

En kritisk gjennomgang av Jacobsen og Naug sin modell for hva som driver boligprisene.

av

Heidi Fredriksen

Veileder: Professor Frode Steen

Utredning i fordypningsområdet Økonomisk analyse

NORGES HANDELSHØYSKOLE

Denne utredningen er gjennomført som et ledd i masterstudiet i økonomisk-administrative fag ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at høyskolen innestår for de metoder som er anvendt, de resultater som er fremkommet eller de konklusjoner som er trukket i arbeidet.

Sammendrag.

I denne utredningen kommenterer og kritiserer jeg en modell for drivere bak boligprisene som Dag Henning Jacobsen og Bjørn E. Naug har bygd. I tillegg prøver jeg å komme med forslag til forbedring.

I første del av utredningen har jeg diskutert boligpriser generelt og dessuten argumentert for mulige forklaringsfaktorer for boligprisene. Deretter gikk jeg over til det teoretiske rammeverket, før jeg gikk løs på selve modellen.

Jeg finner at det er en rekke mangler og feil i Jacobsen og Naug sin modell for hva som driver boligprisene, deriblant inneholder modellen autokorrelasjon, parametrene mangler stabilitet og mulige problemer med endogene variabler er ikke diskutert.

I mitt forslag til forbedret modell har jeg fjernet systematikken i feilleddene, parametrene er blitt mer stabile og jeg har inkludert en trendfaktor som bør være i en modell med nominelle priser. Modellen jeg foreslår er imidlertid ikke så bra at man kan trekke gode, holdbare slutninger om hvorvidt boligprisene er overvurdert. Dette skyldes blant annet at datagrunnlaget ikke er godt nok til å ta hensyn til det mulige simultanitetsproblemet som boligmassen og kanskje renten skaper.

Forord.

Denne utredningen er siste ledd i en mastergrad i økonomisk analyse ved Norges Handelshøyskole.

Arbeidet med oppgaven har vært spennende, og jeg har funnet problemer i en modell som har blitt benyttet som kilde i taler fra Norges Bank, diverse utredninger og avisoppslag, i tillegg til å ha blitt publisert i rapporten "Finansiell stabilitet" nr. 1/2004 som publiseres to ganger i året av Norges Bank, samt i Penger og Kreditt nr. 4/2004 som er Norges Banks kvartalsskrift.

I forbindelse med utredningsarbeidet vil jeg spesielt takke professor Frode Steen for god veiledning. Dessuten vil jeg takke professor Svein-Arne Persson for økonomiske midler og god støtte i startfasen, samt postdoktor Arngrim Hunnes og stipendiat Trond M. Døskeland for å ha kommet med oppmuntrende ord og gode innspill. Jeg vil i tillegg rette en stor takk til Dag Henning Jacobsen som har bidratt med datasettene og dessuten svart raskt på alle mailer med spørsmål. Og sist, men ikke minst takker jeg min gode venninne Katrine Løken for gode diskusjoner underveis og for å ha lest gjennom oppgaven min.

Bergen 4.juni 2007

Heidi Fredriksen

Innholdsfortegnelse.

1. INNLEDNING	7
2. BOLIGMARKEDET OG PRISDANNELSEN	9
2.1. Fakta om boligkjøp	9
2.2. Hva bestemmer tilbud og etterspørsel?.....	9
2.3. Historisk prisutvikling.....	10
2.4. Mulige forklaringsfaktorer bak boligprisen og deres utsikter.....	11
2.5. Boligprisboble?.....	16
2.6. Andre mulige forklaringsfaktorer.....	17
2.7. Avsluttende kommentar.....	18
3. STATISTISK TEORI	19
3.1. Modell spesifisering.....	19
3.2. Minste kvadraters metode.....	19
3.3. Stasjonaritet.....	20
3.3.1. Begrepet	20
3.3.2. Dickey-Fuller-testen	21
3.4. Kointegrasjon.....	23
3.5. Statiske og dynamiske modeller.....	23
3.6. Feiljusteringsmodeller.....	24
3.6.1. Generelt	24
3.6.2. Engle-Granger 2-steps metode.....	25
3.6.3. 1-steps ECM.....	26
3.7. Autokorrelasjon og tester for avdekke problemet.....	27
3.7.1. Hva er autokorrelasjon?.....	27
3.7.2. Grafiske tester.....	28
3.7.3. Durbin-Watson testen.....	29
3.7.4. Box-Pierce og Ljung-Box testene.....	30
3.8. Statiske konsekvenser av spesifikasjonsfeil.....	30
3.9. Data mining	31
3.10. Prediksjon.....	32

4. REESTIMERING OG VURDERING AV JACOBSEN OG NAUG SIN MODELL FOR BOLIGPRISVEKST.	33
4.1. Bakgrunn.	33
4.2. Tidsseriene.	36
4.2.1. Mine forventninger om tidsseriene.	36
4.2.2. Dickey-Fuller testene.	40
4.3. Den konstruerte forventningsvariabelen.	41
4.3.1. Bakgrunn.	41
4.3.2. Jacobsen og Naug sin modell for forventningsvariabelen.	42
4.3.3. Vedrørende rentevriabelen.	43
4.4. Jacobsen og Naug sin modell for boligprisveksten.	43
4.4.1. Reestimering.	43
4.4.2. Tolkning av koeffisientene fra reestimeringen.	45
4.4.3. Testing for autokorrelasjon.	48
4.4.3. Er restleddene normalfordelte?	50
4.4.4. Diskusjon av den pålagte restriksjonen.	50
4.5. Data Mining?	51
4.6. For få observasjoner?	51
4.7. Er boligprisene overvurdert?	53
4.8. Kritikk av modellen.	54
4.9. Konklusjon.	55
5. EN ALTERNATIV BOLIGPRISMODELL.	56
5.1. Innledning.	56
5.2. Tidsseriene.	56
5.2.1. Forventninger.	56
5.2.2. DF-testene for stasjonaritet.	58
5.3. Den konstruerte forventningsvariabelen.	58
5.4. Hva driver boligprisene?	61
5.4.1. En alternativ modell.	61
5.4.1. Tolkning av elastisitetene og justeringsparameteren.	63
5.4.2. Er modellen gyldig?	66
5.4.3. Videre tester.	66
5.5. Multikolinearitet.	67
5.6. Prediksjon.	68
5.7. Er parametrene stabile?	69
5.8. Simultanitetsproblemer.	71
5.10. Konklusjon.	72

6. AVSLUTNING.....	73
REFERANSER.....	74
VEDLEGG 1. SUMMARY STATISTICS.....	77
VEDLEGG 2. REGRESJONSRESULTATER.....	78
VEDLEGG 3. KORRELASJONSMATRISER.....	82

Tabeller.

<i>Tabell 1: Resultatene fra DF-testene.</i>	40
<i>Tabell 2: Modell for husholdningenes forventninger til egen og landets økonomi.</i>	42
<i>Tabell 3: Jacobsen og Naug's feiljusteringsmodell. Deres estimater og nye reestimerte estimater. Koeffisientene i langtidslikningen har blitt endret til de som inngår i likevekten.</i> .	44
<i>Tabell 4: En sammenligning av estimatene fra Jacobsen og Naug's modell for ulike utvalg.</i>	52
<i>Tabell 5: DF-testene.</i>	58
<i>Tabell 6: Modell for husholdningenes forventninger til egen og landets økonomi.</i>	59
<i>Tabell 7: En alternativ feiljusteringsmodell for logaritmen til boligprisene.</i>	62
<i>Tabell 8: En sammenligning av estimatene fra modellen med trend når jeg ser på to forskjellige utvalg, det ene utvalgene er en utvidelse av det andre.</i>	70

Figurer.

<i>Figur 1: Utviklingen i den nominelle kvadratmeterprisen siden 1985, målt i 1000 kr.</i>	11
<i>Figur 2: Registrert arbeidsledighet. Figuren er hentet fra sentralbanksjefens årstale.</i>	12
<i>Figur 3: Norges Banks styringsrente. Figuren er hentet fra sentralbanksjefens årstale.</i>	13
<i>Figur 4: Befolkningsstruktur. Figuren er hentet fra sentralbanksjefens årstale.</i>	14
<i>Figur 5: Positiv Autokorrelasjon.</i>	28
<i>Figur 6: Negativ autokorrelasjon.</i>	28
<i>Figur 7: Ingen autokorrelasjon.</i>	29
<i>Figur 8: Tidsseriene.</i>	39
<i>Figur 9: Plott av restleddene fra tabell 3.</i>	48
<i>Figur 10: Et histogram over feilleddene fra tabell 3.</i>	50
<i>Figur 11: Tidsseriene med det utvidede datasettet.</i>	57
<i>Figur 12: Den konstruerte forventningsindikatoren.</i>	60
<i>Figur 13: Plott av feilleddene fra tabell 7.</i>	66
<i>Figur 14: Et histogram av feilleddene fra tabell 7.</i>	67
<i>Figur 15: Sann boligpris og predikert boligpris for de 8 kvartalene i 2005 og 2006.</i>	68
<i>Figur 16: Sann boligpris, predikert boligpris uten trend og predikert boligpris med trend.</i> ..	69

1. Innledning.

Prisveksten i boligmarkedet har vært enorm de siste 14 årene, og mange spør seg om hvor det skal stoppe. Utviklingen i boligprisene berører mange nordmenn, og spesielt førstegangsetablerende og spekulanter er utsatt for svingningene i boligprisen. Eksperter har uttalt i både aviser og tidsskrifter at boligmarkedet har vært opphetet i lengre tid. På tross av slike uttalelser i media så har boligprisene fortsatt å stige.

I denne utredningen vil jeg sette fokus på Jacobsen og Naug sitt arbeid med en modell for drivkreftene bak boligprisene. Denne modellen ble publisert i artikkelen "Hva driver boligprisene?" i Norges Banks kvartalsskrift Penger og Kreditt, nr. 4/2004. Jeg vil først reestimere og diskutere deres modell før jeg avrunder analysen med å komme med et forslag på en alternativ modell. Jeg har benyttet statistikkprogrammet STATA til alle økonometriske analyser.

På grunn av at arbeidet med utredningen skal ferdigstilles på ett semester, så har jeg vært nødt til å gjøre avgrensninger i oppgaven. Det tok lang tid å få tak i det korrekte datasettet til analysen i kapittel 4, og derfor begrenset jeg analysen i kapittel 5 til den oppdaterte versjonen av dette datasettet. Det ville ha vært interessant å undersøke effektene fra andre variabler, både de som ble diskutert, men ikke benyttet i Jacobsen og Naug sin analyse, og andre faktorer utover disse.

Jeg har kommet frem til at Jacobsen og Naug sin modell inneholder autokorrelasjon, og at estimatene er svært ustabile, både når vi ser på et noe endret utvalg i den samme perioden og når vi så utvider dette. Disse problemene har jeg klart å løse når jeg foreslår en alternativ modell, men et problem gjenstår. Boligmassen blir antakelig påvirket av boligprisene, og da oppstår et simultanitetsproblem som gjør at estimatet for boligmasse verken er forventningsrett eller konsistent, dvs. jeg kan ikke bruke dette estimatet til å trekke noen slutninger om hvilken effekt boligmassen har på boligprisene.

Utredningen er delt inn i 6 kapitler. I kapittel 2 vil jeg se nærmere på boligpriser generelt og mulige forklaringsfaktorer for boligprisveksten. Deretter vil jeg i kapittel 3 behandle det teoretiske rammeverket som trengs for å kunne bygge og vurdere tidsseriemodeller. I kapittel

4 reestimerer og vurderer jeg Jacobsen og Naug sin modell for drivkreftene bak boligprisene, før jeg i kapittel 5 bygger en egen modell. I siste kapittel kommenterer jeg resultatene.

2. Boligmarkedet og prisdannelsen.

2.1. Fakta om boligkjøp.

For de fleste husholdninger er et boligkjøp den største enkeltinvesteringen de gjør i løpet av livet. Boligen er et nødvendighetsgode, og man betrakter en bolig både som et varig forbruksgode og et formuesobjekt. En stor del av Norges befolkning har derfor mesteparten av formuen sin plassert i egen bolig. I Norge eier ca 80 % av husholdningene sin egen bolig (Sentralbanksjefens årstale 2007). I tillegg eier 10 % av husholdningene en ekstra bolig som enten brukes av husholdningen selv, leies ut eller står tom (NoU 2002).

Boligkjøpene er ofte lånefinansiert. I følge kredittilsynets boliglånsundersøkelse for 2006 hadde 77 % av kundene under 35 år en belåningsgrad på boliglånene på over 80 %. 37 % av lånene oversteg også boligens verdigrunnlag. Ytterligere eksponering for boligmarkedet får de som investerer penger i eiendomsfond og de som kjøper hytte. Det er derfor mange personer som sitter med en udiversifisert risiko og er svært utsatt for svingninger i boligprisene. Det er viktig å merke seg at så lenge man har råd til å sitte på boligen, slik at man kan selge den når en selv ønsker det, så betyr ikke svingningene i boligprisene så mye for enkeltpersonen. Da selger man boligen i en periode som gir gevinst.

2.2. Hva bestemmer tilbud og etterspørsel?

Boligprisene bestemmes av tilbud og etterspørsel. På kort sikt vil boligprisene hovedsakelig variere med endringene i etterspørselen. Dette fordi det tar tid å bygge nye boliger, slik at tilbudet er rimelig konstant på kort sikt. Husholdningenes etterspørsel bestemmes av deres betalingsvilje og dermed også deres betalingsevne. Deres inntekt, formue, fremtidige forventninger både på inntekts- og utgiftssiden og ikke minst preferanser er sentrale faktorer bak etterspørselen. På bakgrunn av dette må store prisbevegelser som skjer over kort tid ha sammenheng med endringer i etterspørselen. Ettersom behov og prioriteringer er rimelig stabilt for befolkningen som helhet, må prishoppet skyldes store endringer i betalingsevnen eller i vurderingen av boligen som investeringsobjekt. (NoU 2002).

På lengre sikt vil faktorer som påvirker tilbudssiden også virke inn. Spesielt byggekostnadene, men også tomteprisene har stor betydning for hvor mye ressurser som settes inn i

nybyggingen. Andre langsiktige forklaringsfaktorer er demografi, flytting og sentralisering. Disse påvirker hvor i landet boliger etterspørres, samt hvilke typer boliger som er ettertraktet.

2.3. Historisk prisutvikling.

I etterkrigstiden var deler av boligmarkedet regulert. Dette innebar at boligprisene lå betydelig under markedspris i pressområder, og det var en jevn prisutvikling frem til slutten av 1970-tallet. På 1980-tallet ble boligmarkedet gradvis liberalisert, men det skjedde allikevel over en ganske kort tidshorison. Dereguleringene av kreditt- og boligmarkedene førte sammen med lave låneutgifter til en sterkere utlånsvekst (NoU 2002). Dette resulterte i økt boligkonsum og en sterk realprisstigning, noe som igjen ga større variasjon i boligprisene. Boligprisene nådde et foreløpig toppnivå i 1988.

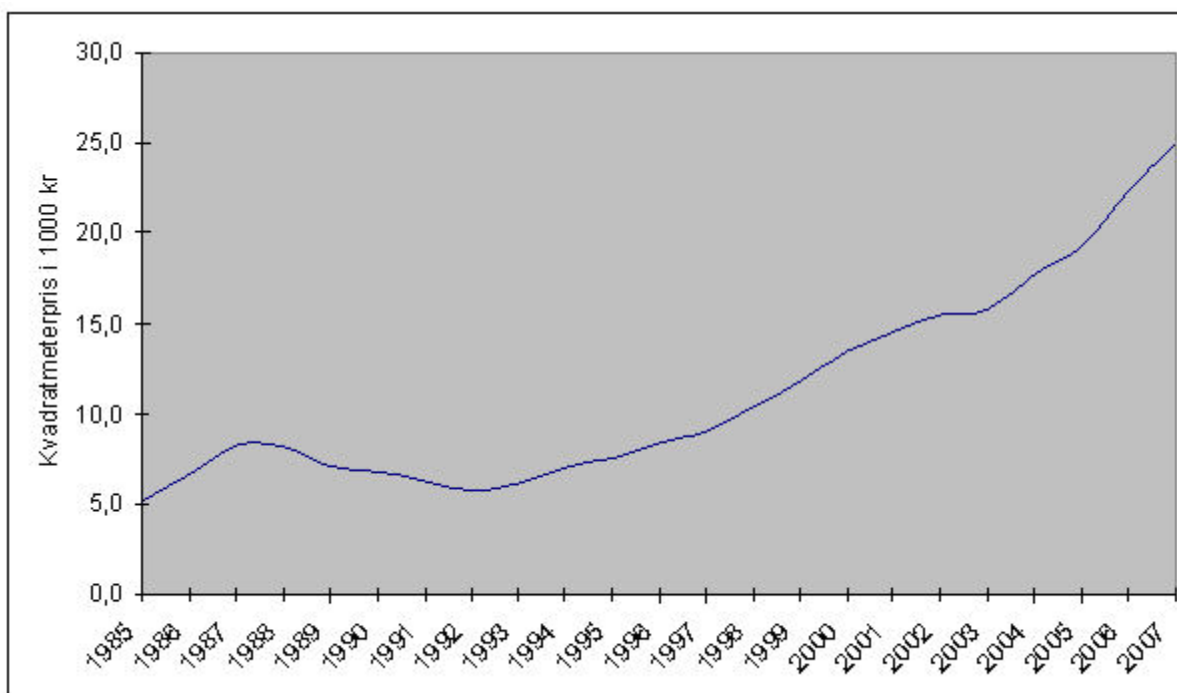
Fra 1988 til 1992 gjorde økende renter, lavere inflasjon og skatteomlegging at låneutgiftene økte betydelig slik at mange husholdninger fikk betalingsproblemer. I perioden falt de nominelle boligprisene med 20 % (NoU 2002). Siden den gang (bunnåret 1992) har prisveksten vært på utrolige 336 % per april 2007.¹ I forhold til forrige toppår, 1987, har gjennomsnittlig boligpris økt med 205 %² (NEF, EFF, ECON, FINN.no, april 2007). Prisene har faktisk økt uavbrutt siden 1992, med unntak av noen måneder i 2003. Det har vært en lignende prisøkning i de andre nordiske landene og flere andre europeiske land. Sammenligner vi Norden forøvrig, så har svingningene vært størst i Finland, og Danmark har vært først ute både med prisoppgangen utover 1980-tallet og den påfølgende nedgangen (NoU 2002).

Grafen under viser hvordan den nominelle kvadratmeterprisen, målt i 1000 kr, har utviklet seg siden 1985. Tallene er hentet fra Eiendomsmeidlerbransjens boligprisstatistikk (NEF, EFF, ECON og FINN.no). Denne prisindeksen gjelder for bruktboliger under ett og måler den gjennomsnittlige boligprisen per kvadratmeter, justert for effekter av boligens størrelse, type og beliggenhet. Det betyr at boligprisen ikke er en størrelse som eksisterer på linje med oljeprisen og valutakurser, men en teoretisk størrelse vi finner ved hjelp av regresjonsanalyse (ECON 2004).

¹ 227 % korrigerert for inflasjon.

² 86 % korrigerert for inflasjon.

Figur 1: Utviklingen i den nominelle kvadratmeterprisen siden 1985, målt i 1000 kr.



Den sterke veksten i boligprisene de siste 14 årene innebærer at de som har eid egen bolig har fått en stor formuesøkning. I noen tilfeller kan man trekke det langt og si at de har fått betalt for å eie, siden deres verdistigning mer enn oppveier for utgiftene de har hatt på renter, vedlikehold osv. Førstegangskjøpere derimot, må ut med mer enn sine forgjengere for å komme inn på boligmarkedet. Allikevel så er det ikke sikkert at prisstigning er en fordel for de som eier selv. De som vil opp på neste trinn i boligkarrieren, for eksempel unge barnefamilier, kan oppleve at prisstigningen på dagens bolig ikke veier opp for prisstigningen på deres neste hjem (NoU 2002). Derfor er det vel for det meste boligspekulanter som kan tjene/tape på boligpriser.

2.4. Mulige forklaringsfaktorer bak boligprisen og deres utsikter.³

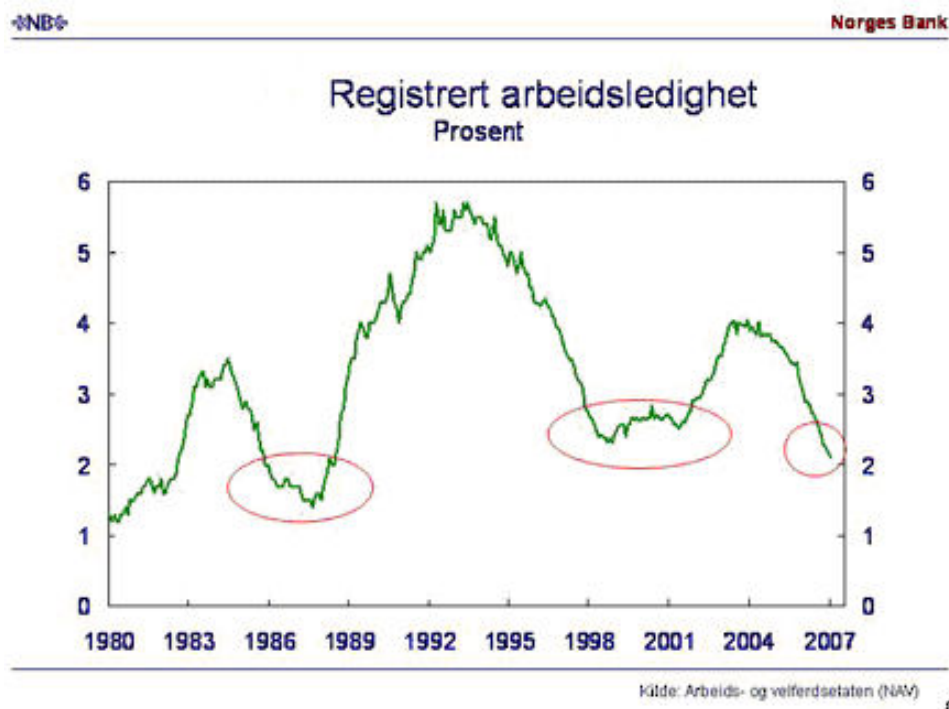
Dette avsnittet bygger på forklaringsfaktorer som Jacobsen og Naug vurderte i sin analyse.

Norsk økonomi er i en høykonjunktur etter en sterk oppgang siden sommeren 2003. Oppgangen avløste en periode med dempet vekst fra 1998 og et mildt tilbakeslag i 2002 og inn i 2003. Selv om konjunktorene har skiftet har veksten vært høy de siste 15 årene.

³ Avsnittet bygger på sentralbanksjefens årstale 2007.

Arbeidsledigheten har falt kraftig de siste kvartalene, og utviklingen har fellestrekk med konjunkturtoppene på midten av 80-tallet og på slutten av 90-tallet. Begge disse høykonjunktorene førte med seg kraftig lønns- og kostnadsvekst, og det har kommet signaler på at lønnsveksten nå tar seg opp. Arbeidsledigheten ligger an til å holde seg lav. Blant annet viser en undersøkelse at flere bedrifter planlegger å øke bemanningen enn antallet som sier de skal nedbemanne.⁴

Figur 2: Registrert arbeidsledighet. Figuren er hentet fra sentralbanksjefens årstale.

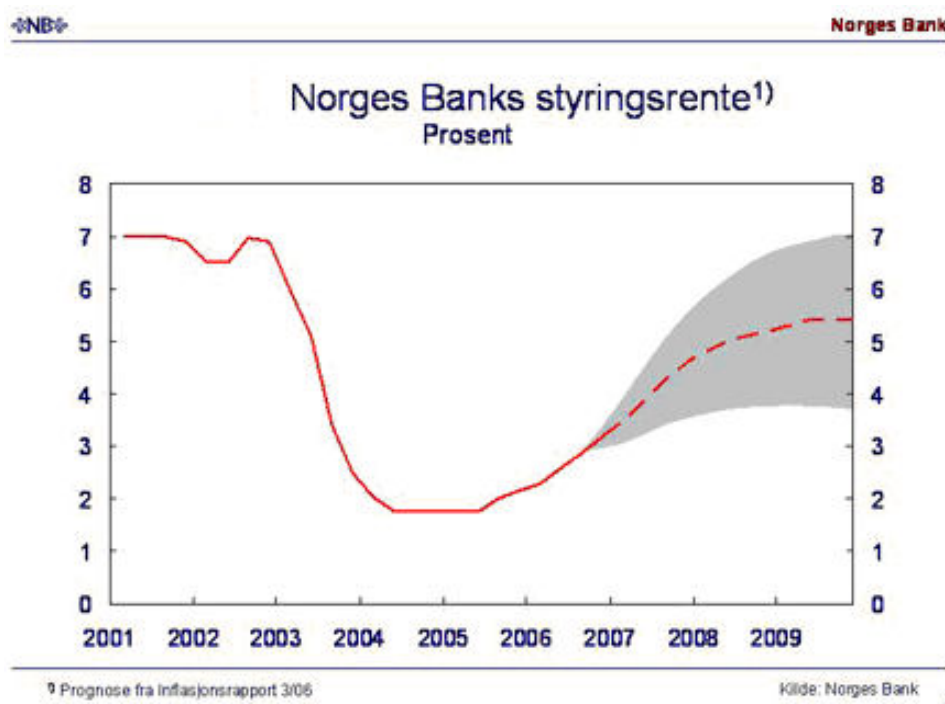


Rentene har begynt å øke igjen etter å ha ligget på et meget lavt nivå siden 2003. Siden da har etterspørselen og produksjonen tatt seg opp, sysselsettingen økt og ledigheten falt. Oppgangen er nå i en moden fase. På grunn av et høyt aktivitetsnivå og utsikter til vekst i inflasjonen så har Gjedrem sagt at han ser for seg en gradvis økning av styringsrenta opp mot et mer normalt nivå på noe over 5 %. På forrige rentemøte, 16. mars, ble styringsrenten satt opp til 4 %. Utlånsrenten ligger omtrent ett prosentpoeng over styringsrenten avhengig av sikkerheten på lånet, men den sterke konkurransen blant bankene kan redusere dette gapet noe. Under er renten og renteutsiktene illustrert, og det grå feltet viser med 70 % sannsynlighet hva renten kommer til å bli.

