

# Norges boligmarked: En makroøkonometrisk prisvurdering

av

Erik Simensen



Mastergradsoppgave i Samfunnsøkonomi

30 Studiepoeng

Handelshøgskolen i Tromsø

Universitetet i Tromsø

Mai 2010

## **I Forord**

Denne oppgaven markerer slutten på fem flotte år med studier ved Handelshøgskolen i Tromsø, så vel som starten på en ny periode som yrkesaktiv. Alle fingre krysses for at neste periode skal preges av like mye spennende teorier, hyggelig samarbeid og mestringsfølelse som studietiden.

Arbeidet med oppgaven har vært spennende, lærerikt og tidvis noe frustrerende. Det er ikke til å legge skjul på at tanker og til og med drømmer i stor grad har dreid seg om spesifisering og estimering av modeller enkelte deler av dette semesteret. Heldigvis for det meste med tilhørende "Eureka"- øyeblikk, gjerne når man minst har ventet det.

I forbindelse med oppgaven rettes naturligvis en stor takk til førsteamanuensis Espen Sirnes for god veiledning og fruktbare diskusjoner rundt oppgaven. Videre vil jeg også takke stipendiat Sverre Braathen Thyholdt for god hjelp med STATA, Håvard Hungnes ved SSB for datasett og nyttige tips til analysen og Christian Lindh for gjennomlesing. Takk også til mine klassekamerater for fantastisk miljø og samarbeid de siste to årene. Mastertiden ville aldri blitt det samme uten dere!

Helt sist vil jeg også takke mine kjære foreldre Gunn og Bjørn. Uten deres oppfølging og støtte ville jeg nok aldri en gang fått oppleve det å være student.

Tromsø, 14. mai 2010.

Erik Simensen

<b>II</b>	<b>Innholdsfortegnelse</b>	
<b>I</b>	<b>Forord</b>	<b>s. 2</b>
<b>II</b>	<b>Innholdsfortegnelse</b>	<b>s. 3</b>
<b>III</b>	<b>Figurliste</b>	<b>s. 5</b>
<b>IV</b>	<b>Tabeller</b>	<b>s. 5</b>
<b>V</b>	<b>Sammendrag</b>	<b>s. 7</b>
<b>1</b>	<b>Innledning</b>	<b>s. 8</b>
<b>2</b>	<b>Norges boligmarked i et historisk perspektiv</b>	<b>s. 10</b>
<b>3</b>	<b>Boligbobler</b>	<b>s. 13</b>
	<b>3.1 Bobledeteksjon</b>	<b>s. 14</b>
	<b>3.2 Fundamentale verdier og spørreundersøkelse</b>	<b>s. 15</b>
	<b>3.3 Price/Earnings, Price/Rent og brukerkostnader</b>	<b>s. 18</b>
<b>4</b>	<b>Stock-flowmodeller</b>	<b>s. 20</b>
	<b>4.1 De første modellene</b>	<b>s. 20</b>
	<b>4.2 Livssyklusmodeller og brukerkostnader</b>	<b>s. 21</b>
	<b>4.3 Prisjustering, forventninger og eiendomspriser</b>	<b>s. 22</b>
	<b>4.4 Vesentlige konklusjoner</b>	<b>s. 24</b>
	<b>4.5 En modell for Norges boligmarked</b>	<b>s. 25</b>
	<b>4.5.1 Langsiktig likevekt</b>	<b>s. 25</b>
	<b>4.5.2 Kortsiktig likevekt</b>	<b>s. 25</b>
	<b>4.6 Vurdering av prisnivå</b>	<b>s. 26</b>
<b>5</b>	<b>Vektor autoregressive modeller - en oversikt</b>	<b>s. 28</b>
	<b>5.1 Forskjellige varianter</b>	<b>s. 28</b>

5.2	Databeskrivelse	s. 30
5.3	Bruksområder	s. 31
6	Statistisk teoridel	s. 33
6.1	Stasjonære og ikke- stasjonære data	s. 33
6.2	Dickey- Fuller testen	s. 36
6.3	Spurious regresjon	s. 38
6.4	Kointegrasjon	s. 38
6.5	Vector Error Correcting Models	s. 40
6.6	Spesifisering	s. 42
6.6.1	Modellens autoregressive orden	s. 42
6.6.2	Kointegrerende rang	s. 44
7	Analyse	s. 45
7.1	Datasett med kommentarer	s. 45
7.2	Dickey- Fuller testene	s. 50
7.3	Spesifisering	s. 51
7.3.1	Vanlige brukerkostnader	s. 51
7.3.2	Alternative brukerkostnader	s. 53
7.4	Estimering	s. 55
8	Resultatbasert vurdering av boligprisnivået	s. 59
9	Avslutning	s. 62
	Litteraturliste	s. 64
	Vedlegg	s. 71

### III Figurliste

<b>Figur 1:</b> Beregnet kvadratmeterpris 1819 – 1939.	s. 10
<b>Figur 2a:</b> Beregnet kvadratmeterpris 1940 – 1969.	s. 12
<b>Figur 2b:</b> Beregnet kvadratmeterpris 1970 – 2010.	s. 12
<b>Figur 3:</b> Realboligpriser og konjunkturer i OECD.	s. 15
<b>Figur 4:</b> Boligprisutvikling og boligprisvekst 1978:4 – 2009:4.	s. 46
<b>Figur 5:</b> Justert konsum 1978:4 – 2009:4, og konsum på endringsform.	s. 46
<b>Figur 6:</b> Boligmassen 1978:1 – 2009:4, og boligmassen på endringsform.	s. 47
<b>Figur 7:</b> Vanlige brukerkostnader 1981:1 – 2009:4, og brukerkostnader på endringsform.	s. 48
<b>Figur 8:</b> Alternative brukerkostnader 1981:1 – 2009:4, og alternative brukerkostnader på endringsform.	s. 49
<b>Figur 9:</b> Bruttoinvesteringer i boliger 1978:1 – 2009:4, og investeringer på endringsform.	s. 49
<b>Figur 10:</b> Byggekostnader 1978:1 – 2009:4, og byggekostnader på endringsform.	s. 50
<b>Figur 11:</b> Langsiktig tilbud og etterspørsel 1985:1 – 2009:4.	s. 59
<b>Figur 12:</b> Utvikling i observerte og fundamentale priser, og $(p_t - p_t^{d^*})$ i perioden 1985:1 – 2009:4.	s. 60

### IV Tabeller

<b>Tabell 1:</b> Regulering av det norske boligmarkedet.	s. 11
<b>Tabell 2:</b> Forholdstall mellom gjennomsnittlige boligpriser og privat inntekt per innbygger samt resultater av regresjon for å forklare boligprisutvikling, 1985:1 – 2002:3.	s. 16
<b>Tabell 3:</b> Kritiske verdier for Dickey- Fuller testene.	s. 37
<b>Tabell 4:</b> Resultatene fra Dickey- Fuller testene.	s. 50
<b>Tabell 5:</b> Seleksjonskriterier for $p_t, i_t, cc_t, ufl_t, c_j$ og $h_t$ .	s. 51
<b>Tabell 6:</b> Johansens Trace Statistic for kointegrasjon i systemet, $HQ(p) = SC(p) = 2$ .	s. 52
<b>Tabell 7:</b> Johansens Trace Statistic for kointegrasjon i systemet, $AIC(p) = 10$ .	s. 52
<b>Tabell 8:</b> Seleksjonskriterier for $p_t, i_t, cc_t, ufa_t, c_j$ og $h_t$ .	s. 53
<b>Tabell 9:</b> Johansens Trace Statistic for kointegrasjon i systemet, $HQ(p) = SC(p) = 2$ .	s. 54
<b>Tabell 10:</b> Johansens Trace Statistic for kointegrasjon i systemet, $AIC(p) = 10$ .	s. 54
<b>Tabell 11:</b> Identifiserende restriksjoner på den kointegrerende matrisen.	s. 55
<b>Tabell 12:</b> Estimerer for de to kointegrerende vektorene, standard brukerkostnader.	s. 56
<b>Tabell 13:</b> Estimerer for de to kointegrerende vektorene, alternative brukerkostnader.	s. 57
<b>Tabell 14:</b> Johansens Trace Statistic for kointegrasjon i systemet, $HQ(p) = SC(p) = 2$ , med hhv. konstant og trend variabel.	s. 71
<b>Tabell 15a, b, c:</b> Estimering av modellen over perioden 1981:1 – 2009:4.	s. 72

**Tabell 16:** Den estimerte modellen for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufl_t$ ,  $cj_t$  og  $h_t$  1985:1 – 2009:4.

s. 74

**Tabell 17:** Den estimerte modellen for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufa_t$ ,  $cj_t$  og  $h_t$  1985:1 – 2009:4.

s. 75

## V Sammendrag

[I]f the reason that the price is high today is *only* because investors believe that the selling price will be high tomorrow – when “fundamental” factors do not seem to justify such a price – then a bubble exists. – Stiglitz (1990)

Med utgangspunkt i bobledefinisjonen til Stiglitz setter denne oppgaven seg fore å vurdere boligprisnivået i Norge over perioden 1985:1 – 2009:4. Ved å estimere den langsiktige delen av en stock-flow boligmarkedsmodell i STATA sammenlignes estimerte priser, som kan forankres i utviklingen i relevante fundamentale verdier, med observerte boligpriser. Eventuelle boligbobler kan således observeres om det foreligger store avvik mellom estimerte og faktiske boligpriser. Stock-flow modellen estimeres som del av en Vector Error Correction Modell.

Den estimerte modellen påviser en tydelig boligboble på slutten av 80- tallet, og impliserer at boligpriser var overvurderte størsteparten av 2000- tallet, riktignok i langt mindre grad enn på slutten av 80- tallet. I følge modellen var boligpriser på et fundamentalt støttet nivå ved utgangen av 2008, mens boligpriser tyder på å ha blitt overvurdert igjen i løpet av 2009.

*Nøkkelord:* boligmarked, bobligboble, stock-flow, vector error correction modell, vector autoregression modell, tidsserieanalyse, VECM, VAR.

# 1 Innledning

Etter mer enn 15 år med voldsom boligprisvekst i de aller fleste vestlige land snudde utviklingen brått og boligprisene falt dramatisk fra høsten 2007 til høsten 2008. I USA med omtrent 20 %, i England nærmere 14 % mens Latvia har rekorden for noe vestlig land med mer enn 38 % prisfall fra første kvartal 2007 til første kvartal 2008. I Norge nådde boligprisene en historisk topp august 2007 før de falt nesten 14 % til desember 2008.

Høsten 2009, omtrent ett år etter at Lehman Brothers gikk over ende, hadde støvet begynte å legge seg etter Subprime- krisen. Mens boligprisveksten var beskjedent i de fleste vestlige land kunne man igjen observere relativt sterk prisvekst i Norge, med priser som faktisk overgikk rekordnivået august 2007. Flere har advart mot den særnorske boligprisutviklingen og hevder at det åpenbart er en boble i det norske boligmarkedet<sup>1</sup>. Denne oppgaven vil derfor ha som målsetning å analysere det norske boligmarkedet for å avgjøre om det er bobletendenser, med følgende problemstilling: *Er norske boligpriser overvurdert?*

Lignende problemstillinger er tidligere belyst i blant annet Jacobsen og Naug (2004) og Grytten (2009), men denne oppgaven vil skille seg fra disse gjennom den analytiske tilnærmingen. Analysen i oppgaven vil forankres i en stock-flow boligmarkedsmodell, som estimeres som en del av en Vector Error Correction Model (VECM). Modellen omfatter data på permanentinntekt, boligbeholdning, brukerkostnader, boliginvesteringer og byggekostnader, og kan således brukes til å estimere et prisnivå som kan forklares ut ifra fundamentale faktorer. Dette prisnivået kan sammenlignes med de faktiske boligprisene for å vurdere boligprisnivået i Norge i perioden 1985:1 – 2009:4.

