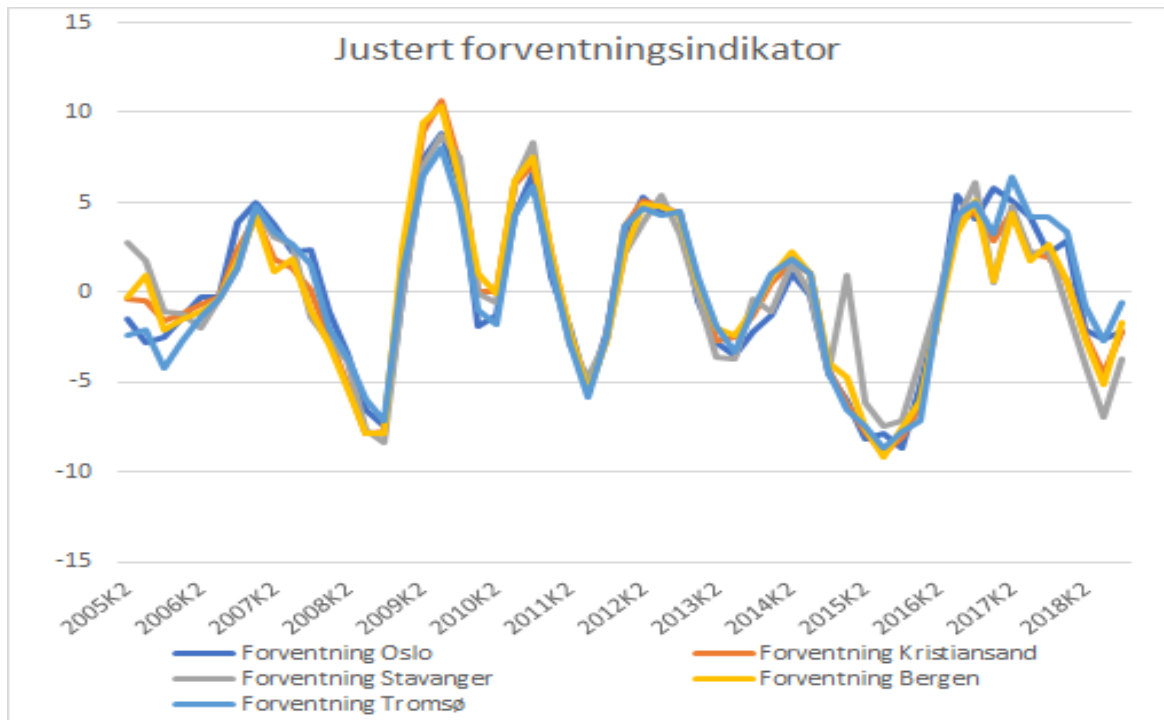
Den andre nedgangen som oppstod i perioden etter 2014 har i stor grad sammenheng med oljekrisen som inntraff da oljeprisene falt drastisk i 2014. Dette var ikke i like stor grad et nasjonalt sjokk og det kan tenkes at ulike regioner ble påvirket ulikt. Spesielt Stavanger, som tidligere nevnt hadde en stor andel ansatte i petroleumsnæringen, ble i større grad påvirket av oljekrisen. Det er derfor viktig å presisere at forventningsbarometeret i figur 4.7 er en landsdekkende variabel og tar dermed ikke høyde for de regionale forskjellene som oppstår, noe vi ønsker å ha med i variabelen vår. Et annet problem med forventningsbarometeret er dets naturlige tilknytning til arbeidsledighet og rente. Dette er som forklart tidligere egne forklaringsvariabler og korrelasjonen mellom disse kan føre til problemer med multikollinearitet når vi skal bruke de i modellen vår.

Problemet med multikollinearitet kan dog løses ved å korrigere forventningsindikatoren til Finans Norge/Kantar AS for effekten av arbeidsledighet og rente. Dette gjøres ved å først estimere en modell for forventningsindikatoren basert på data om rente og arbeidsledighet. Vi kan deretter beregne avviket mellom den anslåtte verdien i forhold til den faktiske verdien. Dette avviket forklarer endringer i forventninger som ikke skyldes arbeidsledighet eller rente. Dette gjør at vi også kan observere regionale endringer i forventninger ettersom den beregnede forventningsindikatoren er basert på regionale data om arbeidsledighet og ikke nasjonale data. Denne fremgangsmåten for å beregne en forventningsindikator ble først gjennomført av Jacobsen og Naug (2004), og vi vil i hovedsak ta i bruk deres metode når vi estimerer vår egen forventningsindikator.

4.8 Utledning av forventningsindikator

I utledningen av vår egen forventningsindikator for de ulike regionene tar vi utgangspunkt i fremgangsmåten til Jacobsen og Naug. Det er dog viktig å legge merke til at Jacobsen og Naugs modellering har blitt utsatt for kritikk. Fredriksen (2007) skrev i sin masteroppgave en gjennomgående kritikk av fremgangsmåten til Jacobsen og Naug og deriblant ble utledningen av forventningsindikatoren kritisert. Fredriksen stiller spørsmål til deres egenkomponerte formel hvor de transformerer deres egen forventningsindikator ved å summere feilleddet de fikk i regresjonen over to kvartaler for så å putte denne verdien inn i formelen $FORV_t = (\varepsilon_t + \varepsilon_{t-1}) + 100 * (\varepsilon_t + \varepsilon_{t-1})^3$. Jacobsen og Naug forklarer ikke i sin artikkel hvorfor de tar med det siste leddet og refererer heller ikke til andre som har tatt i bruk denne metoden. Når vi utleder vår egen forventningsindikator, velger vi derfor å unnlate å transformere forventningsindikatoren.

I utledningen av vår egen forventningsindikator estimerer vi først forventningsvariabelen for hver by med arbeidsledighet og rente som forklaringsvariabler. Vi bruker deretter feilleddene vi får for hver by og bruker disse direkte i vår egen modell for boligpriser. Figur 4.8 viser utviklingen til vår egen konstruerte forventningsindikator for hver enkelt by fra andre kvartal 2005 til fjerde kvartal 2018.



Figur 4.8: Justert forventningsindikator etter region, 2005K2-2018K4

Kilde: Egen utledning av forventningsindikator, basert på forventningsindikatoren fra Finans Norge/Kantar AS, samt tall fra SSB

Som vi ser korrelerer forventningene for byene sterkt med hverandre, og forventningene følger i stor grad samme mønster. Dette tyder på at endringer i forventninger i større grad skyldes endringer på nasjonalt nivå og regionale forskjeller i forventning virker å være små i perioden vi undersøker. Vi ser dog at det er forskjeller, og spesielt Stavanger skiller seg ut i både 2005 og 2015.

4.9 Transformasjon av nominelle verdier til reelle verdier

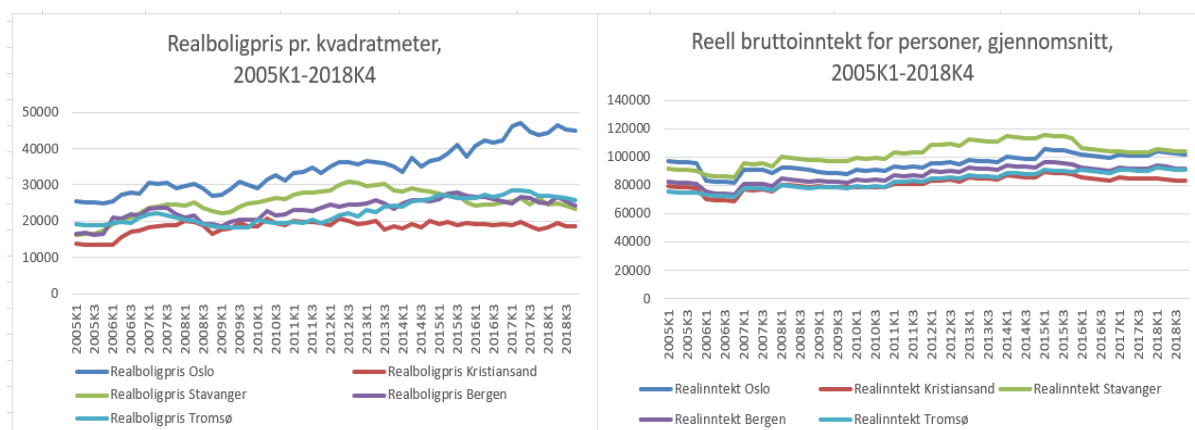
Majoriteten av økonomisk oppførsel antas å bli påvirket av reelle og ikke nominelle variabler. En prisindeks som KPI brukes ikke bare til å kalkulere inflasjonsraten, men er i tillegg nødvendig for å transformere en tidsserie målt i nominelle (nåværende) tall til reelle (konstante) tall. At en variabel er reell vil si at den er justert for den generelle prisveksten på varer og tjenester i økonomien. For å transformere variablene i en tidsserie til reelle verdier, multipliserer vi variabelen med 100 og dividerer med verdien for konsumprisindeksen for hver periode t (Wooldridge, 2012):

$$\frac{\text{Variabel}_t * 100}{KPI_t} \quad (4.1)$$

Fordelen med å transformere nominelle verdier til realverdier i tidsserier er at vi justerer serien for endringer som forårsakes av inflasjon. Dette gjør at vi observerer rene verdiendringer i variabelen, og ikke verdiendringer som er resultat av den generelle prisveksten i økonomien.

Ettersom vi i vår modell undersøker boligprisutvikling og regionale forskjeller vil det være mest naturlig å anvende reelle verdier, for å best fange opp rene verdiendringer som ikke skyldes generell prisvekst i økonomien. Variablene vi transformerer om til reelle størrelser er boligpris, inntekt og rente som henholdsvis beskrives i delkapitlene 4.1, 4.2 og 4.3. Dette gjør vi for å bedre fange opp sammenhenger mellom boligprisutviklingen og endringer i de nevnte variablene, ettersom vi ikke lenger behøver å ta høyde for at boligprisutviklingen kan skyldes den generelle prisveksten i økonomien. Data for KPI ble presentert i månedlig frekvens i delkapittel 3.2.5, og det er denne data som danner grunnlaget når vi justerer for KPI. Vi transformerer denne dataen til kvartalsvise verdier ved å regne ut gjennomsnittsverdien hver 3-måneders periode som definerer hvert respektive kvartal.

Vi anvender formelen i likning 4.1 for å justere data for boligpris og inntekt med KPI for å transformere seriene om til realboligpris og realinntekt. Figur 4.9 gir en presentasjon av de reelle transformasjonene vi foretar på boligpris og inntekt.



Figur 4.9: Reelle transformasjoner av nominelle boligpris- og inntektsdata, 2005K1-2018K4

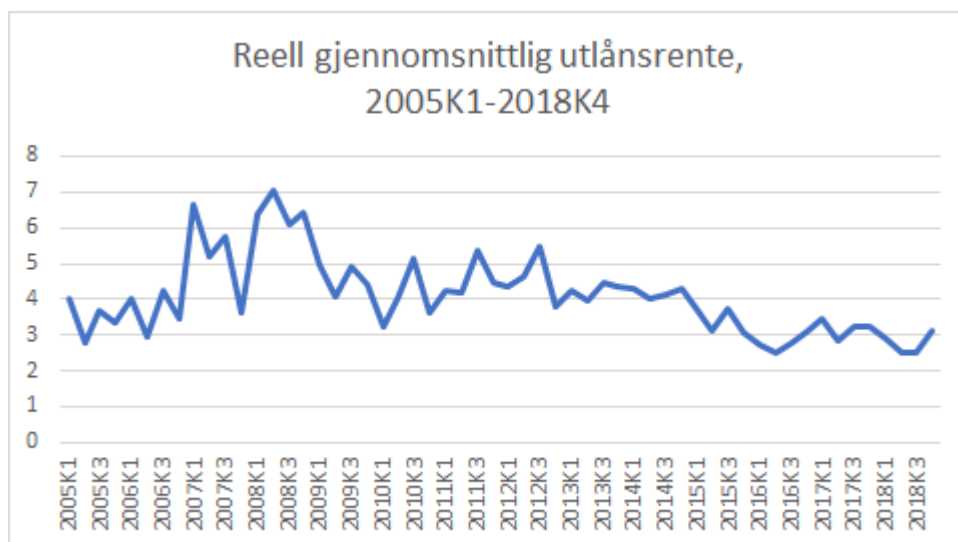
Kilde: Statistisk sentralbyrå, Statistikkbanken, egne beregninger

Videre transformerer vi også bankenes gjennomsnittlige utlånsrente om til realrente. Realrenten er renten utover prisveksten, og angir dermed hvor mye renten reelt sett gir i økt kjøpekraft (Sirnes, 2018). Realrenten blir her dermed differansen mellom den nominelle utlånsrenten målt i prosent og inflasjonen målt i prosent. Med andre ord vil dette si at vi justerer utlånsrenten for inflasjon. Kvartalsvis inflasjon i prosent blir regnet ut som endringen i KPI mellom kvartaler og vi anvender følgende formel:

$$\text{Inflasjon}(\%) = \left(\frac{KPI_{t+1} - KPI_t}{KPI_t} \right) * 100 \quad (4.2)$$

Videre konstruerer vi realrente i kvartal t med følgende formel:

$$\text{Nominell utlånsrente}_t - \text{Inflasjon}_t = \text{Realrente}_t \quad (4.3)$$



Figur 4.10: Reell transformasjon av bankenes nominelle utlånsrente, 2005K1-2018K4

Kilde: Statistisk sentralbyrå, Statistikkbanken, egne beregninger

4.10 Utelatte variabler

I tillegg til de variablene vi inkluderer i vår modell finnes det også andre faktorer som kan påvirke boligpris. Faktorer som husleie, gjeld, bo- og byggekostnader er i litteraturen forklart som faktorer som påvirker boligpris, men grunnet det begrensede datautvalget vi har for disse variablene på et regionalt nivå, ser vi oss nødt til å ekskludere disse. Det hadde også vært interessant å inkludere mål på sentralitet og næringsinndeling i de ulike regionene, men igjen finnes det ikke tilstrekkelige tidsserier på dette som vil kunne bidra til å forbedre modellen

vår. I tillegg hadde det også vært interessant å inkludere demografiske forhold som variabel i analysen vår for å undersøke disse forholdene. Demografiske forhold som eksempelvis aldersfordeling, spesielt med tanke på endring i antall førstegangskjøpere over tid, hadde vært veldig interessant å inkludere som variabel. Dette blir dog for komplisert å konstruere til én variabel og i tillegg la den fungere som en tidsserie, slik at dette blir utenfor vår rekkevidde i oppgaven. Det hadde også vært gunstig å undersøke en lengre tidsperiode, men det begrensede datautvalget gjorde dette umulig i denne omgang.

5. Metode

I dette kapittelet vil vi presentere aktuell teori og metode, som vil fungere som et fundament til utførelsen av de empiriske analysene våre. Vi presenterer først en kort argumentasjon for valget av forskningsdesign, metode og datainnsamling. Deretter gjennomgår vi den metodiske tilnærmingen som sammenfaller best med vår innsamlede data fra kapittel 4. Vi vil gjennomgå betingelser og forutsetninger for at modellene skal gi presise og forventningsrette resultater, samt fremgangsmåtene vi anvender for å undersøke om disse forutsetningene er innfridd. I løpet av dette kapittelet vil de endelige modellene bli utarbeidet på et teoretisk plan, før de blir utført økonometrisk i kapittel 6.

5.1 Modellspesifikasjon

Når vi skal estimere og tolke statistiske modeller som omhandler økonomi gjør vi dette på grunnlag av teorier og metode innenfor statistikk og økonomi. Denne typen metode kalles økonometri. For å identifisere de relevante forklaringsvariablene som skal inkluderes i modellen har vi tatt utgangspunkt i teori om boligpriser og boligmarkedet. Det er også viktig å forstå hvordan boligmarkedet fungerer for å vite hvilken form de ulike variablene skal ha. For å forstå hvordan vi skal estimere modellen legges økonometrisk metode til grunn. Intuisjon på grunnlag av økonometrisk teori legges til grunn i avgjørelsen om hvilken variabel som skal være avhengig og dermed variabelen som blir målt.

5.2 Metodisk tilnærming

I dette delkapittelet gir vi en kort begrunnelse for valget av forskningsdesign, metodisk tilnærming, og datainnsamlingsmetode.

5.2.1 Forskningsdesign

I oppgaven vår har vi valgt et forklarende design. Det vil si at formålet med studiet er å finne kausale sammenhenger mellom variabler. Vektleggingen her er å studere et problem eller en situasjon slik at man kan forklare forholdet mellom variabler (Saunders, Lewis, & Thornhill, 2012). I oppgaven vår er problemet/situasjonen de regionale forskjellene i boligprisutvikling og målet med oppgaven blir å undersøke om vi kan finne kausale sammenhenger som kan forklare årsaken til disse forskjellene.

5.2.2 Valg av metode

Gjennom oppgaven har vi valgt en kvantitativ tilnærming til problemstillingen vår. Denne forskningsmetoden er vanligvis assosiert med en deduktiv tilnærming der fokuset er å bruke data til å teste teori. Kvantitativ forskning kjennetegnes ved at man undersøker sammenhenger mellom variabler som måles numerisk og som analyseres ved hjelp av ulike statistiske metoder (Saunders et al., 2012). Kvantitativ forskning kan også brukes på en induktiv måte der målet er å utvikle ny teori. Dette er imidlertid ikke vårt mål, og gjennom denne oppgaven tar vi i bruk allerede etablert teori for å belyse problemstillingen. Vi bruker deretter denne teorien til å undersøke og beskrive sammenhenger mellom variabler som kan forklare de regionale forskjellene i boligpriser. Vi kommer til å bruke ulike statistiske metoder for å belyse disse sammenhengene og disse vil bli videre forklart gjennom dette kapitlet. Metodene vil bli utført i den generelle statistiske programvarepakken STATA 16.

5.2.3 Datainnsamling

En viktig årsak til valg av metode var tilgjengeligheten på sekundærdata. Det vil si data som allerede har blitt samlet inn av andre aktører til ulike formål. Vi har samlet store mengder data som er egnet til å kvantifiseres i målbare verdier, noe som tilsier at en kvantitativ tilnærming er naturlig. En fordel med bruk av sekundærdata er at det er tidsbesparende og at en får samlet inn og ta i bruk relativt store mengder data på kort tid. Alternativt vurderte vi bruken av primærdata, men ettersom dette både er tidkrevende og mengden data man får samlet inn er begrenset, valgte vi å utelate denne type data. Når vi ønsker å studere sammenhenger mellom utviklingen i den avhengige variabelen og forklaringsvariabler vil mengden data vi observerer være avgjørende for forklaringskraften.

5.3 Regresjonsanalyse og minste kvadraters metode (OLS)

Regresjonsanalyse er en sentral metodisk tilnærming i kvantitativ forskning. En regresjonsanalyse er en statistisk analyse som anvendes for å forklare sammenhengen mellom to eller flere variabler. Formålet med regresjonsanalyse er å undersøke om en eller flere uavhengige variabler har en signifikant påvirkning på den avhengige variabelen, og i hvilken grad de uavhengige variablene påvirker den avhengige variabelen. En regresjonsanalyse skaper et tilnærmet uttrykk for sammenhengen mellom variablene, og gir dermed en prediksjon av det antatte forholdet mellom variablene.

En enkel regresjonsmodell med avhengig variabel y og forklaringsvariabel x kan presenteres på følgende måte (Wooldridge, 2012, s. 22-23):

$$y = \beta_0 + \beta_1 x + u \quad (5.1)$$

hvor

β_0 = konstantleddet, verdien til y når $x=0$

β_1 = koeffisient, forteller hvor mye verdien til y endres av én enhets endring i x

u = feilledd, et restledd som fanger opp all variasjon i y som ikke forklares av den uavhengige variabelen som inngår i modellen.

Ulempen med å bruke en enkel regresjonsanalyse til empirisk arbeid er at det er utfordrende å trekke ceteris paribus konklusjoner om hvordan den uavhengige variabelen påvirker den avhengige variabelen. Ceteris paribus betyr å holde alt annet likt. Dette er på grunnlag av at antakelsen om at feilleddet ikke korrelerer med den uavhengige variabelen x ofte er urealistisk i en enkel regresjonsanalyse.

En multivariat regresjonsanalyse er derimot mer mottakelig for ceteris paribus analyse, fordi den tillater oss å eksplisitt kontrollere mange faktorer som antas å samtidig påvirke den avhengige variabelen. En multivariat regresjonsmodell er en modell som består av flere enn én forklaringsvariabel. Ved å inkludere flere nyttige forklaringsvariabler, kan dermed en større andel av variasjonen i den avhengige variabelen bli forklart (Wooldridge, 2012, s. 68). Dette vil bygge en bedre modell for å predikere den avhengige variabelen, og følgelig noe som passer bedre inn i denne masterutredningen.