⁴

http://www.manpower.no/MPNet3/Content.asp?NodeRef=42317&Ref=NORWAY_JOB_BOARD&LangID=no

Figur 3: Norges Banks styringsrente. Figuren er hentet fra sentralbanksjefens årstale.



Hva angår markedsrentene har de kortsiktige rentene vært på et uvanlig lavt nivå de siste årene, noe som antakelig har bidratt til å øke boligprisveksten. Dette vil trolig snu når rentene økes igjen. De lange rentene gir informasjon om fremtidsutsiktene for den kortsiktige renten. De lange rentene er svært lave i dag, og vil sannsynligvis stige i årene som kommer.⁵

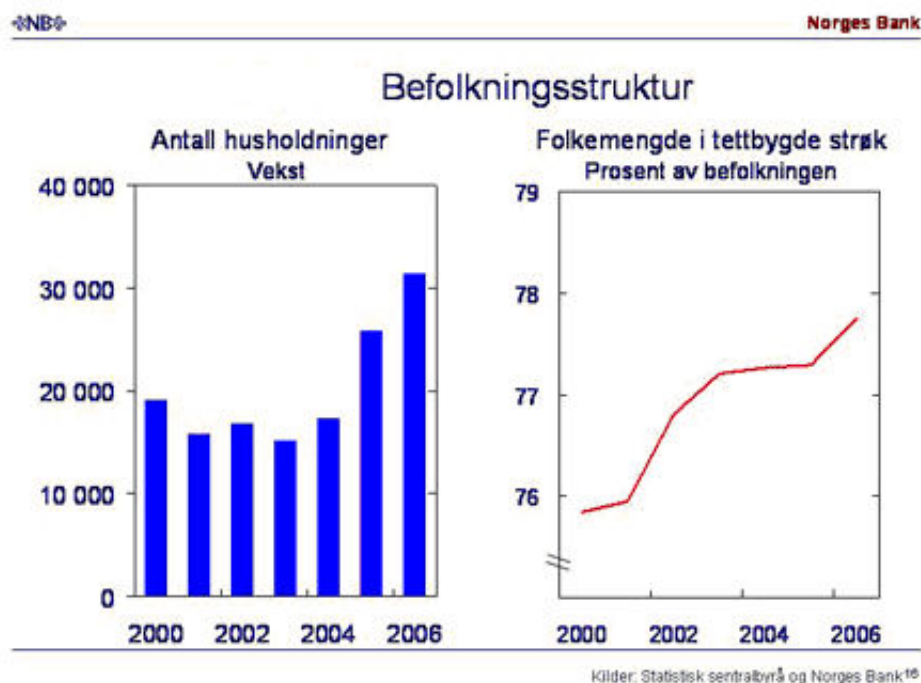
Antall boliger i Norge er ca 2,2 millioner, og endres lite fra år til år. I fjor ble det bygd i overkant av 30000 nye boliger, og boligmassen økte med litt over 1 %. I tillegg til at det tar tid å bygge nye boliger, så preger også den begrensede tilgang på håndverkere og materialer markedet. For eksempel informerer Glava AS om full produksjon og ekstra import av isolasjonsmateriale utover sommeren. Nye byggeforskrifter og økte krav til standard, i tillegg til presset på ressursene, bidrar til at boligprisene holder seg oppe. Ved siden av byggekostnadene, spiller selvsagt også tomteprisene en rolle. Er det knapphet på tomter, vil tomteprisene stige, og dermed stiger også boligprisene.

Høye boligpriser og høye bokostnader er i utgangspunktet en motvirkende kraft mot for sterke flyttestrømmer til sentrale strøk, samtidig som gunstige arbeidsmarkedsforhold i sentrale strøk bidrar til å stimulere flyttingen. Det er en klar trend for sentralisering i flyttemønsteret, og det

⁵ Per 31/05/2007 ligger rentenivået på 10 års norske statsobligasjoner på 5 %.

har vist seg at med høykonjunkturer følger vanligvis at flere enn ellers flytter til mer sentrale strøk. Figuren under viser at i 2006 bor nesten 78 % av befolkningen i tettbygde strøk.

Figur 4: Befolkningsstruktur. Figuren er hentet fra sentralbanksjefens årstale.



Det er rimelig å anta at forholdet mellom boligpriser og markedsleier er relativt stabilt på lang sikt. Øker boligprisene i større grad enn leiekostnadene, så blir det relativt mer lønnsomt å leie fremfor å eie. Dette bidrar til å presse husleiene opp og boligprisene ned. Særlig dersom kostnadene ved å eie i vesentlig grad overstiger kostnadene ved å leie, vil mange nyetablerende velge å leie midlertidig. Når det blir færre etterspørere så tar det lengre tid før boliger blir solgt, og behovet for å få inn pengene fra salget for eksempel for å betale for neste bolig gjør at en del selgere aksepterer lavere priser. Et problem er imidlertid at husleier har vært sterkt regulert i lang tid hvilket innebærer at et forholdstall kan gi misvisende resultater når husleiene nå har gått over til å bli mer og mer markedsbestemte. Det eksisterer heller ikke lange tidsserier for markedsleier, for eksempel går markedsleien for private utleieboliger i Oslo kun tilbake til 4.kvartal 2003.⁶

Tilbakedatert vekst kan være en årsak til videre vekst, og hvis dette er den eneste vesentlige grunnen til boligprisvekst kan det oppstå en oppadgående boligprisspiral. I en slik spiral så kjøper alle en bolig fordi alle andre gjør det. Dette er sammenfallende med en boligprisboble.

⁶ <http://boligbygg.reeltime.no/>

Med en slik kjøpsatferd kommer neppe et prisfall siden selv de rasjonelle aktørene blir med på kjøpefesten. Fordi noen planlegger å være med på oppturen for så å hoppe av i tide, kan oppgangen forsterkes. Men når disse spekulantene får kalde føtter og begynner å selge seg ut, kan resultatet bli et fall i boligprisene.

Noe som er bekymringsfullt er den økende gjelden hos husstandene i dag. Særlig unge mennesker tar opp lån til pipa, uten særlig mulighet til å klare et betydelig hopp i renta. Faktisk er gjelden i norske husholdninger snart 2 ganger disponibel inntekt. Den økte gjelden gjør husholdningene sårbare og vil etter hvert begrense betalingsviljen i boligmarkedet. Derfor kan for eksempel gjeldsgraden (gjeld/egenkapital) være et mål som muligens bidrar til å forklare utviklingen i boligprisene.

Mye av den prisstigningen som har vært i boligprisene er innhenting av prisfallet som var på slutten av 80- og begynnelsen av 90-tallet. I tillegg har det vært gode tider for norsk økonomi. Men for 2006 tyder det på at en betydelig del av boligprisveksten ikke kan forklares av tidligere mønstre, hvor veksten i stor grad ble forklart av inntekter, arbeidsledighet, renter og boligbygging. Selv om tilbudet har begynt å øke, og de korte rentene stiger, så har dette ikke gitt utslag i boligprisene enda. Dette kan blant annet skyldes at rentene har vært på et meget lavt nivå, og når de stiger til et mer normalt nivå på 5 % - 6 % vil vi trolig se et utslag.

Selv om tidligere mønstre ikke lenger følges i samme grad, så trenger det ikke å komme et krakk, boligmarkedet går nemlig i lange sykluser, og dermed går også endringene relativt tregt (DN privat økonomi, mars 2007). Allikevel er det verdt å nevne at målt i forhold til gjennomsnittsinntektene nærmer vi oss et rekordnivå, og er like ved der markedet kollapset sist. I forhold til byggekostnader, leiepriser og den generelle prisveksten har boligprisene aldri vært høyere (DN privat økonomi, mars 2007).

Psykologi er viktig hva angår forventningene, og kan være en av årsakene til at innstramningen i pengepolitikken ikke har gitt utslag i boligprisene enda. I dag er det en generell oppfatning om at det er en "god investering" å kjøpe egen bolig. I tillegg tror de fleste at de vil ha større kjøpekraft om 5 år enn i dag. Om forventningene endres så vil boligene skiftes ut sjeldnere, noe som kan medføre et fall i boligprisene. I aviser og tidsskrifter går flere eksperter ut i disse dager og advarer mot et opphetet boligmarked. Hvis dette fører til at flere avventer kjøp, og dersom mange nok tenker slik, kan det bli et prisfall.

Det er rimelig å forvente at aktørene i boligmarkedet baserer sine valg om boligkjøp på en gjennomtenkt vurdering av den kommende utviklingen i kostnadene ved å eie. Allikevel er mange overoptimistiske og overvurderer markedet, de tror at renten vil holde seg lav, og at prisveksten vil fortsette i den samme takten som i de seneste årene. Gjedrem uttrykte i årstalen at boligmarkedet kunne være preget av eufori. Vi ser i dag at husholdningen tar større sjanser enn før og at mange opptrer fullstendig uten bekymring. I tillegg råder optimismen og mange tror kanskje at boligprisene bare vil fortsette å stige. Dersom forventningene blir overvurdert kan det resultere i en boligprisboble.

2.5. Boligprisboble?

En boligprisboble kan oppstå dersom mange ønsker å kjøpe bolig i dag fordi de forventer at boligprisene skal stige fremover, og disse forventningene ikke er knyttet til fundamentale forhold. Dersom det er en prisboble i boligmarkedet, kan prisene falle kraftig hvis prisforventningene skifter (Jacobsen og Naug, 2004).

I Sentralbanksjefens årstale sammenligner Gjedrem dagens situasjon med trekk fra tidligere perioder, hvor det har vært stigende boligpriser etterfulgt av et krakk. Det er nyttig å betrakte tidligere situasjoner med felles trekk fra dagens situasjon for å observere og lære av tidligere erfaringer. Nedgangen i boligprisene rundt 1990 skyldtes de høye realrentene som behøvdes for å temme inflasjonen og stabilisere valutakursen, og har få fellestrekk med i dag. Perioden på slutten av 1890-tallet har derimot enkelte fellestrekk med det vi opplever i dag. Perioden var preget av stigende boligpriser, høy boligbygging i Kristiania, økende aksjekurser og sterk vekst i kreditt. Samtidig var prisnivået på varer og tjenester stabilt. Etter boligkrakket falt prisene i Kristiania med over 50 % i løpet av 5 år, mens det på landsbasis var mer moderate utslag i prisene. Spørsmålet blir da om vi bør forvente et krakk i dag? Er boligprisene overvurdert i forhold til en fundamentalverdi bestemt av faktorer som for eksempel rente, inntekt, arbeidsledighet og nybygging? Dette diskuteres i Jacobsen og Naug sin modell for boligprisveksten som jeg vil gå i dybden på i kapittel 4.

2.6. Andre mulige forklaringsfaktorer.

Dette avsnittet ser på en undersøkelse som benytter noen faktorer som Jacobsen og Naug ikke vurderte i sin modell.

International monetary fund (IMF) publiserte i september 2004 en artikkel kalt "The global house price boom" som blant annet studerte fluktuasjonene i realboligprisene i industrialiserte land. Estimeringsperioden er fra 1971 til 2003, og uvalget består av 18 OECD-land. Modellen inneholder de signifikante forklaringsfaktorene boligprisvekst forrige år, forholdstallet huspriser delt på inntekt, den reelle disponible inntekten, kortsiktig rente, reell kreditt, reell aksjeprisvekst i forrige år, befolkningsvekst og en dummy for bankkriser.

Lagget boligprisvekst fikk en ganske sterk, positiv effekt på boligprisveksten med en koeffisient på 0,521 i deres modell. Jo større denne koeffisienten er, jo mer vedvarende blir vekstraten i boligprisene. Dersom koeffisienten overstiger 1 i absoluttverdi er vekstraten i boligprisene eksplosiv. Høye, signifikante koeffisienter på denne variabelen kan indikere en fare for boligprisbobler.

Dersom boligprisveksten på lang sikt beveger seg tilbake til et nivå bestemt av fundamentale faktorer, så impliserer dette at boligprisene vil falle når de i en periode har økt relativt mye mer enn inntektene. Forholdstallet boligpris per inntjent krone vil i så fall ha et negativt fortegn, og boligprisene vil avta når for eksempel boligprisen øker mye relativt til inntekten.

I analysen finner de at inntekten påvirker boligprisveksten positivt, mens rente har en negativ effekt. Videre benytter de reell kreditt som en proxy variabel på pantegjeld, vekstraten i aksjeprisene fanger opp muligheten til å restrukturere porteføljen av aktiva, befolkningsveksten brukes som en proxy på veksten i husholdningene og en dummy variabel tar verdien 1 ved bankkriser. Økt reell kreditt eller økt vekstrate i aksjeprisene forrige år vil føre til økt boligprisvekst. Bankkriser assosieres typisk med et fall i boligprisene, og har således en negativ effekt på boligprisen. Demografiske faktorer påvirker også boligprisene, økes populasjonsveksten så økes inflasjonen i realboligprisene.

Av disse forklaringsfaktorene har jeg ikke diskutert boligpris per inntjent krone, pantegjeld, vekstraten i aksjeprisene og bankkriser så langt.

Boligpris delt på inntekt sier noe om hvor mange års inntekt som trengs for å tjene opp det boligen kostet. Forfatterne argumenterer for at dette forholdstallet er et slags mål på fundamentalverdi, og at avvik vil justeres inn med tiden.

Pantegjeld viser til hvor mye av gjelden som er sikret med pant. Pant i for eksempel bolig gjør at bankene aksepterer et høyere lån enn de ellers ville ha gjort. Dette innebærer at etterspørerne får større betalingsevne som igjen kan gi utslag i økte boligpriser. Se for øvrig diskusjonen i avsnitt 2.4. om økt gjeld til husholdningen.

Dersom aksjemarkedet gir høy avkastning vil det være noen som heller vil plassere pengene sine i fond og aksjer fremfor i boliger. Det kan for eksempel hende at avkastningen på midlene plassert i aksjemarkedet fratrukket leiekostnader overstiger verdistigningen på boligen etter at man har tatt hensyn til bokostnadene. Er det tilfellet sitter man igjen med netto økonomisk gevinst hvis man tar pengene ut av boligmarkedet og setter dem inn i aksjemarkedet. I Norge fins det underliggende normer som har gjort det til status å eie egen bolig, og boligkjøp er blitt til ”noe man bare gjør”. Som tidligere nevnt eier ca 80 % av husholdningene sin egen bolig. Derfor er det ikke sikkert at denne variabelen vil være signifikant på norske data. Allikevel kan en slik variabel være nyttig å inkludere fordi den fanger opp spekuleringsaspektet i boligmarkedet.

Det var en bankkrise i Norge på slutten av 80-tallet, og hadde vi hatt data for periodene før og etter dette ville trolig dummy variabelen vært signifikant og hatt en negativ effekt på boligprisene. I dag er imidlertid bankvesenet bedre utformet, og trolig vil ikke en bankkrise hvor boligmarkedet ligger til grunn komme igjen. Skulle det oppstå en bankkrise er det heller andre forhold som for eksempel krig som er årsaken.

2.7. Avsluttende kommentar.

I dette kapittelet har jeg argumentert for at en rekke variabler påvirker boligprisene, og det ville ha vært interessant å ha undersøkt effektene deres i den kommende analysen. På grunn av mangel på tid fikk jeg dessverre ikke samlet inn kvartalsvise data for disse mulige forklaringsfaktorene, og dette innebærer at analysen i kapittel 5 bygger på vurderingene til Jacobsen og Naug om hvilke forklaringsvariabler som påvirker boligprisene.

3. Statistisk teori.

I dette kapitlet vil jeg presentere teori som er nødvendig for å kunne bygge og analysere modellene i senere kapitler.

3.1. Modell spesifikasjon.

God vurderingsevne og erfaring er to sentrale faktorer når man skal estimere økonometriske modeller og tolke dem. Til å velge ut forklaringsfaktorene i modellen vår kan vi først og fremst bruke økonomisk teori. Men for å vite hvilken form de ulike variablene skal ha, dvs. om variabelen skal være lagget, ha dagens verdi eller forventet verdi, er det viktig å forstå hvordan markedet fungerer. Markedsforståelse og økonomisk teori er dessuten viktig for å vite hva som bør være avhengig variabel, om variablene skal transformeres og om det bør konstrueres interaksjonsledd. Dersom økonomisk teori sier at en variabel bør være med i modellen, er det gode grunner for å ha den med uavhengig av hva statistisk teori sier.

Før vi kan inkludere variablene i en modell, så må vi ta en nærmere titt på dataene for å avdekke deres statistiske egenskaper. Tabeller av variablene avslører om det er missing values og om det er blitt gjort store inntastningsfeil. Histogrammer og grafer kan på sin side gi en pekepinn på serienes stasjonaritetsegenskaper, ekstremverdier og sjokk. Det er imidlertid viktig å huske at man ikke må lage en hypotese ut fra et mønster man ser i dataene, men ut fra økonomisk teori. Derfor skal hypotesen være klar før man betrakter dataene.

3.2. Minste kvadraters metode.

Vi tar utgangspunkt i en enkel modell med to variabler, X og Y:

$$Y_t = \hat{\beta}_0 + \hat{\beta}_1 X_t + \varepsilon_t$$

Her betegner β_0 konstantleddet, β_1 viser heldningskoeffisienten, ε er restleddet og t angir hvilket tidspunkt vi ser på.

Dersom vi tror at variabelen vi ønsker å forklare avhenger lineært av en rekke forklaringsfaktorer så kan vi benytte minste kvadraters metode, eng. ordinary least square

(OLS), til å estimere parametrene i modellen. Det er imidlertid fire betingelser som må være oppfylt for at OLS skal gi estimater som er BLUE⁷ (Brooks 2005):

- Forventningen til feilleddene er lik null, $E(\varepsilon_t) = 0$.
- Feilleddene har konstant, endelig varians, $\text{Var}(\varepsilon_t) = \sigma^2 < \infty$.
- Det er ingen kovarians mellom feilleddene over tid, $\text{Cov}(\varepsilon_t, \varepsilon_s) = 0$.
- Det er ingen kovarians mellom feilleddet og forklaringsvariabelen, $\text{Cov}(\varepsilon_t, X_t) = 0$.
Dersom den første antagelsen holder så er $\text{Cov}(\varepsilon_t, X_t) = E(\varepsilon_t X_t) = 0$.

En femte betingelse kreves for at inferenstesting av populasjonsparametrene skal være gyldig:

- Feilleddene er normalfordelt.

Dette har sammenheng med at t - og F -tester antar normalfordelte data.

3.3. Stasjonaritet.⁸

3.3.1. Begrepet.

Ikke-stasjonaritet er et viktig tema hva angår tidsserier siden veldig mange økonomiske variabler er ikke-stasjonære.

En variabel, X_t , er stasjonær når dens statistiske prosesser er konstante, dvs. hvis dens gjennomsnitt, varians og kovarians ikke endres over tid. Når dette er tilfelle skriver vi $X_t \sim I(0)$, som betyr at variabelen er integrert av nulte orden. Hvis variabelen er stasjonær kan den benyttes i en regresjonsanalyse uten problemer.

Vi har imidlertid et problem dersom en eller flere av de statistiske prosessene endres over tid. Serien er da ikke-stasjonær og den vil ikke returnere til gjennomsnittet når tiden går. Variabelen skaper i så fall problemer for regresjonsanalysen, resultatene vi oppnår vil ikke ha noen mening. Vi kan til og med få spuriøse resultater, hvilket betyr at vi finner en sammenheng mellom variabler som ikke eksisterer. Spuriøse regresjoner kan oppstå når variablene har trend eller når de er integrerte, for eksempel hvis begge seriene er random

⁷ BLUE er en forkortelse for Best Linear Unbiased Estimators. Dvs. at dersom de fire betingelsene er oppfylt vil OLS gi oss estimater som er forventningsrette og de beste i betydning lavest varians blant alle lineære estimater.

⁸ Innen statistikken skiller man mellom begrepene streng og svak stasjonaritet. I praktisk økonometri og anvendelse er streng stasjonaritet sjelden oppfylt, men svak stasjonaritet har mindre strenge krav. I denne oppgaven er det svak stasjonaritet som menes når begrepet stasjonaritet nevnes (Helgeland Qvale 2001).

walks⁹ (Wooldridge 2003). Gitt at vi har en ikke-stasjonær variabel vil vi teste om den er integrert av første orden, hvilket skrives som $X_t \sim I(1)$. Dersom vi har en $I(1)$ variabel kan vi differensiere den en gang for å gjøre den stasjonær. Resultatet vil bli meningsfullt når vi foretar en regresjon med første differansen, $\Delta X_t = X_t - X_{t-1}$. Hvis variabelen ikke er $I(1)$ kan vi sjekke om den er integrert av andre orden. De fleste økonomiske variabler er dog integrert av første orden.

Sjokk på systemet vil gradvis dø ut når tiden går for en stasjonær serie. Dette er ikke tilfelle for ikke-stasjonære serier, hvor sjokk fører til en permanent endring i tidsserien slik at vi ikke returnerer til gjennomsnittet. Dette kan demonstreres ved at sjokkene akkumuleres over tid og legges til utgangspunktet X_0 (vi tar også hensyn til en evt. deterministisk trend T):¹⁰

$$X_t = X_0 + \lambda T + \sum_{j=0}^t \varepsilon_j$$

3.3.2. Dickey-Fuller-testen.

Når vi vil teste om en variabel er stasjonær kan vi bruke Dickey-Fuller-testen. DF-testen er en test for enhetsrot. Det første steget er å ta førstedifferansen av variabelen vår.

$$X_t = \mu + \rho X_{t-1} + \lambda T + \varepsilon_t$$

$$X_t - X_{t-1} = \Delta X_t = \mu + (\rho-1)X_{t-1} + \lambda T + \varepsilon_t = \mu + \beta X_{t-1} + \lambda T + \varepsilon_t$$

Her betegner μ driften, T tar hensyn til en evt. deterministisk trend (dersom serien vokser med en fast periodisk rate) og ε betegner restleddet.

Nullhypotesen er at serien er integrert av første orden, $X_t \sim I(1)$, dvs. at den er ikke-stasjonær på nivådata. Vi tester om variabelen er en random walk med drift (μ) og deterministisk trend (T), altså om $\rho=1$. Dette er ekvivalent med at $\beta=0$. Vår alternativhypotese er det ensidige alternativet at β er mindre enn null, da er variabelen stasjonær, $X_t \sim I(0)$. β større enn null er

⁹ Random walk betyr at tidsseriens verdi i neste periode er dagens verdi pluss et ukjent feilledd (Wooldridge 2003). Man vet aldri om verdien går opp eller ned eller størrelsen på hoppet, all historisk informasjon er uten verdi. Serien følger med andre ord en tilfeldig gang.

¹⁰ <http://www.erf.org.eg/html/btrade1.pdf>

ikke realistisk siden vi da vil få en eksplosiv vekst. Dersom vi ikke kan forkaste nullhypotesen så må vi gjøre en ny test hvor H_0 er at $X_t \sim I(2)$ og H_A er at $X_t \sim I(1)$. Forkastningsregionene finner vi i Dickey-Fuller tabellen.

Dickey-Fuller antar at restleddene i modellen er hvitt støy. Hvis vi tror dette kan være et problem i vår modell kan vi velge den utvidede DF-testen, eng. the augmented Dickey-Fuller test (ADF-testen). Denne inkluderer det antall laggede verdier av ΔX_t som skal til for at restleddet blir hvitt støy.

$$\Delta X_t = \mu + \beta X_{t-1} + \lambda T + \gamma_1 \Delta X_{t-1} + \gamma_2 \Delta X_{t-2} + \dots + \gamma_n \Delta X_{t-n} + \varepsilon_t$$

Det kan det være lurt å plote seriene til å begynne med, da ser man raskt om serien inneholder et trendelement, om det bør være med et konstantledd og om man bør trekke ut en sesongfaktor for å nevne noe. Antall lags (n) av ΔX_t vi skal estimere kan vi finne på to måter. Enten kan vi bruke tommelfingerregelen og ta med så mange lags som er signifikant ut fra student t – fordelingen. Alternativet er å bruke informasjonskriteriet Akaike's IC (AIC) eller Schwartz's Bayesian IC (SBIC) for å finne det ut. Begge informasjonskriteriene foretrekker den modellen som gir lavest verdi på kriteriet. SBIC velger asymptotisk sett riktig antall lags (konsistent), mens AIC tenderer til å foretrekke flere lags enn nødvendig. På den andre siden fører valg av modell ut fra SBIC i gjennomsnitt til større varians enn ved bruk av AIC. Ingen av kriteriene er derfor overordnet det andre (Brooks 2005). Det er viktig å prøve å finne det optimale antall lags av forklart variabel som skal inkluderes i regresjonen. Tar man med for få lags vil ikke all autokorrelasjonen bli fjernet og tar man med for mange lags så øker standardavviket til koeffisientene. Denne siste effekten oppstår fordi flere parametere i en regresjon bruker opp frihetsgradene. Gitt alt annet likt vil derfor absoluttverdien av teststatistikken minke, noe som fører til at nullhypotesen om enhetsrot vil bli forkastet for sjeldent (Brooks 2005). Med andre ord vil man konkludere med ikke-stasjonaritet i tilfeller som er stasjonære.