For å se de siste tiårenes prisutvikling i et mer historisk perspektiv begynner oppgaven med en historisk oversikt over prisutviklingen i det norske boligmarked i kapittel 2. I kapittel 3 defineres begrepet prisboble, og forskjellige tilnærminger til bobleanalyse presenteres. Utviklingen av stock-flow teorien legges frem i kapittel 4, der det også vil redegjøres for modellen som den analytiske delen av oppgaven baseres på. Kapittel 5 gir en kort oversikt over Vektor Autoregressive modeller før nødvendig statistisk teori gjennomgås i kapittel 6. I

---

<sup>1</sup> Se feks. "Gjedrem må stanse boligboblen" av Øystein Sjølie og "Kan ikke overse boligprisene" av Olav Chen på e24.no



kapittel 7 presenteres og testes de inkluderte datasettene før to varianter av modellen estimeres og kommenteres. Den endelige analysen fremstilles i kapittel 8, før kapittel 9 konkluderer oppgaven.

Før den historiske oversikten presenteres vil det kort gjøres rede for enkelte nødvendige begrensninger i oppgaven:

- Det vil ikke gjøres rede for matematikken som ligger bak estimeringen av et systems kointegrerende rang, og estimeringen av kointegrerende vektorer.
- Boligmarkedet vil analyseres over ett slik at det ikke tas hensyn til for eksempel geografi eller forskjellige boligtyper.
- Flere variabler kunne vært inkludert i analysen, som blant annet boligeieres gjeldsgrad, arbeidsledighet eller demografisk utvikling<sup>2</sup>.

---

<sup>2</sup> Med unntak av arbeidsledighet finner faktisk ikke Jacobsen og Naug (2004) noe signifikant bevis på at disse variablene forklarer utviklingen i boligpriser.

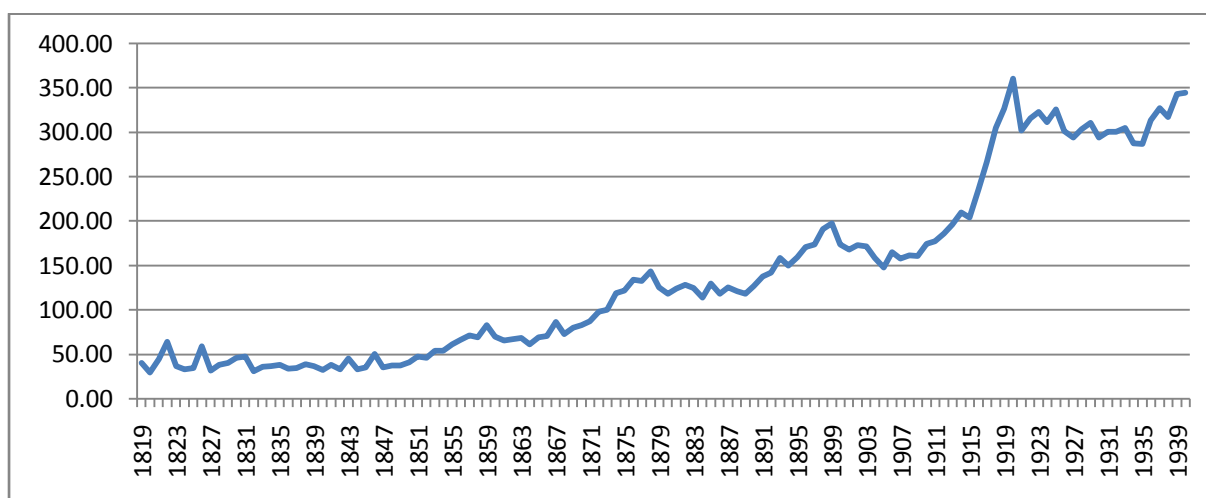
## 2 Norges boligmarked i et historisk perspektiv

Om en ser bort i fra de to verdenskrigene og krisetider har Norges boligpriser stort sett fulgt den generelle prisutviklingen i landet i perioden 1819 – 1954. Deregulering 1954 førte til økt prisvekst frem til 1970- tallet, med enda kraftigere vekst fra 1970- tallet til dags dato.

Utviklingen de siste 40 årene skyldes mindre regulert boligmarkedet, liberalisering av kredittmarkedet i løpet av 1980- årene og solid økonomisk vekst over størsteparten av perioden.

Første nevneverdige boligkrise fant sted i 1899 under det såkalte Kristianiakrakket. De siste ti årene av 1800- tallet var det sterk vekst i hovedstaden. Høye lønninger bidro til omfattende tilflytting, slik at det var enkelt å få seg godt betalt arbeide, men ble vanskelig å finne losji. Kraftig økt etterspørsel etter boliger stimulerte til en overdreven byggeboom, som igjen førte til at det etter noen få år sto omtrent 5000 tomme leiligheter i hovedstaden (tilsvarende omtrent 10 % av alle leiligheter)<sup>3</sup>. I løpet av høykonjunkturen hadde også mange nye banker etablert seg, med relativt liberal utlånspraksis. Det ble for eksempel ganske vanlig å gi boliglån med sikkerhet i aksjer. Hurtig vekst endret seg etter hvert til spekulasjon og da boblen i aksjemarkedet omsider sprakk forplantet det seg nesten umiddelbart til boligmarkedet. Krakket førte til sterk negativ vekst i nominelle boligpriser i Kristiania i flere år, mens boligprisvekst i Bergen, Trondheim og Kristiansand var omtrent null.

**Figur 1:** Beregnet kvadratmeterpris 1819 – 1939.



**Kilde:** Norges Bank, Econ- Pöyry, 2007-priser.

<sup>3</sup> <http://www.ssb.no/emner/00/00/sa39/del8.pdf>

På grunn av ekstrem knapphet på konsumgoder under første verdenskrig var det unormalt høy privat sparing i Norge. Da krigen endte i 1918 og befolkningen kunne begynne å handle for sine oppsparte penger ble resultatet dramatisk økt inflasjon. Dette førte til den neste boligkrisen i Norge på 1920- tallet. Fra 1914 til 1920 steg nominelle boligpriser med 72 %, mens konsumprisindeksen steg med 197 % over samme tidsrom<sup>4</sup>. Den voldsomme inflasjonen var en direkte konsekvens av etterspørselssjokket ved krigens slutt. I 1916 ble det også innført regulering på leiepriser for enkelte boliger, se tabell 1, som også kan tenkes å ha bidratt til redusert prisvekst. I løpet av 1920- årene var det også en nasjonal bankkrise som gjennom redusert kreditt førte til noe fall i boligprisene før de jevnet seg ut i løpet av 1930- tallet.

**Tabell 1:** Regulering av det norske boligmarkedet.

<b>Boligtype</b>	<b>Periode</b>	<b>Form for regulering</b>
Leieboliger	1916 - 1935	Leiekontroll på enkelte leiligheter.
	1940 - 2010	Leiekontroll på enkelte leiligheter.
	1976 - 1983	Forbudt å konvertere leiligheter i borettslag til selveierleiligheter.
Eneboliger	1940 - 1954	Låste priser
	1954 - 1969	Regulerte priser
Borettslag	1940 - 1954	Låste priser
	1954 - 1982/88	Regulerte priser på nye/gamle leiligheter
	1976 - 1983	Forbudt å konvertere leiligheter i borettslag til selveierleiligheter.

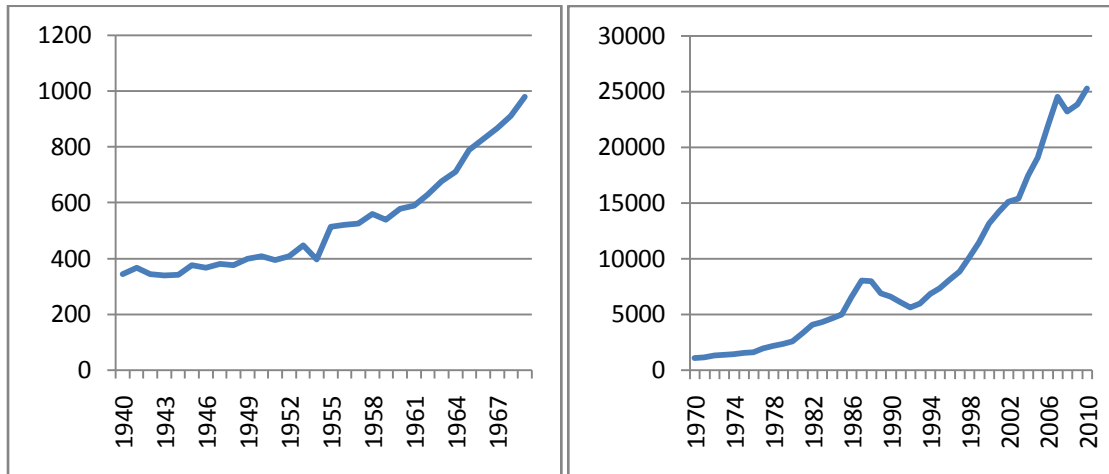
**Kilde:** Eitrheim og Erlandsen (2003)

Det norske boligmarkedet var strengt regulert fra 1940 – 1969. De første 14 årene av denne perioden var boligpriser mer eller mindre låst til nivået de lå på før andre verdenskrig. Dette kan en se i indeksen under der det kun var en ytterst moderat vekst i boligpriser i løpet perioden. Selv om regulering av boligmarkedet fortsatte etter 1954, ser man et tydelig hopp i indeksen fra 1954 til 1955, da denne reguleringsformen opphørte. Ekstraordinær befolkningsvekst i etterkrigstiden kombinert med generell økonomisk optimisme bidro til vekst i boligprisene. Prisveksten skjøt ytterligere fart i begynnelsen av 1980- tallet blant annet grunnet mindre regulerte finansmarkeder verden over. Liberalisering av det norske kredittmarkedet, lav rente samt oljens effekt på både økonomiske forhold og private aktørers forventninger bidro til en kraftig boom, som resulterte i sterkt overopphetet økonomi og til slutt bankkrisen 1987. Boligprisene falt i løpet av den påfølgende lavkonjunkturen frem til

<sup>4</sup> Vekst i konsumprisene er regnet ut med tall fra Grytten (2004).

1992, før de steg dramatisk frem til 2007. Subprime- krisen i USA, og den påfølgende internasjonale finanskrisen, bidro til et midlertidig fall i de norske boligprisene, men begynte igjen å stige i begynnelsen av 2009. Per dags dato er prisene faktisk høyere enn de var i 2007.

**Figur 2a:** Beregnet kvadratmeterpris 1940 – 1969. **Figur 2b:** Beregnet kvadratmeterpris 1970 – 2010.



**Kilde:** Norges Bank, Econ- Pöyry, 2007- priser.

Prisveksten de siste tiårene er uten historisk sidestykke, og har følgelig blitt viet stor oppmerksomhet av økonomer, markedsaktører og media. Flere har fryktet at det har dannet seg en prisboble i boligmarkedet, og det har vært stor fokus på mulige årsaker de siste årene. Det vurderes om det skal endres på skatteforhold for boliger, bankers utlånspraksis og det har til og med vært snakk om at boligpriser skal vektlegges tyngre i Norges Banks rentebeslutning for å bremse boligprisveksten.

I neste kapittel vil det forklares hva som menes med en boble i økonomisk sjargong, og hvorfor bobler vies så mye oppmerksomhet både i media og faglitteraturen.

### 3 Boligbobler

Boblebegrepet er vidt anvendt, men sjelden klart definert. En formell definisjon, som denne oppgaven og flere andre analyser tar utgangspunkt i finnes i Stiglitz (1990):

[I]f the reason that the price is high today is *only* because investors believe that the selling price will be high tomorrow – when “fundamental” factors do not seem to justify such a price – then a bubble exists.

En markedssituasjon der overdreven forventning til fremtidig prisvekst fører til kraftig vekst i markedspris i forhold til andre relevante makroverdier karakteriserer en prisboble. Under en periode med bobledannelse i husmarkedet vil aktører være av den oppfatningen at boliger som ellers ville vært overpriset nå vil være et akseptabelt, eller til og med høyt ettertraktet kjøp på grunn av forventet økning i boligpriser. Siden bolig i de fleste tilfeller utgjør en vesentlig del av en husholdnings formue, oppfattes kraftig boligprisvekst som en form for sparing. For eksempel Skinner (1996) og Case, Quigley og Shiller (2001) viser at privat sparerate ofte faller under en boligboble. Videre frykter ofte førstegangskjøpere å bli priset ut av markedet, og fremskynder gjerne boligkjøp. Forventninger om høy boligprisinflasjon vil også ha kraftig effekt på boliggetterspørsel om den gjennomsnittlige boligkjøper tror det er svært usannsynlig at boligpriser vil falle, siden oppfattet risiko ved boligkjøp vil være veldig lav. Hvis det hovedsakelig er forventninger om sterk vedvarende boligprisvekst som opprettholder boligkjøperes etterspørsel vil huspriser være ustabile. Priser kan ikke stige hurtig til evig tid, og når markedsaktører etter hvert observerer avtakende prisvekst vil etterspørsel reduseres kraftig. Priser får en korreksjon nedover og boblen sprekker.