I en multivariat regresjonsmodell utføres regresjonsanalyse for å finne et tilnærmet uttrykk for hvordan verdien til den avhengige variabelen endrer seg dersom en av forklaringsvariablene endres, mens de andre forklaringsvariablene i modellen holdes uendret. En multivariat regresjonsmodell med k uavhengige variabler kan presenteres på følgende måte (Wooldridge, 2012, s.71):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k + u \quad (5.2)$$

hvor

β_0 = konstant og skjæringspunktet

β_1 = koeffisient, måler endringen i den avhengige variabel y med respekt for x_1 , alle andre faktorer uendret.

β_2 = koeffisient, måler endringen i den avhengige variabel y med respekt for x_2 , alle andre faktorer uendret.

β_k = koeffisient, måler endringen i den avhengige variabel y med respekt for x_k , alle andre faktorer uendret.

u = feilledd, et restledd som fanger opp all variasjon i y som ikke forklares av alle forklaringsvariabler som inngår i modellen.

Dette er en statistisk modell. I vår oppgave ønsker vi å undersøke boligpriser, prisdrivere og hvordan pris og drivere utvikler seg over tid. I vårt tilfelle ønsker vi å modellere boligpris representert ved kvadratmeterpris, med hensyn på ulike fundamentale faktorer representert ved flere ulike variabler. En enkel modell for dette kan følgende ligne:

$$\text{Boligpris}_t = \beta_0 + \beta_1 \text{Forklaringsvariabel1}_t + \beta_2 \text{Forklaringsvariabel2}_t + \dots + \beta_N \text{ForklaringsvariabelN}_t + \varepsilon_t \quad (5.3)$$

der ε_t = feilledd og tidsperiode $t = 1, 2, \dots, T$

Minste Kvadraters Metode, den norske betegnelsen for Ordinary Least Squares (OLS), er en metode for å estimere koeffisienter til modellen, og dermed estimere forholdet mellom den avhengige variabelen og de uavhengige forklaringsvariablene. Koeffisientene i regresjonsanalysen er estimater av de faktiske parametrene til populasjonen. Prinsippet i Minste Kvadraters Metode, heretter MKM, er å estimere koeffisientene som minimerer summen av de kvadrerte restleddene (residualene). Restleddene kan forklares som avviket mellom de faktiske og predikerte verdiene for hver observasjon.

5.3.1 Betingelser/Forutsetninger

Modellen for MKM vil ikke uten videre estimere de mest forventningsrette koeffisientene. Man legger dermed til grunn noen forutsetninger, også kalt betingelser, for at denne modellen kan kalles "den klassiske lineære regresjonsmodellen" (Brooks, 2008, s. 43). Forutsetningene

er knyttet til residualene og er grunnleggende for at modellen skal kunne estimere forventningsrette og presise koeffisienter, og presenteres nedenfor (Brooks, 2008, s. 44):

1. $E(u_t) = 0$

Den første forutsetningen er at den gjennomsnittlige forventningsverdien til residualene er lik null. Ifølge Brooks (2008) vil denne forutsetningen alltid holde dersom man inkluderer et konstantledd i regresjonslikningen.

2. $var(u_t) = \sigma^2 < \infty$

Den andre forutsetningen er at variansen i residualene er konstant og endelig for alle observasjoner. Dette er synonymt med homoskedastisk varians i residualene. Dersom forutsetningen ikke holder vil variansen være heteroskedastisk, og en slik heteroskedastisk varians vil kunne føre til en feilestimering av verdien til den avhengige variabelen. Heteroskedastisitet i residualene vil ikke påvirke om estimatene er forventningsrette, men påvirker derimot standardfeil i den grad at konfidensintervaller og teststatistikker blir ugyldige (Wooldridge, 2012, s. 269).

3. $cov(u_i, u_t) = 0, i \neq j$

Den tredje forutsetningen er at det ikke foreligger kovarians mellom residualet til en uavhengig variabel fra en tidsperiode til en annen. Med andre ord, at residualene er lineært uavhengige av hverandre. Dersom det oppstår en slik korrelasjon i residualene, kalt autokorrelasjon, vil dette føre til ugyldige standardfeil. Dette kan føre til upresise avgjørelser i hypotesetestingen.

4. $cov(u_i, x_t) = 0$

Den fjerde forutsetningen er at perfekt kollinearitet ikke eksisterer. Dette innebærer at det ikke må eksistere korrelasjon mellom de uavhengige variablene og deres respektive residualer. Ingen uavhengig variabel kan være konstant eller perfekt lineært forklart av en annen variabel. Dersom dette oppstår kan det føre til

forstyrrelser i analysen, som vil medføre feilaktige estimater på hvilke uavhengige variabler som faktisk påvirker den avhengige variabelen. Perfekt kollinearitet vil med andre ord medføre forventningsskjevne estimater for koeffisientene.

5. $u_i \sim N(0, \sigma^2)$

Den femte og siste forutsetningen tilsier normalfordelte residualer. Residualene er uavhengige av den respektive uavhengige variabelen og er identisk fordelt som en normal. Denne forutsetningen må være oppfylt for at inferenstesting av parametrene skal være gyldig. Dette fordi F- og t-tester antar normalfordelte data.

5.3.2 Egenskapene til estimatene

I tidsserieregresjoner forutsettes det også at modellen er en stokastisk prosess og lineær i sine parametere. Stokastisk er synonymt for tilfeldig, og tilsier at variablene er tilfeldige basert på en sannsynlighetsfordeling. At parameterne er lineære vil si at leddene i regresjonen enten er konstanter, eller produktet av en parameter og en forklaringsvariabel, og følgelig at den avhengige variabelen blir forklart lineært av forklaringsvariablene. Når dette ligger til grunn, og forutsetninger 1-4 beskrevet ovenfor holder, vil estimatene fra regresjonen være såkalt BLUE. Sammenlagt sammenfaller disse forutsetningene med Gauss-Markov betingelsene for tidsserieregresjoner (Wooldridge, 2012, s.374). Gauss-Markov betingelsene garanterer validiteten til MKM sin evne til å estimere regresjonskoeffisienter.

At en estimator er BLUE (Best Linear Unbiased Estimator), vil si at den besitter ønskelige egenskaper som vil gi forventningsrette koeffisienter. Dersom disse betingelsene ikke er oppfylt vil regresjonen gi oss forventningsskjevne koeffisientestimat, noe som betyr at estimatene for koeffisientene enten vil være over- eller underestimerte i forhold til sann verdi. Ifølge Brooks (2008) får estimatene følgende egenskaper når de nevnte forutsetningene ligger til grunn:

Estimator - $\hat{\alpha}$ og $\hat{\beta}$ er estimatorer av den sanne verdien til α og β

Linear - $\hat{\alpha}$ og $\hat{\beta}$ er lineære estimatorer, det vil si at de er lineære kombinasjoner av variablene

Unbiased - De faktiske verdiene til $\hat{\alpha}$ og $\hat{\beta}$ vil i gjennomsnitt være lik deres sanne verdier

Best - OLS estimatoren $\hat{\beta}$ har den minste variansen i sin klasse av lineære objektive estimatorene som også oppfyller forutsetningene ovenfor.

5.4 Stasjonærhet og Spuriøsitet

Begrepet stasjonærhet innebærer at sannsynlighetsfordelingen til variablene i en tidsserie er konstante over tid. Wooldridge (2012) forklarer dette som at hvis en tar et utvalg tilfeldige variabler og flytter dem h tidsperioder fremover i tid, må den samlede sannsynlighetsfordelingen forbli uendret. Mer formelt defineres dette som: Den stokastiske prosessen $\{x_t: t = 1, 2, \dots\}$ er stasjonær for alle tidsperioder t dersom den samlede sannsynlighetsfordelingen for $(x_{t1}, x_{t2}, \dots, x_{tm})$ utgjør den samme som for $(x_{t+h1}, x_{t+h2}, \dots, x_{t+h_m})$ for alle $h \geq 1$.

For å kunne definere en tidsserie som stasjonær, må følgende forutsetninger oppfylles:

1. $E(x_t) = \mu$ for alle t
2. $Var(x_t) = \sigma^2$ for alle t
3. $Cov(x_t, x_{t+h}) =$ avhenger kun av h og ikke t , for alle t og $h \geq 1$

Første og andre forutsetning krever at variabelens gjennomsnitt og varians må være konstant for alle tidsperioder. Den tredje og siste forutsetningen krever at kovariansen mellom x_t og x_{t+h} kun avhenger av distansen h mellom uttrykkene, og ikke av den innledende beliggenheten i tidsperioden t .

Modellen vi har presentert i likning 5.2 er eksponert for flere potensielle feil. Data anvendt i tidsserier inneholder ofte en felles tidstrend, som vil si at to eller flere tilsynelatende uavhengige variabler stiger eller synker sammen over tid. En slik felles tidstrend kan medføre at en finner høy korrelasjon mellom variabler som faktisk er uavhengige av hverandre, og at de i realiteten kun er relatert gjennom korrelasjon med en tredje variabel. Den påviste korrelasjonen kan med andre ord være helt tilfeldig. Dette defineres som spuriøs regresjon (Wooldridge, 2012, s.644).

Stasjonærhet kommer inn i bildet her, ettersom bruk av ikke-stasjonære variabler kan lede til spuriøs regresjon. Spuriøs regresjon fører ofte til en kunstig høy forklaringskraft (R^2), som

kan være misledende. Bruk av ikke-stasjonære variabler vil også føre til at t-tester og F-tester ikke vil følge riktig t- og F-distribusjon (Brooks, 2008, s.319-320). Det er dermed ønskelig å unngå bruk av ikke-stasjonære variabler og konsekvent unngå spuriøse regresjoner, vi ønsker derfor å avklare om variablene vi anvender er et resultat av stasjonære prosesser.

Dersom forutsetningene presentert ovenfor ikke holder, vil følgelig heller ikke den aktuelle tidsserien være stasjonær. Dersom det er en mulighet for tilstedeværelse av ikke-stasjonære variabler, kan vi transformere tidsserien ved å bruke førstedifferansen av variablene i regresjonsanalysen:

$$\Delta x_t = x_t - x_{t-1} \quad (5.4)$$

Dersom variabelen blir stasjonær etter en slik førstedifferensiering, defineres det som at variabelen x_t er integrert av første orden, $I(1)$. En variabel som i utgangspunktet er stasjonær defineres som $I(0)$.

5.4.1 Testing for stasjonærhet

Stasjonærhet kan testes for ved å anvende en “augmented Dickey-Fuller test” (ADF). Dette er en metode som er hyppig brukt for å identifisere ikke-stasjonære prosesser. Augmented betyr utvidet, og i denne sammenheng tar den utvidede Dickey-Fuller testen, i motsetning til Dickey-Fuller testen (DF), høyde for at residualene kan være autokorrelert. Ikke-stasjonære prosesser kan ha en såkalt enhetsrot, hvor prosessens varians avhenger av, og øker i tid. Et typisk eksempel på dette er en såkalt random walk, med eller uten drift. Dette fører også til spuriøse regresjoner, og når vi tester for stasjonærhet undersøker vi om en tidsserie følger en enhetsrotprosess eller ikke (Wooldridge, 2012).

Vi har en enkel autoregressiv modell (AR(1)) gitt ved:

$$y_t = a + \rho y_{t-1} + e_t, t = 1, 2, \dots, \quad (5.5)$$

hvor y_t er den innledende observerte verdien. Hvis den avhengige variabelen y_t følger denne prosessen har den en enhetsrot hvis $\rho = 1$. Dersom $\alpha = 0$ og $\rho = 1$ følger den avhengige variabelen en random walk. Dersom $\alpha \neq 0$, følger den avhengige variabelen en random walk

med drift, noe som betyr at forventningsverdien til y_t er en lineær funksjon av t . Det er vanlig å la α forbli uspesifisert i Dickey-Fuller tester.

Nullhypotesen blir dermed at y_t har en enhetsrot:

$$H_0 : \rho = 1$$

Den ensidige alternativhypotesen vi vanligvis interesserer oss for er:

$$H_1 : \rho < 1$$

Når vi forkaster nullhypotesen og aksepterer alternativhypotesen, betegnes den avhengige variabelen å være en stabil AR(1) prosess, som betyr at den er svakt avhengig eller asymptotisk ukorrelet. En beleilig fremgangsmåte for å bruke Dickey-Fuller testen er å trekke fra y_{t-1} på begge sider i likning 5.3 og ende opp med:

$$\Delta y_t = \alpha + (\rho - 1)y_{t-1} + e_t \tag{5.6}$$

Dette er en dynamisk komplett modell, og det ser henholdsvis ut som at vi kan gå frem og teste for stasjonæritet. Problemet er at under H_0 vil y_{t-1} være I(1), og følgelig vil ikke det vanlige sentralgrenseteoremet underliggende den standard normalfordelingen for t statistikk gjelde. Dermed kan ikke de vanlige kritiske verdiene anvendes, men nye kritiske verdier har blitt utredet. Derimot kan t -statistikken brukes. Dermed brukes den asymptotiske distribusjonen i t -statistikken kjent som Dickey-Fuller distribusjonen samt de utviklede kritiske verdiene, som samlet former Dickey-Fuller testen for enhetsrot (Wooldridge, 2012).

En sentral antakelse i Dickey-Fuller testen er at det ikke eksisterer autokorrelasjon i residualet. Den utvidede Dickey-Fuller testen tar høyde for dette, og passer til å teste for enhetsrot i modeller med mer kompliserte dynamikker. Intuisjonen her blir at vi legger til flere elementer i den dynamiske modellen vår inntil feilledet er hvit støy, det vil si at det ikke lenger er autokorrelet. Vi utvider den tidligere modellen vi presenterte i likning 5.6 ved å inkludere tidsforskyvninger og får:

$$\Delta y_t = \alpha + \delta t + (\rho - 1)y_{t-1} + y_1 \Delta y_{t-1} + e_t \tag{5.7}$$

Dette forsikrer at Δy_t følger en stabil AR(1) modell under $H_0: \rho = 1$. Hensikten med tidsforskyvninger er å renske opp all form for seriekorrelasjon (autokorrelasjon) i Δy_t .

Lengden på tidsforskyvninger av Δy_t avhenger av frekvensen på data. For årlige data kan en eller to tidsforskyvninger være tilstrekkelig, mens en for månedlige data burde inkludere opp mot 12 tidsforskyvninger. Vi kan også bestemme antall tidsforskyvninger ut ifra ulike informasjonskriterier, for eksempel Schwarz Bayesian Information Criterion. Vi kan også studere t-verdien til de respektive koeffisientene på hver av de laggede variablene, og inkludere de som vises å være signifikante. δt kontrollerer for tidstrend, og inkluderes i tidsserier som har en tydelig tidstrend. I praksis har den utvidede Dickey-Fuller testen holdt ganske bra (Wooldridge, 2012, s. 641-643).

5.5 Autokorrelasjon

Et annet problem vi kan støte på når vi estimerer tidsserien er autokorrelasjon. Autokorrelasjon er at residualet i regresjonen korrelerer over tid. Dersom vi har autokorrelasjon i residualet bryter dette med Gauss-Markov betingelsen om ingen autokorrelasjon, og estimatene vil ikke være BLUE. Autokorrelasjon vil føre til ugyldig standardavvik og teststatistikk i MKM (Wooldridge, 2012, s. 413). Koeffisientestimatene vil fortsatt være objektive (unbiased), men de vil ikke være effisiente. Dette kan føre til at vi trekker feil konklusjon om signifikansen til en forklaringsvariabel, og dermed inkludere eller ekskludere variabler i modellen som ikke burde være der. I tillegg til dette vil forklaringskraften R^2 være oppblåst relativt til "riktig" verdi (Brooks, 2008, s. 149-150). Dette påvirker validiteten til modellen. Autokorrelasjon kan kontrolleres for ved hjelp av robuste standardfeil (Wooldridge, 2009). Robuste standardfeil er større enn vanlige standardfeil, og tar høyde for autokorrelasjon. Autokorrelasjon er ofte et av hovedproblemene i tidsserier, og vi ønsker dermed å benytte ulike tester for å avgjøre om en modell har korrelerende residualer.

5.5.1 Statistisk Test for autokorrelasjon

Durbin-Watson

En vanlig måte å teste for autokorrelasjon av første orden under de klassiske forutsetningene er den såkalte Durbin-Watson testen (Wooldridge, 2012, s. 419). Denne testen krever derimot at regresjonen ikke inneholder tidsforskyvninger av den avhengige variabelen som forklaringsvariabel, og testen er dermed uegnet i dynamiske modeller hvor dette inkluderes. I tillegg er den dårlig egnet for modeller med stokastiske forklaringsvariabler (Brooks, 2008, s. 148).

Breusch-Godfrey

Alternativt til Durbin-Watson testen er Breusch-Godfrey(BG) testen. Denne testen passer bedre for modeller som inkluderer tidsforskyvninger av den avhengige variabelen som forklaringsvariabel. Breusch-Godfrey testen er en mer generell test for autokorrelasjon, som også tar for seg høyere orden enn første orden, opp til den r 'te orden. Modellen for feil under denne testen er (Brooks, 2008, s. 148):

$$u_t = p_1 u_{t-1} + p_2 u_{t-2} + p_3 u_{t-3} + \dots + p_r u_{t-r} + u_t, \quad u_t \sim N(0, \sigma_u^2) \quad (5.8)$$

Null- og alternativhypotesen er under denne testen:

$$H_0 : \rho_1 = 0 \text{ og } \rho_2 = 0 \text{ og } \rho_r = 0$$
$$H_1 : \rho_1 \neq 0 \text{ eller } \rho_2 \neq 0 \text{ eller } \rho_r \neq 0$$

Under nullhypotesen er den nåværende feil på tidspunkt t ukorrelert med alle andre r tidligere verdier. Forklaringskraften R^2 til feilleddet brukes til å beregne BG-teststatistikken:

$$(T - r)R^2 \sim \chi_r^2 \quad (5.9)$$

Testen følger altså en kjikvadratfordeling, og nullhypotesen om ingen autokorrelasjon i residualet forkastes dersom teststatistikken er større enn kritisk verdi. En utfordring med Breusch-Godfrey testen er å fastslå riktig antall lags r av residualet som skal brukes. Dette kan typisk eksperimenteres med frem til mest anvendelig antall lags er funnet, men frekvensen av data spiller også inn (Brooks, 2008, s. 149).

Ljung-Box

Et annet alternativ som anvendes for å teste for autokorrelasjon er Ljung-Box testen. Ljung-Box testen egner seg for å teste autokorrelasjon for mer enn bare første orden, og tester altså også for høyere orden. Dette innebærer at den kan teste om en gruppe autokorrelasjoner i en tidsserie er forskjellig fra null. Ljung-Box testen har gode egenskaper som egner seg for små datautvalg. I tillegg egner testen seg godt i modeller hvor vi inkluderer laggede verdier av den avhengige variabelen som forklaringsvariabel. Grunnet teststatistikken egner testen seg

veldig godt som en generell test for lineær avhengighet i tidsserier (Brooks, 2014, s. 254).

Ljung-Box statistikken er gitt ved:

$$Q^* = T(T + 2) \sum_{k=1}^m \frac{t_k^2}{T-k} \sim \chi_m^2 \quad (5.10)$$

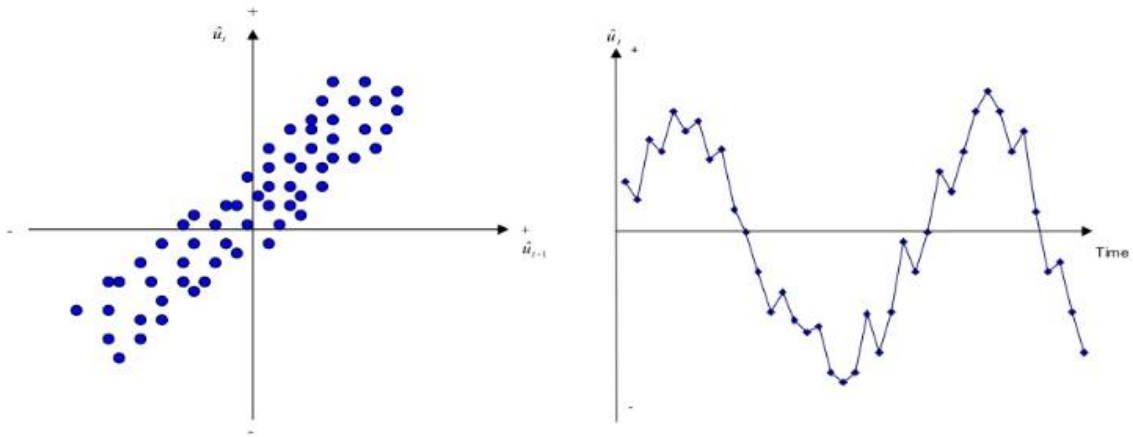
Denne testen følger også en kjikvadratfordeling med M frihetsgrader, hvor k utgjør antall lags og R er den estimerte korrelasjonskoeffisienten. Som vanlig denoteres antall observasjoner med n . Nullhypotesen er at residualene ikke inneholder autokorrelasjon, og at prosessen er som hvit støy å regne (Ljung & Box, 1978). Vi ønsker dermed ikke å forkaste nullhypotesen. Burns (2002) påpeker at testens validitet er avhengig av å velge riktig antall lags, og at det er viktig å holde antall lags relativt få i forhold til utvalgsstørrelse.