Problemet med (A)DF-testene er at de har lav styrke. Den tenderer til å forkaste nullhypotesen når den faktisk er sann og ikke forkaste nullhypotesen når den bør bli forkastet.

3.4. Kointegrasjon.

Dersom man ønsker å se på sammenhengen mellom to eller flere ikke-stasjonære variabler vil man kjøre en regresjon med førstedifferansen til variablene, dvs. på deres endringsform. Resultatet vi oppnår er gyldig, men holder kun på kort sikt. For å finne ut om det eksisterer en langsiktig sammenheng mellom variablene må vi undersøke om de er kointegrert.

For å få et forhold til kointegrasjon kan man tenke seg at kointegrasjon innebærer at to tidsserier er bundet sammen med et bånd, slik at de på lang sikt vil henge sammen, men på kort sikt kan tauet være slakt eller det kan være stramt. Seriene vil aldri kunne løpe ubegrenset fra hverandre, selv om de kan synes å ha en forskjellig utvikling på kort sikt.

Definisjonen av kointegrasjon er at to ikke-stasjonære variable er kointegrert hvis en lineær kombinasjon av dem er stasjonær (Wooldridge 2003). Denne lineære kombinasjonen kalles kointegrasjonsvektoren. Vi kan enten estimere kointegrasjonsvektoren eller definere den med utgangspunkt i økonomisk teori. Når det er flere enn to variabler i et system vil det i prinsippet kunne eksistere mer enn en kointegrasjonsvektor.

3.5. Statistiske og dynamiske modeller.

En statistisk modell tillater kun en sammenheng mellom variabler fra samme tidsperiode. Det betyr at en endring i forklaringsvariabelen på tid t slår umiddelbart og fullt ut i forklart variabel.

$$\text{Eksempel: } y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \varepsilon_t$$

En dynamisk modell vil på den annen side også ta hensyn til tregheter. Det kan hende at en endring i en av forklaringsvariablene i dag (t) ikke endrer forklart variabel så mye i dag, men heller påvirker denne mer i morgen ($t+1$). I så fall bør vi inkludere laggede verdier av forklaringsvariablene og eventuelt forklart variabel. Modeller som har med lags av både forklart variabel og forklaringsfaktorene kalles ”autoregressive distributed lag (ADL) models”

$$\text{Eksempel: } y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \gamma_1 y_{t-1} + \gamma_2 x_{1t-1} + \gamma_3 x_{2t-1} + \varepsilon_t$$

Hvis vi trekker fra y_{t-1} på begge sider av likhetstegnet, samtidig som vi legger til og trekker fra $\beta_1 X_{1t-1}$ og $\beta_2 X_{2t-1}$ på høyre side av ligningen, vil ADL-modellen over kunne se ut som følger:

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_{1t} + \beta_2 \Delta x_{2t} + (\gamma_1 - 1)y_{t-1} + (\gamma_2 + \beta_1)x_{1t-1} + (\gamma_3 + \beta_2)x_{2t-1}$$

En omskrivning av dette gir oss følgende uttrykk:

$$\Delta y_t = \beta_0 + \beta_1 \Delta x_{1t} + \beta_2 \Delta x_{2t} - (1 - \gamma_1) \left[y_{t-1} - \left(\frac{\gamma_2 + \beta_1}{1 - \gamma_1} \right) x_{1t-1} - \left(\frac{\gamma_3 + \beta_2}{1 - \gamma_1} \right) x_{2t-1} \right]$$

Dette er en såkalt feiljusteringsmodell som jeg kommer tilbake til i neste avsnitt.

Når man går over fra statiske til dynamiske modeller så vil eventuell autokorrelasjon reduseres eller forsvinne helt, så lenge autokorrelasjonen ikke skyldes spesifikasjonsfeil. Laggede forklaringsvariabler kan altså fjerne autokorrelasjon, men dette skjer på bekostning av å ha en modell som er lett å tolke. Dessuten, hvis avhengig variabel er lagget brytes forutsetningen om ikke-stokastiske forklaringsfaktorer som er en av betingelsene (nr.4) for at OLS skal gi BLUE. Så lenge autokorrelasjonen fjernes ved å inkludere lagget forklart variabel, så vil estimatene være forventningsskjevne, men konsistente¹¹. Konsistente estimater konvergerer mot den sanne verdien når utvalget går mot uendelig, dvs. asymptotisk, og er derfor et mindre strengt krav enn forventningsrettet.

3.6. Feiljusteringsmodeller.

3.6.1. Generelt.

En feiljusteringsmodell, eng. error correction modell (ECM), er en dynamisk modell som i tillegg til å løse evt. spuriøsitet, også studerer den langsiktige likevekten, og hvor raskt avvik fra likevekt justeres inn. Feiljusteringsmodellen inneholder slik både kortsiktige og langsiktige elementer. En fordel med ECM er at de standard regresjonsteknikkene er gyldige hvis kointegrasjon foreligger, siden alle uttrykkene i modellen er stasjonære. Da kan man også

¹¹ Betingelsene for at estimatoren skal være konsistent er at variansen er konstant og endelig, at forventningen til feilleddet er null og at det ikke er noen kovarians mellom forklaringsvariabelen og feilleddet.

inkludere andre stasjonære variabler i ligningen uten at det skaper problemer. Enda en fordel har vi i Granger's representasjonsteorem, som sier at hvis to variabler er kointegrert, så må det eksistere en ECM og vica versa. En konsekvens av teoremet er at feiljusteringsmodellen er immun mot problemet med spuriøse regresjoner, så lenge variablene kointegrerer (Harris 1995).

Det fins ulike måter å estimere en ECM på. Engle-Granger 2-steps metode er en populær fremgangsmåte som er lett å forstå. Alternativt kan man estimere alt i ett trinn, kalt 1-steps ECM. For mer komplekse modeller foretrekkes ofte Johansens metode, som er mer teknisk. I denne oppgaven kommer jeg til å se på de to første metodene.

3.6.2. Engle-Granger 2-steps metode.

Steg 1- feilkorreksjonsleddet.

Anta at x og y er to $I(1)$ -variabler som vi ønsker å undersøke om er kointegrert. Det første vi gjør er å bruke OLS til å estimere kointegrasjonsvektoren. Det betyr at vi estimerer en regresjon med de $I(1)$ -variablene som vi tror har en sammenheng på lang sikt. La oss si at vi får $y_t = \gamma x_t + \varepsilon_t$. Enhver lineær transformasjon av denne kointegrasjonsvektoren $[1, -\gamma]$ vil også være en potensiell kointegrasjonsvektor (Brooks 2005). Det neste vi gjør er å ta vare på feilleddet fra regresjonen, $\varepsilon_t = y_t - \gamma x_t$, og sjekke om det er stasjonært. Vi sjekker dette ved å benytte en (A)DF-test på feilleddet. Hvis feilleddet er stasjonært er variablene kointegrert med kointegrasjonsvektoren $[1, -\gamma]$ og vi kan gå videre til steg to. Hvis feilleddet ikke er stasjonært så fins det ingen langsiktig likevekt, og vi må nøye oss med den kortsiktige modellen som kun inneholder variablene på differanseform.

Steg 2-ECM.

Hvis vi får at feilleddet er stasjonært i steg en kan vi modellere en ECM for å finne ut hvordan prosessen mot likevekten går. Et eksempel på en slik modell er:

$$\Delta y_t = \beta_1 \Delta x_t + \beta_2 (y_{t-1} - \gamma x_{t-1}) + u_t = \beta_1 \Delta x_t + \beta_2 (\varepsilon_{t-1}) + u_t$$

Parentesen over svarer til det laggede feilleddet fra steg 1. Tolkningen av modellen er at y endres mellom t og $t-1$ etter endringene i x i samme periode. I tillegg korrigeres det for en ulikevekt som eksisterte i forrige periode, $t-1$. γ definerer den langsiktige sammenhengen

mellom x og y , mens β_1 beskriver korttidssammenhengen mellom endringer i x og endringer i y . β_2 sier hvor lang tid det tar før avvik justeres inn, dvs. den måler hvor stor andel av forrige periodes ulikevekt som korrigeres inn (Brooks 2005). β_2 vil alltid ligge mellom 0 og -1. Dersom β_2 er lik -1 betyr det at vi har en momentan tilbakevending til likevekten, og er β_2 lik 0 vil vi aldri komme tilbake til likevekten. Det er lov å gjøre tester på parameterne man oppnår i steg 2 siden alle variablene i denne regresjonen er stasjonære.

Engle-Granger 2-steps metode har en rekke empiriske svakheter. Et problem med denne metoden er at testene om enhetsrot og kointegrasjonstestene mangler styrke. Et annet problem er at valget av forklart variabel og forklaringsvariabel kan påvirke konklusjonen vår. I tillegg kan vi få skjeve estimater på små utvalg. Disse problemene forvinner asymptotisk. Et siste problem er at man ikke kan utføre hypotesetester om den faktiske sammenhengen til kointegrasjonsvektoren i steg 1 siden regresjonen inneholder ikke-stasjonære variabler (Brooks 2005). Dette kan resultere i at misspesifikasjoner i langtidsligningen, steg 1, ikke blir oppdaget og det vil påvirke resultatet i feiljusteringsmodellen i steg 2. Alt som blir gjort i steg 1 er derfor å finne koeffisientestimatene. Problemene forverres når man inkluderer flere enn 2 variabler, blant annet er det ikke sikkert at kointegrasjonsvektoren lenger er unik.

Tross problemene, er Engle-Granger 2-steps metode populær. Fordelene med denne metoden er at den er lett å bruke, at den er superkonsistent¹² og at inferenstesting om koeffisientene i steg 2 er gyldig. Hvis vi ignorerer endogene variabler når alle forklaringsfaktorene er stasjonære, så vil OLS gi forventningsskjeve estimater og dessuten estimater som ikke er konsistente. Restleddet fanger opp eventuelle utelatte variabler som leder til autokorrelasjon. Men dette er ikke et problem ved superkonsistens, for OLS vil gi konsistente estimater likevel. Ved simultanitetsproblemer er derimot forklaringsfaktoren korrelert med feilleddet og estimatoren er ikke lenger konsistent (Harris 1995).

3.6.3. 1-steps ECM.

Denne metoden ligner på Engle-Granger sin 2-steps metode, forskjellen er at man ikke lenger estimerer langtidsløsningen for seg selv. Isteden puttes alle variablene, både de på

¹² Med superkonsistens menes det at OLS-estimatet for to $I(1)$ variable vil nærme seg den sanne verdien mye raskere enn estimatet for OLS med bare stasjonære variabler, når restleddet er stasjonært og utvalget øker. Dette kommer av at $I(1)$ variablene dominerer $I(0)$ variablene asymptotisk.

endringsform og de på nivå, inn i modellen og alle koeffisientene estimeres. Deretter regner man seg tilbake for å finne de langsiktige koeffisientene, dvs. for $I(1)$ -variablene gitt på nivå.

Fordelen med å gjøre alt i ett steg er at man får høyere styrke sammenlignet med 2-steps metoden. I tillegg er 1-steps metoden overlegen 2-steps hva angår skjeve estimater i små utvalg, siden vi modellerer dynamikken og langsiktlikevekten samtidig. I 1-steps metoden så bruker man dessuten signifikansnivået til feilkorreksjonsleddet til å si om variablene er kointegrert, og man trenger ikke kjøre en egen DF-test på feilleddet. Davidson og MacKinnon sier også at inferenstesting om de individuelle koeffisientene er gyldig i 1-steps metoden hvilket er en stor fordel kontra 2-stepsmetoden.

Ulempene med 1-steps metoden er også her at testene om enhetsrot har lav styrke, at valg av avhengig variabel kan påvirke konklusjonen vår og at kointegrasjonsvektoren ikke nødvendigvis er unik. Selv om man kan gjøre inferenstester om de individuelle koeffisientene i modellen, så må man være forsiktig med å gjøre andre inferenstester. For eksempel vil ikke en sammenslått test av to koeffisienter ha en normal kjikvadratfordeling asymptotisk (Davidson, MacKinnon 1993).

3.7. Autokorrelasjon og tester for avdekke problemet.

3.7.1. Hva er autokorrelasjon?

Autokorrelasjon, også kalt seriekorrelasjon, er et av hovedproblemene i tidsserieøkonometri, og innebærer at feilleddet er korrelert mellom ulike tidsperioder. Slik systematikk i feilleddet gjør at koeffisientestimatene vi finner ikke lenger er effisiente, dvs. selv med store utvalg finner vi ikke nødvendigvis den modellen som minimerer variansen. Koeffisientestimatene vil fremdeles være forventningsrette, men estimatene vil ikke lenger være BLUE når vi benytter OLS. Når standardavviket er galt, vil vi kunne ta feil beslutning om relevansen av en variabel. For eksempel predikerer OLS for små standardavvik ved positiv autokorrelasjon, noe som gjør at vi tenderer til å forkaste nullhypotesen feilaktig. En årsak til autokorrelasjon kan være at modellen er feilspesifisert. Løsning er da å respesifisere modellen hvis vi har en modell med fullstendig dynamikk.

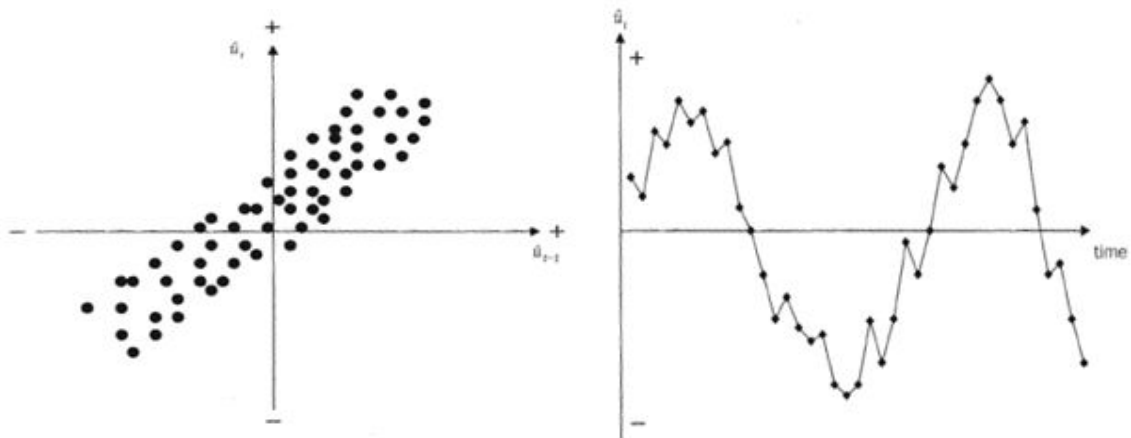
Man kan finne ut om autokorrelasjon er tilstede i modellen ved å betrakte et plott av restleddene eller ved å benytte tester som Durbin-Watson eller Box-Pierce.

3.7.2. Grafiske tester.

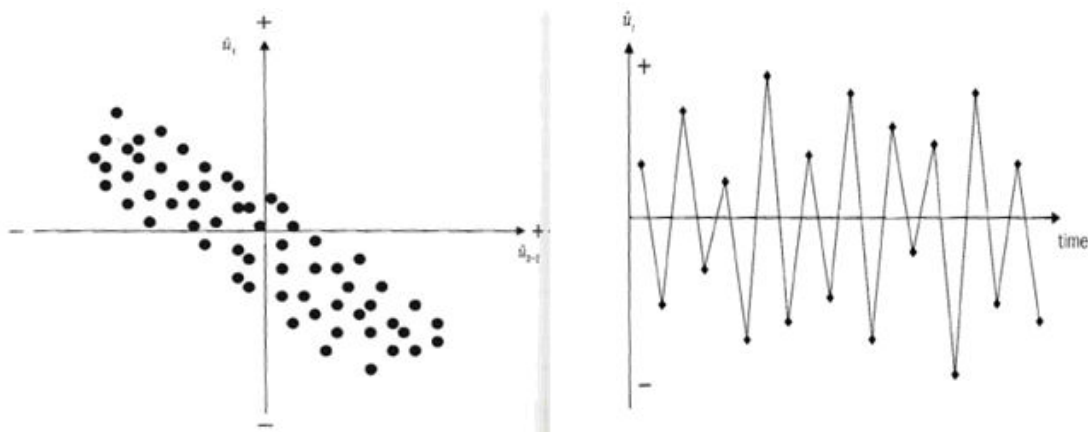
Hvis et plott av feilleddet viser et mønster er det autokorrelasjon til stede i modellen vår. For eksempel har vi positiv autokorrelasjon når restleddene har et syklisk plott over tid. Da vil i gjennomsnitt en positiv verdi på restleddet etterfølges av en positiv, og motsatt. Hvis vi har negativ autokorrelasjon vil det medføre at et restledd med positiv verdi vil etterfølges av et restledd med negativ verdi. Ingen autokorrelasjon vil ha en tilfeldig spredning av feilleddene. Fordi det kan være vanskelig å tolke plott av variabler bør man også utføre statistiske tester.

Illustrert:

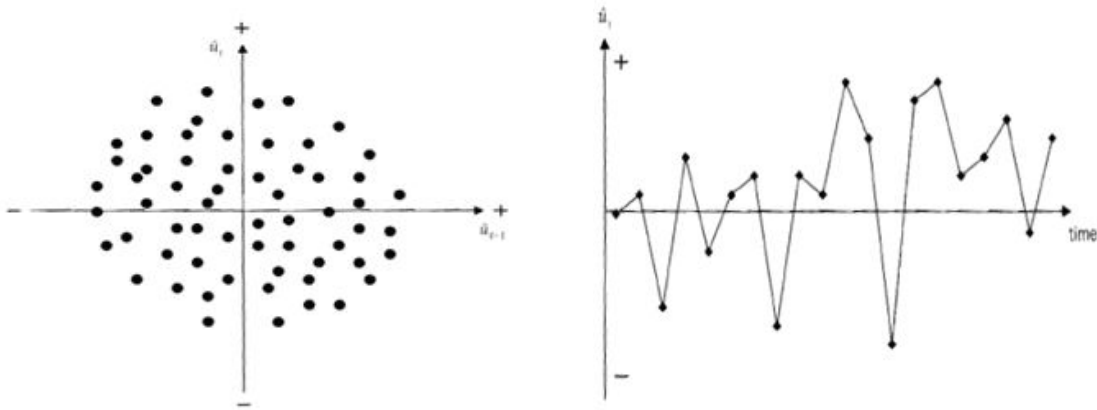
Figur 5: Positiv Autokorrelasjon.



Figur 6: Negativ autokorrelasjon.



Figur 7: Ingen autokorrelasjon.



3.7.3. Durbin-Watson testen.

Durbin-Watson (DW)-testen er en test for første ordens autokorrelasjon, dvs. autokorrelasjon mellom et feilledd og det foregående feilleddet.

Teststatistikken er gitt ved følgende formel:

$$DW = \frac{\sum_{t=2}^T (\hat{\varepsilon}_t - \hat{\varepsilon}_{t-1})^2}{\sum_{t=2}^T \hat{\varepsilon}_t^2}$$

DW- teststatistikken vil alltid ligge mellom 0 og 4. Dersom teststatistikken man oppnår er rundt 2 har vi ingen første ordens autokorrelasjon. Får vi en verdi som er nær 0 er det positiv autokorrelasjon tilstede, og kommer teststatistikken nær 4 har vi negativ autokorrelasjon i modellen. Øvre og nedre kritisk verdi er gitt i en egen tabell, Durbin-Watson tabellen. Disse to, sammen med de to vi får når vi trekker disse verdiene fra tallet 4, er de eksakte grensene som bestemmer når vi kan konkludere med at det er autokorrelasjon eller ikke, og når vi ikke kan avgjøre spørsmålet basert på denne testen.

Kravene for at DW skal være en gyldig test er at (Brooks 2005):

- Regresjonen må inneholde et konstantledd.
- Forklaringsvariablene må ikke være stokastiske.
- Det må ikke være noen lags av forklart variabel i regresjonen.

Dersom den andre eller tredje betingelsen blir brutt så vil teststatistikken bli skjev i retning 2, som tilsier ingen autokorrelasjon i DW-testen.

3.7.4. Box-Pierce og Ljung-Box testene.

En alternativ test for å avdekke potensiell autokorrelasjon er Box-Pierce sin Q-statistikk. Dette er en kombinert test som også sjekker for høyere ordens autokorrelasjon. Nullhypotesen er at det ikke foreligger autokorrelasjon, og Q-statistikken er asymptotisk kjikvadratfordelt med m frihetsgrader.

$$Q = T \sum_{k=1}^m \hat{r}_k^2$$

Her angir T antall observasjoner i utvalget, m er den maksimale lag lengden og r_k angir verdien på den estimerte autokorrelasjonskoeffisienten for et gitt antall lags k.

Box-Pierce testen fungerer ikke godt på små utvalg, den gir oss galt resultat for ofte. En variant av denne testen som fungerer bedre på små utvalg er Ljung-Box testen.

$$Q^* = T(T+2) \sum_{k=1}^m \frac{\hat{r}_k^2}{T-k}$$

Også denne Q^* -statistikken er kjikvadratfordelt med m frihetsgrader. Når utvalget øker mot uendelig vil de to testene være identiske.

3.8. Statistiske konsekvenser av spesifikasjonsfeil.

Utelatte variabler:

Dersom en variabel som er relevant for å forklare variasjonen i avhengig variabel blir utelatt, og denne er autokorrelert, skaper det alvorlige følger. Restleddene i den estimerte modellen vil da være autokorrelert. Estimaten vi oppnår blir forventningsskjeve og variansen er ikke nødvendigvis den laveste blant alle lineære, forventningsrette estimater. Dette gjør igjen at teststatistikken blir ugyldige. En test for utelatte variabler er Ramsey RESET testen.

Irrelevante variabler:

Dette er motstykke til å ta med for få variabler. Det vil si at man inkluderer variabler som ikke har noen relevant effekt på avhengig variabel. Dersom irrelevante variabler inkluderes i en regresjon så vil estimatene fremdeles være forventningsrette, men det gir oss ikke de estimatene som har lavest varians. Man oppdager irrelevante variabler ved at de ikke er signifikante og at de tilfører liten forklaringskraft til modellen.

Utelatte sesongfaktorer:

Dersom avhengig variabel inneholder et sesongmønster eller et syklisk mønster, og dette ikke blir tatt hensyn til i modellen, så vil restleddene i vår estimerte modell typisk være positivt autokorrelert. Dette løser man ved å inkludere dummyvariabler for sesongene. Dersom vi har kvartalsvise data kan vi enten inkludere 3 av de konstruerte dummyvariablene og et konstantledd, eller vi kan ta med en dummy for hver av de 4 sesongene, men da må vi passe på å droppe konstantleddet.

Feil funksjonsform:

Autokorrelasjon kan oppstå hvis vi påtvinger en lineær regresjonsmodell på to variabler som har en ikke-lineær sammenheng. Eksempler på ikke-lineære funksjoner er kvadratiske, logaritmiske og eksponentielle funksjoner. Ofte kan dette løses ved å transformere variablene før vi utfører regresjonen.

3.9. Data mining.

Data mining foregår når noen benytter det samme datasettet til å estimere et stort antall modeller for å finne den ”beste” modellen (Wooldridge 2003).

Hvis man prøver mange forskjellige varianter av modeller for å finne den som har høyest forklaringsgrad og som passer best på koeffisienter og prediksjoner, så kan det hende at man tilpasser seg utvalget av data for mye. Følgene av dette blir i så fall at man forklarer støyen i utvalget i tillegg til de faktiske underliggende effektene. De rapporterte p-verdiene er da for lave, og vil slik føre til at man kanskje ikke forkaster variabler som burde ha blitt fjernet. Poenget er altså at man overtilpasser modellen.

For å unngå data mining bør man velge forklaringsvariabler ut fra økonomisk teori og for eksempel teste modellen på et annet utvalg enn det som ble benyttet ved konstruksjonen. Prediksjoner gir også et hint, spesielt hvis dataene det predikeres ut fra ikke var tilgjengelig da modellen ble bygd. Predikerer modellen godt foreligger det neppe data mining.

3.10. Prediksjon.

Det er ingen nytte i å predikere fremtidige verdier før du vet at du har en god modell. En vanlig fremgangsmåte for å teste prediksjonsegenskapene til en modell er å sammenligne verdien den forutser med virkelige verdier. Dette kan man gjøre ved å først bygge en modell uten alle de tilgjengelige observasjonene, for eksempel ved å utelate de siste 5 årene. Deretter ser man om de estimerte verdiene for disse 5 årene er nær de sanne verdiene. Hvis modellen predikerer godt så kan vi inkludere de siste observasjonene og estimere modellen på nytt.

4. Reestimering og vurdering av Jacobsen og Naug sin modell for boligprisvekst.

4.1. Bakgrunn.

Jeg har fått tilgang til et datasett som er identisk (eller i alle fall svært nær) med datasettet som Jacobsen og Naug brukte i artikkelen ”Hva driver boligprisene?” i Penger & Kreditt 4/04. Med dette datamaterialet har jeg forsøkt å gjenskape resultatene i artikkelen.