Bobler vies generelt mye oppmerksomhet siden det ofte følger en finanskriser i kjølvannet av en boblesprekk. Tulipankrakket i Nederland 1637, Mississippiboblen i Frankrike 1719 og børskrakket på Wall Street 1929 er historiske eksempler på dette. I nyere tid er blant annet bankkrisen i Norge sent på 1980- tallet, Asiakrisen og naturligvis Subprimekrisen gode eksempler. Typisk har slike bobler tre forskjellige faser.

Første fase begynner med finansiell liberalisering, en bevisst beslutning fra sentralbanken om å øke utlån, eller lignende hendelser. Den resulterende kredittveksten følges av prisvekst i eiendeler som bolig og aksjer. Priser vil stige over lengre tid, ofte flere år, mens boblen blåses opp. Under den andre fasen sprekker boblen og prisfall fører til kollaps i bolig- og

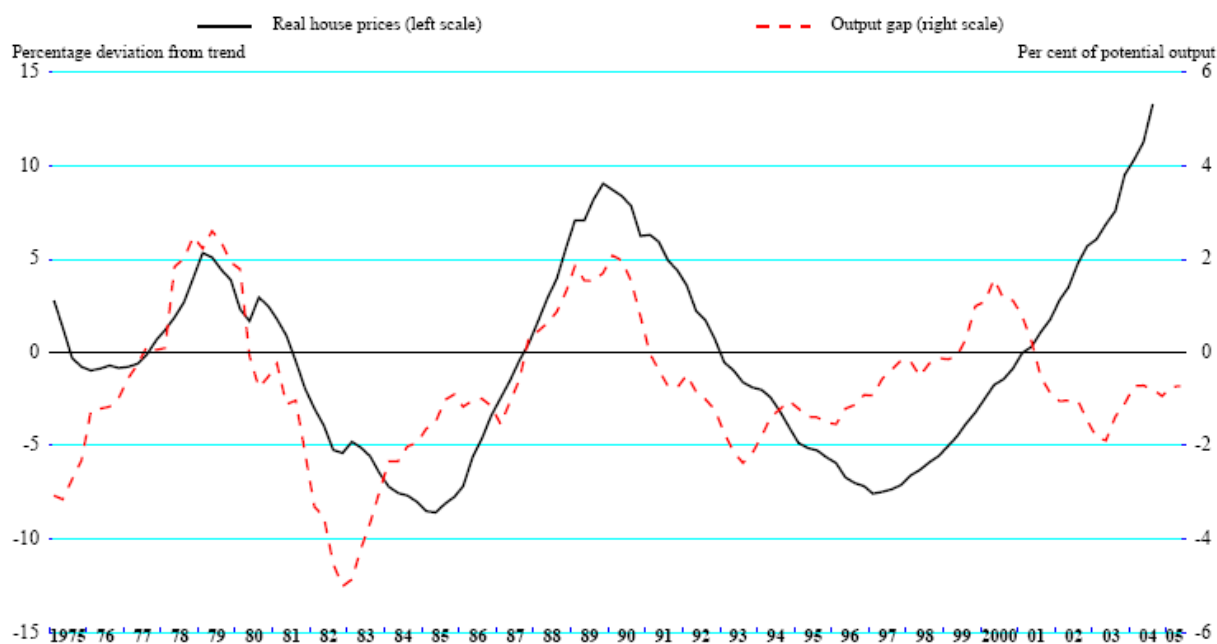
aksjemarkedet. Dette kan skje svært hurtig, som for eksempel etter kollapsen i Lehman Brothers september 2009, eller over lengre tid. I tredje fase klarer ikke lenger aktører å betjene lån som ble tatt opp for å kjøpe høyt prisede eiendeler. Omfattende mislighold av gjeld fører til bankkrise og muligens også valutakrise, avhengig av om sentralbanken prioriterer å forsvare landets valutakurs eller å redusere bankkrisen. Disse problemene smitter over til realsektoren, for eksempel gjennom kredittørke, og kan vare flere år.

De siste tiårene har minst ett aspekt ved boligbobler blitt observert, nemlig en svært kraftig boligprisvekst i Norge. Men høy prisutvikling alene er ikke nok til å hevde det eksisterer en boble i et marked. I stedet er det vanlig å sammenligne prisutvikling med fundamentale faktorer i markedet, som renter, inntektsvekst, byggekostnader og lignende. I neste kapittel presenteres utvalgte tilnærminger til boligmarkedsanalyse.

### **3.1 Bobledeteksjon**

Utgangspunktet for analysene som presenteres i dette kapitlet var den voldsomme stigningen i inflasjonsjusterte boligpriser fra midten av 1990- tallet i de fleste OECD- land. Oppturer i boligmarkeder var for så vidt ikke noe nytt fenomen, men for blant andre Australia, Danmark, Irland, Nederland, Sverige, England, USA og som tidligere nevnt Norge var den samlede veksten i denne perioden langt høyere enn tidligere vekstperioder. Syklusen i boligmarkedene skiller seg også fra tidligere ved hvor langvarig prisoppgangen var. I Australia, Nederland, Norge, Sverige og USA var den faktisk mer enn dobbelt så lang som noen tidligere registrerte oppturer i følge Girouard et. al (2006). Om man sammenligner en indeks for aggregerte realboligpriser med produksjonsgap for hele OECD finner man en nær sammenheng mellom boligprisvekst og konjunkturer i perioden 1970 – 2000. Avviket kom derimot i begynnelsen av 2000- tallet. På tross av en lavkonjunktur, utløst av en bristende IT – boble, fortsatte boligpriser å vokse. Mange argumenterte for at den konjunkturtrossende verdiøkningen i boligmarkedene var et tydelig symptom på bobledannelse. Kanskje spesielt Robert J Shiller og Karl E Case, som står bak den første tilnærmingen til å påvise bobler i boligmarked.

**Figur 3:** Realboligpriser og konjunkturer i OECD.



**Kilde:** Girouard et. al. (2006)

### 3.2 Fundamentale verdier og spørreundersøkelse

I følge Case og Shiller (2003) er det flere kjennetegn ved et overopphetet boligmarked:

- Utbredte forventninger om høy prisvekst.
- Stor oppmerksomhet mot boligprisutvikling i media og private sammenhenger.
- Press på at man bør bli boligeier.
- Boligpriser vokser hurtigere enn lønninger.
- Forenklete oppfatninger av økonomiske sammenhenger i boligmarkedet dominerer.
- Svak risikoforståelse.

For å avgjøre hvilke av disse punktene som kunne observeres i det amerikanske boligmarkedet tidlig på 2000- tallet utviklet Case og Shiller en todelt tilnærming til boligmarkedsanalyse. Shiller, med bakgrunn fra behavioral finance, var mest opptatt av aktørers forventninger og holdninger til markedet, mens Case, med erfaring fra boligmarkedsanalyse, var mer interessert i prisutvikling i forhold til fundamentale verdier.

Analysen deres består derfor av en økonometrisk del, og en spørreundersøkelse blant huskjøpere for å få innsyn i hva som drev amerikansk boligprisutvikling i 2003.

Sentralt i Case og Shillers økonometriske analyse står forholdet mellom pris og inntekt. Det er rimelig greit å anta at en situasjon der priser vokser hurtigere enn inntekt over lang tid er ustabil, slik at det relative prisnivået ikke kan opprettholdes. Dermed kan et slikt forholdstall enkelt gi et innblikk i situasjonen i et boligmarked. I tabellen under kan en se forskjellen mellom de syv statene med mest stabile boligpriser, og de åtte med mest volatile priser.

**Tabell 2:** Forholdstall mellom gjennomsnittlige boligpriser og privat inntekt per innbygger samt resultater av regresjon for å forklare boligprisutvikling, 1985:1 – 2002:3.

Stat	Ratio				R <sup>2</sup> til regresjon av boligpris på:	
	Bunnpunkt	Toppunkt	Standardavvik	Kvartal for toppunkt	Inntekt per innb.	Andre fundamentale verdier *
<i>Stater med mest volatile priser</i>						
Hawaii	7.8	12.5	1.34	1992:3	0.83	0.89
Connecticut	4.5	7.8	1.06	1988:1	0.45	0.69
New Hampshire	4.0	6.6	0.84	1987:2	0.49	0.78
California	6.0	8.6	0.80	1989:4	0.78	0.89
Rhode Island	4.6	7.1	0.75	1988:1	0.65	0.79
Massachusetts	4.3	6.6	0.72	1987:3	0.70	0.88
New Jersey	4.5	6.8	0.68	1987:3	0.73	0.90
New York	3.8	5.6	0.52	1987:3	0.77	0.86
<i>Stater med minst volatile priser</i>						
Nebraska	1.8	2.1	0.09	1985:2	0.96	0.99
Wisconsin	2.1	2.4	0.08	2002:3	0.99	0.99
Illinois	2.6	2.9	0.08	2002:3	0.98	0.99
Kentucky	2.1	2.4	0.08	1985:1	0.99	0.99
Indiana	2.0	2.3	0.06	1986:4	0.99	0.99
Iowa	1.7	1.9	0.06	2002:3	0.98	0.99
Ohio	2.3	2.5	0.04	2002:3	0.99	0.99

**Kilde:** Case og Shiller (2003).

\*: Befolkning, sysselsetting, ledighetsrate, nybygging og boliglånrente utgjør andre fundamentale verdier i de utvidede regresjonene.

Et påfallende mønster fra tabell 2 er at pris/inntekt- ratio er lavere og mer stabilt over hele perioden for de minst volatile statene. En enkel regresjon med pris på inntekt har også god forklaringssevne, som kan sees fra R<sup>2</sup>- verdiene, for disse statene. For de mest ustabile statene er det tydelig at inntekt ikke er en like god forklaringsfaktor, mens andre fundamentale



verdier øker regresjonens forklaringssevne en god del. Faktisk viser Case og Shiller (2003) at inntekt er den viktigste prisdriveren i *alle* stater unntatt de åtte volatile i tabell 2.

Med de utvidede regresjonene over som utgangspunkt estimeres priser for første kvartal 2000 – tredje kvartal 2002. Disse estimatene, basert på fundamentale verdier, kan sammenlignes med observerte priser for å drøfte muligheten for en boble i markedet, i henhold til Stiglitz' bobledefinisjon. Det viser seg at prisene i syv av de åtte volatile statene, ikke Hawaii, ligger over de estimerte prisene. Derfor argumenteres det for at boligprisene i disse statene er høyere enn det de fundamentale verdiene tilsier, som igjen impliserer bobletendenser.

For å utfylle analysen ble det sendt ut en spørreundersøkelse i mars og august 2002 til omtrent 2000 tilfeldige husstander som nylig hadde kjøpt eller solgt bolig, i fire forskjellige byer. Undersøkelsen ble sendt ut i tre byer med voldsom prisvekst, Boston, Los Angeles og San Francisco, samt en by med stabil prisutvikling, Milwaukee, som kontroll. Omtrent 700 spørreskjemaer ble returnert innen sommeren 2003. Spørsmålene var de samme som i en undersøkelse de hadde utført i byene 1988, slik at de kunne sammenligne med en tidligere situasjon da det høyst sannsynlig var bobler i de tre førstnevnte boligmarkedene.

Ved hjelp av en database over aviser og andre tidsskrifter, LexisNexis, viser forfatterne at de var svært heldige med timingen på spørreundersøkelsen. Medias oppmerksomhet mot boligbobler, og det amerikanske boligmarkedet generelt, var på et historisk høyt nivå i oktober 2002. På den måten fikk de analysert aktørenes forventninger, forståelse av boligmarkedet og atferd, i en periode der boligpriser var mye omtalt. Kort sagt mener Case og Shiller at de kan påvise alle de øvrige nevnte punktene ved hjelp av sin todelte analyse for de åtte mest volatile statene, og mener derfor det var tydelige tegn på en boligmarkedsboble i 2003 i enkelte deler av USA.