5.5.2 Grafisk test for autokorrelasjon

Autokorrelasjon kan også påvises grafisk. Dette gjøres ved å skape et plott av residualene over tid, og et systematisk mønster i dette plottet er en indikasjon på autokorrelasjon i modellen. Positiv autokorrelasjon kjennetegnes ved et syklisk mønster i residualet over tid. Residual med positiv autokorrelasjon vil typisk etterfølges av et restledd med samme fortegn i neste periode. Dette betyr at et residual med positivt fortegn i periode $t-1$ typisk vil etterfølges av et residual med positivt fortegn i periode t . Tilsvarende gjelder for residualer med negativt fortegn, ved positiv autokorrelasjon.

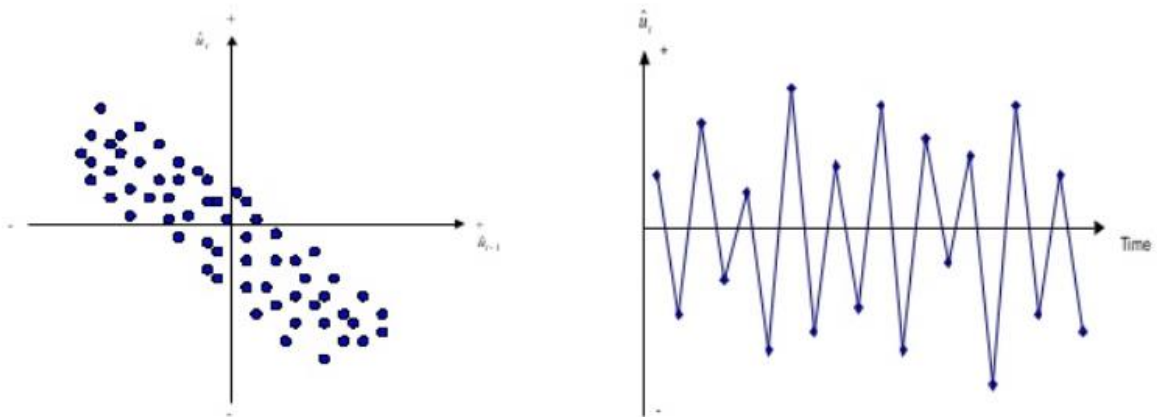
Negativ autokorrelasjon kjennetegnes ved at residualet på tidspunkt t typisk vil ha motsatt fortegn som i forrige periode $t-1$. Dersom residualene har en tilfeldig spredning og et usystematisk mønster antyder dette at det ikke foreligger autokorrelasjon. Ingen autokorrelasjon innebærer at residualet i en tidsserie på tidspunkt t er uavhengige av residualet på tidspunkt $t-1$.

Nedenfor illustreres ulike plott med ulike typer autokorrelasjon, samt et plott med ingen autokorrelasjon. Residualene plottes mot sine laggede verdier i figurene til venstre, og mot tid i figurene til høyre. Det er utfordrende å komme med presise konklusjoner angående autokorrelasjon ved grafisk tolkning alene, og dermed burde grafisk analyse suppleres med statistiske tester som beskrevet i delkapittel 5.5.1.



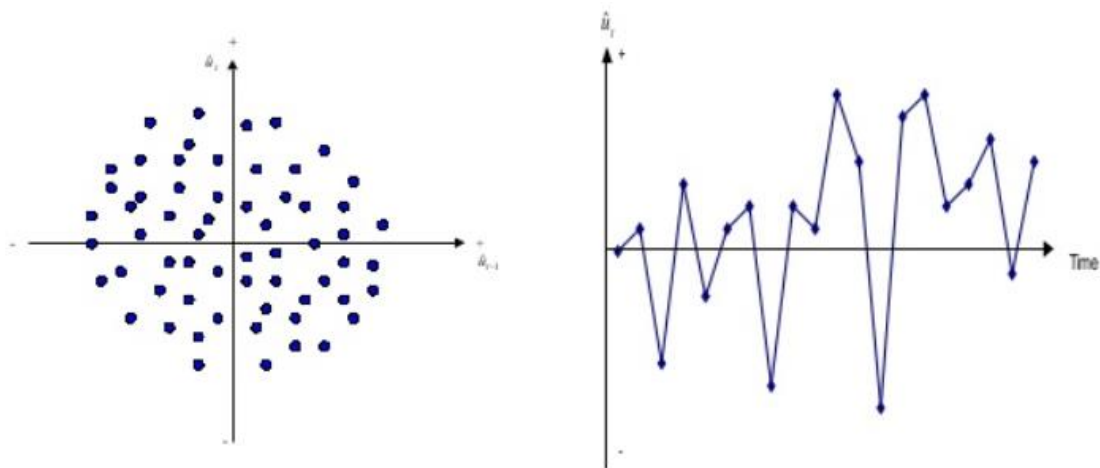
Figur 5.1: Positiv autokorrelasjon

Kilde: Brooks (2015)



Figur 5.2: Negativ autokorrelasjon

Kilde: Brooks (2015)



Figur 5.3: Ingen autokorrelasjon

Kilde: Brooks (2015)

5.6 Heteroskedastisitet

Heteroskedastisitet er et begrep som forklarer modellens residual e_t og dens varians.

Heteroskedastisitet i residualene betyr at variansen ikke er konstant og har ulik spredning over tid. Homoskedastisitet er motsetningen til dette begrepet, og impliserer konstant varians i residualene over tid. Homoskedastisk varians i residualene er en av forutsetningene til Minste Kvadraters Metode, betinget på den avhengige variabelen. Hvis ikke denne forutsetningen holder, vil heteroskedastisitet i residualene kunne medføre unøyaktige standardavvik og feilaktige empiriske resultater. Det vil derfor være av viktighet å teste modellen for heteroskedastisitet for å forsikre oss om at estimatene inneholder minimal spredning, for at konklusjonene skal basere seg på mest mulig presise resultater.

5.6.1 Statistisk Test for Heteroskedastisitet

Før en tester for heteroskedastisitet, må en først ha testet for autokorrelasjon i residualene, ettersom autokorrelasjon vil generelt ugyldiggjøre en test for heteroskedastisitet. Når autokorrelasjon har blitt testet for, tar vi utgangspunkt i følgende likning som ligger til grunn for Breusch-Pagan (BP) testen for heteroskedastisitet (Wooldridge, 2012, s. 435):

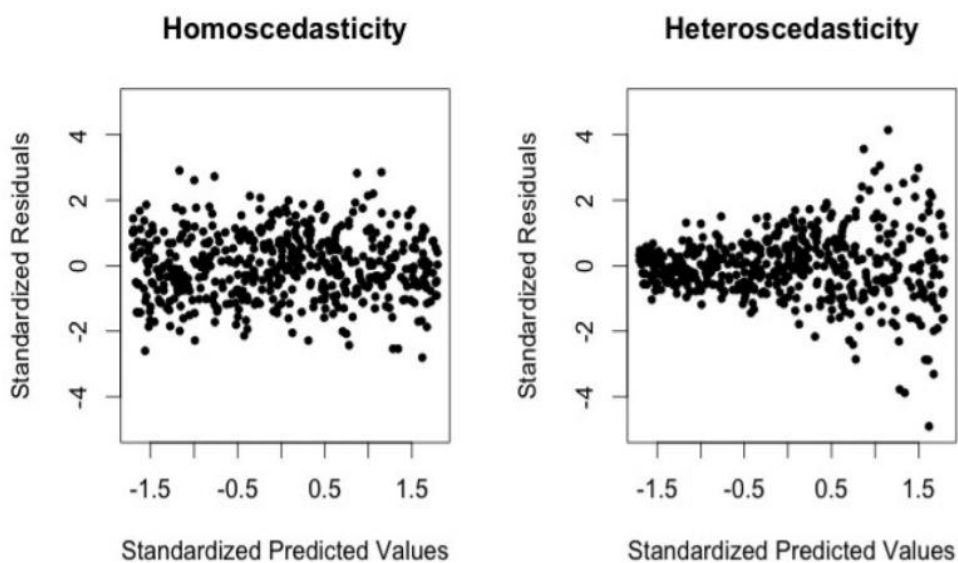
$$u_t^2 = \delta_0 + \delta_1 x_{t1} + \dots + \delta_k x_{tk} + v_t \quad (5.11)$$

Nullhypotesen er at residualene har lik varians, og dermed homoskedastisitet. Den alternative hypotesen er at residualene har ulik varians, og dermed heteroskedastisitet. Testen følger en

kjikkvadratfordeling, og nullhypotesen om homoskedastisitet forkastes på et signifikansnivå på 5%. Dersom heteroskedastisitet blir påvist i residualene, vil dette korrigeres for med White's heteroskedastisk-robuste standardfeil (White, 1980).

5.6.2 Grafisk test for Heteroskedastisitet

Homo- og heteroskedastisitet vil også være mulig å observere grafisk. Dersom man plottet residualene opp mot predikerte verdier fra modellen, vil man kunne identifisere om variansen er konstant eller ei over tid. Heteroskedastisitet kjennetegnes særlig av at spredningen mellom datapunktene øker over tid. Et slikt plott vil synliggjøre en indikasjon på at heteroskedastisitet eksisterer i residualet. Et eksempel på hvordan homo- og heteroskedastisitet typisk ser ut visualiseres nedenfor.



Figur 5.4: Plot av homoskedastisitet og heteroskedastisitet

Kilde: Sage (2015, s.6)

Homoskedastisitet vil typisk ha lik vertikal spredning på tvers av alle residualene fra venstre til høyre. Heteroskedastisitet på sin side vil ha ulik vertikal spredning på tvers av residualene, hvor spredningen typisk øker eller synker fra venstre til høyre. Selv om det er mulig å identifisere variansen i residualet ved grafisk analyse, er dette likevel ikke tilstrekkelig for å kunne trekke en sikker konklusjon. Derfor burde grafisk analyse suppleres med statistiske tester, som beskrevet i forrige delkapittel.

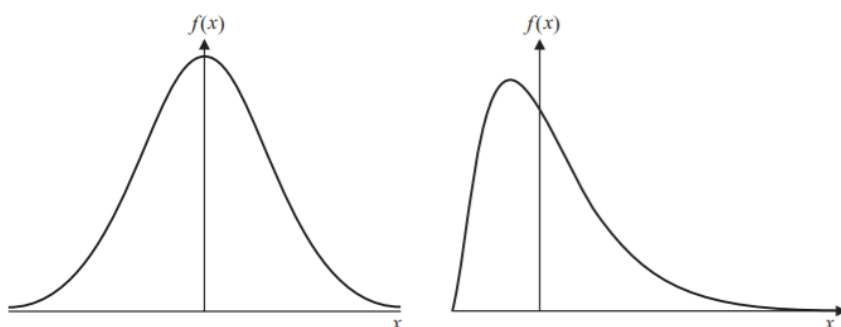
5.7 Normalitet

Når det skal avgjøres om en forklaringsvariabel har en signifikant påvirkning i modellen forutsettes det at residualene følger en normal sannsynlighetsfordeling, som beskrevet i kapittel 5.3.1. Normalfordelte residualer har en gjennomsnittlig forventet verdi lik null, med varians σ^2 . Skjevhet og kurtose er to andre mål på fordelingen til residualene. Skjevhet måler graden av hvor usymmetrisk en fordeling er rundt dens gjennomsnittlige verdi, mens kurtose måler hvor tykke haler fordelingen innehar. En normalfordeling er ikke preget av skjevhet, og defineres å ha en kurtose lik 3. En normalfordeling vil være symmetrisk rundt gjennomsnittet, mens en skjev fordeling vil ha en hale lenger enn den andre (Brooks, 2008, s. 161).

Store avvik fra normalfordeling kan påvirke koeffisientestimatene og dermed svekke troverdigheten til signifikanstesten. Slike avvik kan blant annet være forårsaket av feil i tallmaterialet, eller at modellen har utelatt relevante forklaringsvariabler. Dette er oftest ikke et problem når datasettet inneholder en stor mengde observasjoner, dog et større problem når datasettet er lite.

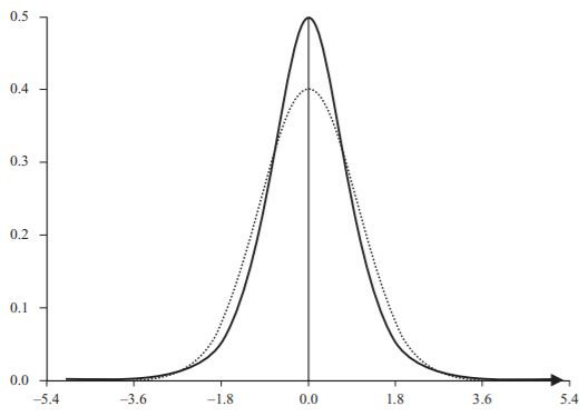
5.7.1 Grafisk test for normalfordeling

Ved å plote residualene i et histogram kan man observere om de følger normalfordelingen eller ikke.



Figur 5.5: Normalfordeling vs. skjevfordeling

Kilde: Brooks (2008, s. 162)



Figur 5.6: Normalfordeling vs. leptokurtisk fordeling

Kilde: Brooks (2008, s. 162)

En leptokurtisk fordeling har tykkere haler og er mer toppet rundt gjennomsnittet enn en normalfordeling. En platykurtisk fordeling er mindre toppet rundt gjennomsnittet, har tynnere haler og innehar mer av fordelingen rundt skuldrene. En normalfordeling defineres å ha en mesokurtisk fordeling (Brooks, 2008, s. 161-162).

5.7.2 Statistisk test for normalfordeling

For å teste om residualene følger en normalfordeling, testes det om koeffisientene til skjevhet og overflødig kurtose er begge lik null. Bera-Jarque testen, formalisert av Bera og Jarque(1981), tar for seg dette. Teststatistikken gis ved:

$$W = T \left[\frac{b_1^2}{6} + \frac{(b_2-3)^2}{24} \right] \quad (5.12)$$

hvor

T= Utvalgets størrelse

b_1 = Skjevhet

b_2 = Kurtose

Testen er gitt av følgende hypoteser:

H_0 : Dataene er normalfordelt

H_1 : Dataene er ikke normalfordelt

Teststatistikken følger en asymptotisk kjikvadratfordeling, og nullhypotesen om normalfordelte data forkastes dersom residualene fra modellen er signifikant skjeve eller signifikant leptokurtisk/platykurtisk.

5.8 Multikollinearitet

En implisitt forutsetning til MKM estimatene er at forklaringsvariablene ikke er korrelert med hverandre. Multikollinearitet eksisterer når to eller flere forklaringsvariabler korrelerer med hverandre. I en situasjon hvor multikollinearitet ikke foreligger, vil man kunne inkludere eller ekskludere en forklaringsvariabel fra regresjonen uten påvirkning på de andre variablene. Det er normalt at variablene i en regresjon korrelerer svakt, men et problem oppstår hvis de korrelerer veldig sterkt.

Det skilles mellom to typer multikollinearitet; perfekt og nær multikollinearitet.

Perfekt multikollinearitet oppstår når det er et eksakt forhold mellom to eller flere variabler. Hvis dette er tilfellet, er det umulig å estimere alle koeffisientene i modellen.

Nær multikollinearitet er mye mer sannsynlig til å oppstå i praksis, og oppstår når det er et ikke-neglisjerbart/ubetydelig, men heller ikke perfekt, forhold mellom to eller flere forklaringsvariabler. Høy korrelasjon mellom forklart variabel og en forklaringsvariabel defineres dog ikke som multikollinearitet (Brooks, 2008, s.171).

5.8.1 Test for multikollinearitet

Nær multikollinearitet vil føre til høy forklaringskraft R^2 . Dette gjør at regresjonen ser bra ut på overflaten, men nær multikollinearitet vil føre til at individuelle koeffisienter ikke er signifikante. Dette fordi det ved nær multikollinearitet er vanskelig å observere de individuelle bidrag fra hver variabel, ettersom de er veldig nærliggende og samvarierer. Regresjonen blir også veldig sensitiv til små endringer, slik at inklusjon og eksklusjon av forklaringsvariabler medfører store endringer i verdiene på koeffisientene eller signifikansen på de andre variablene. Nær multikollinearitet vil medføre store konfidensintervaller for parametrene, og signifikanstester kan dermed gi misledende konklusjoner (Brooks, 2008, s. 172).

En enkel metode å teste for multikollinearitet er ved å sette opp en korrelasjonsmatrise mellom de individuelle forklaringsvariablene. Korrelasjonen mellom disse variablene indikerer graden av multikollinearitet, hvor henholdsvis 1 og -1 indikerer perfekt

multikollinearitet, og regresjonen vil være helt ugyldig. En regel for dette tilsier at verdier over 0,6 kan skade tolkningen av regresjonen. Dersom korrelasjonen mellom to variabler må kunne antas å ses som naturlig vil de ikke måtte forkastes (Brooks, 2008, s.172).

Visuelt kan multikollinearitet oppdages ved å sette opp et plot mellom forklaringsvariabler. I et plot hvor to forklaringsvariabler er høyt korrelerte vil de følge enten den eksakt samme linjen eller befinne seg nær linjen, ved perfekt korrelasjon og nær korrelasjon respektivt. Jo nærmere variablene ligger i plottet, jo sterkere vil korrelasjonen mellom de være (Brooks, 2008, s.171).

5.9 Kointegrasjon & Feiljusteringsmodell

5.9.1 Kointegrasjon

Dersom en teststatistikk viser et signifikant forhold mellom to variabler som egentlig ikke har en direkte kausal sammenheng, kan dette føre til spuriøs regresjon og villedende resultater. Vi ønsker å inkludere variabler som er $I(0)$, altså stasjonære. Generelt sett vil en lineær kombinasjon av to $I(1)$ prosesser resultere i en $I(1)$ prosess. Men, en lineær kombinasjon av variabler som er $I(1)$ kan bli $I(0)$ og dermed stasjonære, dersom variablene er kointegrerte (Brooks, 2008, s. 336). Vi ønsker å undersøke om variablene er kointegrerte for å avdekke om det eksisterer en langsiktig sammenheng mellom variablene. Kointegrasjon vil gjøre regresjoner som involverer ikke-stasjonære $I(1)$ variabler potensielt meningsfulle, og vi unngår problemer med spuriøs regresjon (Wooldridge, 2012, s. 646).

Wooldridge (2012) forklarer kointegrasjon som to eller flere variabler av samme integrasjonsorden som følger en felles langsiktig utvikling. To $I(1)$ prosesser er kointegrerte dersom det finnes en lineær kombinasjon av de to tidsseriene som er $I(0)$.

Wooldridge (2012) definerer kointegrasjon mer formelt:

Dersom $\{y_t: t = 0, 1, \dots\}$ og $\{x_t: t = 0, 1, \dots\}$ er to $I(1)$ prosesser vil som regel $y_t - \beta x_t$ være en $I(1)$ prosess uavhengig av verdien på β . Det finnes derimot en mulighet i noen tilfeller hvor $\beta \neq 0$ for at $y_t - \beta x_t$ blir en $I(0)$ prosess. Hvis en slik β eksisterer, sier vi at y og x er kointegrerte og β er kointegrasjonsparameteren. $I(0)$ prosessen vil dermed ha konstant gjennomsnitt og varians, autokorrelasjoner i feilledet som kun avhenger av tid mellom variablene i serien, og være asymptotisk ukorrelert (Wooldridge, 2012, s. 646). Disse

egenskapene er nødvendig for at regresjonen skal gi forventningsrette og effisiente estimatorer. Kointegrerte variabler gjør dermed at vi kan være sikre på at korrelasjonen vi estimerer ikke er tilfeldig.

Kointegrasjon kan testes for med statistiske tester. I tidsserier kan vi teste for kointegrasjon med en såkalt “Engle-Granger test” etter Engle og Granger (1987). Kointegrasjon kan også testes for med den utvidede Dickey-Fuller testen som tester residualet for enhetsrot.