I artikkelen bygger Jacobsen og Naug en empirisk modell som forsøker å forklare utviklingen i boligprisene. I tillegg til å forklare det som har vært, ønsker de også å kunne predikere boligprisene i nær fremtid. Dessuten blir et viktig tema tatt opp i artikkelen, nemlig om boligprisene er overvurdert i forhold til en fundamentalverdi bestemt av de forklaringsfaktorene de ender opp med som relevante for prisutviklingen. Bakgrunnen for å gjøre denne analysen var at boligprisene hadde hatt en formidabel vekst, i tidsrommet 1992 til 2004 hadde prisene mer enn tredoblet seg.

For å finne ut hva som påvirker boligprisene testet forfatterne for effekter fra en rekke variabler, hele 12 forskjellige før man tar i betraktning at man også kan bruke laggede verdier for å ta hensyn til evt. tregheter. Dette er mange variabler å putte inn i en modell, særlig med en så kort tidsserie som de har for boligprisene de bruker, prisindeksen for bruktboliger under ett.¹³ Denne indeksen har kvartalsdata som går tilbake til 1. kvartal 1990. Løsningen forfatterne brukte på problemet var derfor å estimere en rekke modeller som kun inneholdt en delmengde av variablene, for så å forenkle modellene ved å pålegge restriksjoner som ikke ble forkastet av data og som lettet tolkningen av dynamikken.

De ulike typene av variabler som Jacobsen og Naug testet for var lønnsinntekter, indeksene for betalt husleie og samlet husleie i konsumprisindeksen (KPI), øvrige deler av KPI justert for avgifter og uten energivarer (KPI-JAE), ulike mål på realrenten etter skatt, boligmassen, arbeidsledighetsraten, tilbakedatert vekst i boligprisene, husholdningenes gjeld, totalbefolkningen, andel av befolkningen som er i etableringsfasen det vil si i alderen 20-24

¹³ Denne prisindeksen utgis av Norges Eiendomsmeglerforbund og Eiendomsmeglerforetakenes Forening. Statistikken utarbeides av ECON analyse og er finansiert av FINN.no.

og 25-39 år, ulike mål på flytting/sentralisering og TNS Gallups' indikator for husholdningenes forventninger til egen og landets økonomi.

Lønnsinntektene inngår i den endelige modellen til Jacobsen og Naug. Inntekt påvirker direkte betalingsevnen, og økte inntekter kan derfor resultere i økt betalingsvilje. Når etterspørselen er høy vil mange gå langt for å vinne budrunden, og dermed presses betalingsviljen opp mot betalingsevnen som igjen kan medføre sterk vekst i boligprisene.

Husleiene og andre konsumpriser var generelt ikke signifikante i Jacobsen og Naug sitt datasett. De nevner at grunner til dette kan være at borettslagleie utgjorde en vesentlig andel av husleieindeksen i KPI i estimeringsperioden. I tillegg har en del husleier vært sterkt regulert, noe som kan redusere effekten av samspillet mellom boligpriser og husleier.

Utlånsrenten var svært signifikant i dette datasettet, mens markedsrenten ikke var signifikant når utlånsrenten også var inkludert i den samme modellen. Dette kan skyldes at renten ble brukt til å stabilisere utviklingen i valutaen på store deler av 90-tallet, og at befolkningen derfor benyttet observert rente som anslag på fremtidig rente istedenfor å bruke markedsrenten til dette. Ettersom markedsrenten i noe grad kan fange opp effekter av endrede konjunkturutsikter, så er det grunn til å tro at markedsrenten ble undervurdert som et mål på forventet rente i estimeringsperioden.

Boligmassen vil tilpasse seg til etterspørselen over tid. En langtidsmodell for boligprisene bør derfor inneholde forklaringsfaktorer for utviklingen i boligmassen, som bygge- og tomteknader og prisen på nye boliger. Boligmassen er inkludert i den endelige modellen til Jacobsen og Naug.

Begrunnelsen for å inkludere arbeidsledighet som en forklaringsfaktor i analysen var at økt arbeidsledighet gir forventninger om lavere lønnsvekst og økt usikkerhet om fremtidig inntekt og betalingsevne for en selv og andre. Dette gir igjen redusert betalingsvillighet for eierboliger. Ledigheten kan dessuten være et slags mål på konjunkturutviklingen. Arbeidsledigheten er inkludert i modellen forfatterne ender opp med.

Videre i analysen fant Jacobsen og Naug ingen signifikante effekter av tilbakedatert vekst i boligprisene, hvilket indikerer at husholdningene bare i begrenset grad benytter observert

boligprisvekst som indikator for fremtidig boligprisvekst. En undersøkelse i IMF (2004) konkluderer annerledes om tilbakedatert vekst. Hvis tilbakedatert vekst er uten særlig betydning for fremtidig vekst, så reduseres faren for prisbobler da man ikke nødvendigvis kommer inn i en prisspiral der alle tror at prisene vil fortsette å stige.

Husholdningenes gjeld ga ikke signifikante effekter på boligprisene, verken når de testet for hele estimeringsperioden eller når de kun benyttet perioden 1990-1993. Dette innebærer isolert sett at kreditten til husholdningenes boligkjøp ikke var begrenset av bankenes lønnsomhet. Allikevel er det grunn til å tro at bankenes lønnsomhet i perioden 1990-1993 begrenset utlånene, siden flere banker gikk konkurs under bankkrisen.

Flyttinger og demografiske forhold ga ikke holdepunkter for å ha direkte effekter på boligprisene under ett i Jacobsen og Naug sin analyse. De slår likevel fast at demografiske endringer spiller inn på lønnsinntektene, og i så måte vil påvirke boligprisene. Dessuten endres demografiske forhold sakte over tid, og med den relativt korte tidsperioden som datasettet spenner over kan det bli vanskelig å identifisere slike effekter.

Husholdningenes forventninger til egen og landets økonomi er korrelert med boligprisveksten, men samtidig med rente og arbeidsledighet. Ettersom rente og ledighet er egne forklaringsfaktorer i den endelige modellen, så korrigerer de forventningsindikatoren for effekter fra disse før de benytter den i modellen. Dette kommer jeg nærmere tilbake til i avsnitt 4.3. Jacobsen og Naug nevner videre at forventningsvariabelen de benytter kan fange opp effekter av ikke-fundamentale forhold, men at de ikke finner holdepunkter for at sjokk i forventningene har bidratt til økt vekst i boligprisene i nevneverdig grad.

Dersom man antar at aktørene i boligmarkedet er rasjonelle kan det forsvares å inkludere forventningene som en forklaringsvariabel i modellen. Prisforventningene vil i så fall ikke reflektere andre forventninger enn de som er knyttet til fundamentale endringer. En kontrast til dette finner vi i definisjonen av en boligprisboble som sier at dette er noe som kan oppstå hvis mange ønsker å kjøpe bolig i dag fordi de forventer at boligprisene skal stige fremover, og disse forventningene ikke er knyttet til fundamentale forhold. TNS Gallups' forventningsindikator baseres på en spørreundersøkelse blant ca 1000 husholdninger. Det fins ingen garanti for at husholdningene som deltar i undersøkelsen er rasjonelle aktører, men et stort utvalg gjør undersøkelsen mer troverdig.

Det beste resultatet ga modeller med nominelle verdier på boligpris, lønnsinntekter og rente. Dette er ikke unormalt, jfr. Meen (1990), Hall, Psaradikis og Sola (1997) og modellen i IMF(2004) (Jacobsen og Naug 2004). Modellen forfatterne endte opp med var en feiljusteringsmodell for logaritmen til boligprisene. Den inneholder effekter av samlede lønnsinntekter, boligmasse, arbeidsledighetsraten, bankenes utlånsrente etter skatt og den korrigerede indikatoren for husholdningenes forventninger til egen og landets økonomi.

4.2. Tidsseriene.

4.2.1. Mine forventninger om tidsseriene.

Før jeg kan inkludere variabler i en modell, så er det lurt å forsøke å avdekke eventuelle sjokk og ekstremverdier, og prøve å finne ut om variablene er stasjonære eller ikke. Summary statistics gir en oversikt over datasettet og vises i vedlegg 1. Ved å se på grafer kan jeg rask og greit kunne bestemme om variablene inneholder en trend, om jeg skal inkludere et konstantledd eller om jeg bør trekke ut en sesongfaktor før jeg tester om serien er stasjonær eller integrert av en orden høyere enn null.

Jeg forventer at lønnsinntektene og forventningene vil ha en positiv sammenheng med boligprisene. Høyere inntekt som gir økt kjøpekraft og lånekapasitet og økte forventningene herunder mer optimisme, gjør at mer penger blir brukt på både forbruksvarer og investeringer. Renten, arbeidsledigheten og boligmassen tror jeg vil gi et negativt utslag på boligprisene, da dette øker boligkostnadene, skaper usikkerhet om fremtidig inntekt og øker tilbudet av boligene.

For å begrunne mine antagelser om stasjonaritetsegenskapene til de ulike variablene vil jeg ta utgangspunkt i hva som skjer hvis et sjokk inntreffer.

Et eksempel på sjokk i boligprisene har vi hvis endret politikk, for eksempel en deregulering eller en skatteomlegging, fører til en uventet økning i boligprisene. Vil denne økningen være permanent, eller vil den fjernes med tiden? Mer offensiv skattelegging av eierboliger gjør fordelen ved å eie mindre. Dersom endringen er vedvarende, så er det rimelig å anta at boligprisene vil stabilisere seg på et lavere nivå. Det betyr at boligprisene ikke vil returnere til de boligprisene som var før sjokket, før andre elementer som den innebygde trenden eller nye sjokk fører prisene tilbake. Boligprisene antar jeg derfor at er ikke-stasjonære.

Dersom den nominelle lønnsveksten blir større enn ventet kan dette anses som et sjokk. Spørsmålet blir hva som skjer med inntektene i de neste periodene. Vil de gå tilbake til utgangspunktet eller vil de stabilisere seg rundt et nytt gjennomsnitt? Det skal være ekstremt dårlige tider for at befolkningen skal godta en nedgang i inntektene, så på generell basis vil ikke de nominelle inntektene falle. Isteden vil de bygge dette sjokket inn i forventningene til gjennomsnittsinntekten slik at man i neste periode vil forvente å få inntekt lik foregående periodes inntekt (inkl. sjokket) pluss den forventede økningen i lønnsinntektene. Inntekten er i så fall trendstasjonær frem til det kommer et nytt sjokk. Ettersom sjokk i inntektene ikke dør ut med tiden så antar jeg at inntektene er ikke-stasjonære.

Et sjokk i renten gir en stor og direkte effekt på økonomien. Men vil det nye rentenivå vedvare til det kommer et nytt sjokk? I sentralbanker i den vestlige verden endrer de ikke renten vesentlig uten å tenke på konsekvensene av dette. Settes renten for eksempel til et unormalt høyt nivå, gjøres det fordi det trengs drastiske tiltak for å få økonomien tilbake på fote igjen. Når økonomien etter hvert går tilbake til normale forhold vil rentenivået returnere til et normalt nivå igjen. På bakgrunn av dette tror jeg at renten er en stasjonær variabel.

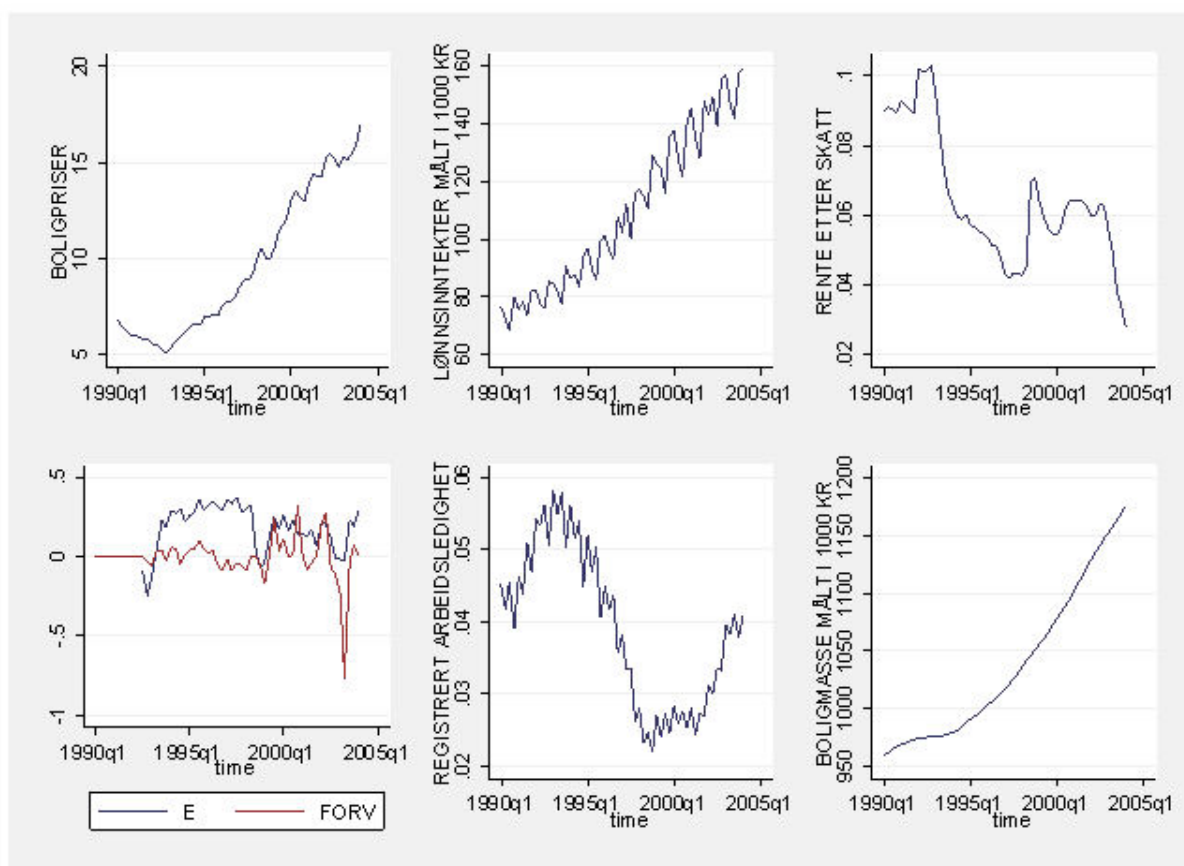
Forventningene til egen og landets økonomi varierer ettersom det er en høykonjunktur eller lavkonjunktur. Hvis landet er i en høykonjunktur, og for eksempel terrorfrykt gjør at oppfatningene om fremtiden endres brått, vil forventningene falle. Men over en lang horisont tror jeg ikke at forventningene vil holde seg på dette nye lave nivået fordi folk er generelt optimistiske og tenderer til å glemme negative episoder ganske raskt. Isteden tror jeg at forventningene vil returnere til gjennomsnittet. Dette innebærer i så fall at forventningene er stasjonære.

Et sjokk i arbeidsledighetsraten kan for eksempel oppstå hvis man åpner for fri flyt av arbeidskraft over landegrensene. Et slikt scenario vil gi en større tilførsel av arbeidskraft uten at avgangen endres i samme grad. Samlet sett kan derfor flere få vanskeligheter med å få seg jobb siden tilbudet av arbeidskraft øker. Vil denne økte arbeidsstyrken medføre et høyere nivå på arbeidsledighetsraten i all fremtid, eller vil man etter hvert returnere til gjennomsnittet som var før sjokket. Teorien om hysteresis sier at høy arbeidsledighet kan binde seg fast slik at ledigheten ikke automatisk vender tilbake til sitt naturlige nivå. Grunnen for at ledigheten holder seg på et høyt nivå kan være at det er vanskelig å få mennesker tilbake til jobb når de

har vært arbeidsløse i lang tid, enten fordi de har lagt til seg andre vaner eller fordi de har mistet produktivitet og kanskje også motivasjonen etter et langtidsfravær. Forsiktige reformer har blitt utprøvd for å skape lavere arbeidsledighet. OECD (1998) har identifisert at land som Danmark, Irland, Nederland og Storbritannia har gjennomført strukturelle reformer med suksess, men nedgangen i ledigheten kan dels forklares ut fra den makroøkonomiske stabiliseringspolitikken de hadde valgt (Ball, 1999). Dessuten har også andre land som Tyskland, Italia og Østerrike gjennomført strukturelle reformer i tråd med OECDs anbefalinger, uten å oppnå de samme gode resultatene. Røed (1996) finner en viss støtte for teorien om hysteresis i de aller fleste europeiske land, herunder Norge, når han tester stasjonaritetsegenskapene til arbeidsledigheten. Disse testene tar imidlertid ikke hensyn til om andre eksogene årsaker som for eksempel strukturelle reformer i arbeidsmarkedet har ført til endret likevektsnivå for arbeidsledigheten (NoU 2000). Arbeidsledigheten kan ut fra denne teorien ha permanente eller langvarige ettervirkninger av sjokk. Ut fra teorien og Røed sin undersøkelse kan det derfor synes som at arbeidsledigheten er ikke-stasjonær.

Hva angår boligmassen, så vil et sjokk på denne kunne oppstå hvis endret klima gjør at mange hus blir ubeboelige, for eksempel hvis vannet stiger slik at mange boliger blir liggende dels under vann. Boligmassen er en fysisk ting, og endringen vil så klart være permanent. Antagelig etterfølges et slikt negativt sjokk av et positivt sjokk som følge av nybygging siden innbyggerne i Norge trenger å bo under tak med det klimaet som er i landet. Siden sjokkene akkumuleres over tid og legges til i nivået vi startet med, så tror jeg at boligmassen er ikke-stasjonær.

Figur 8: Tidsseriene.



Grafen merket E (blå linje) er TNS Gallups' forventningsindikator og FORV (rød linje) er forfatterens egenkonstruerte forventningsvariabel som tar ut effektene fra rente og ledighet. Fra grafene over ser jeg at boligpriser, lønnsinntekter og boligmasse inneholder en trendfaktor. Rente og ledighet har ikke trend i seg hvis man tenker i et langtidsperspektiv, men spesielt renten ser ut til å ha et negativt trendelement i seg i figur 8. Dette skyldes at utvalget av data ikke er stort nok. Det kan videre synes som at den estimerte forventningsvariabelen (FORV) ikke inneholder et konstantledd. Ellers kan det virke som at lønnsinntektene og arbeidsledigheten bør justeres for sesong.

Hva angår sjokk, så ser det ut som at det var et negativt sjokk i 2003 i den konstruerte forventningsvariabelen (FORV). I følge Økonomisk Rapport på nett, orapp.no, skyldes dette pessimisme rundt landets økonomi, og ikke husholdningenes egen økonomi. I 2003 meldte bedrifter om dårlige resultater, for høy kronekurs og nedbemanning, og stadig flere ble arbeidsledige. Dessuten var det lite som tydet på at terrorhandlinger og ustabiliteten i Midt-Østen skulle reduseres. Dette kan være noen av grunnene til pessimismen som rådet.

4.2.2. Dickey-Fuller testene.

Tabell 1: Resultatene fra DF-testene.

teststatistikk fra den (Utvidede) Dickey-Fuller testen.				
Variabel	Antall lag	med trend med konstant	uten trend med konstant	uten trend uten konstant
Boligpris	5	-3,287*	1,584	
Inntekt	3	-2,350	1,374	
Inntekt, justert for sesong	3	-2,313	1,088	
Rente etter skatt	1	-2,075	-1,258	
Forventningsindikator, E	0	-2,860	-2,828*	-1,204
Konstruert forventning, FORV	3	-5,058***	-5,094***	-5,136***
Arbeidsledighet	4	-3,744**	-2,585	
Arbeidsledighet, sesongjustert	4	-3,475*	-2,297	
Boligmasse	3	-2,877	1,237	
Ln(inntekt) – ln(boligmasse)	3	-0,298	-1,166	
Ln(inntekt) – ln(boligmasse), sesongjust.	4	0,308	-1,715	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Hva angår stasjonaritet er resultatene jeg oppnår uavhengig av om jeg justerer for sesong. Boligprisen er oppsiktsvekkende nok stasjonær når jeg tar hensyn til trend, men jeg observerer at teststatistikken er svært nær ved å ikke kunne forkastes på 10 % nivå. Jeg får også at renten etter skatt er ikke-stasjonær, i motsetning til hva jeg forventet. TNS Gallups' forventningsindikator og arbeidsledighetsraten avhenger av om vi tar hensyn til trend eller ikke. Konstantleddet er signifikant for forventningsindikatoren. Trendfaktoren er ikke signifikant for forventningsindikatoren, mens den er signifikant for ledigheten. Også informasjonskriteriene foretrekker at forventningsindikatoren testes uten trend (og med konstantledd) og at ledigheten testes med trend. Basert på informasjonskriteriene og signifikansnivåer får jeg dermed at forventningsindikatoren og arbeidsledigheten er stasjonære variabler i følge DF-testen. Resultatet om at ledigheten er stasjonær støtter ikke teorien om hysteres, men jeg observerer at variabelen er ikke-stasjonær når man tester uten trend, og at trenden trolig er signifikant på grunn av at utvalget består av en kort tidsperiode. Variablene inntekt, FORV og boligmasse har stasjonaritetsegenskaper i tråd med forventningene mine, FORV er stasjonær, mens de to andre er ikke-stasjonære i dette datasettet. Inntekt minus boligmasse er ikke-stasjonær, hvilket betyr at $[1, -1]$ ikke er en

potensiell kointegrasjonsvektor for disse variablene. Det er til slutt verdt å nevne at jeg ser på en relativt kort tidsserie her, vi har kun 57 observasjoner i utgangspunktet. Et lite datasett kan gi misvisende resultater fordi vi kanskje ikke får med de naturlige svingningene i variabelen over tid.

I artikkelen finner de at boligprisene og lønnsinntektene er ikke-stasjonære. At de får boligprisene til å være $I(1)$ kan skyldes en av tre grunner; enten har boligprisene endret seg fra originaldatasettet til det jeg har fått oppgitt, eller så har de benyttet en annen test enn DF-testen for å sjekke det, eller så har forfatterne startet med å sjekke 4 lags ut fra en vurdering om at laggede verdier lenger tilbake enn et år skal man ikke ha med uten å ha en veldig god grunn. Gjør man det finner man at boligprisen er ikke-stasjonær med trend og 4 lags. Allikevel velger jeg å ta med 5 lag på boligprisen, fordi det 5. lagget er klart signifikant i DF-testen. Som nevnt i stad er teststatistikken for boligprisene meget nær ved å ikke kunne forkastes, og vi vet at en svakhet ved DF-testen nettopp er at den kan forkaste nullhypotesen feilaktig. Dette gjør at jeg ikke utelukker at boligprisene kan være ikke-stasjonære, jfr. diskusjon om sjokk på boligprisene.

4.3. Den konstruerte forventningsvariabelen.

4.3.1. Bakgrunn.

Det er rimelig å anta at befolkningens forventninger blir påvirket av rentenivået og arbeidsledigheten i landet. Hvis det er dårlige tider og bedriftene må redusere arbeidsstokken, vil det skape usikkerhet rundt arbeidsplassen. Ved gode tider derimot, når det er et høyt aktivitetsnivå i økonomien og arbeidsledighetsraten kommer på rekordlave nivåer, investerer man som aldri før og forventningene er høye om fremtiden. Norges befolkning har i tillegg en stor andel av sin formue plassert i egen bolig, hvilket gjør dem sensitive overfor endringer i rentenivået. Fordi TNS Gallups' indikator for husholdningenes forventninger til egen og landets økonomi er korrelert med rentenivået og arbeidsledighetsraten¹⁴, så velger forfatterne av artikkelen og korrigerer forventningsindikatoren for disse effektene ettersom de to variablene inngår som egne forklaringsvariable. Ved å heller benytte den konstruerte forventningsvariabelen ser de på om andre faktorer enn rente og ledighet, som påvirker befolkningens forventninger, har effekt på boligprisene.

¹⁴ Se Vedlegg 3

4.3.2. Jacobsen og Naug sin modell for forventningsvariabelen.

Tabell 3 i artikkelen ”Hva driver boligprisene?” gir oss en modell for husholdningenes forventninger til egen og landets økonomi. Modellen er en feiljusteringsmodell som tar med både kortsiktige og langsiktige effekter. I tabellen under presenterer jeg Jacobsen og Naug sin modell og reestimeringen. Små bokstaver betegner at variabelen er på logaritmisk form.

Tabell 2: Modell for husholdningenes forventninger til egen og landets økonomi.

Variabel	J & N sin modell		Reestimeringen		
	Koeffisient	t-verdi	Koeffisient	Standardavvik	t-verdi
$\Delta(\text{RENTE}(1-\tau))_t$	-12,96***	(6,68)	-12,96***	(1,94)	(6,68)
$\Delta\text{ledighet}_t$	-0,43**	(2,47)	-0,43**	(0,17)	(2,47)
E_{t-1}	-0,11	(1,06)	-0,11	(0,10)	(1,06)
$\text{RENTE}(1-\tau)_{t-1}$	-0,40	(0,42)	-0,40	(0,94)	(0,42)
ledighet_{t-1}	-0,03	(0,82)	-0,03	(0,04)	(0,82)
S1	0,21***	(4,57)	0,21***	(0,05)	(4,57)
S2	0,10***	(4,49)	0,10***	(0,02)	(4,49)
S3	0,22***	(5,61)	0,22***	(0,04)	(5,61)
KONSTANTLEDD	-0,07	(0,39)	-0,20	(0,18)	(1,12)
Antall observasjoner	46		46		
R2	0,80		0,8025		
Justert R2			0,7598		
Q*(1)			0,0285		
Q*(4)			4,4423		
DF-teststatistikk			-7,436***		

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Koeffisientene stemmer særdeles godt overens med de i artikkelen. For samtlige variabler unntatt konstantleddet er det ikke avvik på 2. desimal. Estimater på konstanten har blitt endret, fra -0,07 i artikkelen til -0,20, men dette følger av noe ulike data. Modellen viser at rente og ledighet forklarer hele 80 % av variasjonen i TNS Gallups' indikator. De resterende effektene samles i den konstruerte forventningsvariabelen som måler andre skift i forventningene enn de som skyldes rente og ledighet. Dette kan være skift som oppstår i forbindelse med endrede politiske forhold, endrede utsikter for norsk økonomi for eksempel endrede forventninger om arbeidsledigheten, og negative sjokk som krig, terror og børsfall (Jacobsen og Naug 2004).