Selv om Case og Shillers tilnærming omfatter mange viktige variabler, og gir et godt bilde av forventninger i et hett boligmarked, har også analysen deres vært mål for mye kritikk. I hovedsak den økonometriske delen.

I tillegg til utfordringer knyttet til tidsseriedata, som vil drøftes i større detalj senere i oppgaven, er det flere tegn på endogenitetsproblemer i regresjonene. Eksempelvis kan en tenke seg at høy sysselsetting bidrar til økte boligpriser gjennom etterspørselskanaler. Samtidig er det vist i blant annet Case (1986) at det er vanskelig å tiltrekke arbeidskraft til en region med høye boligpriser. Koeffisientene i flere av regresjonene varierer også både i verdi

og fortegn på tvers av statene i analysen. Kanskje mest bemerkelsesverdig er det at boliglånsrente bare får signifikant koeffisient i *en* av de mange regresjonene. En kritiker, Christopher Mayer, mente tilnærmingen led av for mange svakheter. Han slo seg sammen med to andre økonomer, Charles Himmelberg og Todd Sinai, for å finne en forklaring på stadig stigende boligpriser med utgangspunkt i finansteori og Price/Earnings- modellen.

### 3.3 Price/Earnings, Price/Rent og brukerkostnader

Innen finans kan en evaluere nivået på aksjepriser med en Price/Earnings (P/E)- modell. I modellen regner en ut fundamentale P/E- tall ved å dele pris per aksje på estimert fremtidig inntjening de kommende 12 månedene. Man kan også regne ut faktiske P/E- tall ved å dele aksjepris på inntjening de siste 12 månedene. Aksjeprisen kan da vurderes ved å sammenligne fundamental og faktisk P/E- verdi. Med denne metoden som utgangspunkt gjennomfører Himmelberg, Mayer og Sinai (2005) en boligmarkedsanalyse der de ser på forholdet mellom pris og leiekostnader, og knytter dette forholdet opp mot den årlige kostnaden ved å eie en bolig, også kalt boligens brukerkostnader.

I følge Poterba (1984) og Hendershott og Slemrod (1983) består den årlige kostnaden ved å eie en bolig, brukerkostnaden, av seks komponenter som representerer både fordeler og ulemper ved å eie bolig:

$$(3.1) \quad \text{Årlige eierkostnader} = P_t[r_t^{rf} + \tau_t^p - \tau_t^y(r_t^m + \tau_t^p) + \delta_t - g_{t+1} + \gamma_t].$$

Boligpris  $P_t$  ganger risikofri rente  $r_t^{rf}$  tilsvarer alternativkostnaden til boliginvesteringen og er den første komponenten i årlige eierkostnader. Den andre komponenten er utgift til eiendomsskatt pr år, boligpris ganger skattesatsen for eiendom  $P_t\tau_t^p$ . Tredje komponent er en fordel som består av skattefradraget på eiendomsskatt og boliglånsrente. Dette kan estimeres som produktet av boligpris, den effektive inntektsskattesatsen  $\tau_t^y$  og estimert betaling av rente på boliglån og eiendomsskatt:  $P_t\tau_t^y(r_t^m + \tau_t^p)$ .  $\delta_t$  representerer vedlikeholdskostnader som andel av boligens verdi,  $g_{t+1}$  er forventet kapitalgevinst og  $\gamma_t$  er et risikotillegg som kompenserer boligeiere for risikoen de utsetter seg for ved å eie i stedet for å leie.

Siden husleie bør dekke kostnadene og den eventuelle inntjeningen til en bolig impliserer likevekt i boligmarkedet at årlige forventede kostnader ved å eie en bolig ikke kan overskride årlige leiekostnader. Hvis eierkostnader stiger uten at husleie stiger, må boligpriser falle for at

potensielle huskjøpere skal kjøpe i stedet for å leie. Det motsatte skjer om eierkostnader faller. Denne korrigeringsprosessen fungerer som en betingelse om ingen arbitrasje og viser at årlige leieutgifter må være lik årlige eierkostnader:

$$(3.2) \quad R_t = P_t[r_t^{rf} + \tau_t^p - \tau_t^y(r_t^m + \tau_t^p) + \delta_t - g_{t+1} + \gamma_t] = P_t u_t,$$

der  $R_t$  er årlige leiekostnader og  $u_t = [r_t^{rf} + \tau_t^p - \tau_t^y(r_t^m + \tau_t^p) + \delta_t - g_{t+1} + \gamma_t]$ . Ved å omskrive (3.2) til

$$(3.3) \quad P_t/R_t = u_t^{-1},$$

kommer en frem til ligningen som anvendes til boligmarkedsanalyse i Himmelberg, Mayer og Sinai (2005). Ved hjelp av (3.3) kan forholdet mellom boligpriser og husleie knyttes opp mot tidligere nevnte fundamentale faktorer slik at boligprisenivået kan vurderes. På tross av høye priser konkluderer Himmelberg, Mayer og Sinai (2005) med at var det lite som tydet på bobler i det amerikanske boligmarkedet i 2004. De viser at rentenivået i stor grad var årsaken til boligprisenivået.

Selv om de presenterte tilnærmingene er svært interessante, og ofte anvendes<sup>5</sup> vil denne oppgaven ta en makroøkonometrisk vinklet tilnærming. Så vidt undertegnede vet er det ikke tidligere utført en lignende analyse på norske data, slik at dette blir en ny vinkling på en gammel problemstilling. Før modellen drøftes og estimeres er det imidlertid nødvendig med en gjennomgang av noe teori. Med det følgende vil det først redegjøres for makroøkonomisk modellering av boligmarkedet, før det redegjøres for relevante økonometriske emner.

---

<sup>5</sup> Se for eksempel Grytten (2009) eller Norges Bank (2006)

## 4 Stock-flowmodeller

Aggregerte boligmarkedsmodeller har sitt utspring i de makroøkonomiske multisektormodellene som vokste frem i løpet av 1960- tallet. Hovedformålet med de første modellene var opprinnelig å estimere investeringsnivået på nye boliger. Etter hvert vred forskernes interesse seg mot selveierdelen av boligmarkedet, og det var rollen til prisendringer, finansielle institusjoner og kredittmarkeder som sto i fokus. Mer komplette modeller ble utviklet til dette formålet, som inkluderte både tilbud og etterspørsel. Det var nesten uten unntak deler, eller variasjoner av stock-flowmodeller for boligsektoren som ble estimert i denne forskningen.

### 4.1 De første modellene

I de tidligste modellene er etterspørsel avhengig av eksogene variabler  $X_1$ , realprisen på boliger  $P$ , årlig kostnad ved å finansiere boligkjøp  $U$ , og alternativkostnaden ved å leie  $R$ . De eksogene variablene inkluderer blant annet demografiske forhold og permanentinntekt. Frem til begynnelsen av 1980-tallet var bare boliglånsrente med i kostnadsleddet nevnt over. Skattefradrag og forventet avkastning fra boligappresiering kom med senere.

I de første stock-flowmodellene er det en vanlig antakelse at boligmarkedet klareres raskt, og at prisene til enhver tid tilpasser seg for å justere etterspørsel i forhold til eksisterende boligbeholdning (4.1). Dermed blir likevektsprisene fastslått i hver periode som en funksjon av boligbeholdning, etterspørselsfaktorer og rente på boliglån.

Tilbudet av boliger følger derimot en differensialligning, (4.2), der beholdningen sakte slites bort, med raten  $\delta$ , men også vokser med nybygging  $C$ . Nybygging er antatt å være avhengig av boligprisnivået  $P$ , og eksogene variabler  $X_2$ , der faktorpriser og forskjellige renter inkluderes i  $X_2$ .

$$(4.1) \quad D(X_1, P, U, R) = S$$

$$(4.2) \quad \Delta S = C(X_2, P) - \delta S$$

Om  $S$  måles i boligenheter gjenspeiler (4.1) beslutninger om formasjon av husholdninger og vurdering av om husholdningene skal leie eller eie bolig. Måles derimot  $S$  i penger omfattes også konsum av boligjenester i etterspørselsfunksjonen. Tidlige estimater av tilbudsfunksjonen (4.2) brukte data på både nybygging og verdien av nye investeringer i boligmarkedet. De fleste estimatene av (4.1) har derimot tatt utgangspunkt i boligenheter. Det har lenge eksistert ganske gode estimater for antall boliger, mens eiendomspriser ekskluderes i verdien av et lands boligbeholdning. Med andre ord gjøres det kun rede for kapital i disse modellene.

## 4.2 Livssyklusmodeller og brukerkostnader

Tidlig på 1980- tallet kom et viktig bidrag til boligetterspørselsteori gjennom en mer presis definisjon av brukerkostnaden assosiert ved å bo i en bolig. Livssyklusmodeller ble anvendt for å beskrive konsum av boligjenester, og flere forfattere, for eksempel Poterba (1984), hevdet at et korrekt mål på disse brukerkostnadene måtte være justert for skatt og inflasjon. Først ble  $U$ - leddet utvidet til å omfatte kostnaden av gjeld etter skatt, og eiendomsskatt  $(i + \tau^p)(1 - \tau^y)$ . Her er  $i$  nominell rente, og  $\tau^p$  og  $\tau^y$  er marginal skatt på henholdsvis eiendom og inntekt. Inflasjonsjusteringen inkluderte forventet fremtidig boligappresiering i nominelle termer  $E(\Delta P/P)$ . I første omgang ble det en stor diskusjon rundt betingelsene fra finansmarkedet som måtte oppfylles for at konsumenter skulle anse forventet fremtidig avkastning fra bolig som lik kostnaden ved å betjene et boliglån. Nå anerkjenner derimot moderne etterspørselsteori at boligens brukerkostnader må baseres på en ligning som (4.3). Denne blir eksogen i modellen på grunn av prisforventningsleddet.

$$(4.3) \quad U = (i + \tau^p)(1 - \tau^y) - E(\Delta P/P).$$

Første økonom til å inkludere inflasjons og skattejusterte brukerkostnader var Kearl (1979), som inkluderte en ligning lik (4.3) i (4.1). Dermed kunne han løse for boligpriser som en funksjon av etterspørselsfaktorer, boligbeholdning, boliglånsrente, og forventet fremtidig boligpris. Poterba (1984) estimerte en lignende modell, men med fokus på boliginvesteringer. Brukerkostnaden slik den er formulert i Poterba (1984) og (1992) er en fremstilling som benyttes av svært mange i nyere tid. Se for eksempel McCarthy og Peach (2002) og (2004), Himmelberg, Mayer og Sinai (2005) eller Girouard et. al. (2006).

Stock-flowmodeller kan også utvides til å inkludere modellering av faktor og finansmarkeder knyttet til boligmarkedet. I følge for eksempel Smith (1969) kan rente på boliglån gjøres endogen, og fastslås av boligetterspørsel og andre renter i økonomien. Smith foreslår også at byggekostnader kan gjøres endogene og modellerer dem som avhengige av nivået på nybygging og byggekostnader i andre områder av økonomien. Topel og Rosen (1988) utvider igjen dette ved å anta at byggebransjen vet tilbudskurven for nybygging er stigende. Dermed vil de fase inn nybygging på en måte som minimerer intertemporale byggekostnader.

### 4.3 Prisjustering, forventninger og eiendomspriser

Sist vil også noen viktige nyvinninger utviklet av DiPasquale og Wheaton (1994) nevnes. Disse har med tiden blitt vanlige i stock-flowmodeller, og stemmer godt overens med mikroøkonomisk teori:

#### *Gradvis prisjustering.*

Stock-flowmodeller ble lenge estimert med en forutsetning om at tilbud tilpasses over tid, men det er sterke bevis for at også etterspørsel justeres trinnvis. Faktumet at boligpriser ikke er nok til å estimere nybygging, at boligmarkedets ledighetsgrad bidrar til gradvis prisjustering og observasjoner om positiv autokorrelasjon i boligpriser kan alle forklares om det er slik at priser bruker tid på å justere seg etter eksogene sjokk. Det er også flere gode forklaringer på hvorfor markedet bruker lang tid på å klareres. Boliger er langt fra homogene goder, det tar lang tid å finne frem til en bolig en vil kjøpe eller det kan være stor variasjon i formidlingstid. Det er også vanskelig for en enkelt selger å fastslå om vansker med salg kan tilskrives en generell nedgang i boligmarkedet, eller bare uflaks. I et slikt marked det er lite rasjonelt å forvente hurtig prisjustering.