Nullhypotesen er her at tidsseriene y og x ikke er kointegrerte, og under nullhypotesen er regresjonen dermed spuriøs (Wooldridge, 2012, s. 647).

5.9.2 Feiljusteringsmodell

Dersom variablene våre er kointegrerte kan vi benytte oss av en såkalt feiljusteringsmodell. En feiljusteringsmodell tillater oss å studere kortsiktige dynamikker i forholdet mellom kointegrerte avhengig variabel y og uavhengig variabel x (Wooldridge, 2012, s. 651). En feiljusteringsmodell undersøker i tillegg hvordan den avhengige variabelen justeres tilbake mot langsiktig likevekt (Brooks, 2014, s. 37). Feiljusteringsmodellen gis ved:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \mu_t \quad (5.13)$$

Brooks (2014) definerer langsiktig ved at verdiene konvergerer mot langsiktige verdier og at de ikke lenger endres. En feiljusteringsmodell vil løse eventuell spuriøsitet ved inkluderingen av to integrerte variabler av første orden, dersom variablene er kointegrerte. En slik modell er relevant for vår analyse av prisutviklingen i boligmarkedet, da vi vil kunne få informasjon om både kortsiktig dynamikk og en eventuell langsiktig likevektstilpasning, som diskutert i kapittel 2.

La oss ta utgangspunkt i at y_t og x_t er $I(1)$ prosesser og ikke kointegrerte. Vi har den enkle dynamiske modellen i førstedifferanser (Wooldridge, 2012, s. 651):

$$\Delta y_t = a_0 + a_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \varepsilon_t \quad (5.14)$$

Hvis y og x er kointegrerte med parameter β har vi ytterligere en $I(0)$ variabel som kan inkluderes i likningen ovenfor. Vi estimerer et feilledd mellom de kointegrerte variablene:

$$s_t = y_t - \beta x_t \quad (5.15)$$

Det estimerte feilleddet s_t fra de kointegrerte variablene er stasjonært, altså $I(0)$. Vi kan inkludere lagger av dette estimatet i feiljusteringsmodellen, og følgelig utlede den simple feiljusteringsmodellen:

$$\Delta y_t = \alpha_0 + \alpha_1 \Delta y_{t-1} + \gamma_0 \Delta x_t + \gamma_1 \Delta x_{t-1} + \delta (y_{t-1} - \beta x_{t-1}) + \mu_t \quad (5.16)$$

Her blir $\delta (y_{t-1} - \beta x_{t-1})$ kalt feiljusteringsleddet. Parameteren δ bestemmer hastigheten y justeres tilbake mot langsiktig likevekt etter avvik bestemt av x . β definerer det langsiktige forholdet mellom endringer i x og y (Brooks, 2008, s. 339). Denne metoden anvendes når β er kjent. Hvis β derimot ikke er kjent må den estimeres. Dette gjøres ved å erstatte s_t med $\hat{s}_t = y_t - \hat{\beta} x_t$, hvor $\hat{\beta}$ kan være ulike estimatorer av β . Dette kalles Engle-Grangers to-stegs prosedyre (Wooldridge, 2012, s. 652).

Dersom feiljusteringsleddet er korrekt spesifisert, vil justeringsparameteren ligge mellom en og null. Feiljusteringsleddet vil i dette tilfellet inneha en negativ verdi. Jo nærmere null dette leddet ligger, jo tregere vil tilpasningen tilbake etter et avvik fra likevekt være.

5.10 Paneldata

I tillegg til å utføre regresjonsanalyser ved hjelp av minste kvadraters metode ønsker vi også å utføre regresjonsanalyse ved hjelp av paneldata. Målet er da å undersøke regionale forskjeller ved hjelp av paneldata. Metodene knyttet til paneldata og fordelene ved dette forklares intuitivt av Brooks (2014). Paneldata, også kalt longitudinelle data, vil si data med en mengde observasjoner som har blitt målt over tid og som måles over en rekke individer. Disse individene kan for eksempel være firma, husholdninger eller land. I vårt tilfelle gjelder dette byene Oslo, Stavanger, Kristiansand, Bergen og Tromsø i tidsperioden første kvartal 2005 til fjerde kvartal 2018. I et paneldatasett er det viktig at man beholder de samme individene samtidig som man måler noen mengder ved dem over tid. Mengdene vi måler i dette tilfellet blir variablene som vi går igjennom i kapittel 4. I et paneldatasett kombinerer vi dermed tverrsnittsdata med tidsserier.

Fordelene med paneldata kan være mange, men Brooks (2014) peker spesielt på tre ting. Først og fremst kan vi adressere et bredere spekter av problemer i tillegg til å kunne håndtere mer komplekse problemer enn vi kunne gjort dersom vi hadde hatt rene tverrsnittsdata eller tidsserier. For det andre er det ofte interessant å se hvordan variabler, eller forholdet mellom disse, varierer over tid. Dersom man bruker tidsserier alene, må man ha store mengder data for å ha nok observasjoner til å kunne gjøre meningsfulle hypotesetester. Når vi kombinerer tverrsnittsdata og tidsserier derimot, kan vi øke gradene av frihet og dermed også øke testkraften. Den ekstra variansen vi får ved å kombinere dataene på denne måten kan også være med å hindre problemer med multikollinearitet som kan oppstå dersom vi modellerer tidsseriene individuelt. For det tredje, ved å strukturere modellen på denne måten, kan vi fjerne effekten av visse former for utelatte variablers skjevhet i regresjonsresultater.

Vi står ovenfor flere estimeringsmetoder å velge blant når vi skal estimere en paneldatamodell. En enkel estimeringsmetode vi kan anvende er en såkalt pooled OLS modell. I en slik modell samles all data og regresjonen utføres som om det er en tverrsnittsanalyse. Denne estimeringsmetoden har dog sine begrensninger. Brooks (2014) poengterer blant annet at å samle all data sammen på denne måten implisitt forutsetter at gjennomsnittsverdiene til variablene og forholdene mellom dem er konstante over tid og på tvers av alle tverrsnittsenerhetene. Alternativer til pooled OLS modell er å estimere regresjonskoeffisientene ved å bruke fast effekt (FE) eller tilfeldig effekt (RE) estimatorer. På grunnlag av at vi måler en lang tidsperiode og få individer (byer), vil forskjellene mellom disse være små, og verken fast eller tilfeldig effekt estimatorer passer i dette tilfellet til vårt formål. Dermed vil vi ikke gå i detalj på disse estimatorene, og vi refererer derfor til Brooks (2014) for en grundigere gjennomgang.

Alternativt til fast og tilfeldig effekt er at vi kan bruke en generalisert minste kvadraters metode regresjon (FGLS) som passer til paneldata. FGLS beholder panelstrukturen med grupper og tid, og gir oss i tillegg en sammenslått regresjon. Regresjonskoeffisientene i modellen estimeres med vektning basert på minste kvadraters metode, og kontrollerer for eventuell autokorrelasjon i feilledet.

Vi vil benytte oss av en FGLS-estimert paneldatamodell i vår oppgave. Denne modellstrukturen er velegnet til vårt bruk, da vi ønsker å beholde panelstrukturen, og ta hensyn til en eventuell autokorrelasjon i modellen. I tillegg lar denne modellstrukturen seg

lett kombinere med feiljusteringsmodellene vi har laget for de ulike byene vi studerer. Denne modellen gjør også at vi kan inkludere interaksjonsvariabler for å undersøke regionale forskjeller eller eventuelle likhetstrekk. Vi viser til Cameron og Trivedi (2005) for en grundigere gjennomgang av dette.

En interaksjonsvariabel innebærer lineære kombinasjoner av variabler, for eksempel kan vi konstruere interaksjonen mellom en bydummy for Stavanger og data for realinntekt. Vi kan også konstruere interaksjonsleddet mellom bydummyen for Tromsø og data for realrente. Vi får dermed en interaksjonseffekt, og denne effekten oppstår når effekten av en variabel er avhengig av verdiene på den andre variabelen som inkluderes i interaksjonen.

Ved hjelp av slike interaksjonsledd mellom variabler, får vi interaksjonsvariabler som i regresjonsmodellen vil kunne bidra med informasjon om eksempelvis realinntekt eller realrente har en statistisk signifikant forskjellig påvirkning i byen vi måler interaksjonen for, i forhold til de resterende byene vi studerer. Vi studerer altså koeffisientene til interaksjonsvariablene for å påvise en mulig forskjell i helning. Inkludering av interaksjonsledd i regresjonsmodellen vår kan bidra til at vi får en enda bedre forståelse av forholdene mellom ulike variabler.

Vi vil dermed konstruere interaksjoner for alle byene utenom én by vi velger som base. Eksempelvis vil vi lage interaksjoner for Stavanger, Kristiansand, Bergen og Tromsø, og bruke Oslo som basekategori. Hvis vi konstruerer det på en slik måte, vil interaksjonsvariablene fortelle oss om hvorvidt helningen til forklaringsvariabler er ulik helningen til forklaringsvariabler i basekategorien Oslo. Ender vi opp med statistisk signifikante resultater for interaksjonsvariablene, betyr dette at helningen til koeffisientene er signifikant forskjellige fra basekategorien, og det vil implisere at det finnes regionale ulikheter i hvordan boligprisene utvikler seg i de ulike byene basert på den variabelen vi måler. Hvis interaksjonsvariablene viser seg å ikke være statistisk signifikant forskjellige fra basekategorien, vil det implisere at boligprisene utvikler seg likt i byene vi undersøker basert på den variabelen vi måler.

6. Økonometriske tester og analyse

I dette kapitlet utfører vi et utvalg tester som vi redegjorde for i forrige kapittel med den hensikt å undersøke om modellene vil gi forventningsrette og presise resultater basert på forutsetningene og betingelsene vi presenterte. I det etterfølgende estimerer vi tidsseriemodellene for de ulike byene. Vi samler til slutt alle tidsseriemodellene i et paneldatasett, og estimerer denne modellen både med og uten interaksjonsvariabler.

6.1 Testing for Stasjonæritet

I delkapittel 5.4.1 utledet vi den utvidede Dickey-Fuller testen som tester variablene for stasjonæritet. I dette delkapitlet anvender vi Dickey-Fuller testen for å undersøke hvilken integrasjonsorden ikke-stasjonære variabler har, eller om de viser seg å være stasjonære, som er målet. For å finne riktig antall lags som skal benyttes i den utvidede Dickey-Fuller testen bruker vi “Schwarz-Bayesian Information Criterion”(SBIC). Dette er et informasjonskriterium som i STATA vil gi oss det riktige antallet lags å anvende. For en grundigere gjennomgang av dette viser vi til Schwarz (1978).

Vi gjør dette med den hensikt å unngå at variabler som egentlig ikke er stasjonære, består Dickey-Fuller testen for stasjonæritet. Realboligpris, realinntekt og boligbygging er variabler som er konvertert til logaritmisk form, slik som i boligprismodellen vi estimerer.

Tabell 6.1: Testresultater for utvidet Dickey-Fuller test for stasjonæritet, etter by

Utvidet Dickey-Fuller test for stasjonæritet. Signifikans (markert med *) indikerer stasjonær variabel.					
Variabel	Oslo	Stavanger	Kristiansand	Bergen	Tromsø
ln(Realboligpris)	-0.881	-3.157**	-3.182**	-2.593*	-0.913
Antall lags	1	1	1	1	1
ln(Realinntekt)	-1.445	-7.572***	-1.530	-1.164	-0.677
Antall lags	1	0	1	1	4
Realrente	-1.906	-1.906	-1.906	-1.906	-1.906
Antall lags	2	2	2	2	2

Forventning	-4.077***	-4.573***	-3.408**	-4.273***	-3.843***
Antall lags	2	2	2	2	2
Arbeidsledighet	-2.878**	-1.730	-2.612*	-2.676*	-2.501
Antall lags	3	1	3	3	4
ln(Boligbygging)	-3.345**	-6.462***	-5.383***	-3.479***	-6.899***
Antall lags	1	0	0	1	0
Nettoinnflytting	-3.234**	-2.091	-2.749*	-2.399	-3.233**
Antall lags	4	2	4	4	4
Signifikansmerking:	* p < 0,1 ** p<0,05 *** p < 0,01				

Ut ifra testresultatene ser vi at boligbygging og forventning er stasjonære på 1 eller 5% signifikansnivå for samtlige byer. Testresultatene er dog varierte, og de resterende variablene realboligpris, realinntekt, realrente, arbeidsledighet og nettoinnflytting er etter testen å bedømme stasjonære for enkelte byer, mens de for andre byer ikke er stasjonære.

For Oslo blir realboligpris, realinntekt og realrente stasjonære på 1% signifikansnivå etter førstedifferensiering. Dette betyr at disse variablene er integrert av første orden, og følger en I(1) prosess. Dette vil også gjelde for de andre variablene som i det følgende avsnitt blir stasjonære etter førstedifferensiering.

For Stavanger blir realrente, arbeidsledighet og nettoinnflytting stasjonære prosesser på 1% signifikansnivå etter førstedifferensiering. For Kristiansand blir realinntekt og realrente stasjonære på 1% signifikansnivå etter førstedifferensiering. For Bergen blir realinntekt og realrente stasjonære på 1% signifikansnivå etter førstedifferensiering. For Tromsø finner vi at Realboligpris, realinntekt, realrente blir stasjonære på 1% signifikansnivå etter førstedifferensiering, mens arbeidsledighet blir stasjonær på 5% nivå etter førstedifferensiering.

6.2 Feiljusteringsmodell & test for Autokorrelasjon

Vi undersøker som sagt boligprisutviklingen i fem av Norges byer, og følgelig vil vi estimere én modell for hver by med regionale boligpriser og regionale forklaringsvariabler, med unntak av rente som fungerer som en nasjonal forklaringsvariabel for alle modellene. Ved å

lage totalt fem modeller får vi muligheten til å sammenligne forklaringskraft på tvers av modellene. Vi ønsker også å bevise at variablene våre er kointegrerte, slik at vi forsikrer oss om at modellene fanger opp den langsiktige sammenhengen mellom variablene på en forventningsrett måte. For å påvise kointegrerte variabler vil vi utføre en utvidet Dickey-Fuller test der vi tester residualet til hver modell for enhetsrot. Nullhypotesen når vi tester for kointegrasjon gjennom den utvidede Dickey-Fuller testen er at det finnes enhetsrot i de potensielt kointegrerte residualene (Brooks, 2014, s. 377). Signifikante verdier og følgelig at nullhypotesen forkastes, indikerer dermed kointegrerte variabler i den respektive modellen.

Vi etablerte i forrige delkapittel at variablene vi anvender enten er stasjonære eller integrert av første orden. Etter at dette er etablert kan vi konstruere en feiljusteringsmodell slik vi har beskrevet i delkapittel 5.9. Fra likning 5.16 konstruerer vi feiljusteringsmodellen ved å utføre en regresjon på differensiert boligpris på differensierte forklaringsvariabler, samt laggede avhengig og uavhengige variabler.

Vi velger derimot ikke å inkludere differensiert boligbygging i modellen vår, da boligbygging på kort sikt i litteraturen antas å ikke påvirke etterspørsel, og følgelig boligpris. Vi refererer til kapittel 2.1 for nærmere forklaring av boligtilbudet på kort sikt.

Feiljusteringsmodellen vi estimerer for boligpris for hver enkelt by tilsvarer metodeteorien om feiljusteringsmodeller i kapittel 5.9.2 og blir seende ut som følger:

$$\begin{aligned} \Delta \ln(\text{boligpris})_{i,t} = & a_0 + b_1 \Delta \ln(\text{inntekt})_{i,t} + b_2 \Delta \text{rente}_t + b_3 \Delta \text{arbeidsledighet}_{i,t} + \\ & b_4 \Delta \text{forventning}_{i,t} + b_5 \Delta \text{Nettoinnflytting}_{i,t} + c_1 \ln(\text{boligpris})_{i,t-1} \\ & + d_1 \ln(\text{inntekt})_{i,t-1} + d_2 \text{rente}_{t-1} + d_3 \text{arbeidsledighet}_{i,t-1} + d_4 \text{forventning}_{i,t-1} \\ & + d_5 \text{Nettoinnflytting}_{i,t-1} + d_6 \ln(\text{boligbygging})_{i,t-1} \\ & y_1 \text{Sesong}_1 + y_2 \text{Sesong}_2 + y_3 \text{Sesong}_3 + y_4 \text{Finanskrise} + y_5 \text{Oljekrise} + \varepsilon_{i,t} \end{aligned} \quad (6.1)$$

hvor Δ = *førstedifferanse*, \ln = *logaritme* og $\varepsilon_{i,t}$ = *residual til by i for periode t*.

Her er a_0 konstantleddet til modellen, b_i er koeffisientestimatene som reflekterer den kortsiktige dynamikken, c_1 er feiljusteringsparameteren til den avhengige variabelen, og d_i er koeffisientestimatene som reflekterer den langsiktige sammenhengen. Koeffisientestimatene

γ_i reflekterer dummyvariabler som har som hensikt å kontrollere for sesongvariasjoner i modellen vår. Vi har 3 dummyvariabler som hver for seg er lik 1 i det respektive kvartalet de kontrollerer for, fra første til tredje kvartal. Dummyvariabel for fjerde kvartal blir utelatt grunnet kollinearitet, og anvendes som basekategori. Finanskrisen er en dummy som er lik 1 i perioden 2. kvartal 2008- 4.kvartal 2009. Denne inkluderes i et forsøk på å kontrollere for effektene av finanskrisen. Oljekrisen er en dummy som er lik 1 i perioden 3. kvartal 2014 - 4. kvartal 2016. Denne inkluderes i et forsøk på å kontrollere for effektene av oljekrisen.

Vi estimerer de fem modellene og tar vare på residualene. Testresultatene fra regresjonsanalysene vil bli presentert i et senere delkapittel etter vi har påvist om modellene innfrir betingelsene presentert i kapittel 4, og remodellering og justeringer finnes nødvendig eller ei. I delkapittel 5.5 utledet vi Ljung-Box testen for å teste autokorrelasjon i residualene. Nullhypotesen er at residualene ikke inneholder autokorrelasjon, altså fravær av autokorrelasjon, og vi ønsker dermed ikke å forkaste nullhypotesen. Vi velger å bruke 8 lags i testen, altså 8 kvartaler. Dette gjør vi for å inkludere et tilstrekkelig antall lags som forsikrer validiteten til testresultatene, samtidig som vi også ønsker å ikke inkludere for mange lags relativt til utvalgsstørrelsen på 54.

Tabell 6.2: Testresultater for Ljung-Box test, etter by

Ljung-Box (Q*) test for autokorrelasjon		
Variabel	Q-statistikk	p verdi
Residual Oslo	15.2343	0.0547
Residual Stavanger	7.9042	0.4429
Residual Kristiansand	12.7513	0.1207
Residual Bergen	10.7540	0.2160
Residual Tromsø	11.0648	0.1981

Nullhypotesen i Ljung-Box testen er som beskrevet i delkapittel 5.5.1 at residualet ikke er autokorrelert, og at det er som hvit støy å regne. Ljung-Box testen følger en kjikvadratfordeling, og følgelig er det her vi finner de kritiske verdiene. For 8 frihetsgrader er den kritiske verdien 15,507 på et 5% signifikansnivå. Fra testresultatene ser vi at samtlige byer ikke har autokorrelerte residualer på et 5% nivå, og vi beholder nullhypotesen om hvit støy for samtlige residualer. Resultatet for Oslo ligger dog i gråsonen for hva som er akseptabelt, men vi velger å beholde nullhypotesen på et 5% signifikansnivå. Testresultatene indikerer at vi til dels har problemer med autokorrelerte feilledd i modellene våre, men at de aller fleste koeffisientestimatene fra de estimerte regionale modellene er forventningsrette og effisiente, med andre ord BLUE.

6.3 Test for Heteroskedastisitet

I delkapittel 5.6 utledet vi Breusch-Pagan testen. Vi anvender i dette kapittelet denne testen for å kontrollere at residualene i modellene for hver by innehar en homoskedastisk varians. For vurdering av testens resultater er det de kritiske verdiene for kji-kvadratfordelingen som ligger til grunn vedrørende nullhypotesens forkastelse. Nullhypotesen er at residualene har konstant varians, og følgelig ønsker vi å beholde nullhypotesen.

Tabell 6.3: Testresultater for Breusch-Pagan test, etter by

Breusch-Pagan test for heteroskedastisitet		
Variabel	Kjikvadrat-statistikk	p verdi
Residual Oslo	0.02	0.9001
Residual Stavanger	0.46	0.4997
Residual Kristiansand	0.21	0.6507
Residual Bergen	1.66	0.1972
Residual Tromsø	2.81	0.0937

Testresultatene viser at samtlige byer har en homoskedastisk fordeling av residualene, og vi beholder alle nullhypotesene om homoskedastisk varians i residualene. Testresultatene for Tromsø er i gråsonen for hva som er akseptabelt, med en testverdi på 2.81 og en tilhørende p-verdi på 0,0937 vil den forkastes på et 10% signifikansnivå. Vi beholder den derimot på et 5% signifikansnivå, og konkluderer endelig med at residualene til Tromsø er homoskedastisk fordelt.