Forventningsvariabelen som ble brukt i artikkelen fant Jacobsen og Naug ved å ta vare på feilledet fra regresjonen over. Deretter summerte de denne over to kvartaler, dvs. de lagde en

ny variabel som var summen av feilleddet på tid t pluss den forrige verdien $t-1$. Denne puttet de så inn i formelen $FORV_t = (\varepsilon_t + \varepsilon_{t-1}) + 100 * (\varepsilon_t + \varepsilon_{t-1})^3$. Forfatterne sier ikke hvorfor de tar med dette siste leddet. Muligens ønsker de at variabelen skal være tre ganger deriverbar for å fange opp noe om krumningen, dvs. graden av konkavitet eller konveksitet. En annen grunn kan være at de ønsket å glatte ut differanser.

En test for autokorrelasjon i feilleddet gir at det verken foreligger første eller fjerde ordens autokorrelasjon. Kritisk verdi er henholdsvis 3,84 og 9,49 for 1 og 4 frihetsgrader. Tabellen over gir at Ljung-Box teststatistikkene (Q^*) er 0,0285 og 4,4425 som er lavere enn de kritiske verdiene, og dermed kan vi ikke forkaste nullhypotesen om ingen autokorrelasjon. Modellen har ikke systematikk i feilleddene.

4.3.3. Vedrørende rentevARIABLEN.

Dersom resultatet om at renten er ikke-stasjonær stemmer, er regresjonen i tabell 2 trolig ikke gyldig. Modellen inneholder i så fall de stasjonære variablene arbeidsledighet og TNS Gallups' forventningsindikatoren, ved siden av den ikke-stasjonære variabelen rente etter skatt. Dersom rente etter skatt er den eneste ikke-stasjonære variabelen som inngår i modellen kan ikke kombinasjonen av denne og de stasjonære variablene i modellen være stasjonær. Da trengs det minimum to $I(1)$ serier for at jeg skal få en stasjonær modell. En DF-test av restleddet i tabell 3 gir en teststatistikk på -6,852 når jeg tester uten lags og konstantledd, og jeg forkaster nullhypotesen på 1 % signifikansnivå. Restleddet er derfor stasjonært, noe som innebærer at renten antakeligvis er stasjonær. Det kan bety at DF-testen jeg gjorde av rente etter skatt ikke har forkastet nullhypotesen som burde ha blitt forkastet. Alternativt er forventingsvariabelen eller arbeidsledigheten ikke-stasjonær og kointegrerer med renten.

4.4. Jacobsen og Naug sin modell for boligprisveksten.

4.4.1. Reestimering.

Tabell 1 i artikkelen "Hva driver boligprisene?" oppnås ved å kjøre en 1-steps feiljusteringsmodell. Jacobsen og Naug har pålagt at inntekt og boligmasse skal ha samme langtidseffekt, men med motsatt fortegn.¹⁵ Dette får de frem i modellen ved å lage en variabel

¹⁵ Begrunnelsen for dette var at de fant at boligmasse og inntekt var sterkt korrelerte når de justerte for sesong.

”inntekt minus boligmasse” der inntekt og boligmasse begge er gitt på logaritmisk form. Koeffisienter og standardavvik fra denne estimeringen ligger i vedlegg 2. De langsiktige koeffisientene som estimeres her er ikke de som brukes i den langsiktige likevekten, disse må man finne etterpå ved å dele nåværende koeffisient med minus koeffisienten til den laggede boligprisen. I artikkelen til Jacobsen og Naug ble korttidseffektene og den estimerte langtidssammenhengen rapportert i en ligning, og for å kunne sammenligne mine koeffisienter med deres viser jeg begge versjonene i den samme tabellen (små bokstaver betegner logaritmisk form av variabelen, kursiv betegner langtidssammenhengen).

Tabell 3: Jacobsen og Naug’s feiljusteringsmodell. Deres estimater og nye reestimerte estimater. Koeffisientene i langtidsligningen har blitt endret til de som inngår i likevekten.

Variabel	J & N sin versjon		Reestimeringen		
	Koeffisient	t-verdi	Koeffisient	Standardavvik	t-verdi
Δ inntekt _t	0,12*	(1,94)	0,104	(0,066)	(1,57)
Δ (RENTE(1- τ)) _t	-3,16***	(7,04)	-3,180***	(0,468)	(6,80)
Δ (RENTE(1- τ)) _{t-1}	-1,47***	(3,27)	-1,464***	(0,459)	(3,19)
FORV _t	0,04***	(3,09)	0,044***	(0,015)	(3,02)
<i>boligpris</i> _{t-1}	-0,12***	(5,69)	-0,110***	(0,024)	(4,59)
<i>(RENTE(1-τ))</i> _{t-1}	-4,47**	(2,54)	-4,810**	(2,173)	(2,21)
<i>ledighet</i> _t	-0,45***	(3,48)	-0,495***	(0,163)	(3,04)
<i>(inntekt-boligmasse)</i> _{t-1}	1,66***	(8,63)	1,599***	(0,237)	(6,74)
S1	0,04***	(3,35)	0,035***	(0,011)	(3,19)
S2	0,02*	(1,80)	0,020*	(0,011)	(1,74)
S3	0,01	(0,73)	0,007	(0,013)	(0,53)
KONSTANTLEDD	0,56***	(3,42)	0,494**	(0,192)	(2,57)
Antall observasjoner	54		54		
R2	0,8773		0,8696		
Justert R2	ikke rapportert		0,8354		
Durbin-Watson	2,57		2,53		
Q*(1)			5,3434		
Q*(4)			9,5046		

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Denne regresjonen samsvarer ikke like bra som den i tabell 2.¹⁶ Ettersom jeg fikk beskjed om at datasettet var svært nær, men ikke nødvendigvis identisk med det originale datasettet vil jeg forvente noe avvik. Siden resultatene i tabell 2 var så nær de opprinnelige må det bety at

¹⁶ Korrelasjonsmatrisen for variablene som inngår i tabell 3 er vist i vedlegg 3.

variasjonen fra originalsettet må ligge i en eller flere av variablene som ikke var med i tabell 2, dvs. i boligpris, inntekt eller boligmasse.

Jacobsen og Naug sin modell for boligprisveksten er ikke balansert, og innebærer derfor nullrestriksjoner. Dette er ikke nevnt i artikkelen. Utgangspunktet for en feilkorreksjonsmodell er en balansert modell hvor de samme variablene inngår både på endringsform og på lagget form. Dersom ikke alle variablene inngår både på kort og lang sikt, har man testet ned modellen for å komme frem til en modell med for eksempel bare signifikante variabler.

4.4.2. Tolkning av koeffisientene fra reestimeringen.

For å tolke koeffisienten gjør jeg om estimatene til elastisiteter. Der hvor variabelen allerede er på logaritmisk form er koeffisientene lik elastisitetene, men for renten og forventningene må jeg gange snittet av disse med koeffisienten for å finne elastisiteten.

De langsiktige elastisitetene:

Variabel	Koeffisient	Standardavvik
$(RENTE(1-\tau))_{t-1}$	-0,312**	(0,141)
Ledighet _t	-0,495***	(0,163)
$(inntekt-boligmasse)_{t-1}$	1,599***	(0,237)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

De kortsiktige elastisitetene:

Variabel	Koeffisient	Standardavvik
$\Delta inntekt_t$	0,104	(0,066)
$\Delta(RENTE(1-\tau))_t$	-0,206***	(0,030)
FORV _t	-0,0001***	(0,00004)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Hvis renten øker permanent med 1 % vil boligprisene på lang sikt falle med $0,312 \cdot (1-0,28) = 0,225$ % hvis de øvrige forklaringsfaktorene ligger fast. Nettoeffekten av en renteøkning på 1% er nemlig bare $(1-0,28) = 0,72$ % når vi tar hensyn til skatt. Dersom utlånsrenten økes fra

5 % til 6 % innebærer det en økning i renten på $\left(\frac{0,06 - 0,05}{0,05}\right) = 20\%$. En slik økning vil redusere boligprisene med $(20 * 0,225) = 4,5\%$. Renteeffekten er signifikant på 5 % nivå. Størrelsen på effekten på boligprisene virker rimelig, og fortegnet er negativt som forventet.

Når det gjelder arbeidsledigheten så tok de denne på tid t i langtidsløsningen fordi dette ga best føyning.¹⁷ Jeg har ikke sett eksempler på at dagens verdi av en variabel inngår i en langtidslikevekt før, og dette blir litt vanskelig å tolke. Tolket man det direkte så betyr det at hvis ledigheten øker med 1 % på slutten av perioden, så vil boligprisene falle med 0,495 % på starten av perioden. Effekten er signifikant.

Dersom inntekten øker med 1 % vil boligprisene stige med 1,599 % på lang sikt. Dette er en signifikant effekt. Koeffisienten virker for stor, dette fordi økt inntekt for det første ikke benyttes i sin helhet til økt boligkonsum, og for det andre at større andeler enn før må lånefinansieres ved kjøp av bolig. Boligmassen har som nevnt den samme signifikante langtidseffekten som inntekt, bare med motsatt fortegn. Derfor vil en økning i boligmassen på 1 % gjøre at boligprisene faller med 1,599 %. Fortegnet er negativt som forventet, men størrelsen på endringen er stor. Grunnen til at estimatet for inntekt minus boligmasse ikke er i tråd med hva jeg forventer kan være at boligmassen er en endogen variabel. Det er ikke urimelig at boligmassen blir påvirket av boligprisen. Boligmassen avhenger av nybygging og avgang, og boligprisene virker inn på hvor mange hus som bygges. For eksempel kan det hende at noen velger å kjøpe eller leie brukte boliger fremfor å bygge selv hvis prisene er høye, eller at økte boligpriser gir et potensial for fortjeneste for utbyggere slik at de øker aktiviteten. Dersom boligmassen er endogen, så blir koeffisientestimatet forventningsskjevt og ikke konsistent.¹⁸ Dermed er tolkningen for inntekten og boligmassen på lang sikt ikke nødvendigvis riktig.

Justeringsparameteren, som vises som koeffisienten foran lagget boligpris, er signifikant på 1 % nivå, og sier at avvik fra likevekten justeres inn med 11 % hvert kvartal. Det betyr at et avvik på tidspunkt t justeres inn med 11 % etter ett kvartal, etter to kvartaler er omtrent 20,8%

¹⁷ Mail fra Jacobsen 15/05/07.

¹⁸ Mer om simultanitetsproblemet i kapittel 5.8.

justert inn, etter 3 kvartaler er totalt 29,5 % av avviket justert inn osv. Vi ser at det tar tid før vi når tilbake til likevekten, og den omtrentlige tiden det tar er $\frac{1}{0,11} \approx 9$ kvartaler.

På kort sikt vil en økning i inntekten på 1 % fører til at boligprisene øker med 0,104 %. Dette er ikke en signifikant effekt, så gitt datasettet mitt har ikke inntektsvekst noen effekt på boligprisveksten på kort sikt.

Når det gjelder renten så har Jacobsen og Naug tatt med den endringen som skjer i renten denne perioden, og i tillegg et etterslep fra forrige periode. Økes rentene før skatt med 1 % vil dette redusere boligprisene med $(0,206*0,72) = 0,148$ % på kort sikt. Som vist i stad innebærer en renteendring fra 5 % til 6 % i utlånsrenten at renten endres med 20 %. Effekten på boligprisene vil da være et fall på $(0,148*20) = 2,96$ % på kort sikt. Effekten er signifikant på 1 %, og størrelsen på koeffisienten virker rimelig. Vedrørende den laggede renteendringen i modellen så innebærer den at man drar med seg noe fra forrige periode, altså at det er tregheter i økonomien slik at ikke alle reagerer umiddelbart på renteendringen.

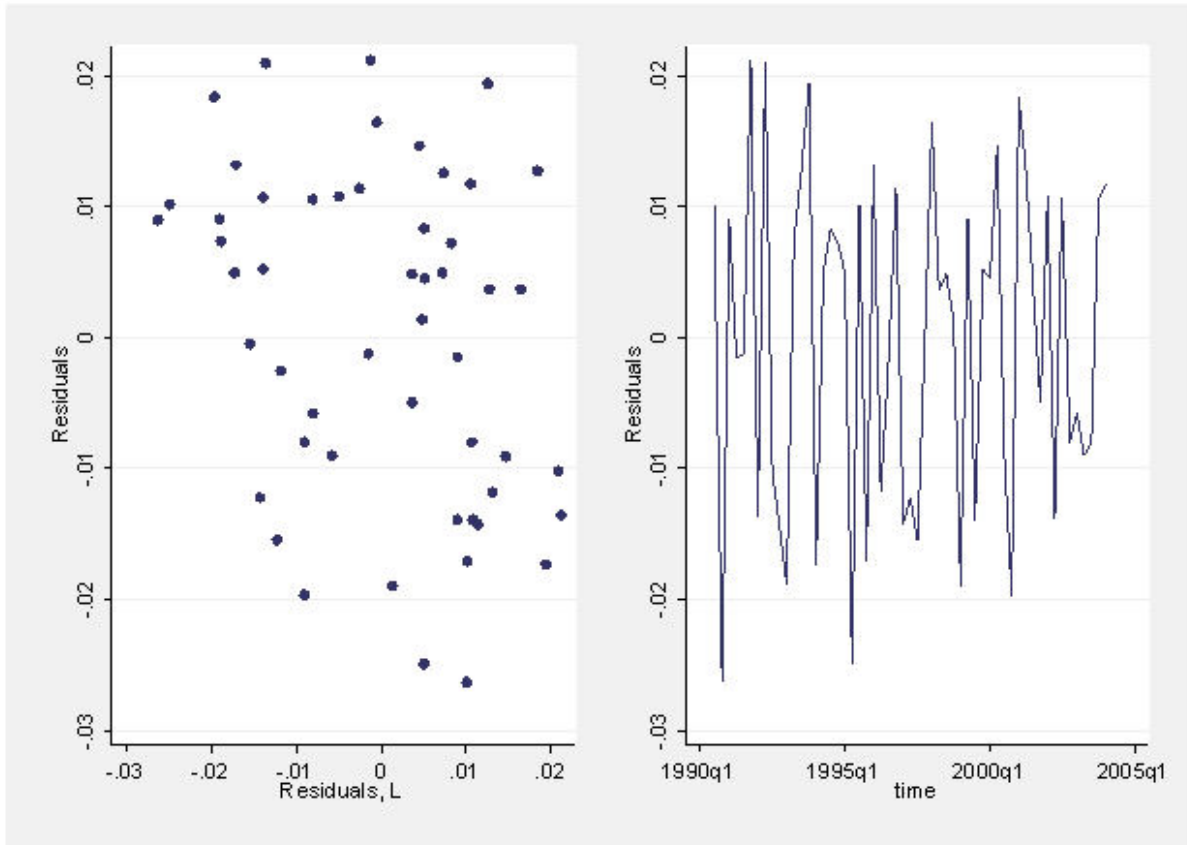
Forventningene (de konstruerte) inngår på nivåform i korttidsløsningen, hvilket er uvanlig for en standard feiljusteringsmodell. Dersom forventningene øker med 1 % på kort sikt, så vil boligprisene avta med 0,0001 %. Effekten er signifikant, men omtrent lik null.

I 1-steps feiljusteringsmodeller sier signifikansen til justeringsparameteren om variablene som inngår i modellen er kointegrert eller ikke. Modellen min gir at denne er signifikant på 1 % signifikansnivå, og derfor er boligpris, inntekt og boligmasse (og muligens rente og ledighet dersom disse er ikke-stasjonære) kointegrert. En Dickey-Fuller test av restleddet bekrefter dette.

4.4.3. Testing for autokorrelasjon.

For at estimatene skal være BLUE er det viktig å sjekke for autokorrelasjon i modellen. Det eksisterer grafiske tester og statistiske tester for å avdekke autokorrelasjon. I figuren under har jeg plottet restleddet fra tabell 3 opp mot dets forrige verdi og opp mot tiden.

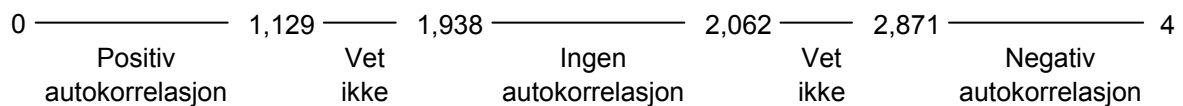
Figur 9: Plott av restleddene fra tabell 3.



Vi ser at det i praksis ikke er like lett å tyde plott som i lærebokseksemplene. I figuren til venstre kan det ved første øyekast se ut til at punktene er tilfeldig spredt over alle fire kvadranter. Undersøker jeg nærmere så finner jeg at det er hhv. 13, 15, 9 og 15 observasjoner i 1., 2., 3. og 4. kvadrant. At det er flest punkter i 2. og 4. kvadrant kan bety at restleddene er negativt autokorrelert, men det er umulig å konkludere med et så lite datamateriale når det ikke er et klarere mønster enn her. I figuren til høyre vises restleddene plottet over tid. Vi kan ikke avgjøre spørsmålet om autokorrelasjon ut fra disse figurene, dette må gjøres med statistiske tester. Hvis det skulle eksisterer autokorrelasjon i modellen tyder figurene på at det er negativ autokorrelasjon.

Ettersom jeg har en multivariat modell hvor lagget avhengig variabel inngår som en forklaringsvariabel (ADL-modell), og ikke-stasjonære variabler er inkludert etc. er Durbin-Watson testen ikke et godt valg for å teste for autokorrelasjon. Siden utvalget dessuten er lite, det er kun 57 observasjoner, velger jeg å heller teste for autokorrelasjon med Ljung-Box testen. Det er aktuelt å teste for første ordens og fjerde ordens autokorrelasjon, fjerde ordens fordi vi har kvartalsdata.

Et oppsiktsvekkende resultat er at jeg finner første ordens autokorrelasjon i tabell 3. Jeg får en Ljung-Box teststatistikk på 5,34 og kritisk verdi er 3,84. Fordi forfatterne benyttet DW testen her, fant de ikke dette resultatet. Blant annet vet vi at DW er forventningsskjev i retning 2 når for eksempel stokastiske variabler er inkludert, og denne modellen inneholder ikke-stasjonære variabler. Tar jeg DW-testen på min reestimerte modell, får jeg en teststatistikk på 2,53. De kritiske verdiene fra DW tabellen er 1,129 og 2,062, og tallinjen under viser de ulike regionene. Både denne DW-teststatistikk og 2,57 som de fant i artikkelen ligger i en region hvor vi ikke kan konkludere hvorvidt det er autokorrelasjon i modellen. I tillegg er begge målene forventningsskjeve mot 2.

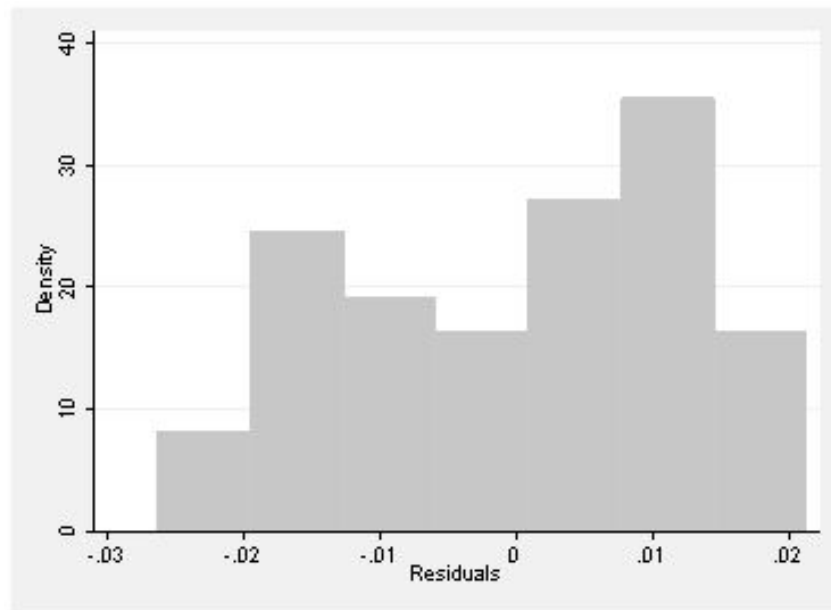


Det kan være flere grunner til at autokorrelasjon er til stede i modellen. Som nevnt i den statistiske teoridelen kan det være utelatt en relevant variabel eller man kan ha en upassende funksjonsform på en eller flere variabler. Kanskje har de ikke tatt med nok lags av den forklarte variabelen eller forklaringsvariablene. Inferenstesting om OLS-estimatene vil ikke nødvendigvis gi riktige konklusjoner siden standardavviket er galt når det er systematikk i feilleddene. Løsningen er å respesifisere modellen.

4.4.3. Er restleddene normalfordelte?

En forutsetning bak t -testene og F -testene er at feilleddene er normalfordelte. Et histogram av feilleddene viser at fordelingen er relativt symmetrisk, men uten en topp på midten. Så lenge det er symmetri, så er OLS estimatene akseptable hva angår normalitetstesten.

Figur 10: Et histogram over feilleddene fra tabell 3.



4.4.4. Diskusjon av den pålagte restriksjonen.

Jacobsen og Naug argumenterer for at inntekt og boligmasse har den samme langtidseffekten med motsatt fortegn. Dette fordi de finner en høy korrelasjon mellom inntekt og boligmasse når de justerer for sesong, og dermed blir estimatene uskarpe hvis begge variablene inngår med separate koeffisienter. De har også testet dette med en F -test. Med mitt datasett finner jeg en korrelasjon mellom inntekt og boligmasse på 0,982 når jeg justerer for sesong. Det Jacobsen og Naug derimot ikke tenkte på er at begge variablene inneholder en trend som burde ha blitt trukket ut før man eventuelt pålegger en restriksjon om estimatene. Når jeg undersøker korrelasjonen mellom inntekt og boligmasse etter at trend og sesonger er trukket ut finner jeg en korrelasjonskoeffisient på 0,6871. Det er fortsatt en relativt høy korrelasjon, men det blir uriktig å pålegge at effektene deres er i samme størrelsesorden, men med motsatt fortegn.

4.5. Data Mining?

Definisjonen av data mining er at det er noe som foregår når noen benytter det samme datasettet til å estimere et stort antall modeller for å finne den ”beste” modellen.

Jacobsen og Naug testet for effekter av en lang liste med forklaringsfaktorer ved å estimere en rekke modeller der de kun inkluderte en delmengde av variablene. Deretter forenklet de disse modellene ved å pålegge restriksjoner som ikke ble forkastet av data og som lettet tolkningen av dynamikken. Om de har testet for mange varianter er vanskelig å si, for det er vanlig å prøve seg frem for å finne en god modell. I artikkelen undersøker de på hva som skjer med boligprisen fremover gitt modellen og de forventninger som er beskrevet i inflasjonsrapporten¹⁹ 3/2004 (perioden etter deres estimeringsvindu). Modellen deres klarte å predikere 2. og 3. kvartal av 2004 rimelig godt. Dette tyder på at data mining ikke er et problem i modellen.

4.6. For få observasjoner?

En enkel metode for å undersøke om man har kommet frem til en holdbar modell er å se om koeffisientene endrer seg mye når datasettet utvides. Dersom estimatene holder seg omtrent på det samme nivået som før har man funnet en robust modell. Endrer koeffisientene seg derimot mye er modellen svak, og den er ikke optimal for å predikere fremtiden. Man har i så fall bygd en modell på for få observasjoner eller så har man feilspesifisert noen av forklaringsfaktorene.

I tillegg til ”originaldatasettet” fikk jeg tilsendt et utvidet sett som tok med observasjoner til og med 4. kvartal 2006. Dette hadde gjennomgått omlegginger og datarevisjoner fra det originale datasettet for å gi mer korrekte verdier, og en nærmere titt viste at enkelte av variablene hadde blitt endret ganske mye. På grunn av dette velger jeg å kjøre den samme testen som er beskrevet over på dette datasettet, først for perioden til og med 1. kvartal 2004, så for hele datasettet. Grunnen til at jeg gjør dette er for å se om estimatene endrer seg mye, både mellom datasettene gitt den samme tidsperioden, og innenfor et datasett når vi legger til flere observasjoner i dette.

¹⁹ Nå pengepolitisk rapport.

Tabell 4: En sammenligning av estimatene fra Jacobsen og Naug's modell for ulike utvalg.