DiPasquale og Wheaton (1994) inkluderer gradvis markedsklarering ved å ta med en enkel prisjusteringsmekanisme. Anta at (4.1) gir en hypotetisk likevektspris  $P^*$ , og faktiske priser konvergerer mot denne likevektsprisen med årlig prosentvis rate  $\lambda$ . Da kan prisjusteringsmekanismen skrives som i (4.4).

$$(4.4) \quad \Delta P = \lambda (P^* - P).$$

### *Forventninger.*

Sheffrin (1983) og Poterba (1984) tok i bruk makroøkonomisk teori om rasjonelle forventninger ved å anta at markedsprognoser anvender all tilgjengelig informasjon om et markeds virkemåte. Faktisk er jo forventede fremtidige priser løsningen av en modell der fremtidige eksogene variabler er usikre. Selv om markeder med rasjonelle forventninger antas å være effektive kan priser fortsatt vise tegn til systematiske mønstre og autokorrelasjon. Om det er systematiske bevegelser i de eksogene variablene til et marked, eller om markedet bare tillater gradvis justering av beholdning, vil selv rasjonelle prisprognoser være høyt korrelerte over tid. Eksempelvis i Poterba (1984) blir resultatet av et positivt etterspørselssjokk både en forventet og umiddelbar prisøkning, og over tid en gradvis autokorrelert prisreduksjon ettersom boliginvesteringer tiltar. Et lignende eksempel kan finnes i Mankiw og Weil (1989), men da som resultat av en babyboom.

Nok en kilde til intertemporal prisavhengighet kan introduseres om forventninger er bakoverskuende. Altså at historisk prisutvikling danner grunnlaget for aktørenes prisprognoser. Både Mankiw og Weil (1989) og Case og Shiller (1989) finner signifikante bevis for at slik ekstrapolering er en vanlig, om enn irrasjonell, form for atferd blant husholdninger. Nåværende priser vil altså være positivt autokorrelert med tidligere priser gitt bakoverskuende forventninger.

### *Boligpriser, nybygging og eiendomsmarkedet.*

I de fleste boligmarkedsmodeller antas aktører i boligindustrien å være små pristakere som står ovenfor stigende faktorkostnadskurver. Økende kostnader skyldes på kort sikt flaskehals, og på lang sikt fra stigende tilbudskurver for både arbeidskraft og materiale. En langsiktig prisøkning vil dermed resultere i en permanent økning i nybygging. Hvor stor denne økningen faktisk blir avhenger av hvor mye faktorkostnadene stiger når nybygging tiltar.

Problemet med denne forklaringen er at en av de viktigste innsatsfaktorene for boligtilbud er eiendommer, og eiendomspriser er avhengig av boligbeholdningen, ikke aktivitetsnivået i byggebransjen eller hvor mange nye boliger som settes opp. Med en faktortilbudsfunksjon av den sorten vil innvirkningen av en økning i boligpriser bli ganske annerledes enn den beskrevet over. Stigende priser vil i utgangspunktet generere meravkastning i byggebransjen, men vil kun bidra til midlertidig økt nybygging. Ettersom boligbeholdningen vokser vil eiendomspriser stige, og etter hvert oppsluke all meravkastningen. Dermed faller boligoppføring tilbake til normalnivået.

#### **4.4 Vesentlige konklusjoner**

Stock-flowmodeller har nå god fartstid innen faget. Til tross for evolusjonen som har skjedd i modellene i løpet av denne tiden er de empiriske resultatene fra modellene bemerkelsesverdig konsistente i sine konklusjoner om boligmarkedet:

*Boligmarkedet følger til en viss grad en forutsigbar syklus med positiv autokorrelasjon på pris.* Flere studier, blant andre Case og Shiller (1989), viser at markedet for eierokkuperte boliger er langt fra effektivt. Dette viser de ved å påpeke at priser slett ikke følger noe random-walk, men heller følger forutsigbare mønstre. Det observeres at ligninger for likevektspriser i tidligere stock-flowmodeller som oftest har signifikant autokorrelasjon, som igjen tilsier et ineffektivt marked.

*Boligmarkedet kan vise signifikante avvik fra likevekt slik at priser ikke er en tilstrekkelig beskrivende statistikk.* Blant andre har Fair (1972) og Rosen og Smith (1983) bevist at bolig og leiepriser justeres gradvis, og at ledighetsgrad fungerer som en slags justeringsmekanisme. På tilbudssiden i markedet inneholder estimerte modeller for nybygging en rekke variabler i tillegg til boligpriser og byggekostnader. Ledighetsrate, boliglånsrente, formidlingstid, og generelle makroøkonomiske faktorer ser alle ut til å bidra betraktelig til å forutse byggeaktivitet. Om markedet ble hurtig klarert og dermed kom i likevekt, ville alle disse faktorene vært fullt omfattet i boligpriser, som da ville vært eneste determinant for tilbudssiden.

*Sammenhengen mellom boligbygging og faktormarkedene er i stor grad uklar.* De fleste ligninger som estimerer boliginvesteringer inneholder en ratio mellom boligpris og byggekostnader som en uavhengig variabel. For det meste er disse variablene signifikante,



men bare for relativt lave konfidensnivåer. Poterba (1984), og senere Topel og Rosen (1988) er blant de første som brukte en ren kostnadsindeks som uavhengig variabel. Ingen av disse fikk noen signifikant koeffisient på byggekostnader. Sammenhengen mellom eiendomsmarkedet og nybygging er i enda større grad uklart siden det finnes svært lite data på eiendomspriser.

#### 4.5 En modell for Norges boligmarked

I denne oppgaven anvendes en variant av stock-flowmodeller basert på modellen i McCarthy og Peach (2002). Opprinnelig ble modellen anvendt for å estimere transmisjonsmekanismer fra pengepolitikk til boliginvesteringer, men er også velegnet til vurdering av boligprisnivå<sup>6</sup>.

##### 4.5.1 Langsiktig likevekt

Den langsiktige etterspørselsfunksjonen vil, for en gitt boligbeholdning  $h_t$ , avgjøre prisen  $p_t^{d*}$  som klarer nåværende boligbeholdning. Nivået på funksjonen avhenger av husholdningenes permanentinntekt, konsum av ikke- varige goder og tjenester  $c_t$  fungerer som proxy, og brukerkostnaden for boligkapital  $u_t$ . På logaritmisk form som kan denne sammenhengen uttrykkes som:

$$(4.5) \quad p_t^{d*} = \alpha_1 c_t + \alpha_2 h_t + \alpha_3 u_t.$$

Siden (4.5) er en etterspørselsfunksjon antas koeffisientene å være negative for boligbeholdning og brukerkostnad, og positiv for konsum.

Med forutsetningen om at bedrifter kostnadsfritt kan etablere seg eller forlate markedet vil langsiktig profitt i byggebransjen være lik null. Gitt byggebransjens kostnadsstruktur vil boligprisen  $p_t^{s*}$  da sørge for en investeringsrate  $(I/H)$ , som er høy nok til å dekke kapitalslit  $\delta$  og nybygging  $\Delta h_t$ ,  $(I/H_t) = \Delta h_t + \delta$ . Dette forholdet kan skrives:

$$(4.6) \quad p_t^{s*} = \gamma_1 (i_t - h_t) + \gamma_2 c c_t,$$

der  $i_t$  er logaritmen av boliginvesteringer, slik at  $(i_t - h_t)$  er logaritmisk investeringsrate, og  $c c_t$  er logaritmen av en byggekostnadsindeks. Begge koeffisientene antas å være positive siden (4.6) er en tilbudsfunksjon.

---

<sup>6</sup> Se for eksempel McCarthy og Peach (2004).

#### 4.5.2 Kortsiktig likevekt

Ved å inkludere en feiljusteringsprosess i tilbuds- og etterspørselsdelen av modellen tas det hensyn til at boligmarkedet justerer seg sakte mot langsiktig likevekt. Om modellen er i langsiktig likevekt og utsettes for et sjokk vil det oppstå en kile mellom faktisk boligpris  $p_t$ , og de langsiktige tilbuds- og etterspørselsprisene  $p_t^{s*}$  og  $p_t^{d*}$ . Om ikke modellen utsettes for nye sjokk, vil avvikene i pris gradvis avta, og forsvinne.

På etterspørselssiden impliserer dette at kortsiktig boligprisvekst vil avta hvis det eksisterer en positiv kile mellom faktiske boligpriser og  $p_t^{d*}$ . I tillegg påvirkes kortsiktig boligprisvekst av endring i permanentinntekt, brukerkostnad, husholdningenes formue  $\Delta w_t$  og husleie  $\Delta p_t^r$ , slik at kortsiktig etterspørsel kan skrives:

$$(4.7) \quad \Delta p_t = \lambda_d(p_{t-1} - p_{t-1}^{d*}) + \beta_0 + \beta_1 \Delta c_t + \beta_2 \Delta u_t + \beta_3 \Delta w_t + \beta_4 \Delta p_t^r + \varepsilon_t,$$

der  $\varepsilon_t$  er feilleddet, og  $\Delta$  betyr at en variabel er på endringsform. Koeffisientene på priskilen og endring i brukerkostnader ventes å være negative, resten positive.

Med utgangspunkt i antakelsene for tilbudssiden av modellen er det rimelig å anta at en positiv priskile mellom faktiske priser og  $p_t^{s*}$  indikerer et godt miljø for byggebransjen og vil stimulere til økte investeringer. Mayer og Sommerville (2000) viser at voksende boligpriser har en positiv effekt på investeringer, mens økte byggekostnader etter all forventning har en negativ effekt. Dermed kan en skrive en kortsiktig tilbudsfunksjon som:

$$(4.8) \quad \Delta(I/H)_t = \lambda_s(p_{t-1} - p_{t-1}^{s*}) + \theta_0 + \theta_1 \Delta p_t + \theta_2 \Delta cc_t + \theta_3 r_t + v_t,$$

der  $r_t$  er realrente og  $v_t$  feilleddet. Priskilen og prisvekst ventes å ha positive koeffisienter, mens byggekostnader og realrente ventes å være negative. En kan også inkludere eiendomspriser og antall boliger til salgs i (4.8). Vekst i disse vil redusere investeringer.

#### 4.6 Vurdering av prisnivå

I modellen bestemmes likevektspris i markedet for en gitt boligbeholdning og vil avgjøre nybygging. Om nybygging er større enn kapitalslit, rivning og lignende vokser

boligbeholdning over tid. I langsiktig likevekt bygges det like mange boliger som det ødelegges, slik at beholdningen er konstant. Den kortsiktige dynamikken i modellen tilsier at markedet justerer seg sakte mot den langsiktige likevekten. Nøkkelvariablene for å avgjøre etterspørsel er permanentinntekt og boligkapitalens brukerkostnad, mens boligpris og byggekostnader bestemmer tilbudet.

Som tidligere nevnt er  $(p_t - p_t^{d*})$  i (4.7) et mål på priskilen mellom faktiske priser, og priser som er implisert av de langsiktige etterspørselsfaktorene. Derfor kan denne kilen benyttes til å vurdere boligprisnivået opp mot de fundamentale faktorene i markedet. Et mulig avvik mellom de estimerte og observerte prisene vil være av stor interesse i denne oppgavens sammenheng, siden en positiv kile vil indikere høyere boligpriser enn det nivået fundamentale verdier støtter. Dermed tyder en positiv kile på en mulig prisboble. For å komme frem til  $p_t^{d*}$  kunne en tenke seg at det var tilstrekkelig med en enkel regresjon av (4.5). I virkeligheten er det derimot noe mer komplisert enn dette siden alle variablene i (4.5) og (4.6) er tidsserier. Resten av oppgaven vil nå i all hovedsak dreie seg om estimeringen av (4.5) og (4.6), og utfordringer knyttet til dette. Før det redegjøres for nødvendig teori om tidsserieanalyse, vil det kort greies ut om bakgrunnen til den metodiske verktøykassen som skal brukes.