Ljung-Box testen for autokorrelasjon vi utførte i 6.2 påviste resultater i gråsonen for hva som er akseptabelt for Oslo. Disse resultatene for autokorrelasjon kan føre til at standardavvikene for Oslo blir til en viss grad misledende, noe som naturligvis også vil gjelde variansen. Dette tatt i betraktning gjør at testresultatene for Oslo kan være noe unøyaktige, og at Oslo kan ha en mer heteroskedastisk varians enn det Breusch-Pagan testen tilsynelatende viser.

6.4 Test for Normalfordeling

I delkapittel 5.7 gjorde vi rede for normalitet og dens viktighet i økonometri. I det følgende vil vi teste residualene til hver modell for å forsikre oss om at modellene innehar normalfordelte data. Dette utføres med en skjevhet-kurtose test. For vurdering av testens resultater er det de kritiske verdiene for kji-kvadrat testen som ligger til grunn vedrørende nullhypotesens forkastelse. Nullhypotesen er at dataene er normalfordelte basert på skjevhet og kurtose, og følgelig er det ønskelig å beholde nullhypotesen.

Tabell 6.4: Testresultater for Skjevhet-Kurtose test, etter by

Skjevhet-Kurtose test for normalfordeling				
Variabel	Pr(Skjevhet)	Pr(Kurtose)	Kjikvadrat-statistikk	p verdi
Residual Oslo	0.4367	0.4667	1.18	0.5543
Residual Stavanger	0.2505	0.1709	3.37	0.1853
Residual Kristiansand	0.7169	0.3229	1.15	0.5619
Residual Bergen	0.9718	0.0576	3.81	0.1489
Residual Tromsø	0.0297	0.7905	4.78	0.0916

Testresultatene viser at samtlige byer følger normalfordelte data. Med andre ord beholdes nullhypotesen om normalitet for samtlige byer. Kurtosen til Bergen ligger derimot i gråsonen for hva som er akseptabelt, men sammenlagt med skjevheten beholder vi nullhypotesen om normalfordelte data for Bergen. Skjevheten til Tromsø er signifikant forskjellig fra skjevheten til en normalfordeling på et 5% signifikansnivå. Sammenlagt innehar Tromsø derimot en teststatistikk på 4.78 og påfølgende p verdi på 0,0916, noe som fører til at vi også beholder nullhypotesen om normalfordelte data for Tromsø. Selv om ikke disse fordelingene er symmetrisk perfekte, så er fordelingen relativt symmetrisk og estimatene til modellen er derfor innenfor en akseptabel sone. Dog er disse testresultatene suboptimale, og bruk av t-test og f-test er derfor til dels gyldig.

6.5 Test for Multikollinearitet

Vi benytter oss av korrelasjonsmatriser for å undersøke om det foreligger forklaringsvariabler i modellen som korrelerer høyt med andre forklaringsvariabler. Hvis vi har korrelasjon mellom forklaringsvariabler med verdier -1 eller 1 har vi perfekt multikollinearitet, noe som vil gjøre regresjonen ugyldig. Nedenfor presenteres korrelasjonsmatriser for samtlige boligprismodeller, som vi i det følgende vil undersøke nærmere. Korrelasjonsmatrisene er hentet fra STATA.

Korrelasjonsmatrise 6.1: Oslo

	dlnRea~s	dlnRea~t	dRealr~e	dForve~g	dArbei~t	lnBoli~g	dNetto~g	lnR~s_L1	lnR~t_L1	Realre~1	Arbeid~1	lnBoli~1	Nettoi~1	Forven~1
dlnRealbol~s	1.0000													
dlnRealinn~t	0.1977	1.0000												
dRealrente	0.3814	0.3262	1.0000											
dForventning	0.1649	-0.1591	-0.1280	1.0000										
dArbeidsle~t	-0.0036	0.1561	0.1955	0.0362	1.0000									
lnBoligbyg~g	-0.1764	-0.1179	-0.2617	-0.0331	-0.1034	1.0000								
dNettoinnf~g	0.0192	0.1458	0.4091	0.0437	0.4233	-0.0000	1.0000							
lnRealboli~1	-0.1493	0.0898	-0.0829	-0.0933	0.0343	0.4211	-0.0131	1.0000						
lnRealinnt~1	-0.0999	-0.2031	-0.1373	-0.0743	0.0326	0.4557	-0.0273	0.7877	1.0000					
Realrente_L1	-0.2998	-0.0504	-0.4478	0.0108	0.0556	-0.2544	-0.1920	-0.4379	-0.4011	1.0000				
Arbeidsled~1	0.1117	-0.1711	-0.0422	0.0398	-0.3156	0.3132	-0.0587	-0.3871	-0.1527	-0.2036	1.0000			
lnBoligbyg~1	0.0423	-0.0801	-0.0138	0.1203	0.0183	0.5877	0.1065	0.3872	0.3997	-0.4757	0.2490	1.0000		
Nettoinnfl~1	0.0009	-0.1751	-0.3204	-0.0199	-0.4978	-0.0624	-0.8673	-0.2457	-0.1413	0.1902	0.2093	-0.0627	1.0000	
Forventnin~1	0.0434	0.2171	0.0860	-0.3654	-0.0355	-0.1909	-0.0202	0.0418	-0.2788	-0.0083	-0.0103	-0.2121	-0.0842	1.0000

Korrelasjonsmatrise 6.2: Stavanger

	dlnRea~s	dlnRea~t	dRealr~e	dForve~g	dArbei~t	lnBoli~g	dNetto~g	lnR~s_L1	lnR~t_L1	Realre~1	Arbeid~1	lnBoli~1	Nettoi~1	Forven~1
dlnRealbol~s	1.0000													
dlnRealinn~t	0.1141	1.0000												
dRealrente	0.1209	0.4527	1.0000											
dForventning	0.0257	-0.0418	-0.0605	1.0000										
dArbeidsle~t	-0.2885	0.1687	0.1489	0.1989	1.0000									
lnBoligbyg~g	0.0490	0.0015	-0.0388	0.0495	-0.2698	1.0000								
dNettoinnf~g	-0.2146	0.0338	0.1176	-0.0559	0.1279	-0.0430	1.0000							
lnRealboli~1	-0.4244	0.2473	-0.0368	-0.1313	0.2194	0.2165	0.1068	1.0000						
lnRealinnt~1	-0.2967	-0.2202	-0.2002	-0.0815	0.1654	0.3515	-0.0018	0.4761	1.0000					
Realrente_L1	-0.1691	-0.0323	-0.4478	0.0110	0.1137	-0.3356	-0.0761	0.1837	-0.2131	1.0000				
Arbeidsled~1	0.0751	-0.2411	-0.0819	0.0152	-0.2513	0.2736	-0.0382	-0.2298	0.3935	-0.5780	1.0000			
lnBoligbyg~1	-0.0503	-0.0388	0.1249	0.1243	0.0828	0.2756	0.0551	0.3227	0.3553	-0.3159	0.1259	1.0000		
Nettoinnfl~1	0.0080	-0.0540	-0.0518	0.0443	-0.0561	0.1575	-0.7893	-0.0419	0.1478	-0.0668	0.0904	0.0976	1.0000	
Forventnin~1	0.1441	0.1543	0.0414	-0.3540	-0.1448	-0.0749	0.0207	0.0363	-0.2360	-0.0079	-0.0077	-0.0278	0.0113	1.0000

Korrelasjonsmatrise 6.3: Kristiansand

	dlnRea~s	dlnRea~t	dRealr~e	dForve~g	dArbei~t	lnBoli~g	dNetto~g	lnR~s_L1	lnR~t_L1	Realre~1	Arbeid~1	lnBoli~1	Nettoi~1	Forven~1
dlnRealbol~s	1.0000													
dlnRealinn~t	0.1141	1.0000												
dRealrente	0.1209	0.4527	1.0000											
dForventning	0.0257	-0.0418	-0.0605	1.0000										
dArbeidsle~t	-0.2885	0.1687	0.1489	0.1989	1.0000									
lnBoligbyg~g	0.0490	0.0015	-0.0388	0.0495	-0.2698	1.0000								
dNettoinnf~g	-0.2146	0.0338	0.1176	-0.0559	0.1279	-0.0430	1.0000							
lnRealboli~1	-0.4244	0.2473	-0.0368	-0.1313	0.2194	0.2165	0.1068	1.0000						
lnRealinnt~1	-0.2967	-0.2202	-0.2002	-0.0815	0.1654	0.3515	-0.0018	0.4761	1.0000					
Realrente_L1	-0.1691	-0.0323	-0.4478	0.0110	0.1137	-0.3356	-0.0761	0.1837	-0.2131	1.0000				
Arbeidsled~1	0.0751	-0.2411	-0.0819	0.0152	-0.2513	0.2736	-0.0382	-0.2298	0.3935	-0.5780	1.0000			
lnBoligbyg~1	-0.0503	-0.0388	0.1249	0.1243	0.0828	0.2756	0.0551	0.3227	0.3553	-0.3159	0.1259	1.0000		
Nettoinnfl~1	0.0080	-0.0540	-0.0518	0.0443	-0.0561	0.1575	-0.7893	-0.0419	0.1478	-0.0668	0.0904	0.0976	1.0000	
Forventnin~1	0.1441	0.1543	0.0414	-0.3540	-0.1448	-0.0749	0.0207	0.0363	-0.2360	-0.0079	-0.0077	-0.0278	0.0113	1.0000

Korrelasjonsmatrise 6.4: Bergen

	dlnRea~s	dlnRea~t	dRealr~e	dForve~g	dArbei~t	lnBoli~g	dNetto~g	lnR~s_L1	lnR~t_L1	Realre~1	Arbeid~1	lnBoli~1	Nettoi~1	Forven~1
dlnRealbol~s	1.0000													
dlnRealinn~t	-0.0134	1.0000												
dRealrente	0.1892	0.5428	1.0000											
dForventning	0.0141	-0.0348	-0.0167	1.0000										
dArbeidsle~t	0.0085	0.3077	0.1483	0.2725	1.0000									
lnBoligbyg~g	0.0620	0.0594	0.0779	0.1015	-0.0611	1.0000								
dNettoinnf~g	-0.1661	0.2075	0.3688	-0.0029	0.3516	0.0070	1.0000							
lnRealboli~1	-0.3436	0.0400	-0.0692	-0.0787	0.0601	-0.2395	0.0429	1.0000						
lnRealinnt~1	-0.1250	-0.2015	-0.1997	-0.0637	0.1966	-0.3062	-0.0437	0.7310	1.0000					
Realrente_L1	-0.1649	-0.1123	-0.4478	0.0083	-0.0565	-0.2526	-0.1433	-0.2797	-0.2824	1.0000				
Arbeidsled~1	0.1807	-0.3040	0.0030	-0.0003	-0.3354	0.4765	-0.0999	-0.0721	0.0780	-0.6732	1.0000			
lnBoligbyg~1	0.1097	-0.2036	0.2454	-0.0349	-0.2862	0.3687	0.0116	-0.2668	-0.2927	-0.2157	0.4929	1.0000		
Nettoinnfl~1	0.1359	-0.1668	-0.3318	-0.0229	-0.3003	-0.0254	-0.8490	-0.2970	-0.1288	0.2820	0.0216	-0.0088	1.0000	
Forventnin~1	0.1268	0.1312	0.0114	-0.3741	-0.2062	-0.0913	0.0147	-0.1569	-0.2173	-0.0063	0.0029	-0.0192	0.1013	1.0000

Korrelasjonsmatrise 6.5: Tromsø

	dlnRea~s	dlnRea~t	dRealr~e	dForve~g	dArbei~t	lnBoli~g	dNetto~g	lnR~s_L1	lnR~t_L1	Realre~1	Arbeid~1	lnBoli~1	Nettoi~1	Forven~1
dlnRealbol~s	1.0000													
dlnRealinn~t	-0.0032	1.0000												
dRealrente	0.0817	0.0990	1.0000											
dForventning	-0.1035	-0.1220	-0.0580	1.0000										
dArbeidsle~t	-0.0448	-0.0583	0.3172	0.1663	1.0000									
lnBoligbyg~g	0.1935	-0.1283	-0.1386	0.1176	-0.0219	1.0000								
dNettoinnf~g	0.0372	0.1407	0.1697	-0.0404	0.0353	-0.2746	1.0000							
lnRealboli~1	-0.1356	-0.0131	-0.0262	-0.0495	-0.0466	0.1492	-0.0206	1.0000						
lnRealinnt~1	-0.0059	-0.3507	-0.1028	-0.0348	0.0257	0.2012	-0.0316	0.7807	1.0000					
Realrente_L1	-0.1431	-0.0270	-0.4478	0.0022	-0.0157	-0.1313	-0.0812	-0.4944	-0.4111	1.0000				
Arbeidsled~1	0.1926	0.0377	-0.1081	0.0424	-0.3110	0.1479	-0.0435	-0.7088	-0.4975	0.1417	1.0000			
lnBoligbyg~1	-0.0211	0.0242	-0.1310	0.0237	-0.2140	0.1246	0.2647	0.1812	0.0963	-0.2518	0.1571	1.0000		
Nettoinnfl~1	0.1067	-0.1241	-0.1145	0.1387	0.0361	0.0163	-0.7387	0.0614	0.0983	-0.0466	-0.0833	-0.3913	1.0000	
Forventnin~1	0.1640	0.0985	0.0286	-0.3666	-0.1355	-0.0782	-0.0377	-0.0563	-0.1043	-0.0023	-0.0036	0.0054	0.1189	1.0000

Felles for alle modellene er at nettoinnflytting og tidsforskyvningen av nettoinnflytting korrelerer høyt. Høye korrelasjonsverdier betegnes som imperfekt multikollinearitet, noe som er tilstedeværende i modellene våre. Korrelasjonen mellom disse forklaringsvariablene er høy nok til at det kan skade tolkningen av regresjonen, gitt korrelasjonsverdier mellom 0,7 til 0,85 i boligprismodellene. Brooks (2008) presiserer at hvis korrelasjonen mellom variabler må kunne anses som naturlig, må de ikke forkastes. Dog anser vi ikke nettoinnflytting og tidsforskyvningen av denne variabelen som naturlig.

Vi velger derfor å forkaste nettoinnflytting samt tidsforskjøvet nettoinnflytting som variabler i modellene våre. Dette vil fjerne de mest alvorlige tilfellene av multikollinearitet i modellene våre, noe som konsekvent fører til mindre misledende konklusjoner.

6.6 Regresjonsresultater

I tabell 6.5 under har vi fordelt forklaringsvariablene i kategorier, henholdsvis kortsiktig sammenheng, feiljusteringsmekanismen, langsiktig sammenheng og dummyvariablene vi har inkludert i modellen. Kategoriene kortsiktig- og langsiktig sammenheng samt feiljusteringsmekanismen tolkes slik det redegjøres for i kapittel 4. Dummyvariablene for sesong tolkes som nivåforskjeller i periodene der de respektive dummyene er lik 1. Dummyvariablene for sesongvariasjon måles som nivåforskjeller opp mot en referansekategori, eller basekategori. Basekategorien for sesongvariasjon er som tidligere nevnt sesong 4, altså det fjerde kvartalet i hvert respektive år. Dummyvariablene for finans- og oljekrise tolkes som nivåforskjeller fra den fullstendige regresjonen i perioden de er lik 1.

Tabell 6.5: Regresjonskoeffisienter fra regresjonsmodell med standardfeil i parentes, etter by

	Variabel	Oslo	Stavanger	Kristiansand	Bergen	Tromsø
kortsiktig sammenheng	$\Delta \ln(\text{Realinntekt})_t$	0.221 (0.2419)	-0.0029 (0.0147)	0.4818 (0.3571)	-0.8248** (0.3447)	0.0451 (0.113)
	$\Delta \text{Realrente}_t$	0.0124 (0.0077)	0.0088 (0.0067)	0.0034 (0.0100)	0.0276** (0.0135)	-0.0016 (0.0067)

	$\Delta\text{Arbeidsledighet}_t$	-0.0275 (0.02985)	-0.008 (0.015)	-0.0125 (0.0338)	-0.016 (0.0467)	-0.0358 (0.0263)
	$\Delta\text{Forventning}_t$	0.0036** (0.00175)	0.0028* (0.0015)	0.00058 (0.0025)	0.00174 (0.002)	0.00048 (0.0016)
Feiljusteringsmekanisme	$\ln(\text{Realboligpris})_{t-1}$	-0.0716 (0.0853)	-0.124*** (0.036)	-0.3262*** (0.108)	-0.2523** (0.1145)	-0.1958* (0.119)
Langsiktig Sammenheng	$\ln(\text{Realinntekt})_{t-1}$	0.1117 (0.187)	-0.0047 (0.0221)	-0.0546 (0.1945)	0.1654 (0.1509)	0.1609 (0.1274)
	Realrente_{t-1}	-0.0089 (0.0079)	0.0015 (0.0088)	-0.0021 (0.0102)	0.0031 (0.0159)	-0.00302 (0.0071)
	$\text{Arbeidsledighet}_{t-1}$	0.0036 (0.0148)	-0.0081 (0.0088)	-0.024 (0.0198)	-0.0042 (0.0368)	-0.0089 (0.0247)
	Forventning_{t-1}	0.0017 (0.0016)	0.004*** (0.0012)	0.0031 (0.0019)	0.00269* (0.0015)	0.00211* (0.0013)
	$\ln(\text{boligbygging})_{t-1}$	0.0012 (0.0153)	-0.0009 (0.0186)	0.0216 (0.0204)	-0.0222 (0.0221)	-0.0037 (0.0078)
Dummy-variabler	Sesong 1	0.048*** (0.0176)	-0.0047 (0.0152)	-0.0059 (0.0212)	0.0588** (0.0224)	0.0422* (0.0169)
	Sesong 2	0.063*** (0.0145)	0.0223 (0.0137)	0.0465** (0.0187)	0.0527*** (0.0162)	0.0193 (0.0156)

	Sesong 3	0.0211 (0.019)	-0.0066 (0.0142)	-0.0195 (0.0234)	-0.0093 (0.2285)	0.0205 (0.0159)
	Finanskrise	0.0143 (0.0233)	-0.0206 (0.0172)	-0.0107 (0.0273)	-0.0428* (0.0234)	-0.0312 (0.0204)
	Oljekrise	0.0087 (0.016)	0.0115 (0.0193)	0.0489* (0.0249)	0.0312 (0.0243)	0.0324* (0.0183)
	Konstantledd	-0.545 (1.557)	1.337*** (0.4517)	3.712** (1.722)	0.8337 (1.413)	0.177 (1.192)
	N	54	54	54	54	54
	R^2	0.5784	0.5065	0.5059	0.5709	0.3851
	Justert R^2	0.4120	0.3116	0.3108	0.4015	0.1423
	DF-statistikk, Residual	-11.713***	-9.854***	-9.837***	-7.442***	-8.600***
	Signifikansmerking: * $p < 0,10$ ** $p < 0,05$ *** $p < 0,01$					

Innledningsvis i tolkingen vil vi påpeke at testobservatoren for de utvidede Dickey-Fuller testene på modellenes residual er statistisk signifikante i nesten samtlige modeller. Dette bekrefter kointegrasjon i modellene vi har konstruert, og følgelig at feiljusteringsparameteren kan anvendes. Det eneste unntaket er Oslo, der vi ikke får et signifikant koeffisientestimat for feiljusteringsparameteren. Dette kan være konsekvensen av at autokorrelasjon ble funnet å eksistere på et 10% signifikansnivå for Oslo. Følgelig ekskluderer vi Oslo i tolkingen av feiljusteringsleddet. Vi tolker feiljusteringsleddet slik det blir forklart i 5.9.2, og følgelig vil koeffisientene som ligger nærmest null justere seg tregest tilbake etter et eventuelt avvik fra

likevekt. Dermed justeres et avvik fra likevekt tregest i Stavanger. Her vil et avvik fra den estimerte langtidssammenhengen justeres inn med 12,4% per kvartal. Et avvik fra den estimerte langtidssammenhengen justeres raskest inn i Kristiansand, hvor et avvik fra likevekt henholdsvis justeres inn med 32,62% per kvartal.