Variabel	Originaldata		Justerte data		Utvidet og justerte data	
	Koeffisient	St.avvik	Koeffisient	St.avvik	Koeffisient	St.avvik
Δ inntekt _t	0,104	(0,066)	0,297	(0,280)	0,092	(0,282)
Δ (RENTE(1- τ)) _t	-3,180***	(0,468)	-2,764***	(0,455)	-3,148***	(0,499)
Δ (RENTE(1- τ)) _{t-1}	-1,464***	(0,459)	-1,439***	(0,460)	-1,052**	(0,489)
FORV _t	0,044***	(0,015)	0,028***	(0,009)	0,017*	(0,010)
boligpris _{t-1}	-0,110***	(0,024)	-0,119***	(0,022)	-0,070***	(0,019)
(RENTE(1- τ)) _{t-1}	-0,530***	(0,172)	-0,451**	(0,186)	-0,748***	(0,183)
ledighet _t	-0,054***	(0,013)	-0,039***	(0,014)	-0,046***	(0,013)
(inntekt-boligmasse) _{t-1}	0,176***	(0,057)	0,223***	(0,054)	0,108**	(0,051)
S1	0,035***	(0,011)	0,042***	(0,006)	0,041***	(0,006)
S2	0,020*	(0,011)	0,016	(0,011)	0,015	(0,011)
S3	0,007	(0,013)	-0,011	(0,010)	-0,003	(0,010)
KONSTANTLEDD	0,494**	(0,192)	0,683***	(0,185)	0,303*	(0,179)
Antall observasjoner	54		56		67	
R2	0,8696		0,8671		0,8023	
Justert R2	0,8354		0,8339		0,7628	
Q*(1)	5,3434		2,4205		0,0065	
Q*(4)	9,5046		17,7139		5,8016	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

De justerte dataene har to observasjoner mer enn de originale. Dette skyldes at TNS Gallups' forventningsindikator ikke ble samlet inn før 3. kvartal 1992, og verdiene på variabelen i originaldatasettet er satt til missing før dette kvartalet. Når den konstruerte forventningsvariabelen deretter blir satt til null til og med 3.kvartal 1992, så oppstår det en missing value for 4.kvartal 1992 i originaldatasettet som skyldes at vi først bruker lagget forventningsindikator i en regresjon for så å også bruke det laggede restleddet fra denne når vi konstruerer forventningen. Fordi vi mangler en verdi så predikerer ikke STATA boligprisen for denne perioden, og dette fører igjen til at vi mister lagget boligpris i neste periode. Fordi de justerte dataene har satt TNS Gallups' forventningsindikator lik null og ikke missing i perioden før 3. kvartal 1992, så oppstår ikke dette problemet for de justerte dataene. Dermed er det to observasjoner mindre i reestimeringen av modellen kontra estimeringen på det justerte datasettet.

Det er en ganske stor forskjell fra originaldataene til de justerte selv om vi ser på den samme tidsperioden foruten to ekstra observasjoner. Når estimatet endres så mye på et "nytt" utvalg, altså når dataene er blitt oppgradert til å være mer nøyaktige verdier (i tillegg til at to ekstra

kvartaler er tatt med), er det ganske oppsiktsvekkende. I tillegg er det også en betydelig forskjell når det justerte datasettet utvides med 11 kvartaler ekstra. Jacobsen og Naug's modell kan dermed synes å være tilpasset utvalget av data og kan derfor potensielt vise for høy forklaringsgrad. Den manglende stabiliteten i koeffisientestimatene betyr at de ikke har funnet en robust modell.

4.7. Er boligprisene overvurdert?

Forfatterne fant ikke holdepunkter for at boligprisene er overvurdert i forhold til en fundamentalverdi bestemt av rente, inntekter, arbeidsledighet og nybygging. Når de testet for effektene av tilbakedatert boligprisvekst var disse ikke signifikante, noe som må bety at husholdningene bare i begrenset grad benytter observert vekst som indikator for fremtidig boligprisvekst.²⁰ Videre kan deres konstruerte forventningsvariabel fange opp effekter av ikke-fundamentale forhold, men de fant ikke holdepunkter for at sjokk i forventningene har bidratt til å trekke opp boligprisene i særlig grad de siste par årene. Modellen impliserer at boligprisene stiger mer på kort enn lang sikt hvis det kommer et rentefall, men selv ved store rentefall er denne overreaksjonen ganske beskjeden, så man kan ikke sidestille dette med en boble.

Den endelige modellen forsøker å forklare boligprisveksten med fundamentale variabler, og derfor kan en eventuell boble i noe grad bli fanget opp av renten og de andre forklaringsvariablene. Som følge av dette gjorde de en undersøkelse hvor de estimerte modellen over en kortere periode, for så å simulere modellen 4 år frem i tid. Dersom det var en boble i boligmarkedet, så ville koeffisientene i modellen vært ustabile og prediksjonen ligge under de virkelige verdiene på boligprisveksten. I prognoseperioden benyttet Jacobsen og Naug de faktiske verdiene på forklaringsfaktorene. Modellen viste å predikere prisnivået og prisveksten ganske godt, og dette støtter hypotesen om at boligprisene ikke er overvurdert.

På grunn av alle problemene rundt Jacobsen og Naug sin modell av boligprisene, er det vanskelig for meg å begrunne verken for eller i mot en prisboble med utgangspunkt i denne modellen.

²⁰ I boligprisligningen for 18 OECD-land i IMF(2004) fant de sterke, positive effekter av tilbakedatert vekst i realboligprisene.

4.8. Kritikk av modellen.

Et av hovedproblemene med modellen er at det er systematikk i feilleddene. Når det er autokorrelasjon i modellen så vil OLS gi gale standardavvik. Inferenstesting om variablene vil ikke være gyldig, og man kan trekke gale slutninger om en variabel bør være med i modellen eller ikke.

I tillegg synes jeg det er merkelig at koeffisientene er så ustabile, både når vi går over til det reviderte datasettet, og når vi så utvider dette med de ekstra observasjonene. Dette tyder på at elastisitetene fra modellen ikke kan brukes til å si noe om effektene forklaringsfaktorene har på boligprisene, spesielt over tid.

Feiljusteringsmodellen som Jacobsen og Naug modellerer ligner på ingen måte de modeller lærebøkene viser. Det er spesielt at de har inkludert arbeidsledigheten på tidspunkt t i langtidssammenhengen med begrunnelse av at det var det som ga best føyning. Vanligvis er likevekten gitt med variabler som er lagget en gang mer enn de variablene som inngår i korttidodynamikken. Dette burde dessuten ført til at likevekten var gitt med $t-2$ variabler siden renten også inngår som lagget på endringsform. Alternativt impliserer spesifikasjonen alle nullrestriksjonene.

En annen ting som Jacobsen og Naug kanskje burde ha tenkt på er å inkludere trend i modellen. Vi både ser ut fra tidsserien, og vet fra erfaring at det er en underliggende trend i flere av de nominelle variablene. Dette kunne ha bedret prediksjonen hvis utvalget av variablene deres er korrekt. Det manglende hensynet til trendfaktoren gjør dessuten at de pålegger en restriksjon om at inntekt og boligmasse har den samme effekten med bare motsatt fortegn på lang sikt. Når vi trekker ut trendfaktoren så ser vi at denne restriksjonen ikke medfører riktighet.

Ved siden av dette synes jeg også at deres konstruerte forventningsvariabel er litt rar i og med at de bruker summen over to kvartaler før de putter dette inn i deres "formel". TNS Gallup rapporterer den kvartalsvise forventningsindikatoren²¹, og jeg kan ikke finne noen kilder på at denne er gitt som sum over to kvartaler.

²¹ http://www.orapp.no/forventningsind_nr_1_2007/20070515/forventningsindikatoren_sesongjustert/
Første kolonne er identisk med tallene i datasettet mitt.

De nevner ingenting i artikkelen om eventuelle endogene variabler eller simultanitetsproblemer. Dersom det er et høyt press i boligmarkedet kan det tenkes at sentralbanken vil bruke renten som et virkemiddel for å dempe boligprisveksten. Dessuten vil jeg anta at boligmassen er endogen ettersom det er rimelig at nybygging blir bestemt av priser på bruktboliger i tillegg til andre faktorer som bygge- og tomtekostnader. Endogenitet er et forhold som burde ha blitt tatt i betraktning.

Til slutt vil jeg påpeke at datagrunnlaget ikke er optimalt for å gjøre analyser om hva som driver boligprisene. Når vi betrakter observasjonene våre, er det ingen tall fra forløpet til forrige store nedgangsperiode i boligmarkedet, vi har kun med slutten av nedgangsperioden som endte i 1993. Begrensningen ligger i at boligprisene ikke ble samlet inn og lagret før 1985 av Eiendomsmeglerbransjens boligprisstatistikk (årlige innsamlinger frem til 1990), og før 1991 av Statistisk sentralbyrå. Ettersom det ikke har vært noen stokastiske sjokk i perioden vi ser på er det lite å modellere. Vi vet nemlig ikke hvilke variabler som er viktige drivere for sjokkene. Norsk økonomi har hatt høy vekst de siste 15 årene, og hvordan kan vi vite at boligprisveksten henger sammen med nevnte faktorer i stedet for at alle bestemmes av noe underliggende som driver norsk økonomi?

4.9. Konklusjon.

Modellen til Jacobsen og Naug inneholder autokorrelasjon, trolig som følge av spesifikasjonsfeil. Dette gjør at konklusjonene man trekker om hvorvidt en variabel har effekt eller ikke på boligprisene kan være feil. Selv om autokorrelasjon ikke gjør koeffisientene forventningsskjevne, så bør man løse problemet slik at inferenstesting bli gyldig. Jeg vil likevel understreke at mine observasjoner ikke er identisk deres, da jeg får noe forskjellige estimater.

På tross av en dårlig modell bidrar artikkelen med gode forslag på hva om kan påvirke boligprisene. Selv om de finner at en rekke av variablene ikke er signifikante med deres datasett, kunne det vært interessant å teste disse på nytt i et større datasett, hvor svingningene har vært større i flere av variablene.

Siden jeg finner at modellen i dette avsnittet ikke er spesifisert korrekt vil jeg i neste kapittel forsøke å komme med et forslag på forbedring.

5. En alternativ boligprismodell.

5.1. Innledning.

I NoU 2002 ble sentrale faktorer som spiller inn på boligprisene på kort sikt sagt å være disponibel inntekt og inntektsforventninger, skattefordel ved å eie egen bolig, det reelle rentenivået, forventninger om fremtidig boligpris og priser på drift og vedlikehold av boligen. Av disse valgte utvalget å gå dypere inn på virkningene av endret inntekt, renter og forventninger. Disse tre er derfor forklaringsfaktorer jeg vil ta med i min modell. Dessuten kan arbeidsledighetsraten være naturlig å inkludere i modellen, fordi hvis det blir dårligere tider med økt arbeidsledighet vil boligprisene trolig begynne å falle. På lang sikt er boligmasse en viktig variabel. Dersom det bygges flere boliger vil tilbudet økes, og etter hvert dempes prisveksten. I mitt modellforslag vil jeg med bakgrunn i dette starte med å inkludere disse fem variablene for å forklare utviklingen i boligprisene siden 1990.

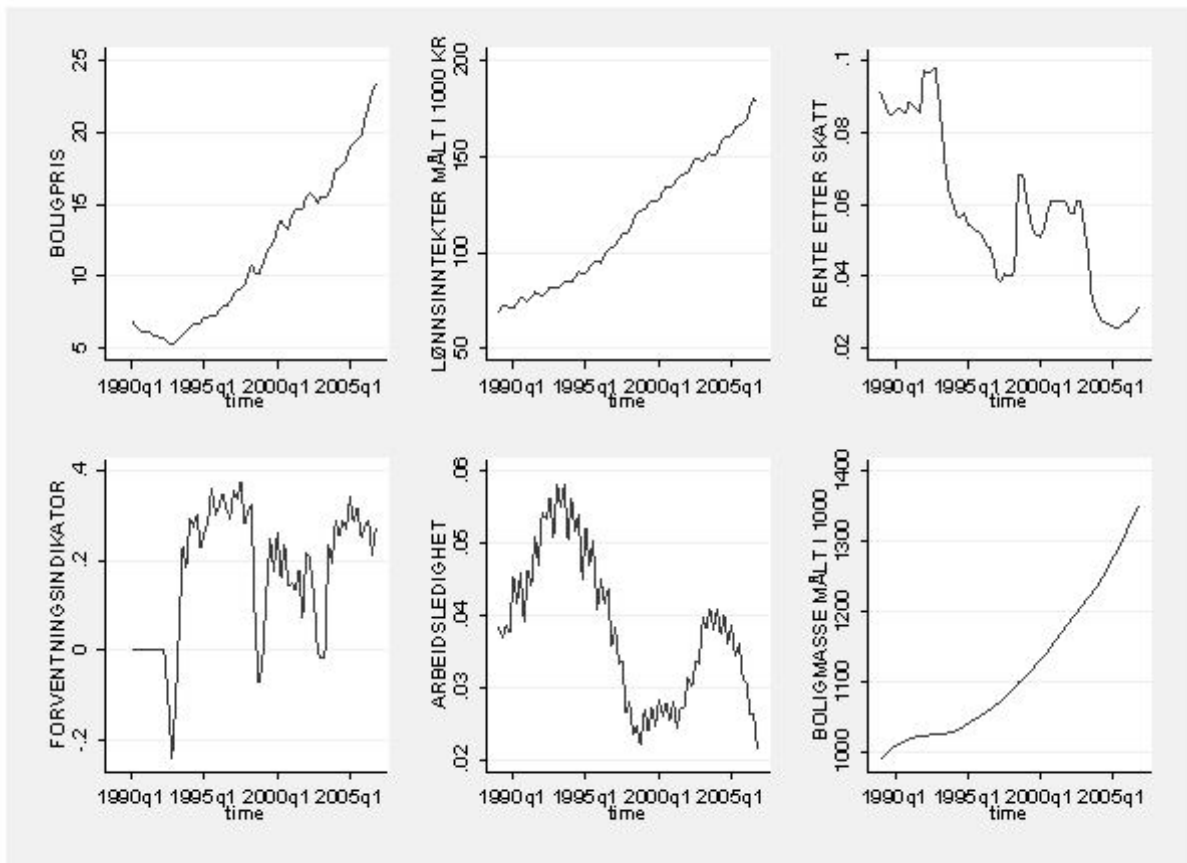
5.2. Tidsseriene.

5.2.1. Forventninger.

Jacobsen ga meg i tillegg til ”originaldataene” som jeg benyttet for å reestimere deres modell i forrige kapittel, også et utvidet sett som tok med data frem til og med 4.kvartal 2006. I forhold til originaldatasettet var det foretatt datarevisjoner og omlegginger, og både boligpris, rente, ledighet, inntekt og boligmasse hadde blitt revidert. Summary statistics er vist i vedlegg 1. Enkelte av variablene hadde relativt store endringer, og på grunn av dette tester jeg seriene for stasjonaritet på nytt. Særlig kan en serie som jeg for et mindre datasett fant å være stasjonær, nå bli ikke-stasjonær etter utvidelsen.

Vedrørende mine forventninger om variablene henviser jeg til kapittel 4.2, hvor samtlige variabler ble diskutert.

Figur 11: Tidsseriene med det utvidede datasettet.



Grafene over viser at boligpris, inntekt og boligmasse inneholder en trendfaktor. De samme variablene synes dessuten å se eksponentielle ut, og derfor er det nødvendig å transformere de til logaritmisk form før jeg tar de med i en lineær regresjonsmodell. Videre så bør nok arbeidsledigheten korrigeres for sesong siden den fremstår som hakkete. Det er også vanlig at lønnsinntektene korrigeres for sesong siden disse typisk har en oppgang i 3.kvartal. Forventningsindikatoren skissert her er den som TNS Gallup publiserer. Jeg kommer tilbake til den eventuelle egenkonstruerte forventningsvariabelen. Det synes ikke som at det er noen sjokk i perioden etter 1.kvartal 2004 som var da originaldatasettet stoppet. Det som er nytt er at rentenivået har begynt å stige igjen, opp mot et mer normalt nivå.

5.2.2. DF-testene for stasjonaritet.

Tabell 5: DF-testene.

Variabel	teststatistikk fra (A)DF-testen.		
	Antall lag	med trend med konstant	uten trend med konstant
Boligpris	5	-1,333	2,744
Inntekt	5	-2,010	1,770
Inntekt, justert for sesong	5	-1,849	1,803
Rente etter skatt	1	-2,473	-1,484
Forventningsindikator, E	3	-2,202	-2,139
Arbeidsledighet	4	-3,269*	-2,230
Arb.ledighet, sesongjustert	4	-3,078	-2,122
Boligmasse	5	0,637	3,012

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Når jeg utvider datamaterialet går boligprisen over til å bli ikke-stasjonær, hvilket er i tråd med forventningene mine. For inntekten har den signifikante lag-lengden økt fra det tidligere datasettet, men konklusjonen om ikke-stasjonaritet er uendret. Trenden er signifikant for rentevariabelen i mitt datasett. Dette er ikke en rimelig faktor å ta med for renten i et lengre løp, da renten ikke vil fortsette å stige eller falle i evig tid. Allikevel skjønner man ved å se på grafen av renten hvorfor statistiske tester gir at det er en negativ trend i dette datamaterialet. Uavhengig om jeg tar med trend eller ikke får jeg at renten er ikke-stasjonær i dette datasettet, og dette er ikke hva jeg forventer av renten. Forventningsindikatoren har gått fra å være stasjonær til å bli ikke-stasjonær. Jeg noterer meg at det er på marginen at forventningen ikke er stasjonær når den testes uten trend gitt et 10 % nivå. Arbeidsledighetens stasjonaritetsegenskaper avhenger om av om vi korrigerer for sesong. I tillegg så gir tallmaterialet at arbeidsledigheten bør inneholde en trendfaktor siden denne er signifikant. Dette er ikke rimelig på lengre sikt, da arbeidsledighetsraten svinger opp og ned med konjunktorene. Også grafen indikerer at denne ikke skal ha trend i seg. Uten trend blir variabelen ikke-stasjonær uavhengig av sesongkorreksjon. Med trend er arbeidsledigheten stasjonær når vi ikke justerer for sesong. Boligmassen er klart ikke-stasjonær som forventet.

5.3. Den konstruerte forventningsvariabelen.

Jeg bruker det samme prinsippet som i artikkelen til Jacobsen og Naug for å finne forventinger som skyldes andre forhold enn rente og ledighet siden disse inngår som egne

forklaringsvariable i modellen min. Her ser jeg på datamaterialet i perioden 1.kvartal 1990 til 4.kvartal 2004. I all kommende analyse betyr små bokstaver at en variabel er på logaritmisk form. Δ betyr at variabelen er på endringsform.

Tabell 6: Modell for husholdningenes forventninger til egen og landets økonomi.

Variabel	Koeffisient	Standardavvik
$\Delta(\text{RENTE}(1-\tau))_t$	-9,598***	(1,549)
$\Delta\text{ledighet}_t$	-0,350**	(0,169)
E_{t-1}	-0,212**	(0,096)
$(\text{RENTE}(1-\tau))_{t-1}$	-1,143	(0,738)
Ledighet_{t-1}	0,013	(0,031)
S1	0,189***	(0,044)
S2	0,077***	(0,020)
S3	0,172***	(0,038)
KONSTANTLEDD	0,029	(0,145)
Antall observasjoner	59	
R2	0,7086	
Justert R2	0,6620	
Q*(1)	0,1325	
Q*(4)	3,0997	
DF-teststatistikk	-7,436***	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Jeg tar vare på restleddet fra regresjonen over og tester dette for autokorrelasjon. Ljung-Box statistikken er kjikvadratfordelt med m frihetsgrader, der m angir den maksimale lag lengden. Nullhypotesen er ingen autokorrelasjon i modellen. For 1 frihetsgrad er kritisk verdi 3.84, og for 4 frihetsgrader er kritisk verdi 9.49. I tabellen over ser vi at begge teststatistikkene er langt under kritisk verdi, og ergo kan vi ikke forkaste nullhypotesen. Autokorrelasjon er ikke et problem i denne modellen.

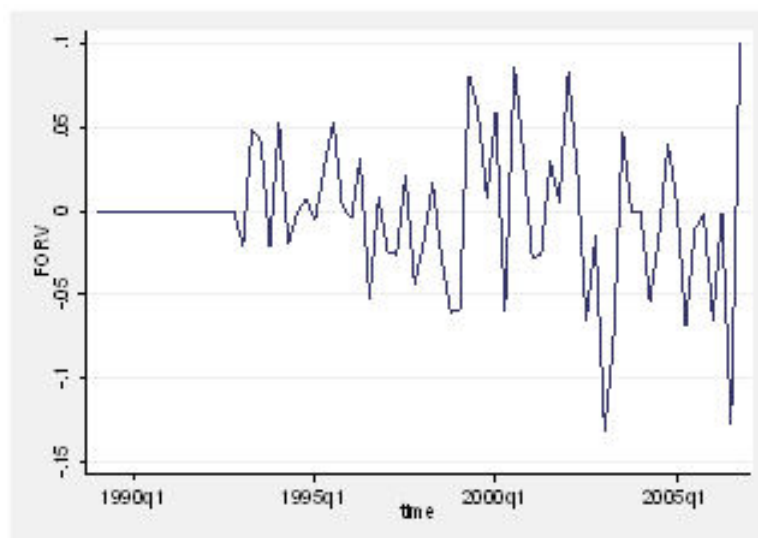
Siden forventningsindikatoren, renten og til dels arbeidsledigheten var ikke-stasjonære i følge DF-testen, så må vi undersøke om de er kointegrert. Den laggede forventningsindikatoren, E, er signifikant i tabell 6, og dette betyr at variablene er kointegrert. DF-testen av restleddet gir det samme resultatet.

Videre bruker jeg restleddet fra tabellen over til å konstruere en forventningsvariabel som fanger opp andre effekter enn de som skyldes rente og ledighet. Jeg vil ta vekk disse effektene

fordi nevnte variabler inngår i hovedregresjonen som egne forklaringsvariabler. Jeg velger å konstruere den nye forventningsvariabelen på en annen måte enn det som ble gjort i artikkelen. Tallene fra TNS Gallups' trendindikator er gitt på kvartalsbasis og er ikke summen over to kvartaler. Derfor ser jeg ingen grunn til å bruke summen over to kvartaler av restleddet, noe som Jacobsen og Naug gjorde i sin artikkel. I tillegg sier de ikke eksplisitt hvorfor de legger til et ledd med restleddet opphøyd i tredje. Det skyldes muligens at de ønsker at variabelen skal være tre ganger deriverbar for å fange opp noe om krumningen av konkavitet eller konveksitet. Dette valget er ikke opplagt for meg, så jeg velger å utelate dette ekstra leddet. Jeg kaller derfor restleddet fra tabell 6 for FORV, og dette er variabelen jeg skal benytte i den videre analysen.

Hva angår min konstruerte forventningsvariabel så velger jeg å gjøre som Jacobsen og Naug og sette denne lik null i perioden frem til og med 4. kvartal 1992. Bakgrunnen for dette valget er at denne variabelen er flaskehalsen i tallmaterialet, og ved å anta at variabelen er nøytral så øker størrelsen på datasettet. Når man antar at en variabel er nøytral, så vil den ikke ha noen effekt på modellen. Forskjellen fra å ikke ha noen observasjoner for FORV i det hele tatt er at vi nå tillater å bruke observasjonene fra de andre forklaringsvariablene som ellers ikke ville ha blitt tatt med. Selv om det kan virke ukorrekt å si at variabelen har en verdi den ikke har (null) i en periode, så kan det forsvares ved at tallmaterialet er såpass lite at fordelene ved å benytte et større datasett overstiger ulempen ved å "feilaktig" sette verdien på variabelen lik null.

Figur 12: Den konstruerte forventningsindikatoren.



I figur 12 er den konstruerte forventningsvariabelen skissert. Den ser ut til å være uten konstantledd. Vi ser at det fremdeles er variasjon i tallmaterialet etter at vi har tatt bort effektene fra rente og ledighet, men siden variabelen er konstruert vil den også inneholde støy. Sjøkket i 2003 kommer frem i grafen, og vi ser at det også var et negativt sjokk i 3. kvartal 2006. I denne perioden falt børsen ganske kraftig, så dette kan være en av grunnene til de negative forventningene vi observerer. Børsfallet varte kort, og i 4.kvartal var aksjekursene på vei opp igjen. Frykten for nedgangstider var trolig glemt og forventningene var tilbake på et høyt nivå. En DF-test av FORV gir at variabelen er stasjonær gitt 1 % signifikansnivå, teststatistikken for en test uten lags og konstantledd er på -8.393.

5.4. Hva driver boligprisene?

5.4.1. En alternativ modell.

Jeg velger å bruke en dynamisk modell fordi boligmarkedet er et saktebevegende marked, hvor det kan være tregheter som må tas hensyn til. Fordi jeg vil undersøke både korttids og langtidseffektene i boligprisene, så modellerer jeg en såkalt feiljusteringsmodell. Datamaterialet mitt består av observasjoner av rente etter skatt, inntekt, arbeidsledighet, forventninger, boligmasse og boligpriser. Av forklaringsvariablene forventer jeg at inntekt og forventninger vil ha en positiv sammenheng med boligprisene, mens økte renter, ledighet og boligmasse vil legge en demper på boligprisene.