## 5 Vektor autoregressive modeller - en oversikt

Hva skjer med inflasjonsraten i løpet av et år om Norges Bank setter opp styringsrenten med 50 basispunkter? Hvor godt predikerer Phillipskurven inflasjon? Hvor mye må styringsrenten reduseres for å motvirke en økt forventet arbeidsledighet på en halv prosent? Alle disse spørsmålene er typisk eksempler på makroøkonomiske problemstillinger analytikere gjerne skal kunne svare på. Helst innenfor et smalest mulig konfidensintervall. På 1970- tallet var det flere metoder som ble anvendt for å se på problemstillinger som de ovenfor. Disse omfattet alt fra enorme systemer med flere hundre ligninger til enkle univariate modeller som bare betrakter en enkelt tidsserie. Etter det økonomiske kaoset mot slutten av 1970- tallet mistet derimot mange av disse modellene sin troverdighet. I Sims (1980) ble det presentert et nytt makroøkonometrisk rammeverk spesialdesignet for slike problemstillinger: Vektor autoregressive (VAR)- modeller.

En univariat autoregresjon er en lineær modell med en ligning og en variabel, der verdien av variabelen i dag forklares ut ifra dens egne historiske verdier. En VAR er en lineær modell med  $n$ -ligninger og  $n$ -variabler der hver variabel forklares av sine egne historiske verdier, samt dagens og historiske verdier på de andre ( $n-1$ ) variablene. Dette rammeverket bidrar sånn sett med en systematisk metode ment til å fange opp dynamikken mellom flere tidsserier.

### 5.1 Forskjellige varianter

Det finnes tre utgaver av VAR: redusert form, rekursiv og strukturell. VAR på redusert form uttrykker hver variabel som en lineær funksjon av egne historiske verdier, historiske verdier for de andre variablene og et feilledd som ikke er seriekorrelet. For å illustrere kan man tenke seg en VAR som ser på sammenhengen mellom inflasjon ( $\pi_t$ ), arbeidsledighet ( $l_t$ ) og styringsrente ( $r_t$ ). I en VAR på redusert form vil en da ha tre ligninger:

$$(5.1) \quad \pi_t = f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^\pi,$$

$$(5.2) \quad l_t = f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^l,$$

$$(5.3) \quad r_t = f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^r.$$

(5.1) viser inflasjon i periode  $t$  som en funksjon av tidligere verdier for inflasjon, ledighetsrate og styringsrente, (5.2) arbeidsledighet som en funksjon av tidligere verdier for ledighetsrate, inflasjon og styringsrente og (5.3) det tilsvarende for styringsrente. Hver ligning estimeres med minste kvadraters metode, og  $\varepsilon_t^\pi$ ,  $\varepsilon_t^l$  og  $\varepsilon_t^r$  er feilleddene. Hvor mange tidligere verdier som skal inkluderes i hver ligning kan fastslås ved hjelp av flere metoder, som vil forklares i neste kapittel. Bevegelsene i parametrene som ikke gjøres rede for med de inkluderte historiske verdier, omfattes i feilleddene. Om de forskjellige variablene er korrelerte, som er vanlig for tidsseriedata, vil også feilleddene i den reduserte VARen være korrelert på tvers av ligningene.

En rekursiv VAR konstruerer feilleddene i hver regresjonslikning slik at de ikke er korrelert med feilleddene i de øvrige likningene. Dette gjennomføres ved å ta med noen forklarende verdier fra samme tidsperiode. Om en eksempelvis tenker seg en VAR med tre variabler rangert som 1) inflasjon, 2) arbeidsledighet og 3) styringsrente, ser den slik ut:

$$(5.4) \quad \pi_t = f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^\pi,$$

$$(5.5) \quad l_t = \pi_t + f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^l,$$

$$(5.6) \quad r_t = \pi_t + l_t + f(\text{lagede verdier av } \pi, l, r) + \varepsilon_t^r.$$

I likningen (5.4) er inflasjon den avhengige variabelen og forsinkede verdier på alle de tre variablene er uavhengige variabler. I (5.5) er ledighetsraten den avhengige variabelen og uavhengige variabler er forsinkede verdier av alle tre variablene og dagens verdi for inflasjonsrate. Styringsrente er avhengig variabel i (5.6) der forsinkede verdier for alle tre variabler, nåværende verdi for inflasjon og for ledighetsraten er de uavhengige. Om hver likning estimeres med minste kvadraters metode vil feilleddene,  $\varepsilon_t^\pi$ ,  $\varepsilon_t^l$  og  $\varepsilon_t^r$ , ikke være korrelert på tvers av likningene. Rekkefølgen man rangerer variablene vil påvirke resultatene man får, og dermed likningene, koeffisientene og residualene. Det finnes  $n!$  forskjellige rekursive VAR, som representerer alle rangeringer. I tilfelle her finnes det seks forskjellige alternativer.

I strukturelle VAR er det økonomisk teori som legges til grunn for å finne sammenhenger mellom variablene. Strukturelle VAR krever identifiserende forutsetninger som tillater at man

tolker korrelasjon i modellen kausalt. Disse antakelsene kan involvere hele VARen, slik at alle kausale sammenhenger er spesifisert, eller bare en enkelt likning, der bare en spesifikk kausal sammenheng identifiseres. Dette produserer instrumenter som kan benyttes til instrumental variables regression. Med denne regresjonen kan gjeldende sammenhenger mellom variablene estimeres. Antall strukturelle VAR avhenger bare av forskerens oppfinnsomhet.

Innen eksempler knyttet til inflasjon, arbeidsledighet og styringsrente kan en tenke seg to forskjellige strukturelle VAR- modeller, hver med forskjellige antakelser som bestemmer den kausale påvirkningen pengepolitikk har på rentenivå, ledighet og inflasjon. Den første kan baseres på en utgave av Taylors renteregulering, der sentralbanken modelleres som reaktiv. Altså at de setter styringsrenten med utgangspunkt i tidligere verdier for rente, inflasjon og ledighet. (5.7) er en variasjon av Taylor renteregelen i og med at produksjonsgap er erstattet med Okuns lov. I denne modellen vil sentralbanken sette styringsrenta  $r_t$  i henhold til regelen i (5.7).

$$(5.7) \quad r_t = r^* + 1,5(\bar{\pi}_t - \pi^*) - 1,25(\bar{l}_t - l^*) + f(\text{laggede verdier av } \pi, l, r) + \varepsilon_t,$$

der  $r^*$  er målet for realrenten,  $\bar{\pi}_t$  og  $\bar{l}_t$  er gjennomsnittelige verdier for inflasjon og ledighet de siste fire årene,  $\pi^*$  og  $l^*$  representerer inflasjonsmålet og ønsket ledighetsnivå, og  $\varepsilon_t$  er feilleddet. Taylors renteregulering impliserer at sentralbanken bare reagerer på tidligere informasjon, via  $\bar{\pi}_t$  og  $\bar{l}_t$ . Dette bestrides av flere forskere, som hevder sentralbankens atferd beskrives bedre som fremoverskuende. Derfor kan en også tenke seg en strukturell VAR, som i (5.7), der  $\bar{\pi}_t$  og  $\bar{l}_t$  erstattes med prognoser fire kvartaler fremover for ledighet og inflasjon. Prognosene kan estimeres fra VARen på redusert form.

## 5.2 Databeskrivelse

I VAR- analyse er det vanlig å rapportere resultatene fra Granger-causality tester, impulsrespons og dekomposisjon av variansen til prognosefeilen. På grunn av at VAR ofte har meget kompliserte dynamikk, er det i mange tilfeller mer informativt å se på resultatene fra de resultatene disse testene, enn  $R^2$  verdi eller de forskjellige koeffisientene.

En Granger-causality test undersøker om forsinkede verdier av en tidsserie bidrar til å predikere verdien på en annen. Om for eksempel arbeidsledighet ikke vil bidra til å forutsi inflasjonsrate i den reduserte VAR-en over vil koeffisientene på de historiske verdiene for arbeidsledighet ikke være signifikant forskjellig fra null i ligning (.1). Det er vanlig å rapportere resultatene fra en Granger-causality test med  $F$ -testenes tilhørende  $p$ -verdier.  $F$ -testene undersøker om de relevante koeffisientene er forskjellig fra null.

Impulsrespons anvendes for å se hvordan dagens og fremtidige verdier av hver enkelt variabel responderer når dagens verdi for et av feilleddene økes med en enhet. Det antas i testen at feilleddet går tilbake til null i de neste periodene, og at alle andre feilledd er lik null.

Tankeeksperimentet assosiert med å endre et feilledd, *ceteris paribus*, gir mest mening når de forskjellige ligningenes feilledd ikke er korrelerte med hverandre. Impulsrespons er derfor vanligst i rekursive og strukturelle VAR-modeller.

Dekomposisjonen av prognosefeil er prosentandel av variansen til feilestimeringen av en variabel som stammer fra et spesifikt sjokk over en gitt periode. Et muligens oppklarende eksempel: Hvor stor prosentandel av en estimeringsfeil i inflasjon som stammer fra et sjokk i feilleddet i ledighetsligningen over to år. På den måten blir dekomposisjonen av prognosefeil en slags  $R^2$ - verdi for feilen ved estimeringen av en variabel.

### 5.3 Bruksområder

I hovedsak er det fire forskjellige ting makroøkonomer gjør: Beskriver makroøkonomiske data, lager økonomiske prognoser, kvantifiserer det man vet og ikke vet om den underliggende makroøkonomiske strukturen og kommer med råd til utforming av makroøkonomisk politikk. Stock og Watson (2001) evaluerer VAR- modeller til bruk innen disse fire områdene. I følge de er VAR- modeller imponerende verktøy til bruk innen databeskrivelse og multivariate prognoser. Når det gjelder analyse av struktur og politikk er det imidlertid mer komplisert. Selv om en strukturell VAR kan fange opp mange dynamiske egenskaper av flere tidsserie data, er de strukturelle implikasjonene bare så gode som de identifiserende antakelsene som anvendes. Selv om det ifølge Stock og Watson finnes eksempler på gjennomtenkte identifiserende forutsetninger, er normen faktisk at det ignoreres. I enkelte felter innen økonomi, som arbeidsøkonomi og offentlig finans, kan slike forutsetninger finnes fra naturlige eksperimenter som tillater noe eksogen variasjon, men for

det meste er slike naturlige eksperimenter svært sjeldne innen makroøkonomi. Det er her verdt å merke seg at det ikke bare er VAR- modeller som har problemer med strukturelle implikasjoner eller analyse av økonomisk politikk. Dynamiske stokastiske generelle likevekts modeller som er kalibrert er svært presise når det gjelder kausale sammenhenger og forventninger, og kan derfor bidra med modeller til politisk analyse som er teoretisk gode. Ulempen er at de ofte passer data dårlig. I det andre ytterpunkt har en enkle modeller med bare en ligning, som er lette å estimere og kan komme med gode prognoser. Problemet er at det er enda vanskeligere å atskille korrelasjon og kausalitet i en slik modell enn det er i en VAR. En slik enkel modell kan jo ansees som en av ligningene i en større VAR. Det gjenstår altså arbeid med å utvide VAR- modellene, hovedsakelig til å omfatte flere dimensjoner, og ikke- lineære strukturer, men selv uten disse utvidelsene er VAR- modeller en verdifull del av makroøkonomers verktøykasse. Før verktøyene kan tas i bruk er det imidlertid nødvendig med noe grunnleggende tidsserie teori, som vil gjøres rede for i neste kapitell.