Forklaringskraften til modellene er relativt høy, men med fordel hadde det vært ønskelig at variablene forklarte en enda større andel av variasjonen i boligpriser. Modellene for Oslo og Bergen har en R^2 verdi på 58% og 57% respektivt, og det er disse modellene som har høyest forklaringskraft. Med andre ord forklares ca 57% av variasjonen i boligpris i Oslo og Bergen av forklaringsvariablene inkludert i modellen. Modellene for Stavanger og Kristiansand slår noe dårligere ut enn Oslo og Bergen i henhold til forklaringskraft, med en R^2 verdi på rundt 50% i begge byer. Modellen slår dårligst ut i Tromsø, hvor kun 38% av variasjonen i boligprisene forklares av variablene i modellen.

Dette kan være en indikasjon på at boligprisene i Tromsø i større grad påvirkes av andre forhold enn variablene i modellen sammenlignet med de 4 andre byene. En mulig forklaring på dette er Tromsø sin geografiske beliggenhet som innebærer et helt annet klima enn de andre regionene, i tillegg til forskjeller i demografiske forhold som viker fra de andre byene. Tromsø har blitt preget av sterk befolkningstilvekst i perioden vi undersøker. Fra tabell 01222 i Statistikkbanken til SSB finner vi at befolkningen ved utgangen av fjerde kvartal 2004 til fjerde kvartal 2018 har økt fra 62 541 til 76 649. Dette tilsvarer en økning i befolkning på 22,56%. Denne sterke befolkningsveksten kan være en mulig forklaringsfaktor til boligprisene i Tromsø, noe som ikke blir fanget opp av modellen vår.

Vi har mange insignifikante resultater. Vi må ta i kritikk av oppgaven at vi har et kort datasett, hvor enkelte variabler ble hentet inn med årlig frekvens og gjort om til kvartalsvise tall. I tillegg påviste vi en del feil i testene vi foretok oss tidligere i dette kapittelet, som følgelig får sine konsekvenser her. Regresjonsanalysene utføres under omstendigheter som ikke er optimale for presise resultater, og dette kan være noe av grunnen til at regresjonen påviser få signifikante resultater. I delkapittel 6.5 presenterte vi korrelasjonsmatriser for hver respektive bymodell, og vi finner til dels relativt høye korrelasjonskoeffisienter blant flere ulike variabler i flere modeller. Brooks og Tsolacos (2010) presiserer at en konsekvens av nær multikollinearitet er at når variablene er nært relaterte oppstår det vanskeligheter med å

observere hver variabels individuelle bidrag til modellen, som igjen fører til at individuelle variabler ikke er signifikante i modellen. På bakgrunn av dette vil vi ikke kunne trekke for bastante konklusjoner videre i oppgaven.

Vi får nesten ingen signifikante koeffisientestimat for realinntekt, verken i kortsiktig eller langsiktig sammenheng. Dette er ikke uventet med tanke på egenskapene til inntekt i modellen. Vi har årlige data som vi har transformert til kvartalsvise data, noe som medfører lite til ingen variasjon mellom kvartaler innad i et år. Transformasjonen til realinntekt vil føre til litt variasjon grunnet ulik inflasjon i kvartalene. Det er ikke noe overraskende at vi ender opp med få signifikante resultater grunnet disse egenskapene, men på tross av dette mener vi at det blir feil å utelate data for inntekt i analysen, ettersom det er hjemlet i litteraturen at inntekt har en påvirkning på boligpriser.

Det som er litt overraskende, er at koeffisientestimatet for realinntekt til Bergen i den kortsiktige sammenheng viser seg å ha en signifikant effekt på boligpriser. Dog har koeffisienten feil fortegn, og resultatet for Bergen impliserer at en nivåendring i realinntekt tilsvarende ett prosentpoeng vil redusere boligpriser med 82,5% i den kortsiktige sammenheng. Dette er ikke fornuftig ut ifra litteratur om boligpriser, og følgelig ikke et parameter å legge tiltro i. At Bergen får et slikt estimat er uventet, og en mulig forklaring er at regresjonen har funnet en sammenheng som ikke egentlig eksisterer. En annen mulig forklaring er at vi i delkapittel 6.4 testet for normalfordelte data, og testobservatoren for Bergen lå helt i gråsonen når det gjaldt kurtosen, og en konsekvens av at data ikke følger en normalfordeling kan føre til ugyldig signifikans i parametrene slik vi gjorde rede for i delkapittel 5.3.1.

Ingen av byene får signifikante koeffisientestimat for realrente i den langsiktige sammenheng. I den kortsiktige sammenheng har realrenten en signifikant påvirkning på boligprisene i Bergen. Påvirkningskraften er svak, men positiv. Resultatet for Bergen impliserer at en nivåendring i realrente tilsvarende ett prosentpoeng vil øke boligprisene med 2,76%. Hvis vi justerer for marginal skatterate som forklart i kapittel 2, som i dag er på 28%, vil dette tilsvare en økning i boligpriser på omtrent 2% etter skatt.

Ingen av byene får signifikante koeffisientestimat for arbeidsledighet, verken i kortsiktig eller langsiktig sammenheng. Dette er verken et overraskende eller uventet resultat, ettersom

variabelen for arbeidsledighet fra 2015 til 2018 kun oppgis for november hvert respektive år. Dette medfører liten til ingen variasjon mellom kvartalene i disse årene. I tillegg meldte SSB om et tidsbrudd i serien for år 2018, noe som ytterligere senket kvaliteten på data vi hadde til disposisjon for arbeidsledighet. Likevel kan en se, for nesten alle byer, at koeffisientestimatene for arbeidsledighet både i kortsiktig og langsiktig sammenheng får de negative fortegnene vi forventer ut ifra økonomisk teori. Dette vil si at økt arbeidsledighet har en negativ påvirkning på boligpriser, nemlig at boligprisene synker. Det eneste unntaket er den langsiktige sammenhengen i Oslo. Til tross for begrenset tilgang på data av god kvalitet når det gjelder arbeidsledighet, mener vi det blir feil å utelate data for arbeidsledighet i analysen vår, ettersom arbeidsledighet i litteraturen regnes å være en kjent faktor som påvirker boligpriser. Vi mener fortsatt at arbeidsledighet påvirker boligpriser, til tross for disse motsigende resultatene.

Boligbygging på lang sikt viser seg å være insignifikant for alle regioner. Antall igangsatte boliger sier kun noe om det nye tilbudet på markedet, og ikke allerede eksisterende boligmasse. Dermed fører begrensninger i data til at kun det nye tilbudet inkluderes, og ikke det totale tilbudet. Dette kan være en forklaring på insignifikante koeffisienter i den langsiktige sammenheng for boligbygging.

Husholdningers forventninger til egen og landets økonomi er den variabelen i vår modell som hyppigst viser seg å ha en signifikant påvirkning på boligpriser i flest regioner, både i den kortsiktige og langsiktige sammenhengen. I den kortsiktige sammenhengen har forventning en svak, men statistisk signifikant positiv påvirkning på boligpriser i Oslo og Stavanger. En endring i kortsiktige forventninger til egen og landets økonomi tilsvarende ett prosentpoeng vil i Oslo medføre en økning i boligpris på 0,36%, sammenlignet med 0,28% i Stavanger. I de resterende byene, henholdsvis Kristiansand, Bergen og Tromsø, vil effektene av en endring i forventninger være enda svakere enn i Oslo og Stavanger, og heller ikke signifikante.

I den langsiktige sammenhengen har forventning også her en svak, men statistisk signifikant positiv påvirkning på boligpriser. Forventning er statistisk signifikant i Stavanger, Bergen og Tromsø. En endring i langsiktige forventninger til egen og landets økonomi tilsvarende ett prosentpoeng vil medføre økninger i boligpris på 0,4%, 0,27% og 0,21% i Stavanger, Bergen og Tromsø respektivt. Dette samsvarer med den deskriptive presentasjonen i delkapittel 4.8 av regionale forventninger, hvor forventningene for det meste er positive med unntak av

krisene som inntreffer i løpet av tidsperioden. Det virker som dummyvariablene for finanskrisen og oljekrisen til en viss grad lykkes i å fange opp noe av variasjonen i forventninger i de periodene som preges av kriser og dermed negative forventninger til egen og landets økonomi. For øvrig får dummyvariabelen for finanskrisen signifikant effekt i Bergen, mens dummyvariabelen for oljekrisen får en signifikant effekt i Tromsø.

Avslutningsvis kan vi bruke feiljusteringsleddet til å beregne halveringstid. Halveringstid er hvor lang tid halvparten av et eventuelt avvik fra den estimerte langtidssammenhengen bruker på å justeres inn igjen. Bernhardsen og Røisland (2000) presenterer følgende formel for å beregne halveringstid:

$$\frac{\ln(0,5)}{\ln(1 - \theta_i)} \quad (6.2)$$

der θ_i er koeffisienten foran feiljusteringsleddet i de respektive regionene.

Vi anvender denne formelen og beregner halveringstiden til de respektive byene. Vi finner at Kristiansand innehar den raskeste halveringstiden på ca. 1.8 kvartal. Dette tilsvarer 5,4 måneder, og betyr at det tar i underkant av et halvt år før et avvik fra boligpris i likevekt på 1% har blitt redusert til 0,5%, ceteris paribus. Stavanger innehar den tregeste halveringstiden, der det tar 5.4 kvartaler før et avvik fra boligpris i likevekt har halvert seg. Dette tilsvarer omtrent 16 måneder. Halveringstiden i Bergen og Tromsø er henholdsvis 2,4 og 3.18 kvartaler, som tilsvarer henholdsvis omtrent 7,2 og 9,5 måneder. Følgelig har Kristiansand den raskeste halveringstiden, og dette kan muligens tale for at boligmarkedet i Kristiansand er minst volatil, og følgelig det mest stabile boligmarkedet. Dette underbygges av de deskriptive figurene vi presenterte i kapittel 4, hvor vi ser at spesielt boligpris, inntekt, arbeidsledighet og boligbygging har utviklet seg stabilt i Kristiansand.

6.7 Regresjonsresultater fra paneldatamodell

Den modellstrukturen vi har anvendt så langt egner seg bra til å undersøke boligprisutvikling individuelt i byene, men vi kan ikke sammenligne resultatene mellom byene uten videre. For å sammenligne resultatene fra regresjonen mellom byene anvender vi derfor en paneldatamodell med interaksjonsvariabler for å sammenligne utviklingen mellom byene. Vi vil her utføre en paneldataregresjon for å undersøke om boligprisene mellom regioner

utvikler seg forskjellig eller om de følger felles utviklingstrekk. Dette ga vi en intuitiv forklaring på i delkapittel 5.10, der vi også argumenterte for bruken av en FGLS-estimert paneldatamodell. Denne typen estimering tar hensyn til en eventuell autokorrelasjon mellom grupper, og ettersom noen av testobservatorene for autokorrelasjon i noen av modellene var helt i gråsonen for hva som er akseptabelt, er dette en bra løsning. I tillegg er denne type estimering velegnet for bruken av interaksjonsvariabler.

Vi vil som sagt se etter ulikheter eller eventuelle likheter ved hjelp av interaksjonsvariabler. I delkapittel 5.10 ga vi en forklaring av bruken av og fremgangsmåten ved interaksjonsvariabler. Vi velger dermed å gjøre slik vi eksemplifiserte det i 5.10, ved å bruke Oslo som basekategori ettersom det både er den største byen vi måler, samt hovedstaden. Dermed sammenligner vi utviklingen i henholdsvis Stavanger, Kristiansand, Bergen og Tromsø opp mot utviklingen i Oslo. Vi konstruerer interaksjonsvariabler for forklaringsvariablene våre, som henholdsvis er boligpris, realinntekt, realrente, arbeidsledighet, forventning og boligbygging. Vi er i større grad interessert i de langsiktige sammenhengene enn de kortsiktige, og konstruerer dermed interaksjonsvariabler med laggede forklaringsvariabler.

Paneldatamodellen innehar de samme likhetstrekkene som tidsseriemodellene vi konstruerte tidligere. Forskjellen er her at vi som nevnt inkluderer interaksjonsvariabler for å påvise en eventuell forskjell i utvikling på et regionalt nivå. I tillegg konstruerer vi en dummyvariabel for alle byer med unntak av Oslo. Når vi utfører regresjonen der vi inkluderer dummyvariablene for byer ekskludert interaksjonsvariablene, vil vi kunne påvise eventuelle nivåforskjeller mellom hver respektive bydummy målt opp mot Oslo. Disse resultatene samles sammen med regresjonen vi utfører der vi inkluderer interaksjonsvariabler, og presenteres i tabell 6.6 nedenfor.

Tabell 6.6: Regresjonsresultater fra paneldataregresjon, med og uten interaksjoner

	Variabel	Panelmodell uten interaksjoner		Panelmodell med interaksjoner	
		Koeffisient	Standardfeil	Koeffisient	Standardfeil
Kortsiktig sammenheng	$\Delta \ln(\text{Realinntekt})_t$	0.0103009	(0.016212)	0.00874	(0.015923)
	$\Delta \text{Realrente}_t$	0.004555	(0.003143)	0.006662**	(0.00324)
	$\Delta \text{Arbeidsledighet}_t$	-0.0139547	(0.0088849)	-0.01572*	(0.009202)
	$\Delta \text{Forventning}_t$	0.001466*	(0.0007812)	0.001545**	(0.000769)
Feiljusteringsmekanisme	$\ln(\text{Realboligpris})_t$	-0.1191333***	(0.0215529)	-0.079051	(0.061103)
Langsiktig sammenheng	$\ln(\text{Realinntekt})_{t-1}$	0.0157364	(0.0231666)	-0.003204	(0.068781)
	Realrente_{t-1}	-0.0099167***	(0.0031799)	-0.016212***	(0.006287)
	$\text{Arbeidsledighet}_{t-1}$	-0.0100058**	(0.0043786)	-0.00511	(0.012978)
	Forventning_{t-1}	0.0018913***	(0.0005814)	0.001136	(0.001406)
	$\ln(\text{boligbygging})_t$	0.0037045	(0.0045737)	-0.000299	(0.013121)
Dummyvariabler	Sesong 1	0.024747***	(0.0073465)	0.027858***	(0.007436)

	Sesong 2	0.0393266*** (0.0064896)	0.0418*** (0.006635)
	Sesong 3	-0.0006347 (0.007531)	0.001434 (0.007527)
	Finanskrise	-0.0063285 (0.0162445)	0.009387 (0.02031)
	Oljekrise	0.0196234 (0.0137495)	0.01044 (0.01526)
	Stavanger	-0.0409939*** (0.0120045)	
	Kristiansand	-0.0700678*** (0.0153101)	
	Bergen	-0.0487783*** (0.0123445)	
	Tromsø	-0.0537184*** (0.0156576)	
Interaksjonsvariabler	Stavanger* $\ln(\text{Realboligpris})_t$		-0.015709 (0.07143)
	Kristiansand* $\ln(\text{Realboligpris})_t$		-0.13031 (0.09125)
	Bergen* $\ln(\text{Realboligpris})_t$		-0.13565* (0.08365)
	Tromsø* $\ln(\text{Realboligpris})_t$		-0.00928 (0.09965)

	Stavanger* $\ln(\text{Realinntekt})_{t-1}$		0.01711	(0.0681)
	Kristiansand* $\ln(\text{Realinntekt})_{t-1}$		0.10505	(0.08099)
	Bergen* $\ln(\text{Realinntekt})_{t-1}$		0.12199	(0.07897)
	Tromsø* $\ln(\text{Realinntekt})_{t-1}$		-0.0052	(0.09128)
	Stavanger* Realrente_{t-1}		0.014676*	(0.00863)
	Kristiansand* Realrente_{t-1}		0.01245	(0.00831)
	Bergen* Realrente_{t-1}		0.005456	(0.00896)
	Tromsø* Realrente_{t-1}		0.01011	(0.00779)
	Stavanger* $\text{Arbeidsledighet}_{t-1}$		-0.001366	(0.01411)
	Kristiansand* $\text{Arbeidsledighet}_{t-1}$		-0.00467	(0.016729)
	Bergen* $\text{Arbeidsledighet}_{t-1}$		0.006186	(0.02026)
	Tromsø* $\text{Arbeidsledighet}_{t-1}$		-0.00056	(0.021)

	Stavanger* <i>Forventning</i> _{t-1}		0.001947 (0.001837)
	Kristiansand* <i>Forventning</i> _{t-1}		0.001286 (0.001785)
	Bergen* <i>Forventning</i> _{t-1}		-0.00006 (0.001783)
	Tromsø* <i>Forventning</i> _{t-1}		0.000226 (0.0018)
	Stavanger* <i>ln(boligbygging)</i> _t		-0.01414 (0.02175)
	Kristiansand* <i>ln(boligbygging)</i> _t		0.00047 (0.01889)
	Bergen* <i>ln(boligbygging)</i> _t		-0.01087 (0.02022)
	Tromsø* <i>ln(boligbygging)</i> _t		0.00921 (0.01467)
	Konstantledd	1.092184*** (0.2910318)	0.9368105** (0.379024)
	Observasjoner	270	270
	Grupper	5	5
	Tidsperioder	54	54

Signifikansmerking: * $p < 0,10$ ** $p < 0,05$ *** $p < 0,01$

Regresjonen gir statistisk signifikant negative koeffisienter for alle fire bydummyene. Dette tolkes som at Oslo innehar den høyeste reelle boligprisen, og koeffisientene sin negative verdi gjenspeiler hvor mye mindre boligprisene er i de andre byene. Koeffisientestimatet for Kristiansand er sterkest negativt, noe som impliserer at avstanden i boligpris mellom Oslo og Kristiansand er den største. Tromsø har et negativt koeffisientestimat med mindre styrke enn Kristiansand igjen. Til slutt er koeffisientestimatene for Stavanger og Bergen de estimatene med den svakeste negative styrken, som impliserer at avstanden i boligprisene mellom Oslo og Stavanger og Bergen er minst. Disse nivåforskjellene samsvarer med forskjellene i regionenes reelle boligpris, presentert i figur 4.9.

Regresjonen gir få statistisk signifikante resultater for interaksjonsvariablene vi inkluderte mellom de respektive byene og variablene vi måler. Dette taler for at boligprisene i regionene påvirkes likt av de variablene vi måler, dog med litt forskjellig helning. Vi vil til tross for lite signifikante resultater tolke regresjonen i lys av koeffisientene, og vil i det følgende tolke helningen og fortegnene til de estimerte koeffisientene for å forsøke å si noe om regionale forskjeller.

Vi finner ikke signifikante forskjeller i helning i koeffisientestimatene for lagget boligpris, som i praksis betyr at justering tilbake mot likevekt etter avvik ikke er signifikant ulik mellom byene. I tillegg er ikke feiljusteringsparameteren signifikant i modellen med interaksjonsvariabler, slik at vi ikke kan dra noen bastante konklusjoner her, ettersom dette tyder på at paneldatamodellen med interaksjoner ikke er kointegrert.

Regresjonskoeffisientene for interaksjonsvariablene mellom de ulike regionene med realinntekt taler for at boligprisene i de ulike regionene påvirkes likt av endringer i realinntekt, ettersom regresjonen viser at ingen av byene påvirkes signifikant ulikt av realinntekt. Det samme kan sies om henholdsvis arbeidsledighet, forventningsvariabelen og boligbygging. Også her taler modellens resultater for at boligprisene påvirkes likt i de ulike regionene av endringer i disse faktorene.

Regresjonen gir et statistisk signifikant koeffisientestimat for interaksjonen mellom realrente og dummyvariabel Stavanger. Koeffisienten er svak og positiv og tolkes som at realrenten på et signifikant nivå påvirker boligprisene i Stavanger ulikt enn i Oslo. Her vil en nivåendring i realrente tilsvarende ett prosentpoeng medføre en omtrent 1,5% høyere økning i boligprisene i Stavanger enn i Oslo. Dette taler for at renten har en sterkere påvirkningskraft i Stavanger enn i Oslo. Regresjonen viser også at realrente i et helhetlig perspektiv har en langsiktig signifikant påvirkning på boligpriser, uavhengig av region. Dette resultatet påvises både i regresjonen uten interaksjoner, og i regresjonen med interaksjoner.

7. Konklusjon

Gjennom oppgaven har vi forsøkt å forklare hva som kan være årsaken til forskjellene i regional boligprisutvikling. For å forsøke å besvare dette har vi valgt å studere utviklingen i fundamentale faktorer og hvordan disse påvirker boligpriser i ulik grad i ulike regioner.

Vi har konstruert modeller for regionene vi studerer, og gjennomført regresjonsanalyser i et forsøk på å avdekke hvilke fundamentale faktorer som har signifikant påvirkning i de ulike regionene. Fra regresjonsanalysene vi gjennomførte fant vi at påvirkningskraften til ulike fundamentale faktorer varierer mellom ulike regioner. Spesielt den langsiktige sammenhengen mellom realinntekt og boligpriser skilte seg ut, og regresjonskoeffisientene varierte stort mellom regionene. Resultatene vi finner fra regresjonsanalysen må dog ses i sammenheng med forklaringskraften til modellen vår og hvorvidt verdiene er signifikante. Vi endte opp med få signifikante koeffisienter for forklaringsvariablene vi inkluderte.