Jeg starter med å bygge en balansert modell hvor de samme variablene inngår både i korttids og langtidseffektene. I tillegg inkluderer jeg sesonger, da enkelte av variablene inneholder en sesongfaktor. Fordi jeg ønsker å se hvor godt modellen predikerer boligprisene så utelater jeg de to siste årene med observasjoner når jeg bygger modellen. Resultatet av denne første regresjonen er gitt i vedlegg 2. Jeg finner at boligmassen på endringsform ikke er signifikant, noe jeg heller ikke hadde regnet med. Boligpristeori sier at tilbudet er konstant på kort sikt, og boligmassen måler tilbudet av boliger. Siden både økonomisk teori og statistikken sier at boligmassen ikke er relevant på kort sikt, så velger jeg å utelate boligmassen på endringsform i den videre analysen. Dermed ender jeg opp med en feiljusteringsmodell som ut fra observasjonene i perioden 1990 til 2004 gir følgende estimater og standardavvik:

Tabell 7: En alternativ feiljusteringsmodell for logaritmen til boligprisene.

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,056	(0,296)
Δ (RENTE(1- τ)) _t	-3,305***	(0,471)
Δ FORV _t	0,114**	(0,053)
Δ ledighet _t	-0,086	(0,060)
boligpris _{t-1}	-0,190***	(0,048)
inntekt _{t-1}	0,093	(0,067)
(RENTE(1- τ)) _{t-1}	-0,936***	(0,242)
FORV _{t-1}	0,275***	(0,072)
ledighet _{t-1}	-0,076***	(0,025)
boligmasse _{t-1}	0,522	(0,325)
S1	0,052***	(0,017)
S2	0,026**	(0,012)
S3	0,010	(0,017)
KONSTANTLEDD	-8,118**	(3,994)
Antall observasjoner	59	
R2	0,8489	
Justert R2	0,8052	
Q*(1)	0,9812	
Q*(4)	1,2016	
DF-teststatistikk	-9,005***	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Dette er en 1-steps feiljusteringsmodell, og koeffisientene til de laggede variablene i regresjonen over er ikke identisk med koeffisientestimatene i den langsiktige likevekten. Disse må man regne ut for seg selv ved å dele estimatet på minus estimatet til den laggede boligprisen. I tabellen under viser jeg de langsiktige elastisitetene for tolkningsformål. Der hvor jeg i modellen har variabler som ikke er på logaritmisk skala har jeg passet på å gange inn snittet av forklaringsvariabelen slik at estimatet blir det samme som elastisiteten:

Langsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
inntekt _{t-1}	0,491	(0,392)
(RENTE(1- τ)) _{t-1}	-0,303***	(0,088)
FORV _{t-1}	0,0002***	(0,00006)
ledighet _{t-1}	-0,401***	(0,091)
boligmasse _{t-1}	2,743**	(1,218)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Også for de kortsiktige estimatene må vi huske på å gange inn snittet av rente og forventning slik at vi får koeffisientene uttrykt som elastisiteter:

Kortsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
$\Delta \text{inntekt}_t$	0,056	(0,296)
$\Delta(\text{RENTE}(1-\tau))_t$	-0,204***	(0,029)
ΔFORV_t	0,00001**	0,000007
$\Delta \text{ledighet}_t$	-0,086	(0,060)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

5.4.1. Tolkning av elastisitetene og justeringsparameteren.

De langsiktige elastisitetene:

Når vi befinner oss i likevekt, langtidsløsningen, så vil en 1 % økning i inntekt føre til at boligprisene øker med 0,49 %. Inntekt er ikke signifikant i modellen, men fordi økonomisk teori sier at inntekt skal være med, så velger jeg å beholde den som en forklaringsfaktor. Fortegnet er i tråd med mine forventninger.

Dersom renten øker med 1 % så vil boligprisene reduseres med $0,303 \cdot (0,72) = 0,218$ %. Rente etter skatt er signifikant i modellen. Dersom renten øker fra for eksempel 5 % til 6 % så vil dette tilsvare en økning på 20 %. Boligprisene vil med en slik økning i utlånsrenten reduseres med $(20 \cdot 0,218) = 4,36$ %. At boligprisene faller med litt over 4 % når renten øker med ett prosentpoeng er en rimelig størrelsesorden.

Hvis forventningene øker med 1 % så vil boligprisene øke med 0,0002 %. Dette er en svært liten økning, men variabelen er signifikant på 1 % nivå så jeg velger å beholde forventninger som en forklaringsvariabel i modellen. Det må skje store, varige skift i forventningene før det gjenspeiles i boligprisene i likevekt.

Elastisiteten til boligprisene med hensyn på arbeidsledighet er -0,401, og er signifikant med 1 % signifikansnivå. Dette betyr at hvis ledigheten øker med 1 % så faller boligprisene med 0,401 %. Både fortegnet og størrelsen på effekten virker rimelig siden økt arbeidsledighet henger sammen med mindre gode tider, og i usikre tider tør ikke folk flest å investere i egen bolig i samme grad som før.

Hvis boligmassen øker med 1 % vil boligprisene øke med 2,743 % i følge dette datasettet. Dette stemmer ikke overens med mine forventninger om at økt tilbud i en viss grad demper prispresset. I tillegg ser jeg at standardavviket er relativt stort, noe som indikerer at estimatet ikke er spesielt presist. Et 95 % konfidensintervall for boligmasse er [0.289, 5.197] når kritisk t -verdi med 45 frihetsgrader er 2,0141. Den positive effekten kan skyldes at det faktisk er en positiv sammenheng mellom boligmasse og boligpriser, eller at det foreligger et simultanitetsproblem. For eksempel kan en mulig forklaring på den positive effekten være at boligmassen tilpasser seg etterspørselen over tid, og derfor øker boligprisene når etterspørselen (og dermed boligmassen) øker. Fordi både tilbud og etterspørsel bestemmer boligmassen, klarer jeg ikke å skille ut tilbudseffekten på boligprisene ved å bare inkludere boligmassen som en variabel. Siden økt etterspørsel øker prisveksten, mens økt tilbud demper prisveksten, vil totaleffekten fra boligmasse kunne bli positiv. I den videre analysen velger jeg å jobbe videre med denne modellen, og heller komme tilbake til simultanitetsproblemet hvis tiden tillater det. De resterende koeffisientestimatene i modellen vil fremdeles være konsistente, det er kun for den eventuelle endogene variabelen at estimatene ikke gir noen mening.

Hvor raskt justeres avvik inn?

Likevekten i dette datasettet er gitt ved de langsiktige elastisitetene vist over;

$$\text{boligpris} = 0,491 * \text{inntekt} - 0,303 * \ln(\text{RENTE}(1 - \tau)) + 0,0002 * \ln(\text{FORV}) - 0,401 * \text{ledighet} + 2,743 * \text{boligmasse}$$

Dersom det oppstår en ulikevekt i langtidssammenhengen, vil modellen hver periode justere inn 19 % av avviket fra likevekten. Dette finner vi fra tabell 7 som koeffisienten foran lagget

boligpris. Det tar $\left(\frac{1}{0,19}\right) \approx 5$ kvartaler før vi kommer tilbake i likevekt. Dersom det kommer et

sjokk på tidspunkt t vil 19 % av avviket justeres inn etter 1 kvartal, 34,4 % vil være justert etter 2 kvartaler, 46,9 % av avviket vil være justert inn etter 3 kvartaler osv. Denne justeringsparameteren er svært signifikant, og dette betyr at parametrene i den langsiktige likevekten kointegrerer. Boligpris, inntekt, rente, ledighet og boligmasse henger med andre ord sammen på lang sikt.

Kortsiktige elastisiteter:

Dersom inntekten øker med 1 % så vil boligprisene på kort sikt øke med 0,056 %. Igjen er inntektsestimatet ikke signifikant forskjellig fra null, men fordi det er en rimelig antagelse at større endringer i inntekt kan slå raskt ut i betalingsviljen for boliger, så velger jeg å beholde variabelen også i korttidssammenhengen til boligprisveksten.

En renteendring er signifikant på et 1 % nivå, og er derfor en relevant variabel for å si noe om boligprisveksten på kort sikt. Dersom utlånsrenten øker med 1 %, så vil boligprisene reduseres med $0,72 \cdot 0,204 \% = 0,147 \%$ på kort sikt når vi tar hensyn til skatt. Dersom renten økes fra 5 % til 6 % (20% økning) vil boligprisene reduseres med $(20 \cdot 0,147) = 2,94 \%$ når vi tar hensyn til skattefordelen ved å eie egen bolig. At boligprisene faller med nesten 3 % på kort sikt hvis renten økes fra 5 % til 6 % virker fornuftig. Ser vi derimot på en renteendring fra 2 % til 3 % tilsvarer dette en endring på 50 %. Modellen sier da at en slik økning i renten vil gjøre at boligprisene faller med $(50 \cdot 0,147) = 7,35 \%$. Dette virker kanskje ikke så rimelig, da det er normalt å anta at en renteøkning har større effekt på boligprisene når man nærmer seg marginen for hva man kan betale enn når renten er på et unormalt lavt nivå.

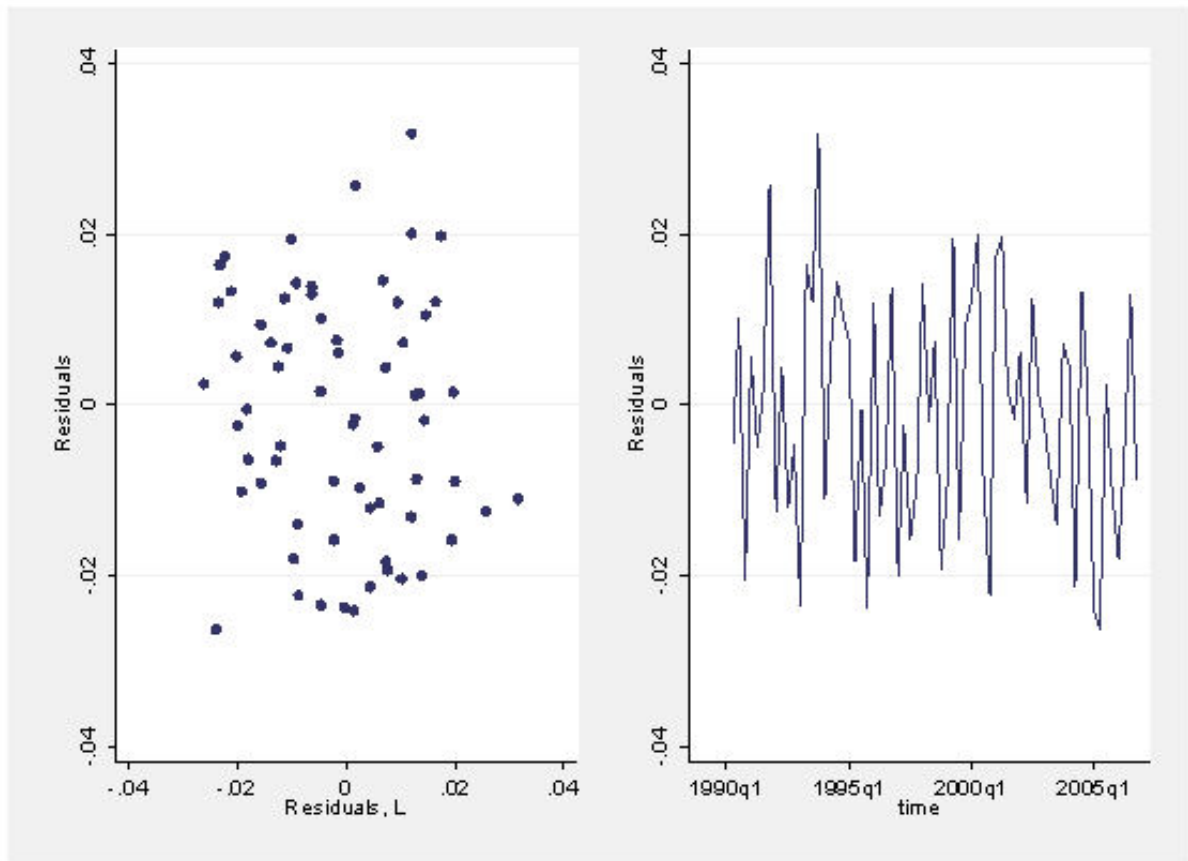
Den konstruerte forventningsvariabelen er igjen signifikant, men svært nær null. Det er riktig å ta den med i modellen, men effekten av at forventningene endres med 1 % er omtrent null. Derfor vil kun veldig høye eller veldig lave forventninger slå ut i boligprisene.

Arbeidsledigheten er ikke signifikant på kort sikt. Det er en avveining om endringer i arbeidsledighet virker inn på endringer i boligpris på kort sikt, som regel antar man at det skjer en gradvis endring i arbeidsledighetsraten som først merkes på lengre sikt. Jeg velger likevel å beholde den i korttidssammenhengen, for hvis det skulle komme store endringer i arbeidsledighetsraten i løpet av et kort tidsrom så får det trolig ringvirkninger til boligmarkedet i løpet av kort tid. Datasettet mitt gir at dersom arbeidsledigheten øker med 1% så vil boligprisene avta med 0,086 %. Øker arbeidsledigheten fra 3 % til 4 % tilsvarer dette en økning på $\left(\frac{0,04 - 0,03}{0,03} \right) = 33,3\%$ Med en slik endring i arbeidsledigheten vil boligprisene falle med $(33,3 \cdot 0,086) = 2,87 \%$. Dette virker som en rimelig reduksjon i boligprisene når arbeidsledighetsraten øker fra 3 % til 4 %.

5.4.2. Er modellen gyldig?

En grafisk test av feilleddene fra tabell 7 er vist i figur 12:

Figur 13: Plott av feilleddene fra tabell 7.



Ut fra figuren over ser det ut som at feilleddene er tilfeldig spredt innenfor de fire kvadrantene og at de over tiden beveger seg over x-aksen med jevne mellomrom. Om de passerer x-aksen for ofte, er vanskelig å si ut fra figuren. Derfor må vi gå over til statistiske tester.

Ljung-Box testen forkaster førsteordens og fjerdeordens autokorrelasjon i modellen. Vi husker fra i stad at kritisk verdi for 1 frihetsgrad er 3.84, og for 4 frihetsgrader 9.49, og teststatistikkene Q^* i tabell 7 ligger langt under disse.

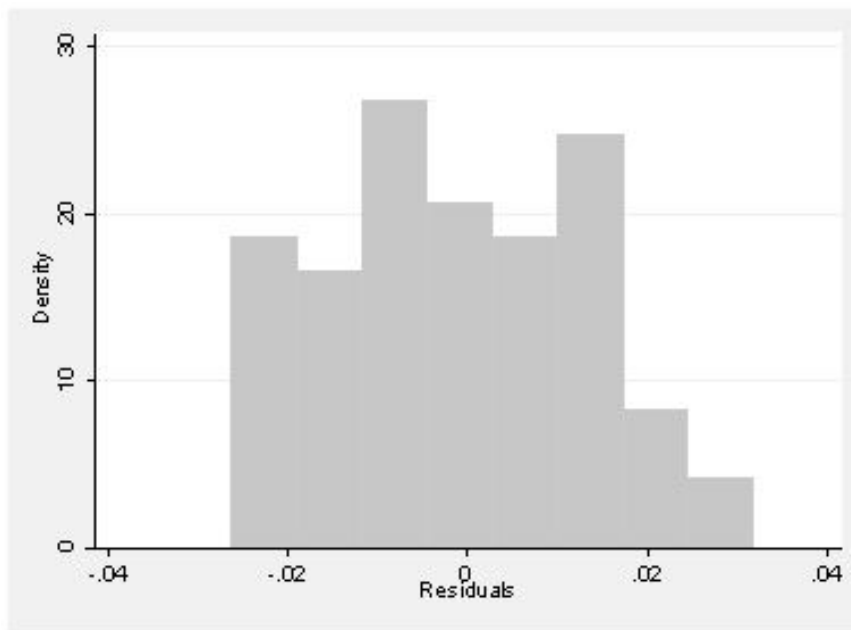
5.4.3. Videre tester.

Dersom man foretar en test for utelatte variabler kalt Ramsey RESET test så får man en F-verdi på 0.17, og man kan ikke forkaste nullhypotesen om at det ikke er utelatt sentrale

variabler. Breush-Pagan testen for heteroskedastisitet gir en teststatistikk på 0.44, da kan vi ikke forkaste nullhypotesen om konstant varians med en kjikvadratfordeling med 1 frihetsgrad. Dette tyder på at betingelsen om konstant varians holder og at vi ikke har glemt noen viktige forklaringsvariabler.

Det neste jeg vil sjekke er om feilleddene er tilnærmet normalfordelte.

Figur 14: Et histogram av feilleddene fra tabell 7.



Feilleddene er noenlunde symmetriske, slik at OLS estimatene jeg får er akseptable å bruke. Det er ikke en klar normalfordeling, men de høyeste verdiene ligger i midtsjiktet. Et feilledd som ikke er normalfordelt gjør at forutsetningen bak t - og F -testene blir brutt. Inferenstestene kan i så fall bli uriktige. Spesielt fordi det ikke er særlig skjevhet, så godtar jeg dette som en tilnærmet normalfordeling.

5.5. Multikolaritet.

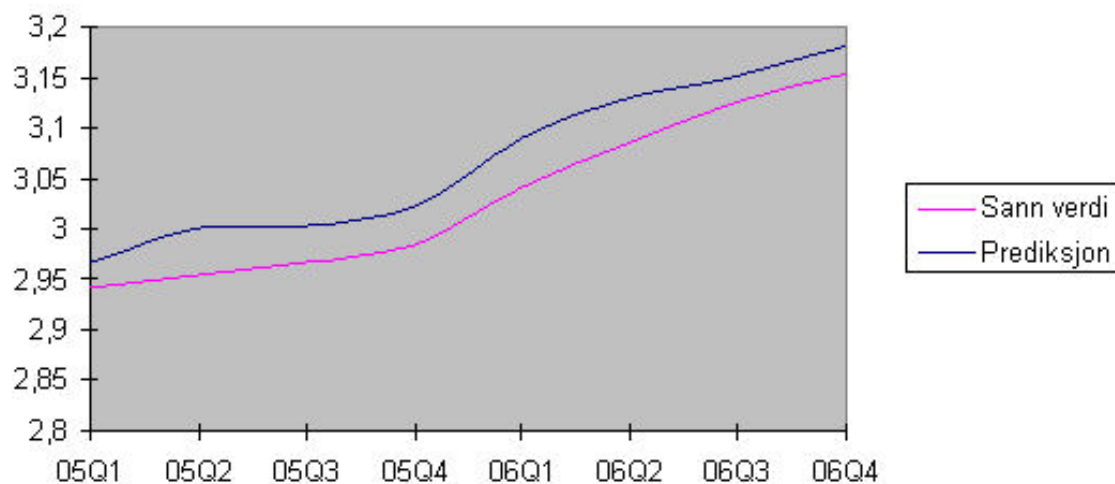
Datsettet mitt gir en korrelasjonskoeffisient på 0,9639 mellom boligmasse og inntekt, jf. korrelasjonsmatrisen i vedlegg 3. Dette er en meget høy korrelasjon, og jeg ble litt bekymret for nær multikolaritet. Tidligere i oppgaven har jeg vist grafer over tidsseriene, og vi ser at både boligmassen og inntekten inneholder en klar trend. Spørsmålet er da om det er denne trenden som utløser den høye korrelasjonskoeffisienten. Jeg har trend som en

forklaringsvariabel i min endelige modell, så for å justere boligmassen og inntekten for trend, kjører jeg en regresjon på inntekt forklart av trend og sesonger, og en tilsvarende regresjon på boligmasse. Jeg tar vare på restleddene fra regresjonene, da disse inneholder variasjoner i inntekt og boligmasse som skyldes andre forhold enn trend og sesonger. Når jeg så sjekker korrelasjonen mellom disse to finner jeg en korrelasjonskoeffisient på -0.0116 , dvs. at det er knapt noen korrelasjon i det hele tatt mellom disse to variablene. Multikolaritet er derfor neppe et problem i denne modellen.

5.6. Prediksjon.

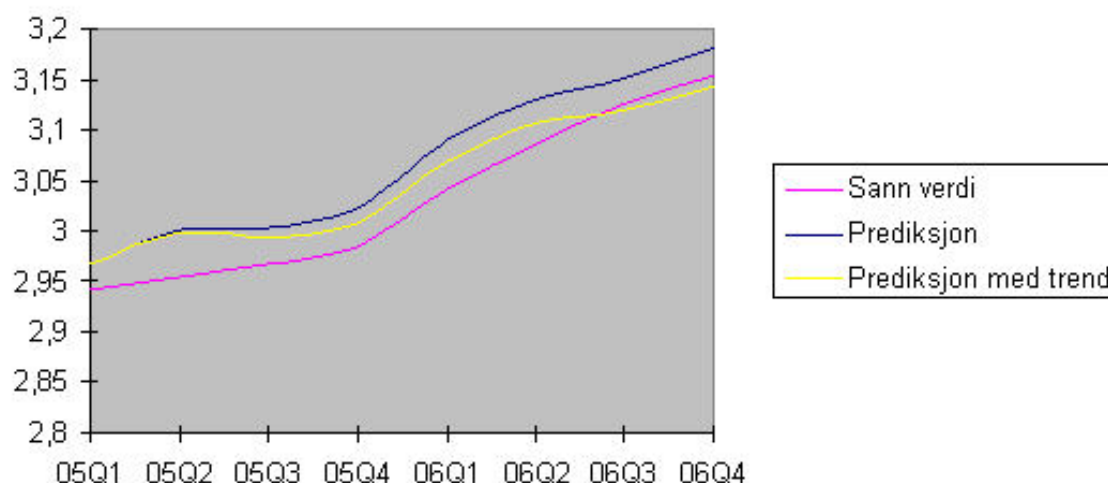
Modellen over synes å forklare variasjonene i boligprisveksten ganske godt, og det neste jeg vil gjøre er å undersøke om den predikerer de to årene som jeg hittil har utelatt godt. Figuren under viser den sanne verdien på logaritmen til boligprisene og den predikerte boligprisen i samme diagram:

Figur 15: Sann boligpris og predikert boligpris for de 8 kvartalene i 2005 og 2006.



Modellen jeg har kommet frem til predikerer forholdsvis godt, men vi ser i figuren at den overpredikerer boligprisene i hele perioden. En løsning på dette kan være å inkludere trend som en egen forklaringsvariabel i modellen. Dette er fornuftig, da det er en underliggende trend i boligprisene som vi bør korrigere for.

Figur 16: Sann boligpris, predikert boligpris uten trend og predikert boligpris med trend.



Når jeg predikerer boligprisene med trend inkludert finner jeg at modellen overpredikerer i starten, mens den de siste par kvartalene underpredikerer boligprisene. Det er ikke store avvik det er snakk om. Sammenligner jeg de to prediksjonene så ligger den med trend nærmest den sanne verdien på boligprisene. Dette kan man se grafisk, evt. så kan man benytte det samme prinsippet som for OLS med å minimere kvadrerte avvik. På bakgrunn av prediksjonsegenskapene så velger jeg å gå videre med modellen som også tar med trend som en forklaringsfaktor. Denne modellen er vist i vedlegg 2, og tolkningene blir på samme måte som vist før.

5.7. Er parametrene stabile?

Nå som jeg har funnet en modell som predikerer relativt godt, kan jeg utvide datasettet til å inkludere alle observasjonene jeg har tilgjengelig. Jeg må også estimere forventningsvariabelen om igjen i tillegg til selve feiljusteringsmodellen. Denne estimeringen av forventningsvariabelen ligger i vedlegg 2.

Det kan hende at modellen er estimert på et for lite datasett eller at jeg har feilspesifisert noen av variablene uten at tester for autokorrelasjon har fanget opp dette, og en måte å avdekke dette på kan være å se om koeffisientestimatene endrer seg mye når jeg inkluderer de to siste årene i modellen. Under viser jeg derfor både resultatene fra modellen der jeg ekskluderte de to siste årene, og modellen for hele datasettet, år 1990 til 2006 i den samme tabellen:

Tabell 8: En sammenligning av estimatene fra modellen med trend når jeg ser på to forskjellige utvalg, det ene utvalgene er en utvidelse av det andre.

Variabel	Uten 2005 og 2006		Hele datasettet	
	Koeffisient	Standardavvik	Koeffisient	Standardavvik
Δ inntekt _t	0,218	(0,374)	0,049	(0,312)
Δ (RENTE(1- τ)) _t	-3,289***	(0,474)	-3,437***	(0,460)
Δ FORV _t	0,121**	(0,054)	0,100**	(0,046)
Δ ledighet _t	-0,074	(0,063)	-0,066	(0,053)
boligpris _{t-1}	-0,205***	(0,053)	-0,193***	(0,049)
inntekt _{t-1}	0,357	(0,376)	0,155	(0,261)
(RENTE(1- τ)) _{t-1}	-1,115***	(0,349)	-0,945***	(0,316)
FORV _{t-1}	0,287***	(0,075)	0,242***	(0,070)
ledighet _{t-1}	-0,053	(0,042)	-0,077***	(0,023)
boligmasse _{t-1}	0,504	(0,328)	0,472**	(0,205)
S1	0,055***	(0,017)	0,047***	(0,016)
S2	0,028**	(0,012)	0,025**	(0,011)
S3	0,008	(0,017)	0,008	(0,015)
TREND	-0,003	(0,005)	-0,001	(0,004)
KONSTANTLEDD	-10,686*	(5,395)	-8,107	(4,934)
Antall observasjoner	59		67	
R2	0,8506		0,8380	
Justert R2	0,8031		0,7944	
Q*(1)	0,8409		1,1065	
Q*(4)	1,3291		1,5681	
DF-teststatistikk	-8,989***		-9,151***	

Langsiktige elastisiteter:

Variabel	Uten 2005 og 2006		Hele datasettet	
	Koeffisient	Standardavvik	Koeffisient	Standardavvik
inntekt _{t-1}	1,747	(1,737)	0,806	(1,228)
(RENTE(1- τ)) _{t-1}	-0,336***	(0,097)	-0,284***	(0,080)
FORV _{t-1}	0,0002***	(0,00006)	0,001***	(0,0003)
ledighet _{t-1}	-0,258	(0,208)	-0,397**	(0,162)
boligmasse _{t-1}	2,464*	(1,233)	2,446***	(0,657)

Kortsiktige elastisiteter:

Variabel	Uten 2005 og 2006		Hele datasettet	
	Koeffisient	Standardavvik	Koeffisient	Standardavvik
Δ inntekt _t	0,218	(0,374)	0,049	(0,312)
Δ (RENTE(1- τ)) _t	-0,203***	(0,029)	-0,199***	(0,027)
Δ FORV _t	0,00002**	(0,000007)	0,0001**	(0,00003)
Δ ledighet _t	-0,074	(0,063)	-0,066	(0,053)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Inntekt er den variabelen som endrer seg mest når jeg utvider datasettet, men inntekt er ikke signifikant forskjellig fra null i noen av utvalgene. Estimaten for samtlige variabler ligger innenfor konfidensintervallet i det andre utvalget.²² Resultatene synes å være ganske robuste siden estimatene ikke er veldig sensitive når datasettet utvides, dette på tross av jeg har ganske få observasjoner.