## 6 Statistisk teoridel

Ideen om at enkelte økonomiske variabler på lang sikt vil henge sammen, og ikke skal avvike fra hverandre i for stor grad, kan hevdes å ligge i kjernen av økonomisk teori. På kort sikt kan slike variabler drive fra hverandre eksempelvis på grunn av sesongbetingede faktorer eller sjokk. Fortsetter de derimot å være for langt fra hverandre over lengre tid vil økonomiske mekanismer føre disse variablene sammen igjen. Slike variabler kan tenkes å være renter på verdipapirer med forskjellige forfallsdato, priser på goder i forskjellige deler av et land, offentlige inntekter og utgifter eller salgspriser og produksjonskostnader. Andre eksempler er pris- og lønnsutvikling, import og eksport, systematisk risiko og avkastning på aksjer eller spotpris og fremtidig pris på et gode. I noen tilfeller kan en økonomisk teori med likevektskonsepter tilsi nær langsiktig sammenheng også med mulighet for å legge til flere variabler. Hvorvidt det faktisk finnes en langsiktig sammenheng mellom forskjellige sett med variabler er derimot et empirisk spørsmål. Den underliggende ideen om kointegrasjon muliggjør spesifisering av modeller som kan fange opp slike eventuelle langsiktige sammenhenger. Særlig for en type variabel som er svært vanlig i makroøkonomi. Siden ”lang sikt” er et økonomisk konsept som i høyeste grad er dynamisk, er det derfor teori og analyse innen tidsserie- variabler som blir det logiske fagområdet. I denne delen av oppgaven vil det redegjøres for noen vesentlige emner innenfor tidsserieteori slik at den planlagte analysen kan gjennomføres korrekt.

### 6.1 Stasjonære og ikke- stasjonære data

Betrakt en enkelt tidsserie  $y_t$ , som måles ved like tidsintervaller. Tidsserieanalyse begynner med å vurdere seriens genererende mekanisme, datagenereringsprosessen. Denne prosessen bør kunne generere alle de statistiske egenskapene for serien, eller i det minste gjennomsnitt, varians og seriekorrelasjon. Altså seriens lineære egenskaper, gitt tidligere observerte data. Om de lineære egenskapene eksisterer og er invariante for tid sier en at serien har egenskapen å være stasjonær. Det vil si at gjennomsnitt og varians er konstant over tid, og kovarians

mellom to observasjoner fra serien kun avhenger av intervallet mellom de to observasjonene, og ikke når observasjonene fant sted. Om  $y_t$  er en stasjonær tidsserie må følgende gjelde:

$$(6.1) \quad E(y_t) = \mu \quad (\text{konstant gjennomsnitt})$$

$$(6.2) \quad \text{var}(y_t) = \sigma^2 \quad (\text{konstant varians})$$

$$(6.3) \quad \text{cov}(y_b, y_{t+s}) = \text{cov}(y_b, y_{t-s}) = \gamma_s \quad (\text{kovarians avhenger av } s, \text{ ikke } t)$$

Videre vil fokus derimot ligge på en noe svakere, men langt mer teknisk betingelse, nemlig at serien har et endelig spektrum som er forskjellig fra null ved alle frekvenser. Noe forenklet betyr dette at serien har endelige verdier, *ikke* konvergerer mot uendelig og har et endelig integral. En serie med slike egenskapene kan anvendes i analyser som skal omtales under og kalles  $I(0)$ , som betegner at den er integrert av nulte grad. At en serie må være  $I(0)$  er en nødvendig, men ikke tilstrekkelig betingelse for stasjonærhet. Dermed er alle stasjonære serier  $I(0)$ , men ikke nødvendigvis motsatt, selv om det er vanlig å bruke disse begrepene om hverander. Enkelte serier må differensieres for å oppnå slike egenskaper, og vil sies å være integrert av første grad. Serien vil da angis  $y_t \sim I(1)$ . For å generalisere kan en si at en serie som må differensieres  $d$  ganger for å bli  $I(0)$ , vil være integrert av grad  $d$  og skrives  $y_t \sim I(d)$ . La  $\Delta^b$  bety at man tar differansen av en serie  $b$  ganger. Hvis  $y_t \sim I(d)$ , og en tar differansen  $b$  ganger vil serien  $\Delta^b y_t$  være  $I(d-b)$ .

Det enkleste eksempelet på en  $I(0)$  serie er såkalt white noise  $\varepsilon_t$ . En slik serie vil oppfylle (6.1) – (6.3), og dermed ha lineære egenskaper som er invariante for tid. Utover dette har white noise også egenskapen at  $\text{corr}(\varepsilon_b, \varepsilon_{t-s}) = 0$  for alle  $s \neq 0$ .

Som et annet eksempel kan man betrakte en stasjonær  $AR(1)$  prosess  $y_t$  som genereres av

$$(6.4) \quad y_t = \rho y_{t-1} + \varepsilon_t,$$

der  $|\rho| < 1$ , og kalles den autoregressive koeffisienten. Denne koeffisienten viser påvirkningskraften av forsinkede variabler på variabelen i periode  $t$ .  $\varepsilon_t$  er fortsatt er white noise med null i gjennomsnitt.  $AR(1)$  modellen viser at hver realisering av den tilfeldige variabelen  $y_t$  inneholder en del  $\rho$  av forrige periodes verdi  $y_{t-1}$  samt feilleddet  $\varepsilon_t$  som trekkes fra en distribusjon med null i snittverdi og varians  $\sigma_\varepsilon^2$ . Siden det bare er *en* forsinket variabel i (6.4) kalles det en førsteordens autoregressiv modell,  $AR(1)$ , mens en modell av samme form med  $p$  forsinkede ledd, fra  $y_t$  til  $y_{t-p}$  kalles en  $AR(p)$  eller  $p$ -ordens autoregressiv modell. Endrer en (6.4) slik at  $|\rho| = 1$  vil man få en såkalt random walk der  $y_t$  genereres av

$$(6.5) \quad y_t = y_{t-1} + \varepsilon_t,$$

som teoretisk sett ville være tilfellet for en pris som fastsettes i et marked med full informasjon. En random walk er det enkleste eksemplet på en  $I(1)$  prosess, og det er verdt å merke seg at førstedifferansen er white noise, som dermed følgelig er stasjonær. De mest generelle  $I(1)$  seriene erstatter  $\varepsilon_t$  i (6.5) med en hvilken som helst  $I(0)$ , som ikke nødvendigvis må ha gjennomsnittlig verdi lik null, så lenge snittverdien er konstant.

Ellers er det flere viktige forskjeller på  $I(0)$  og  $I(1)$  serier. En  $I(0)$  serie har en gjennomsnittsverdi, og serien har en tendens til å returnere til denne. Dermed er det typisk for en  $I(0)$  serie å fluktuere rundt snittverdien, med hyppig kryssing av den, men serien også ha omfattende avstikkere. Autokorrelasjon avtar hurtig ettersom forsinkelsen øker, slik at prosessen vil gi liten vekt til hendelser på mellomlang og lang sikt. Dermed har serien også praktisk sett endelig minne (finite memory). En  $I(1)$  prosess uten drift vil være relativt jevn, vil vandre bredt, og bare ytterst sjelden vende tilbake til en tidligere verdi. For en gitt vilkårlig verdi i en random walk prosess er faktisk den forventede tiden det tar før serien igjen vender tilbake til denne samme verdien uendelig. Det betyr ikke at serien aldri vil vende tilbake til samme verdi, men at den forventede tiden før retur til tidligere verdi har en svært bred distribusjon. Autokorrelasjonene,  $\rho_s$ , for en  $I(1)$  serie er alle nær en, selv for stor  $s$ . En innovasjon, eller et sjokk, til prosessen påvirker alle senere verdier slik at serien har et uendelig minne. For å se dette kan en betrakte random walk-serien i (6.5), og foreta litt rekursiv substitusjon:

$$(6.5a) \quad y_1 = y_0 + \varepsilon_1$$

$$(6.5b) \quad y_2 = y_1 + \varepsilon_2 = (y_0 + \varepsilon_1) + \varepsilon_2 = y_0 + \sum_{s=1}^2 \varepsilon_s$$

$$(6.5c) \quad y_t = y_{t-1} + \varepsilon_t = y_0 + \sum_{s=1}^t \varepsilon_s.$$

Random walk-serien inneholder en opprinnelig verdi  $y_0$  og en komponent som utgjøres av summen av alle tidligere stokastiske ledd  $\sum_{s=1}^t \varepsilon_s$ .  $y_0$  settes ofte lik null siden observasjonen intraff for så lenge siden at dens effekt på  $y_t$  er nær ubetydelig. Den siste komponenten kalles den stokastiske trenden siden et stokastisk ledd  $\varepsilon_t$  legges til i hver periode og dermed får

serien til å følge en uforutsigbar trend. Det er også verdt å merke seg at variansen til  $y_t$  kan finnes fra (6.5c):

$$(6.6) \quad \text{var}(y_t) = \text{var}\left(\sum_{s=1}^t \varepsilon_s\right) = t\sigma_\varepsilon^2,$$

som avhenger av tid. Dermed bryter den med betingelsen i (6.2) og  $\text{var}(y_t)$  vil konvergere mot uendelig ettersom  $t$  øker.

## 6.2 Dickey- Fuller testen

Det er flere tester som kan benyttes for å fastslå om en tidsserie er stasjonær eller ikke. Den mest utbredte er Dickey- Fuller testen, som vil anvendes i denne oppgaven. Om en igjen ser på en  $AR(1)$  prosess,  $y_t = \rho y_{t-1} + \varepsilon_t$ , vil  $y_t$ , som nevnt over, være stasjonær når  $|\rho| < 1$ , men endres til en ikke- stasjonær random walk prosess når  $\rho = 1$ . Sånn sett kan en altså avgjøre om en serie er stasjonær ved å teste om  $\rho$  er lik en, eller signifikant forskjellig fra en. Tester av denne sorten kalles enhetsrottester, siden de påviser en eventuell enhetsrot for polynomet som modellen utgjøres av.

Med utgangspunkt i  $AR(1)$ - modellen

$$(6.7) \quad y_t = \rho y_{t-1} + \varepsilon_t,$$

tester man for ikke- stasjonæritet ved å sette nullhypotesen at  $\rho = 1$  mot alternativhypotesen at  $|\rho| < 1$ . Praktisk sett ser en på om  $\rho < 1$ , siden  $\rho > 1$  vil føre til eksplosiv vekst. Ved å trekke fra  $y_{t-1}$  på begge sider av (6.7) oppnår man en mer praktisk form:

$$(6.8) \quad \Delta y_t = \gamma y_{t-1} + \varepsilon_t,$$

der  $\gamma = \rho - 1$ . Dermed blir hypotesene:

$$H_0 : \rho = 1 \quad \text{mot} \quad H_1 : \rho < 1,$$

eller

$$H_0 : \gamma = 0 \quad \text{mot} \quad H_1 : \gamma < 0.$$

Nullhypotesen spesifiserer at tidsserien er ikke- stasjonær og alternativhypotesen at serien er stasjonær. Det er verdt å merke seg at testen er en enhalet test, og at (6.8) enkelt kan utvides til å omfatte et konstantledd, eller en tidsavhengig trend:

$$(6.9) \quad \Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t,$$

og

$$(6.10) \quad \Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + \varepsilon_t,$$

der  $\alpha$  er et konstantledd og  $\lambda t$  er en lineær trend.

Ved testing for enhetsrot er det ofte nyttig å plote dataseriene i en graf. Da kan man fastslå hvilken av testene (6.8)- (6.10) som er egnet til evaluering av serien en betrakter. I alle tre tilfellene over estimerer en testligningen (6.8), (6.9) eller (6.10) ved minste kvadraters metode, og inspiserer  $t$ -verdien til hypotesen at  $\rho = 1$ . Denne  $t$ -verdien har ikke lenger en vanlig  $t$ -fordeling siden variansen til  $y_t$  vil øke som en funksjon av tid dersom nullhypotesen er sann, og  $y_t$  ikke er stasjonær. Denne økte variansen endrer dermed på  $t$ -fordelingen når  $H_0$  er sann. Derfor anvendes en verdi kalt en  $\tau$ -verdi, som avhenger av hvilket av tilfellene (6.8)- (6.10) en ser på. Disse verdiene ble opprinnelig utregnet av David Dickey og Wayne Fuller, slik at alle enhetsrotstester som anvender disse testverdiene, eller verdier regnet ut på tilsvarende måter, kalles Dickey- Fuller tester. Kritiske verdier for Dickey- Fuller testen vises i tabell 3 under.

**Tabell 3:** Kritiske verdier for Dickey- Fuller testen.

Modell	1 %	5 %	10 %
$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$	-2.56	-1.94	-1.62
$\Delta y_t = \alpha + \gamma y_{t-1} + \varepsilon_t$	-3.43	-2.86	-2.57
$\Delta y_t = \alpha + \lambda t + \gamma y_{t-1} + \varepsilon_t$	-3.96	-3.41	-3.13

**Kilde:** Davidson og MacKinnon (1993)

Om utregnet testverdi er lavere enn kritisk  $\tau$ -verdi,  $\tau \leq \tau_c$ , kan en fastslå at serien er stasjonær, mens  $\tau > \tau_c$  tilsier det at dataene er ikke- stasjonære.