Forklaringskraften til modellen vår for de respektive byene ble presentert i delkapittel 6.6 og varierte noe. Variablene vi har inkludert forklarer omtrent 57% av variasjonen i boligpriser i Oslo og Bergen, 50% i Stavanger og Kristiansand, mens de forklarer bare 38% av variasjonen i Tromsø. Det vil si at 43% eller mer av variasjonen i boligpriser i samtlige regioner, forklares av faktorer som vi ikke har inkludert i modellen vår. Dette kan blant annet være geografiske, demografiske og psykologiske forhold. Dette svekker naturligvis resultatene våre ettersom mye av variasjonen i boligprisutvikling mellom regioner kan være forklart av faktorer vi ikke har inkludert i modellen vår. I tillegg fant vi gjennom de økonometriske testene at datamaterialet i seg selv inneholder ulike feil, som bidrar til at modellene vi utledet ikke gir forventningsrette og presise estimater. Dette betyr at modellene vi har konstruert til dels ikke lykkes i å forklare utviklingen i boligpriser, og vi kan dermed ikke være for bastante i konklusjonene vi trekker.

At modellen vår gir såpass ulik forklaringskraft i de ulike regionene, er i seg selv et interessant funn. Dette kan indikere at ulike regioner blir påvirket i ulik grad av variablene vi har inkludert i modellen vår. Modellen for Tromsø har betydelig lavere forklaringskraft og er fra et geografisk perspektiv den byen som skiller seg mest ut av regionene vi studerer. I

hvilken grad slike faktorer påvirker boligpriser kan vi ikke si noe sikkert om, men slike ulikheter kan potensielt være en mulig årsak til hvorfor Tromsø sin modell har en lavere forklaringskraft.

Fra regresjonsresultatene vi presenterte i delkapittel 6.6 og 6.7 ser vi at de fleste koeffisientestimatene ikke er statistisk signifikante. Det kan være mange grunner til dette, og et lavt antall observasjoner kombinert med dårlig kvalitet på noe av datamaterialet vi bruker, er høyst sannsynlig medvirkende årsaker. Noen av koeffisientene er dog signifikante. Vi får relativt lave p-verdier for forventningsvariabelen i den langsiktige sammenhengen, hvor forventningsvariabelen for både Bergen og Tromsø er statistisk signifikante på et 10% signifikansnivå, og for Stavanger er den statistisk signifikant på et 1% signifikansnivå. Koeffisientene er alle positive og indikerer at en økning i forventninger til egen og landets økonomi bidrar til å øke boligprisene. Koeffisientene er dog svakt positive og relativt like i henhold til påvirkningskraft.

Vi finner at Kristiansand er byen hvor et avvik fra likevekt justeres raskest inn igjen, i tillegg til at dette er byen hvor et avvik fra likevekt raskest halveres. Vi finner også at Stavanger er den byen hvor et avvik fra likevekt i boligpris halveres tregest.

Vi har utført paneldataregresjon i et forsøk på å avdekke hva som forårsaker boligprisforskjellene. Paneldataregresjonen bekreftet at det er signifikante nivåforskjeller i boligpris mellom regionene. Paneldataregresjonen avdekket at realrenten har en sterkere positiv påvirkning i Stavanger enn i Oslo. Resultatene fra paneldataregresjonen ellers ga oss ikke tilstrekkelig bevis til å konkludere om hva som forårsaker forskjeller i boligprisutvikling mellom byene, ettersom koeffisientestimatene for interaksjonsvariablene vi konstruerte ikke var signifikante.

Våre data og modeller gir oss ikke tilstrekkelig grunnlag til å konkludere med hva som kan forklare de store regionale forskjellene i boligprisutvikling vi har sett i Norge fra 2005 til 2018. Fundamentale faktorer forklarer i snitt halvparten av variasjonen i boligpriser, og få signifikante resultater blir påvist, trolig grunnet begrenset datautvalg og ulike feil i data. Vi mener derfor heller ikke at oppgaven gir tilstrekkelig grunnlag til å forkaste etablert teori om hva som påvirker utvikling i boligpriser.

Det hadde vært av stor interesse å studere utviklingen i boligpriser med utvidet datagrunnlag. En lengre undersøkelsesperiode samt bedre kvalitet på data ville sannsynligvis hevet oppgaven. I tillegg ville det vært interessant å inkludere de variablene vi har valgt å utelate fra modellene våre, dersom kvantitative data på disse som kan måles over tid blir tilgjengelig på et senere tidspunkt.

Referanseliste

- Aamo, B. S. (2019). Boligmarkedet som kilde til finansielle kriser. *Tidsskrift for boligforskning*, 2(2-2019), 67-74.
<https://doi.org/10.18261/issn.2535-5988-2019-02-02>
- Anundsen, A. K., Jansen, E. S. (2013). *Self-reinforcing effects between housing prices and credit: an extended version* (SSB rapport 756/2013). Hentet fra https://www.ssb.no/en/forskning/discussion-papers/_attachment/142946?ts=141a239d4a8
- Bernhardsen, T. & Røisland, Ø. (2000). Hvilke faktorer påvirker kronekursen?. *Penger og Kreditt*, 29(3).
- Brooks, C. (2008). *Introductory Econometrics for Finance*. Cambridge University Press, Cambridge.
- Brooks, C. & Tsolacos, S. (2010). *Real Estate Modelling and Forecasting*. Cambridge University Press, Cambridge.
- Brooks, C. (2014). *Introductory Econometrics for Finance* (3rd ed.). Cambridge University Press, Cambridge.
- Brooks, C. (2015, 2. mai). Classical linear regression model assumptions and diagnostics. Hentet fra <https://www.slideshare.net/IdrisAlMalki/20150404-rm-autocorrelation>
- Burns, P. J., (2002, oktober 6). *Robustness of the Ljung-Box Test and its Rank Equivalent*. Hentet fra <https://ssrn.com/abstract=443560>
- Bye, T. & Hægeland, T. (2014). KPI 100 år. *Økonomiske analyser*, 5(2014). Hentet fra https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/_attachment/203138?ts=14957fe6838
- Cameron, A. C. & Trivedi, P. K. (2005). *Microeconometrics: Methods and Applications*. Cambridge University Press. DOI: [10.1017/CBO9780511811241](https://doi.org/10.1017/CBO9780511811241)
- E24. (2017, 04. februar). Flere unge eier sin egen bolig.
- Ekeland, A. (2014). *Sysselsatte i petroleumsnæringene og relaterte næringer 2012* (SSB rapport 12/2014). Hentet fra <https://www.norskoljeoggass.no/contentassets/ad330a6c651f4428ac43f2a6e90f6e2a/sysselsatte-i-petroleumsnaringen-2012.pdf>
- Emblem, A. W., Theisen, T., Aamo, B. S. (2017). *Samfunnsøkonomen*, 131(6), 14-25. Hentet fra <https://samfunnsokonomene.no/wp-content/uploads/2019/04/Samfunns%C3%B8konomen-nr.-6-2017-2.pdf>

Engle, R. F. & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276. Hentet fra <https://www.jstor.org/stable/1913236?seq=1>

E24. (2017, 4. februar). Flere unge eier sin egen bolig. Hentet fra <https://e24.no/norsk-oekonomi/i/wP3pOM/flere-unge-eier-sin-egen-bolig>

Finans Norge. (2014). Utvikling i boligprisindeksen for Norge, i forhold til konsumprisveksten. Hentet fra <https://www.finansnorge.no/contentassets/f2ad4fbbfe38459f80c21970d24c1a63/boligmarked-og-boligfinansiering.pdf>

Finans Norge. (u.å.). Bakgrunn og formål med undersøkelsen. Hentet 10. mars 2020 fra <https://www.finansnorge.no/aktuelt/nyheter/forventningsbarometeret/bakgrunn-og-formal-med-undersokelsen/>

Fredriksen, H. (2007). *En kritisk gjennomgang av Jacobsen og Naug sin modell for hva som driver boligprisene*. (Mastergradsavhandling, Norges Handelshøyskole). Hentet fra <https://openaccess.nhh.no/nhh-xmlui/handle/11250/167773>

Gjedrem, S. (2008, 16. september). *Pengepolitikk i et historisk perspektiv*. Innlegg presentert ved Samfunnsøkonomenes jubileumskonferanse, Oslo. Sammendrag hentet fra <https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Foredrag-og-taler/2008/Gjedrem-2008-09-16/>

Gjedrem, S. (2009, 30. september). *Erfaringer fra finanskrisen*. Innlegg presentert i regi av Centre for Monetary Economics ved Handelshøyskolen BI. Sammendrag hentet fra <https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Foredrag-og-taler/2009/Erfaringer-fra-finanskrisen/>

Halvorsen, B. (2019). Gjelden til norske familier vokser mindre enn før. Hentet fra <https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/gjelden-til-norske-familier-vokser-mindre-enn-for>

Hendry, D. F. & Wallis, K. F. (1984). *Econometrics and quantitative economics*. Blackwell Publishers, Oxford.

Høiby, H. (2017). *Nordisk sammenlikning av byggekostnadsindeksene for boliger*, Notater 2017/35. Statistisk sentralbyrå. Hentet fra <https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/attachment/326873?ts=15f4ce98fc0>

Jacobsen, D. H. & Naug, B. E. (2004). Hva driver boligprisene. *Penger og Kreditt*, 32(4), 229-240. Hentet fra https://www.norges-bank.no/globalassets/upload/publikasjoner/penger_og_kreditt/2004-04/jacobsen.pdf

Jarque, C. M. & Bera, A. K. (1981). Efficient Tests for Normality, Homoscedasticity and Serial Independence of Regression Residuals: Monte Carlo Evidence. *Economics Letters*, 7(4), 313-318. [https://doi.org/10.1016/0165-1765\(81\)90035-5](https://doi.org/10.1016/0165-1765(81)90035-5)

Johannessen, R. (2014). Konsumprisindeksen- en levekostnadsindeks. *Økonomiske analyser*, 5(2014). Hentet fra <https://www.ssb.no/priser-og-prisindekser/artikler-og-publikasjoner/attachment/203142?ts=1495b28c170>

Leamer, E. E. (2002). Bubble trouble? Your Home Has a P/E Ratio Too. UCLA Anderson Forecast.

Ljung, G. M. & Box, G. E. P. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika*, 65(2), 297-303. <https://doi.org/10.1093/biomet/65.2.297>

Melsom, E.T. (2019). Økt sparing i husholdningene. Hentet fra <https://www.ssb.no/nasjonalregnskap-og-konjunkturer/artikler-og-publikasjoner/okt-sparing-i-husholdningene>

Norges Bank. (2008, 1. desember). Norges Banks vurdering av stabiliteten i det finansielle systemet- høsten 2008. Hentet fra <https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Brev-og-uttalelser/2008/Brev-2008-12-01/>

Norges Bank. (2019, 19. desember). Pengepolitisk rapport med vurdering av finansiell stabilitet 4/19. Hentet fra <https://www.norges-bank.no/aktuelt/nyheter-og-hendelser/Publikasjoner/Pengepolitisk-rapport-med-vurdering-av-finansiell-stabilitet/2019/ppr-419/>

NOU 2004:2. (2004). Effekter og effektivitet: *Effekter av statlig innsats for regional utvikling og distriktpolitiske mål*. Hentet fra <https://www.regjeringen.no/no/dokumenter/nou-2004-2/id383676/>

NOU 2015:1. (2015). *Produktivitet- grunnlag for vekst og velferd*. Hentet fra <https://www.regjeringen.no/no/dokumenter/nou-2015-1/id2395258/>

Pirounakis, N. G. (2013). *Real Estate Economics* (1st ed.). Taylor and Francis.

Reinhart, C. M., Rogoff, K. S. (2009). *This time is different*. Princeton University Press.

Revold, M. K., Sandvik, L. & With, M. L. (2018). *Bolig og boforhold - for befolkningen og utsatte grupper* (SSB rapport 13/2018). Hentet fra <https://www.ssb.no/bygg-bolig-og-eiendom/artikler-og-publikasjoner/attachment/346817?ts=162d8bb3be0>

Sage. (2015). Learn to Test for Heteroscedasticity in SPSS With Data From the Early Childhood Longitudinal Study (1998). Hentet fra <http://methods.sagepub.com/base/download/DatasetStudentGuide/heteroscedasticity-in-eclsk-1998>

Samfunnsøkonomisk analyse. (2018). *Bokostnadsindeksen for norske husholdninger* (rapport 23/2018). Hentet fra <https://static1.squarespace.com/static/576280dd6b8f5b9b197512ef/t/5b9ba5734d7a9ccfbee2171a/1536927097148/R23-2018+Bokostnadsindeksen+for+norske+husholdninger.pdf>

Samfunnsøkonomisk analyse AS. (2020). *Førstegangskjøpere og sekundærboliger*. Hentet fra <https://www.nef.no/nyheter/flere-forstegangskjopere-nedgang-i-sekundaerboliger/>

Saunders, M., Thornhill, A., & Lewis, P. (2012). *Research Methods for Business Students*. Harlow, Pearson.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2), 461-464.

Sirnes, E. (2018, 18. desember). realrente. Hentet 3. mars 2020 fra <https://snl.no/realrente>

Statistisk sentralbyrå. (2017, 17. februar). Byggeareal, 2016. Hentet fra <https://www.ssb.no/bygg-bolig-og-eiendom/statistikker/byggeareal/aar/2017-02-17#content>

Statistisk sentralbyrå. (2019a). *Konjunkturtendensene med økonomisk utsyn over året 2018* (SSB rapport 1/2019). Hentet fra <https://www.ssb.no/inntekt-og-forbruk/artikler-og-publikasjoner/attachment/380156?ts=16958558ee0>

Statistisk sentralbyrå. (2019b, 25. juni). Familier og Husholdninger. Hentet 7. mai 2020 fra <https://www.ssb.no/befolkning/statistikker/familie>

Statistisk sentralbyrå. (2020, 1. april). Boforhold, registerbasert. Hentet fra <https://www.ssb.no/bygg-bolig-og-eiendom/statistikker/boforhold/aar>

Stiglitz, J. E. (1990). Symposium on Bubbles. *Journal of Economic Perspectives*, 4(2), 13-18.

Torsvik, R. M. (1999, 18. juni). Bankkrisen. Hentet fra <https://www.ssb.no/bank-og-finansmarked/artikler-og-publikasjoner/bankkrisen>

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica*, 48(4), 817-838. DOI: 10.2307/1912934

Wooldridge, J. M. (2009). *Introductory econometrics: a modern approach*. (4th ed.). Mason, Ohio: South-Western Cengage Learning.

Wooldridge, J. M. (2012). *Introductory econometrics: a modern approach* (5th ed.). Mason, Ohio: South-Western Cengage Learning.

Øksnes, K. (2019, 20. august). Rentestatistikk fra SSB: Gjennomsnittlig boliglånsrente nærmer seg tre-tallet. *Pengenytt*. Hentet fra <https://www.pengenytt.no/rentestatistikk-fra-ssb-gjennomsnittlig-boliglansrente-naermer-seg-tre-tallet/>

Appendiks

Appendiks A: Tabell med regresjonskoeffisienter fra forventningsmodell, etter region.

Variabel	Oslo	Stavanger	Kristiansand	Bergen	Tromsø
$\Delta Realrente$	-0.0743 (-0.09)	-1.1588 (-1.27)	-0.1834 (-0.20)	-0.4022 (-0.43)	-0.0082 (-0.01)
$\Delta Arbeidsledighet$	3.6548 (1.05)	-4.0268** (-2.46)	-1.4720 (-0.48)	-4.6278 (-1.21)	-1.3413 (-0.41)
$Forventning_{t-1}$	-0.0712 (-1.17)	-0.1538** (-2.44)	-0.0728 (-1.27)	-0.0894 (-1.58)	-0.0984* (-1.82)
$Arbeidsledighet_{t-1}$	1.9323 (1.58)	-1.9245* (-1.86)	0.2541 (0.17)	-0.3542 (-0.17)	2.8827* (1.99)
$Realrente_{t-1}$	-1.1197 (-1.51)	-2.7418** (-2.52)	-1.0582 (-1.12)	-1.3213 (-1.25)	-1.2253* (-1.80)
<i>Sesong 1</i>	-0.7815 (-0.37)	0.5010 (0.26)	0.2253 (0.11)	1.7718 (0.74)	0.9095 (0.41)
<i>Sesong 2</i>	-0.4036 (-0.22)	0.0546 (0.03)	-0.1513 (-0.08)	0.4285 (0.23)	0.1362 (0.07)
<i>Sesong 3</i>	-1.7094 (-0.72)	0.1383 (0.07)	0.1646 (0.07)	1.1675 (0.50)	0.1254 (0.06)
<i>Konstantledd</i>	0.2220 (0.04)	18.0549** (2.44)	4.3093 (0.56)	6.5247 (0.72)	-0.2474 (-0.06)
<i>Observasjoner</i>	55	55	55	55	55
<i>R²</i>	0.1780	0.2357	0.1275	0.1528	0.2073
<i>Justert R²</i>	0.0351	0.1028	-0.0242	0.0054	0.0694
Signifikansmerking: * p<0,10 ** p< 0,05 *** p<0,01					

Appendiks B: Refleksjonsnotat Andreas Rasmussen

Sammendrag

I denne masterutredningen undersøker vi boligprisutviklingen i det norske boligmarkedet. Vi undersøker boligprisutviklingen på et regionalt nivå, og vi studerer utviklingen i henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø. Utviklingen i boligpriser har variert vidt mellom disse byene, og vårt forskningsspørsmål er dermed «Hva kan forklare de store regionale forskjellene i boligprisutvikling?». Målet med oppgaven er å undersøke om fundamentale faktorer kan bidra til å forklare disse regionale forskjellene. Vi forsøker å forklare utvikling i boligpris med følgende fundamentale faktorer: inntekt, rente, arbeidsledighet, boligbygging og en forventningsindikator.

Vi konstruerer en boligprismodell for hver av byene basert på økonometrisk metode og teori, og estimerer disse modellene for å forsøke å forklare hva boligprisene påvirkes av.

Resultatene viser at 38-58% av variasjonen i boligpriser blir forklart av de fundamentale faktorene vi bruker som forklaringsvariabler i modellene våre. Vi finner at modellen for Tromsø har lavest forklaringskraft (38,5%), mens boligprisene i Bergen og Oslo forklares best av modellen (57-58%). Datamaterialet vi hadde til disposisjon var begrenset, og vi måtte nøye oss med å undersøke en relativt kort periode. Grunnet disse begrensningene, samt at datamaterialet inneholdt sentrale feil en ønsker å unngå i økonometri, ender vi opp med få signifikante koeffisientestimater for forklaringsvariablene. Vi forkaster dog ikke etablert teori om hvilke faktorer som antas å ha en påvirkning på boligpriser.

Internasjonalisering

I likhet med Norge, så har alle land i verden et boligmarked. De har et boligmarked hvor boligpris bestemmes av tilbud og etterspørsel, og hvor det også antas at det er sentrale faktorer som driver boligprisen. I vår oppgave studerer vi inntekt, rente, arbeidsledighet, forventning til egen og landets økonomi og nybygg av boliger. Dette er faktorer som i litteraturen anses å påvirke boligpriser. I andre boligmarkeder i andre land vil også dette gjelde. Dermed vil disse fundamentale faktorene vi studerer også være fundamentale faktorer som driver boligpriser i andre land. Dermed finnes det en viss overførbarhet av vårt studium av boligpriser over landegrensene.

I en verden hvor både kulturelle grenser og landegrenser åpnes, påvirkes ethvert land av aktiviteter i andre land. Welsh og Luostarinen (1988) definerer i økonomisammenheng internasjonalisering som prosessen av økende involvering av virksomheters aktivitet på tvers av landegrenser.

Analyseenhetene våre er store byer i Norge, henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø. Disse byene drives av langt mer enn kun nasjonal aktivitet. Import og eksport av varer og tjenester på tvers av landegrenser er sentrale aktiviteter som driver samtlige byer, og Norge for øvrig. Følgelig er disse byene preget av internasjonalisering, og påvirkes dermed av internasjonal aktivitet.