5.8. Simultanitetsproblemer.

Et simultanitetsproblem oppstår når en av forklaringsfaktorene bestemmes samtidig med forklart variabel. Et eksempel kan være at boligmassen er en endogen variabel som blir påvirket av boligprisen, slik at forklart variabel også påvirker en av forklaringsvariablene i modellen. Da vil boligmassen være korrelert med feilleddet, og en av antakelsene bak OLS er brutt. Implikasjonen av et simultanitetsproblem er at estimatene blir forventningsskjeve og dessuten ikke konsistente. Det betyr at estimatet på boligmassen ikke er forventningsrett selv asymptotisk.

I tillegg til at boligmassen trolig er endogen, så kan det hende at renten ikke er fullstendig eksogen. I dag er det et enormt press i boligmarkedet, og dette kan gjøre at renten blir brukt som et virkemiddel mot boligprisveksten. Dersom dette skjer er ikke estimatet på renten konsistent.

En løsning på endogene variabler er å finne et instrument som jeg kan benytte isteden. Det som er viktig er at instrumentet er sterkt korrelert med variabelen det skal erstatte, og samtidig at det ikke er korrelert med feilleddet. Den første betingelsen kan vi sjekke, men den andre må vi basere på for eksempel økonomisk teori.

Dersom vi finner et godt instrument kan vi benytte 2SLS²³ til å konstruere en ny variabel, og denne kan vi sette inn i regresjonen istedenfor variabelen vi mistenker at er endogen. For eksempel kan kanskje byggekostnadene være et instrument for boligmassen, da dette påvirker om folk bygger nye hus eller om de heller kjøper brukte boliger eller leier. På den annen side kan man lure på om byggekostnadene er noe som også bestemmes av boligprisene, for da er ikke disse uavhengige av feilleddet. Å finne et passende instrument krever tid og kreativitet,

²² Se vedlegg 2.

²³ Two-stage-least-square; en metode for å estimere en instrumentvariabel.

og bare det kan være en oppgave i seg selv. Derfor vil jeg ikke gå videre inn på instrumentvariabler her, da dette ligger utenfor rammen av denne oppgaven. En IV-regresjon vil derimot være en naturlig oppfølging av denne modellen.

5.10. Konklusjon.

Jeg har forsøkt å bygge en modell som forbedrer det vi så i forrige kapittel. Allikevel er jeg klar over at det er mangler, og særlig det eventuelle simultanitetsproblemet ville jeg gått nærmere inn på om tiden hadde tillatt det.

Jeg baserer fremdeles analysen på et lite utvalg, kun 67 observasjoner, uten nevneverdige observasjoner på sjokk i økonomien. Derfor er det ubesvart om effektene jeg måler er de som bestemmer boligprisen når den faktisk svinger, eller om det er andre underliggende effekter som har blitt utelatt i analysen. Det som er sikker er at tallmaterialet er for lite til at det kan trekkes bastante konklusjoner.

6. Avslutning.

I starten drøftet jeg mulige forklaringsfaktorer bak boligprisen. Faktorene som driver etterspørselen påvirker boligprisene på kort sikt, det være seg rente, forventninger og markedsleie for å nevne noen. På lengre sikt vil også boligmasse og demografi kunne være noen av faktorene som forklarer utviklingen i boligprisene.

Jeg fant at modellen til Jacobsen og Naug inneholder første ordens autokorrelasjon hvilket gjør at gale beslutninger om hvilke variabler som er relevante for å forklare boligprisutviklingen kan ha blitt tatt. Det kan for eksempel hende at variabler som egentlig burde ha vært med i modellen har blitt utelatt. I tillegg er den manglende stabiliteten i koeffisientene et dårlig tegn, hvilket kan bety at modellen er tilpasset utvalget av data i stor grad. Ved siden av dette så er det ikke diskutert problematikken med endogene variabler, verken for boligmassen, renten eller noen av de andre variablene.

I mitt forslag til modell så har jeg forsøkt å løse en del av problemene jeg fant i Jacobsen og Naug's modell. Jeg finner en modell som er uten autokorrelasjon, som ikke pålegger restriksjoner på to av variablene, som har relativt stabile parametere og som tar hensyn til trend som definitivt forklarer noe av utviklingen vi ser i boligprisene og i de andre tidsseriene. Men på grunn av manglende tid og data har jeg ikke funnet et instrument for boligmasse, og estimatene mine for denne variabelen er antakelig verken forventningsrette eller konsistente.

Konklusjon: På tross av at min modell har svakheter i og med at endogenitetsproblemet ikke er løst, så syns jeg at jeg har kommet frem til en forbedret versjon av Jacobsen og Naug sin modell. Jeg har løst en del av de statistiske problemene, og dessuten funnet en modell som er mer intuitiv å tolke.

Referanser.

Aly, Hassan Y. og Strazicich, Mark C.: "Terrorism and Tourism: Is the Impact Permanent or Transitory? Time Series Evidence From Some MENA Countries", s.4.

Internettadresse: <http://www.erf.org.eg/html/btrade1.pdf> [Lesedato 27/05/07]

Ball, L.: Aggregate Demand and Long-Run Unemployment. Paper presented at the Brookings Panel on Economic Activity, September 1999.

Boligbygg: Prisstatisikk for markedsleie, 2007.

Internettadresse: <http://boligbygg.reeltime.no/> [Lesedato 27/05/07]

Brooks, Chris: "Introductory econometrics for finance", Cambridge University Press, 2005.

Dagens Næringsliv: DN privatøkonomi, mars 2007, s.15

Davidson, Russell og MacKinnon, James G.: "Estimation and Inference in Econometrics", Oxford University Press, 1993, s. 724-725.

ECON Analyse: "Justeringer i Eiendomsmeglerbransjens boligprisstatistikk", Notat 2004-007

Internettadresse: http://www.nef.no/asset/44/1/44_1.pdf [Lesedato 07/02/07]

Glava AS: Kundebrev fra Glava AS: Leveringssituasjonen, april 2007.

Internettadresse: <http://www.glava.no/default.asp?page=93&article=296> [Lesedato 27/05/07]

Hall, S., Psaradikis, Z. og Sola, M.: "Switching error-correction models of house prices in the United Kingdom", Economic Modelling 14, 1997, s.549-565.

Harris, Richard: "Using Cointegration Analysis in Econometric Modelling", Prentice Hall, 1995, s. 25, 53.

International monetary fund (IMF): "The global house price boom", september 2004 s.71-76

Internettadresse: <http://www.imf.org/external/pubs/ft/weo/2004/02/pdf/chapter2.pdf> [Lesedato 28/05/07]

Jacobsen, Dag Henning og Naug, Bjørn E.: ”Hva driver boligprisene?”, Penger og Kreditt nr. 4/2004.

Kredittilsynet: ”Boliglånsundersøkelsen 2006”, februar 2007.

Internettadresse: http://www.kredittilsynet.no/archive/f-avd_word/01/07/Bolig040.doc

[Lesedato 25/04/07]

Manpower: ”Høy sysselsettingsvekst og rekordlav ledighet: Mot alvorlig luksusproblem for bedriftene”, september 2006. Internettadresse:

http://www.manpower.no/MPNet3/Content.asp?NodeRef=42317&Ref=NORWAY_JOB_BOARD&LangID=no [Lesedato 31/05/07]

Mean, Geoffrey P.: ”The removal of mortgage market constraints and the implications for econometric modelling of UK house prices”, Oxford Bulletin of Economics and Statistics 52, 1990, s.1-23.

Norges Eiendomsmeglerforbund, Eiendomsmeglerforetakenes Forening, ECON, FINN.no: Eiendomsmeglerbransjens boligprisstatistikk, april 2007.

Internettadresse: http://www.nef.no/asset/1503/1/1503_1.pdf [Lesedato 15/05/07]

Norges Eiendomsmeglerforbund, Eiendomsmeglerforetakenes Forening, ECON, FINN.no: Eiendomsmeglerbransjens boligprisstatistikk, historikk boligpriser (1985-2007).

Internettadresse: <http://www.nef.no/index.gan?id=1614&subid=0> [Lesedato 14/05/07]

Norges offentlige utredninger, NoU 2000:21: ”En strategi for sysselsetting og verdiskaping”, vedlegg 2: kapittel 2.2.4.

Internettadresse: <http://www.regjeringen.no/nb/dep/fin/dok/NOUer/2000/NOU-2000-21/16.html?id=360358> [Lesedato 31/05/07]

Norges offentlige utredninger, NoU 2002:2: ”Boligmarkedene og boligpolitikken” Kapittel 3 og kapittel 5.

Internettadresse: <http://www.regjeringen.no/nb/dep/krd/dok/NOUer/2002/NOU-2002-2.html?id=145338> [Lesedato 19/04/07]

OECD (1998) The OECD Jobs Strategy: Progress Report on Implementation of Country-Specific Recommendations. Economics Department, Working Papers No. 196.

Qvale, Jon Helgeland: ”En økonometrisk analyse av det europeiske markedet for blåskjell, østers og kamskjell”, SNF-Rapport nr. 55/2001, s.81

Røed, K. (1996) Unemployment Hysteresis – Macro Evidence from 16 OECD Countries. Empirical Economics, Vol 21, 589–600.

Sentralbanksjefens årstale 2007: Økonomiske perspektiver: Foredrag av sentralbanksjef Svein Gjedrem på Norges Banks representantskapsmøte torsdag 15. februar 2007.

Internettadresse: <http://www.petroleumsfondet.no/front/pakke/no/foredrag/2007/2007-02-15/>
[Lesedato 17/04/07]

Wooldridge, Jeffrey M.: ”Introductory Econometrics, A Modern Approach”, Thomson South-Western, 2003

Økonomisk Rapport på nett: ”Forventningsindikatoren”, mai 2007. Internettadresse: http://www.orapp.no/forventningsind_nr_1_2007/20070515/forventningsindikatoren_sesongj_ustert/ [Lesedato 21/05/07]

Økonomisk Rapport på nett: ”Verste måling på 10 år”, mai 2003.

Internettadresse: http://www.orapp.no/verste_mal/ [Lesedato 18/05/07]

Vedlegg 1. Summary statistics.

Et sammendrag av de mest vanlige statistiske egenskapene til hver av forklaringsvariablene:

Husk at små bokstaver betegner variabler på logaritmisk skala.

- Summary statistics, originaldatasettet.

	boligpris	inntekt	RENTE(1- τ)	FORV	ledighet	(inntekt – boligmasse)
Antall observasjoner	57	57	57	56	57	57
Gjennomsnitt	2,193	11,570	0,065	-0,003	-3,281	-2,281
Standardavvik	0,384	0,252	0,019	0,139	0,292	0,193
Minimumsverdi	1,638	11,132	0,028	-0,760	-3,817	-2,648
Maksimumsverdi	2,833	11,977	0,103	0,325	-2,842	-1,991

- Summary statistics, utvidet datasett eksklusiv observasjonene i 2005 og 2006.

	boligpris	inntekt	RENTE(1- τ)	ledighet	boligmasse	FORV
Antall observasjoner	60	64	64	64	64	64
Gjennomsnitt	2,248	11,561	0,062	-3,280	13,902	0,0001
Standardavvik	0,402	0,268	0,020	0,275	0,071	0,040
Minimumsverdi	1,660	11,142	0,027	-3,816	13,805	-0,131
Maksimumsverdi	2,898	11,983	0,098	-2,841	14,051	0,086

- Summary statistics, utvidet datasett alle observasjoner.

	boligpris	inntekt	RENTE(1- τ)	ledighet	boligmasse	FORV
Antall observasjoner	68	72	72	72	72	72
Gjennomsnitt	2,340	11,615	0,058	-3,305	13,922	0,0010
Standardavvik	0,455	0,295	0,022	0,276	0,090	0,043
Minimumsverdi	1,660	11,142	0,025	-3,847	13,805	-0,129
Maksimumsverdi	3,152	12,100	0,098	-2,841	14,117	0,099

Vedlegg 2. Regresjonsresultater.

- Estimeringsresultatene fra 1-stegs feiljusteringsmodellen i kapittel 4.4.1:

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,104	(0,066)
Δ (RENTE(1- τ)) _t	-3,180***	(0,468)
Δ (RENTE(1- τ)) _{t-1}	-1,464***	(0,459)
FORV _t	0,044***	(0,015)
boligpris _{t-1}	-0,110***	(0,024)
(RENTE(1- τ)) _{t-1}	-0,530***	(0,172)
ledighet _t	-0,054***	(0,013)
(inntekt-boligmasse) _{t-1}	0,176***	(0,057)
S1	0,035***	(0,011)
S2	0,020*	(0,011)
S3	0,007	(0,013)
KONSTANTLEDD	0,494**	(0,192)
Antall observasjoner	54	
R2	0,8696	
Justert R2	0,8354	
Durbin-Watson	2,53	
Q*(1)	5,3434	
Q*(4)	9,5046	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

- Balansert modell nevnt i kapittel 5.4.1. der alle variablene inngår både på endringsform og på lagget form. Under vises både regresjonsresultater og elastisiteter (uten 2005 og 2006):

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,041	(0,301)
Δ (RENTE(1- τ)) _t	-3,264***	(0,486)
Δ FORV _t	0,114**	(0,053)
Δ ledighet _t	-0,081	(0,062)
Δ boligmasse _t	1,903	(4,619)
boligpris _{t-1}	-0,204***	(0,059)
inntekt _{t-1}	0,101	(0,070)
(RENTE(1- τ)) _{t-1}	-0,897***	(0,262)
FORV _{t-1}	0,277***	(0,073)
ledighet _{t-1}	-0,074***	(0,026)
boligmasse _{t-1}	0,537	(0,331)
S1	0,051***	(0,017)
S2	0,028**	(0,013)
S3	0,010	(0,017)
KONSTANTLEDD	-8,389**	(4,084)
Antall observasjoner	59	
R2	0,8495	
Justert R2	0,8016	
Q*(1)	0,4176	
Q*(4)	0,6989	
DF-teststatistikk	-8,619***	

Langsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
inntekt _{t-1}	0,493	(0,369)
(RENTE(1- τ)) _{t-1}	-0,271**	(0,108)
FORV _{t-1}	0,0002***	(0,00006)
ledighet _{t-1}	-0,364***	(0,120)
boligmasse _{t-1}	2,631**	(1,189)

Kortsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,041	(0,301)
Δ (RENTE(1- τ)) _t	-0,201***	(0,030)
Δ FORV _t	0,00001**	(0,000006)
Δ ledighet _t	-0,081	(0,062)
Δ boligmasse _t	1,903	(4,619)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

- Modellen nevnt i kapittel 5.6. når trend er inkludert som en forklaringsfaktor, uten observasjonene i 2005 og 2006:

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,218	(0,374)
Δ (RENTE(1- τ)) _t	-3,289***	(0,474)
Δ FORV _t	0,121**	(0,054)
Δ ledighet _t	-0,074	(0,063)
boligpris _{t-1}	-0,205***	(0,053)
inntekt _{t-1}	0,357	(0,376)
(RENTE(1- τ)) _{t-1}	-1,115***	(0,349)
FORV _{t-1}	0,287***	(0,075)
ledighet _{t-1}	-0,053	(0,042)
boligmasse _{t-1}	0,504	(0,328)
S1	0,055***	(0,017)
S2	0,028**	(0,012)
S3	0,008	(0,017)
TREND	-0,003	(0,005)
KONSTANTLEDD	-10,686*	(5,395)
Antall observasjoner	59	
R2	0,8506	
Justert R2	0,8031	
Q*(1)	0,8409	
Q*(4)	1,3291	
DF-teststatistikk	-8,989***	

Langsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
inntekt _{t-1}	1,747	(1,737)
(RENTE(1- τ)) _{t-1}	-0,336***	(0,097)
FORV _{t-1}	0,0002***	(0,00006)
ledighet _{t-1}	-0,258	(0,208)
boligmasse _{t-1}	2,464*	(1,233)

Kortsiktige elastisiteter:

Variabel	Koeffisient	Standardavvik
Δ inntekt _t	0,218	(0,374)
Δ (RENTE(1- τ)) _t	-0,205***	(0,029)
Δ FORV _t	0,00001**	(0,000007)
Δ ledighet _t	-0,074	(0,063)

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

- Estimering av forventningsvariabelen på hele datasettet, jf. kapittel 5.7.

Variabel	Koeffisient	Standardavvik
$\Delta(\text{RENTE}(1-\tau))_t$	-9,433***	(1,556)
$\Delta\text{ledighet}_t$	-0,320**	(0,157)
E_{t-1}	-0,219**	(0,089)
$(\text{RENTE}(1-\tau))_{t-1}$	-1,027	(0,649)
ledighet_{t-1}	0,011	(0,031)
S1	0,172***	(0,042)
S2	0,066***	(0,019)
S3	0,152***	(0,036)
KONSTANTLEDD	0,029	(0,140)
Antall observasjoner	67	
R2	0,6573	
Justert R2	0,6101	
Q*(1)	0.1106	
Q*(4)	4.1381	
DF-teststatistikk	-7.498***	

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

- Estimerer og 95 % konfidensintervall fra tabell 7, jf. kapittel 5.7.

Variabel	Uten 2005 og 2006			Hele datasettet		
	Estimat	95 % konfidensintervall		estimat	95 % konfidensintervall	
$\Delta\text{inntekt}$	0,218	-0,536	0,972	0,049	-0,577	0,675
$\Delta(\text{RENTE}(1-\tau))$	-3,289***	-4,245	-2,333	-3,437***	-4,360	-2,513
ΔFORV	0,121**	0,012	0,231	0,100**	0,007	0,192
$\Delta\text{ledighet}$	-0,074	-0,200	0,053	-0,066	-0,173	0,040
boligpris	-0,205***	-0,311	-0,098	-0,193***	-0,290	-0,095
inntekt	0,357	-0,401	1,115	0,155	-0,368	0,679
$(\text{RENTE}(1-\tau))$	-1,115***	-1,819	-0,411	-0,945***	-1,578	-0,312
FORV	0,287***	0,137	0,438	0,242***	0,101	0,383
ledighet	-0,053	-0,137	0,031	-0,077***	-0,122	-0,031
boligmasse	0,504	-0,158	1,165	0,472**	0,060	0,884

***/ signifikant på 1 % nivå, **/signifikant på 5 % nivå, */ signifikant på 10 % nivå.

Vedlegg 3. Korrelasjonsmatriser.

- Korrelasjonsmatrise for regresjonen i tabell 2. Korrelasjon mellom e , rente og ledighet er uthevet, jf. kapittel 4.3.1. Endringene i rente og endringene i ledighet er relativt høyt korrelert med endringene i forventningene, og rente og forventninger er høyt korrelerte på nivå.

	ΔE_t	$\Delta(\text{RENTE}(1-\tau))_t$	$\Delta\text{ledighet}_t$	E_{t-1}	$(\text{RENTE}(1-\tau))_{t-1}$	Ledighet_{t-1}
ΔE_t	1					
$\Delta(\text{RENTE}(1-\tau))_t$	-0,6835	1				
$\Delta\text{ledighet}_t$	0,4276	-0,1175	1			
E_{t-1}	-0,3921	0,4799	-0,3493	1		
$(\text{RENTE}(1-\tau))_{t-1}$	0,1275	-0,3194	0,0970	-0,7317	1	
Ledighet_{t-1}	0,1426	-0,3914	-0,2737	-0,0178	0,3772	1

- Korrelasjonsmatrisen for Jacobsen og Naug's modell for boligprisveksten, jf. tabell 3:

	$\Delta\text{boligpris}_t$	$\Delta\text{inntekt}_t$	$\Delta(\text{RENTE}(1-\tau))_t$	$\Delta(\text{RENTE}(1-\tau))_{t-1}$	FORV_t
$\Delta\text{boligpris}_t$	1				
$\Delta\text{inntekt}_t$	-0,0192	1			
$\Delta(\text{RENTE}(1-\tau))_t$	-0,5015	-0,0206	1		
$\Delta(\text{RENTE}(1-\tau))_{t-1}$	-0,3251	0,1607	0,3736	1	
FORV_t	0,0900	0,1129	0,1524	0,0285	1
boligpris_{t-1}	0,1307	0,0214	0,0535	0,0266	-0,0893
$(\text{RENTE}(1-\tau))_{t-1}$	-0,3712	-0,0670	-0,1073	0,1775	0,0298
ledighet_t	-0,1557	-0,1846	-0,3023	-0,3061	-0,0502
$(\text{innt-b.masse})_{t-1}$	0,3473	-0,1802	-0,0140	-0,0088	-0,1358
	boligpris_{t-1}	$(\text{RENTE}(1-\tau))_{t-1}$	ledighet_t	$(\text{innt-b.masse})_{t-1}$	
$\Delta\text{boligpris}_t$					
$\Delta\text{inntekt}_t$					
$\Delta(\text{RENTE}(1-\tau))_t$					
$\Delta(\text{RENTE}(1-\tau))_{t-1}$					
FORV_t					
boligpris_{t-1}	1				
$(\text{RENTE}(1-\tau))_{t-1}$	-0,5913	1			
ledighet_t	-0,7255	0,4992	1		
$(\text{innt-b.masse})_{t-1}$	0,9310	-0,6259	-0,6752	1	

Under vises korrelasjonsmatrisen for variablene som inngår i modellene i kapittel 5.

- Tabell 6:

	ΔE_t	$\Delta(\text{RENTE}(1-\tau))_t$	$\Delta \text{ledighet}_t$	E_{t-1}	$(\text{RENTE}(1-\tau))_{t-1}$	Ledighet_{t-1}
ΔE_t	1					
$\Delta(\text{RENTE}(1-\tau))_t$	-0,6138	1				
$\Delta \text{ledighet}_t$	0,3632	-0,0019	1			
E_{t-1}	-0,3072	0,3123	-0,3441	1		
$(\text{RENTE}(1-\tau))_{t-1}$	0,0435	-0,1410	0,1931	-0,7892	1	
Ledighet_{t-1}	0,0874	-0,2856	-0,1763	-0,1994	0,4697	1

- Tabell 7:

	$\Delta \text{boligpris}_t$	$\Delta \text{inntekt}_t$	$\Delta(\text{RENTE}(1-\tau))_t$	ΔFORV_t	$\Delta \text{ledighet}_t$	
$\Delta \text{boligpris}_t$	1					
$\Delta \text{inntekt}_t$	-0.0635	1				
$\Delta(\text{RENTE}(1-\tau))_t$	-0.4577	0.0780	1			
ΔFORV_t	0.0712	-0.0543	-0.1261	1		
$\Delta \text{ledighet}_t$	0.0402	-0.0682	-0.0218	0.0373	1	
boligpris_{t-1}	0.2126	-0.0090	0.1048	0.0389	-0.0646	
inntekt_{t-1}	0.2890	-0.0693	0.0503	0.0391	-0.0601	
$(\text{RENTE}(1-\tau))_{t-1}$	-0.3926	-0.0824	-0.1410	-0.0191	0.2141	
FORV_{t-1}	-0.0276	0.0489	0.1348	-0.7164	-0.0522	
ledighet_{t-1}	-0.2333	-0.0161	-0.2896	-0.0469	-0.1521	
boligmasse_{t-1}	0.2295	-0.0583	0.0372	0.0541	-0.0736	
Korrelasjonsmatrise	boligpris_{t-1}	inntekt_{t-1}	$(\text{RENTE}(1-\tau))_{t-1}$	FORV_{t-1}	ledighet_{t-1}	boligmasse_{t-1}
$\Delta \text{boligpris}_t$						
$\Delta \text{inntekt}_t$						
$\Delta(\text{RENTE}(1-\tau))_t$						
ΔFORV_t						
$\Delta \text{ledighet}_t$						
boligpris_{t-1}	1					
inntekt_{t-1}	0.9834	1				
$(\text{RENTE}(1-\tau))_{t-1}$	-0.7595	-0.7747	1			
FORV_{t-1}	-0.1873	-0.1899	0.0866	1		
ledighet_{t-1}	-0.6712	-0.6664	0.4217	0.1110	1	
boligmasse_{t-1}	0.9716	0.9639	-0.7428	-0.2332	-0.5330	1