På grunnlag av at disse byene drives av import og eksport, vil internasjonal aktivitet påvirke dem i stor grad. Fluktuasjoner i valutakurser vil eksempelvis ha en stor effekt på disse byene. En appresiering av kronen vil medføre virkninger for import i disse byene. Appresiering av kronen betyr at kronekursen øker, altså at kronen blir mer verdt målt opp mot andre valutaer, og dette vil ha en rekke virkninger på norsk økonomi, og også på de regionale økonomiene i byene vi analyserer. Holden (2010) gir en intuitiv forklaring av direkte og indirekte virkninger av en appresiering av kronen. Når kronen appresieres, vil kostnadsnivået i forhold til utlandet øke. Dette rammer konkurranseutsatt sektor, altså de virksomheter som konkurrerer utenlandsk, og vil redusere innenlandsk etterspørsel. Dette vil føre til redusert pris på importvarer og dermed lavere konsumpriser, og videre vil dette medføre lavere lønnsomhet i eksportnæringer og importkonkurrerende næringer.

Holden (2010) argumenterer at en indirekte virkning av appresiering av kronen er reduserte priser på importvarer som kan medføre at norske produsenters priser reduseres, og at arbeidstakernes lønn følgelig reduseres. Dette kan dermed påvirke inntektsvariabelen vi inkluderer i masteroppgaven vår, og som følge av redusert lønn i byene vi studerer vil boligprisene her også reduseres ut ifra litteraturen om boligpriser som vi baserer oss på. En annen indirekte konsekvens av en appresiering av kronen er at sentralbanken kan redusere renten for å motvirke reduksjon i inflasjonen.

I oppgaven vår anvender vi konsumprisindeksen for å transformere nominelle verdier til reelle verdier. Vi transformerer blant annet boligpris (målt ved kvadratmeterpris), inntekt og bankenes utlånsrente om til reelle verdier. En appresiering av kronen vil også påvirke

konsumprisindeksen ettersom importerte konsumvarer utgjør ca. 20% av konsumprisindeksen. Holden (2010) forklarer at en appresiering av kronen vil redusere konsumprisindeksen dersom prisene på importvarene holdes uendret målt i utenlandsk valuta. En reduksjon i konsumprisindeksen vil dermed øke verdien av realboligpris, realinntekt og realrenten. Dermed vil en appresiering av kronen medføre både økninger og reduksjoner av boligpris samt et flertall av faktorer som påvirker boligpris i byene vi studerer.

Holden (2015) presiserer også at en langvarig appresiering vil gi lavere lønnsomhet i eksportnæringer. Innenlandsk produksjon vil bli dyrere sammenlignet med utenlandsk produksjon, og sysselsetting og investering i Norge vil reduseres.

I oppgaven vår forsøker vi å forklare utvikling i boligpris i enhetene vi analyserer basert på utviklingen i arbeidsledigheten. At sysselsetting reduseres impliserer at graden av arbeidsledighet øker. At arbeidsledigheten øker betyr at husholdningers lønnsinntekter bortfaller som følge medfører redusert kjøpekraft, og videre vil dette føre til at boligprisen reduseres i områdene hvor arbeidsledighet øker. Dermed vil en ringvirkning av en langsiktig appresiering av kronen være at inntekt og boligpriser synker. Lavere lønnsomhet i eksportbedrifter vil også ha en negativ virkning på resten av økonomien. Ifølge en rapport fra Menon (2018) er Oslo, Hordaland og Rogaland de 3 fylkene som står for størst andel eksport i Norge målt i milliarder kroner i 2017. Agder fylke innehar sjetteplassen og Troms fylke eksporterer 4. minst av fylkene i Norge. Det er dermed i henholdsvis Oslo, Bergen og Stavanger disse virkningene vil inntreffe i størst grad.

Ifølge en rapport fra Menon (2015) utgjorde eksport nesten 40% av Norges bruttonasjonalprodukt i 2013. Norsk næringsliv og sysselsetting er dermed svært avhengig av utenlandsk etterspørsel. Eksportvirksomhet skaper store inntekter for landet som helhet og er også viktig på et regionalt nivå. Selv om Oslo, Bergen og Stavanger eksporterer mest i nominelle kroner, er det Stavanger, Kristiansand og Bergen som er mest avhengig av eksportvirksomheten. Tall fra rapporten til Menon (2015) viser at 10-25% av sysselsettingen i disse byene er relatert til eksportvirksomhet. Det er dermed disse tre byene fra oppgaven vår som vil være mest sårbare ovenfor et plutselig sjokk fra utlandet. I tillegg viser tall fra Menon (2018) at EU/EØS er Norge sitt desidert viktigste og største eksportmarked per 2017, sett bort fra olje- og gasseksport. Dette betyr at flere av byene vi studerer i vår oppgave er avhengige av aktiviteten i og etterspørselen fra de 28 landene som utgjør EU/EØS. Byene vi studerer er

derfor også avhengig av valutakursen på Euroen. Hvis Euroen appresieres i relasjon til kronen er dette svært gunstig for eksportvirksomheten i Norge, ettersom dette vil medføre økte eksportinntekter og økt sysselsetting i de største eksportbyene vi studerer. Følgelig vil dette drive boligprisene oppover i de respektive byene. En depresiering av Euroen vil ha motsatt effekt og virkning, liknende de effekter en appresiering av kronen har.

Innovasjon

Boligpriser, og boligmarkedet for øvrig, er noe som kontinuerlig forskes på, både i Norge og globalt. Det forskes på jevnlig på utvikling i boligpriser i forskjellige regioner i Norge. Det forskes og rapporteres også hyppig om utviklingen i de faktorene vi anvender som forklaringsvariabler i utredningen vår. For eksempel oppdateres det jevnlig rundt rentenivået og arbeidsledighetsnivået i forskjellige deler av landet. Det føres detaljert statistikk på inntektsnivået til den norske befolkning, samt hvor mye det bygges i forskjellige landsdeler. Slike faktorer blir studert både på årlig, kvartalsvis og månedlig basis.

Jeg mener at begrepet innovasjon må tolkes som noe helt nytt, noe som forandrer tidligere fremgangsmåter, oppfatninger eller syn, og som bidrar til at noe gjøres annerledes fra det punktet innovasjon finner sted og videre fremover i tid. Jeg mener at begrepet innovasjon omhandler en forbedring eller effektivisering av måten en utfører noe på, enten om det skulle være en prosess for å skape noe eller en prosess for å komme frem til noe. Jeg mener at begrepet innovasjon omhandler et nytt og mer effektivt syn på hvordan noe skal produseres, eller hvordan noe skal utføres. Jeg mener at innovasjon er noe nyskapende.

Med det sagt så mener jeg at vår masteroppgave ikke bidrar til at en får et nytt syn på boligmarkedet, eller hva som påvirker og driver boligpriser. Våre forskningsresultater bidrar ikke til ny informasjon angående boligpriser og faktorer som påvirker boligpriser. Våre forskningsresultater er heller ikke motsigende allerede etablert teori om dette temaet. Studiet vårt av boligpriser er til en viss grad en replikasjon av tidligere studier hvor i større grad de samme framgangsmåter anvendes, og de samme forklaringsvariablene anvendes. Vi undersøker dog en annen tidsperiode, en nyere tidsperiode, enn det som har blitt gjort tidligere. Jeg mener at denne utredningen ikke bidrar til at nye syn og oppfatninger skapes om hva som påvirker og driver boligprisutvikling i forskjellige regioner i Norge, og det gir oss

heller ikke en ny og forbedret måte å måle det på. Dermed mener jeg at konseptet innovasjon er irrelevant for vår masterutredning.

Ansvar

Når det kommer til ansvar og etiske utfordringer som måtte oppstå, så finner ikke jeg at forskning på regional boligprisutvikling reiser noen etiske spørsmål underveis. Forskningen vi gjennomfører på boligprisutvikling i forskjellige regioner i Norge er et rent kvantitativt studium, og vi har ikke anvendt noen form for primærkilder i datainnsamlingsprosessen i form av spørreundersøkelse, intervju osv.

Når jeg tolker ansvar i lys av etiske utfordringer relatert til forskningsprosessen med tanke på for eksempel personvern, så mener jeg at dette hovedsakelig angår kvalitativ data hentet gjennom slike metoder for datainnsamling jeg nevnte ovenfor. Når slike metoder som spørreundersøkelser og intervjuer anvendes for å samle inn data, oppstår det flere utfordringer enn i vårt tilfelle. I slike tilfeller, der spørreundersøkelser og intervjuer anvendes, er det essensielt å verne personen(e) som stilles til observasjon. Her fremkommer det oftere at sensitive spørsmål og temaer som intervjuobjektene kan være skeptiske til å besvare blir brakt frem i lys, og det er akkurat her det er utrolig viktig å forsikre intervjuobjektets anonymitet. Særlig når det omhandler spørsmål om atferd og personlige meninger eller erfaringer er det ytterst viktig å verne intervjuobjektene. Dette vil være nødvendig for å unngå at det oppstår en etisk konflikt mellom intervjuer og intervjuobjektet.

I vårt tilfelle har vår data blitt innsamlet fra statistiske databaser som er offentlig tilgjengelige, og betegnes dermed som sekundærdata. Vi samler hovedsakelig inn data fra Statistisk Sentralbyrå, som har hovedansvaret for å dekke behovet for statistikk om det norske samfunnet. Denne typen statistikk er ment å være tilgjengelig for alle, og kan fritt brukes til ulike formål så lenge en følger retningslinjene for kildesitering og -henvisning. Dette er ikke sensitiv informasjon, og følgelig mener jeg at vår oppgave ikke møter på noen etiske utfordringer relatert til datainnsamlings- og forskningsprosessen. Dermed vil jeg konkludere med at konseptet med etisk ansvar er irrelevant for vår masterutredning.

Oppsummering

I dette refleksjonsnotatet kommer jeg frem til at internasjonalisering er høyst relevant for denne masterutredningen. Jeg kommer frem til at fluktuasjoner i valutakurser i stor grad påvirker import og eksport i Norge og i byene vi analyserer, som igjen kan ha store konsekvenser for boligprisutviklingen og de underliggende faktorene som driver boligpris. Jeg konkluderer med at masterutredningen ikke bidrar til innovasjon, samt at det ikke foreligger ansvar i relasjon til etiske utfordringer i utarbeidelsen og gjennomføringen av denne masterutredningen. Avslutningsvis vil jeg konkludere med at det har vært både interessant og lærerikt å reflektere rundt disse konseptene i relasjon til masterutredningen vår.

Referanseliste:

Holden, S. (2010, desember). Valuta og valutamarked. Forelesningsnotat nr 9. Hentet fra <http://folk.uio.no/sholden/E1310/fnotat9-valuta.pdf>

Holden, S. (2015, september). Kapittel 14 Valuta og valutamarked. Hentet fra <http://folk.uio.no/sholden/E1310/1310-V16-sh.html>

Menon. (2015). *Eksport fra norske regioner – et regionalt perspektiv på norsk eksportvirksomhet* (Menon rapport nr. 9/2015). Hentet fra <https://www.regjeringen.no/contentassets/518427735d7b4a18847355af689430fa/rapport-norsk-eksport-fra-regioner-23-mars---endelig.pdf>

Menon. (2018). *Fylkes- og kommunefordelt eksport i 2017 – betydning for sysselsetting* (Menon rapport nr. 101/2018). Hentet fra <https://www.nho.no/siteassets/publikasjoner/menon-hovedrapport-ny.pdf>

Welch, L. & Luostarinen, R. (1988). Internationalization: Evolution of a concept. *Journal of General Management*, 14(2):155-171

Appendiks C: Refleksjonsnotat Johannes Oseland

Sammendrag

Boligprisene i Norge har økt over en lengre periode. I en artikkel skrevet av Emblem, Theisen & Aamo (2017, s.14) peker de på den senere boligprisutviklingen vi har sett i Norge og hvordan utviklingen i boligpriser har variert vidt mellom regioner. Dette ga oss inspirasjon til å undersøke temaet nærmere og med artikkelen som utgangspunkt stilte vi problemstillingen:

Hva kan forklare de regionale forskjellene i boligprisutvikling?

For å kunne svare på dette har vi studert det som i litteraturen kalles fundamentale faktorer. Dette er ulike faktorer som antas å ha stor påvirkning på boligpriser og vi har derfor studert utviklingen til disse faktorene og undersøkt hvorvidt de påvirker ulike regioner i ulik grad. Vi forsøker å forklare utvikling i boligpris med følgende fundamentale faktorer: inntekt, rente, arbeidsledighet, boligbygging i tillegg til en forventningsindikator som vi selv konstruerer. Vi tok utgangspunkt i 5 store regioner i Norge, henholdsvis Oslo, Stavanger, Kristiansand, Bergen og Tromsø og har studert utviklingen i boligpriser for disse regionene sett i lys av utviklingen av fundamentale faktorer.

I den første delen av oppgaven presenterer vi sentral teori om emnet. Vi går gjennom tilbud, etterspørsel og hvordan likevekt i boligmarkedet oppstår, samt introduserer boligprismodellen til Jacobsen og Naug (2004) som blir sentral for hvordan vi løser problemstillingen. Videre tar vi for oss utviklingen i boligmarkedet i Norge. Vi ser på utviklingen i boligpriser over tid og hvordan utviklingen i sentrale faktorer har vært over tid. Vi gjennomgår deretter datagrunnlaget for oppgaven hvor vi presenter utviklingen i de utvalgte fundamentale faktorene i de valgte regionene. Vi går også gjennom økonometrisk metode som benyttes i de empiriske analysene senere. Med utgangspunkt i de presenterte data konstruerer vi en boligprismodell for hver av de respektive regionene med utgangspunkt i den empiriske metoden samt boligprismodellen til Jacobsen og Naug (2004).

Vi bruker de konstruerte boligprismodellene til å gjennomføre økonometriske analyser av de regionale boligmarkedene i perioden første kvartal 2005 til fjerde kvartal 2018. Med utgangspunkt i boligprismodellene vi konstruerer fant vi at 38-58% av variasjonen i boligpriser kan forklares ved hjelp av de fundamentale variablene vi har inkludert. Tromsø

hadde den laveste forklaringskraften med 38%, modellen for Stavanger og Kristiansand hadde en forklaringskraft på omtrent 50% mens Oslo og Bergen hadde den høyeste forklaringskraften der variablene forklarer omtrent 57% av variasjonen i boligpriser. Ifølge modellen vår forklares derfor Tromsøs boligpriser i adskillig høyere grad av faktorer vi ikke har inkludert og er et interessant funn i seg selv.

Regresjonsresultatene viste få signifikante koeffisientverdier. Dette gjør det vanskelig å konkludere i hvilken grad ulike fundamentale faktorer påvirker ulikt i ulike regioner. Vi fikk dog positive koeffisientverdier for den langsiktige sammenhengen mellom forventninger og boligpriser for samtlige regioner der koeffisientene for Tromsø og Bergen var statistisk signifikante på et 10% signifikansnivå, og koeffisienten for Stavanger var statistisk signifikant på et 1% signifikansnivå. Dette tyder på at økte forventninger til egen og landets økonomi er knyttet til en økning i boligpris. Effekten var dog liten, og differansen mellom regionene var små.

Internasjonale trender

I masterutredningen vår ser vi på utviklingen i regionale boligpriser sett i lys av fundamentale faktorer. Hvordan internasjonale trender har en effekt på disse fundamentale faktorene, og dermed også boligpriser, er høyst interessant.

I utredningen vår kommenterte vi blant annet effektene av finanskrisen som startet i 2007 og hvordan dette var med på å påvirke de fundamentale faktorene vi undersøker. Det er liten tvil om at finanskrisen hadde en global effekt som påvirket store deler av verden og det er heller ingen tvil om at de fundamentale faktorene vi har inkludert i modellen vår ble påvirket av dette. Den generelle trenden i verden så vel som i Norge, var lavere inntekter og høyere arbeidsledighet. I utledningen av forventningsindikatoren vår så vi også et tydelig dupp i fremtidige forventninger til egen og landets økonomi i 2009 som høyst sannsynlig var et resultat av finanskrisen.

Det er også interessant å se hvordan internasjonale trender kan ha ulike virkninger på ulike regioner. Oljekrisen som oppstod i 2014 var et resultat av internasjonale forhold og hadde blant annet en helt annen effekt på Stavanger sammenlignet med de fire andre regionene vi inkluderte i analysen vår, og Stavanger var i denne perioden preget av økt arbeidsledighet og lavere inntekt. Dette fordi Stavanger er en av kommunene med høyest andel sysselsatte i

petroleumsnæringen, som i 2012 var på 15,4 prosent (Ekeland, 2014, s.12). Hvordan ulike næringer blir påvirket av nasjonale trender er derfor interessant. Dette fordi ulike regioner er preget av ulike næringssammensetninger slik vi ser med Stavanger, og internasjonale trender som påvirker næringer i Norge kan ha ulik effekt i de ulike regionene.

Innovasjon

Temaet for oppgaven vår er på ingen måte et nytt tema, og har vært forsket på i en rekke tidligere vitenskapelige artikler og masterutredninger. Innfallsvinkelen til oppgaven vår har i høy grad vært inspirert av boligprismodellen som Jacobsen og Naug (2004) konstruerte og boligprismodellen vi konstruerer selv er ikke revolusjonerende, og oppgaven er i så måte ikke preget av høy innovasjon. Innovasjon trenger dog ikke være noe helt nyskapende, og dersom man klarer å trekke inn noen nye momenter og gjøre et vitenskapelig arbeid som kan bidra til noen funn eller videre forskning, kan man argumentere for at det har vært et innovativt arbeid, uten at utredningen i seg selv er revolusjonerende.

I utredningen vår finner vi blant annet at Tromsøs boligpriser påvirkes i lavere grad av de fundamentale faktorene enn de resterende regionene vi studerer, basert på regresjonsresultatene våre. Hva som skiller de ulike regionene fra hverandre og hva som gjør at Tromsø sin boligprismodell har lavere forklaringskraft kan vi ikke konkludere med i denne oppgaven og kan være interessant å studere videre. De spesielle særegenheter og egenskaper til de ulike regioner er noe vi bare kort har presentert og er noe som med fordel kunne vært studert videre. Vi fant som nevnt tidligere at Stavanger ble påvirket i høyere grad av oljekrisen i 2014 grunnet den høye andelen ansatte i petroleumsindustrien. For videre forskning hadde det vært spesielt interessant og gjort en nøye studie av hele næringssammensetningen til de ulike regionene og hvordan dette kan ha betydning for utviklingen i de fundamentale faktorer, og videre hva slags effekt dette kan ha på utviklingen i boligpriser i ulike regioner. Dette kunne i seg selv vært et innovativt bidrag til videre forskning av problemstillingen vår.

Ansvar

Når man skal skrive en masterutredning, eller gjøre et annet vitenskapelig arbeid kan det oppstå etiske problemer man må ta stilling til. En masteroppgave er ofte forbundet med spørreundersøkelser og intervju og man har et ansvar for å ivareta personvernet til de

involverte. Dersom man samler inn sensitive data fra et selskap som ikke vil identifiseres har man også et ansvar for å presentere data på en måte som ikke kan angi hvem de/den involverte er. For å møte slike problemstillinger kan man anonymisere spørreundersøkelser og utelate navn og informasjon om intervjuobjekter. Man kan også gi mulighet til intervjuobjekter å lese gjennom sine egne utsagn for å kontrollere at utsagnene stemmer for å unngå uoverensstemmelser i ettertid.

I vår masterutredning har vi ikke måttet ta hensyn til slike problemstillinger ettersom vi har tatt i bruk offentlig tilgjengelig informasjon og data, og har ikke jobbet med sensitiv informasjon. Dette har gjort at ansvaret vi har hatt overfor andre aktører har vært minimalt. Det har uansett vært viktig å kildeføre riktig slik at informasjon og data vi tar i bruk fra andre kilder ikke blir presentert som eget materiale.

Oppsummering

Å fullføre en masterutredning har vært en krevende, men også lærerik prosess. Når vi nå fullfører oppgaven vår, har det vært interessant å reflektere rundt arbeidet vårt og hvordan dette er knyttet opp mot sentrale verdier til handelshøyskolen ved UiA. Oppgaven kan kanskje ikke knyttes direkte opp mot problemstillinger forbundet med ansvar, men det var interessant å reflektere både rundt hvordan oppgaven vår kan knyttes opp mot internasjonale trender, i tillegg til å reflektere hvorvidt vår oppgave har bidratt til å innovere.

Referanser

- Ekeland, A. (2014). Sysselsatte i petroleumsnæringene og relaterte næringer 2012 (SSB rapport 12/2014). Hentet fra <https://www.norskoljeoggass.no/contentassets/ad330a6c651f4428ac43f2a6e90f6e2a/sysselsatte-i-petroleumsnaringen-2012.pdf>
- Emblem, A.W., Theisen, T., Aamo, B.S. (2017). *Samfunnsøkonomen*, 131(6), 14-25. Hentet fra <https://samfunnsokonomene.no/wp-content/uploads/2019/04/Samfunns%C3%B8konomen-nr.-6-2017-2.pdf>