En kan også enkelt utvide testen til å korrigere for autokorrelasjon ved å legge til forsinkede verdier av serien en betrakter. En slik test kalles en augmented Dickey- Fuller test, og det

bestemmes hvor mange forsinkede verdier ved hjelp av informasjonskriterier, som vil drøftes nærmere under.

### 6.3 Spurious regresjon

Den viktigste motivasjonen for å finne ut om en tidsserie er stasjonær eller ikke, før en begir seg ut på en analyse, er faren ved å komme frem til tilsynelatende signifikante resultater for data som faktisk er totalt uavhengige.

Om en for eksempel kjører en regresjonsanalyse på to forskjellige random walk serier med stigende trend, risikerer en å ende opp med tilsynelatende signifikante koeffisienter, selv om disse vil være meningsløse. Årsaken til dette er at en da vil ha relatert en serie med en stokastisk trend, til en annen serie med en helt annen stokastisk trend. Dermed vil seriene ikke ha noe som helst til felles og heller ikke være kausalt forbundet på noen som helst måte, selv om en graf kan tyde på at de to er korrelert.

Bruken av ikke-stasjonære data i en regresjonsmodell vil altså føre til resultater som feilaktig tilsier et signifikant forhold mellom variablene i modellen. Siden mange makroøkonomiske tidsserier ikke er stasjonære er det viktig å være varsom når en skal estimere modeller med slike. For ikke å få problemer med feilaktig regresjon bør en sann sett ha som en tommelfingerregel å unngå ikke-stasjonære data i en regresjonsmodell. Det er derimot et meget viktig unntak fra denne regelen; kointegrasjon.

### 6.4 Kointegrasjon

Om  $y_t$  og  $x_t$  er to ikke-stasjonære  $I(1)$  variabler vil det generelt sett forventes at enhver lineær kombinasjon av disse to seriene også vil være  $I(1)$ . Finnes det derimot en konstant  $A$ , slik at

$$(6.11) \quad z_t = x_t - Ay_t$$

er  $I(0)$ , impliserer dette at  $y_t$  og  $x_t$  har en felles stokastisk trend. I dette tilfellet eksisterer det en spesielt sterk kobling mellom de to variablene, som kan være svært interessant fra et økonomisk perspektiv. Etter Granger (1981) og Engle og Granger (1987) heter det seg at

variabler er kointegrerte om de har en felles stokastisk trend. I tilfellet over sier en at  $y_t$  og  $x_t$  er kointegrerte, og  $A$  kalles den kointegrerende parameteren.

En kan også relativt enkelt utvide kointegrasjonsbegrepet til å omfatte vektorer. La  $\mathbf{y}_t$  være en  $(n \times 1)$  tidsserievektor der hver av de  $n$  seriene for seg er  $I(1)$ .  $\mathbf{y}_t$  sies å være kointegrert dersom det finnes en lineær kombinasjon av seriene  $\mathbf{a}'\mathbf{y}_t$  som er  $I(0)$ , når  $\mathbf{a}$  er en  $(n \times 1)$  vektor forskjellig fra nullvektoren. I det tilfellet vil  $\mathbf{a}$  være en kointegrerende vektor for seriene som utgjør  $\mathbf{y}_t$ . Det er verdt å merke seg at  $\mathbf{a}$  ikke vil være unik, for om  $\mathbf{a}'\mathbf{y}_t$  er stasjonær, vil også  $b\mathbf{a}'\mathbf{y}_t$  være det, der  $b$  er en skalar forskjellig fra null. Om  $\mathbf{a}$  er en kointegrerende vektor vil også  $b\mathbf{a}$  være det. Når en snakker om verdien til den kointegrerende vektoren må en vilkårlig normalisering gjennomføres, slik at det første elementet i  $\mathbf{a}$  er én. Dette vil utdypes ytterligere i neste del av kapittelet.

Hvis det er flere enn to variabler i  $\mathbf{y}_t$  kan det finnes to  $(n \times 1)$  vektorer  $\mathbf{a}_1$  og  $\mathbf{a}_2$ , der  $\mathbf{a}_1, \mathbf{a}_2 \neq \mathbf{0}$ , slik at både  $\mathbf{a}_1'\mathbf{y}_t$  og  $\mathbf{a}_2'\mathbf{y}_t$  er stasjonære.  $\mathbf{a}_1$  og  $\mathbf{a}_2$  må også være lineært uavhengige, som betyr at det *ikke* finnes noen skalar  $b$  slik at  $\mathbf{a}_1 = b\mathbf{a}_2$ . Faktisk kan det eksistere  $h < n$  lineært uavhengige  $(n \times 1)$  vektorer  $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_h)$  slik at  $\mathbf{A}'\mathbf{y}_t$  er en stasjonær  $(h \times 1)$  vektor, der  $\mathbf{A}'$  er følgende  $(h \times n)$  matrise:

$$(6.12) \quad \mathbf{A}' = \begin{pmatrix} \mathbf{a}'_1 \\ \vdots \\ \mathbf{a}'_h \end{pmatrix}$$

Heller ikke i dette tilfellet vil vektorene  $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_h)$  være unike. Om  $\mathbf{A}'\mathbf{y}_t$  er stasjonær, og  $\mathbf{b}'$  er en  $(1 \times h)$  vektor forskjellig fra nullvektoren, vil også skalaren  $\mathbf{b}'\mathbf{A}'\mathbf{y}_t$  også være stasjonær. Dermed vil også  $(n \times 1)$  vektoren  $\boldsymbol{\pi}$ , definert som  $\boldsymbol{\pi}' = \mathbf{b}'\mathbf{A}'$ , være en kointegrerende vektor.

La  $\mathbf{A}'$  være en  $(h \times n)$  matrise med lineært uavhengige rader slik at  $\mathbf{A}'\mathbf{y}_t$  er en stasjonær  $(h \times 1)$  vektor. La så  $\mathbf{c}'$  være enhver  $(1 \times n)$  vektor som er lineært uavhengig av radene i  $\mathbf{A}'$ . Da vil  $\mathbf{c}'\mathbf{y}_t$  være en ikke-stasjonær skalar. Det betyr at det ikke finnes flere vektorer som både er lineært uavhengige av radene i  $\mathbf{A}'$ , og gir et stasjonært produkt med  $\mathbf{y}_t$ , og en kan slutte at det finnes nøyaktig  $h$  kointegrerende relasjoner mellom elementene i  $\mathbf{y}_t$ . Teknisk sett heter det seg da at  $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_h)$  utgjør en basis for vektorrommet til de kointegrerende vektorene.

I kapittel 5 ble det VAR- modeller og deres bruksområder presentert. For å se hvordan en kan anvende disse begrepene til tidsserieanalyse vil det nå tas utgangspunkt i en noe mer teknisk fremstilling av VAR-modellene.

## 6.5 Vector Error Correcting Models

Det er nå nødvendig å ta utgangspunkt i en litt mer teknisk fremstilling av VAR enn tidligere. For et sett av  $n$  tidsserier  $\mathbf{y}_t = (y_{1t}, \dots, y_{nt})'$  vil en VAR-modell fange opp det dynamiske samspillet mellom seriene. En enkel  $p$ -ordens model, VAR( $p$ ), ser slik ut:

$$(6.13) \quad \mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t,$$

der  $\mathbf{A}_i$  er  $(n \times n)$  koeffisientmatriser for  $i = 1, \dots, p$  og  $\mathbf{u}_t = (u_{1t}, \dots, u_{nt})'$  er et ikke-observerbart feilledd. Vanligvis antas  $\mathbf{u}_t$  å være white noise, som altså betyr at leddet har null i gjennomsnittsverdi og en tidsinvariant positiv definitte kovariansmatrise  $E(\mathbf{u}_t \mathbf{u}_t') = \Sigma_u$ . Med andre ord er altså  $\mathbf{u}_t$  en uavhengig stokastisk vektor med  $\mathbf{u}_t \sim (\mathbf{0}, \Sigma_u)$ .

Prosessen i (6.13) er stabil om

$$(6.14) \quad \det(\mathbf{I}_n - \mathbf{A}_1 z - \dots - \mathbf{A}_p z^p) \neq 0 \text{ for } |z| \leq 1.$$

Matematisk sett betyr det at polynomet som defineres av determinanten til den autoregressive operatoren,  $\det(\mathbf{I}_n - \mathbf{A}_1 z - \dots - \mathbf{A}_p z^p)$ , må ha reelle røtter. Med forutsetningen at serienes datagenereringsprosess begynte for uendelig lenge siden ( $t = 0, \pm 1, \pm 2, \dots$ ) vil den generere stasjonære tidsserier som har lineære egenskaper som er tidsinvariante. Om derimot polynomet i (6.14) har en enhetsrot, slik at determinanten er lik null for  $z = 1$ , impliserer det at noen, eller alle variablene er integrerte. For å forenkle antas det med det følgende at variablene på det meste vil være  $I(1)$ . Som nevnt ovenfor er det mulig at forskjellige serier kan ha felles stokastisk trend, og at det kan eksistere lineære kombinasjoner av de som er  $I(0)$ . At de altså er kointegrerte. Det kan også være nyttig å betrakte systemer med både  $I(1)$  og  $I(0)$  variabler. På den måten utvides kointegrasjonskonseptet ved å kalle enhver  $I(0)$  lineær kombinasjon for en kointegrerende relasjon. Det er viktig å notere seg at denne terminologien ikke er helt i tråd med det opprinnelige begrepet siden det kan skje at en lineær kombinasjon av  $I(0)$  variabler sies å være en kointegrerende relasjon.



Hvis det er kointegrerende variabler tilstede i et system av variabler er ikke VAR- modellen optimal for analyse. I det tilfellet er det fordelaktig å anvende en mer spesifikk parametrisering som støtter analysen av den kointegrerende strukturen. Om en trekker fra  $y_{t-1}$  på begge sider av (6.13):

$$(6.13a) \mathbf{y}_t - \mathbf{y}_{t-1} = \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t - \mathbf{y}_{t-1}$$

$$(6.13b) \quad \Delta \mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p} + \mathbf{u}_t - \mathbf{y}_{t-1} + \mathbf{A}_2 \mathbf{y}_{t-1} - \mathbf{A}_2 \mathbf{y}_{t-1} + \dots + \mathbf{A}_p \mathbf{y}_{t-p+1} - \mathbf{A}_p \mathbf{y}_{t-p+1}$$

$$(6.13c) \quad \Delta \mathbf{y}_t = -(\mathbf{I}_n - \mathbf{A}_1 - \dots - \mathbf{A}_p) \mathbf{y}_{t-1} - (\mathbf{A}_2 + \dots + \mathbf{A}_p) \Delta \mathbf{y}_{t-1} - \dots - (\mathbf{A}_{i+1} + \dots + \mathbf{A}_p) \Delta \mathbf{y}_{t-i}$$

La  $\Pi = -(\mathbf{I}_n - \mathbf{A}_1 - \dots - \mathbf{A}_p)$  og  $\Gamma_i = -(\mathbf{A}_{i+1} + \dots + \mathbf{A}_p)$  for  $i = 1, \dots, p-1$  slik at (6.5.1c) kan skrives

$$(6.15) \quad \Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \Gamma_1 \Delta \mathbf{y}_{t-1} + \dots + \Gamma_{p-1} \Delta \mathbf{y}_{t-p+1}$$

Ligning (6.15) er best kjent som en Vector Error Correction Model (VECM). På grunn av antakelsen om at variabler maksimalt er  $I(1)$  vil  $\Delta \mathbf{y}_t$  være  $I(0)$ . Derfor er  $\Pi \mathbf{y}_{t-1}$  det eneste leddet som inneholder ikke-stasjonære variabler. Imidlertid må  $\Pi \mathbf{y}_{t-1}$  må være  $I(0)$  for at (6.15) skal gjelde. For at det skal være tilfellet må  $\Pi \mathbf{y}_{t-1}$  inneholde de kointegrerende relasjonene, og blir følgelig spesielt interessant i analysen. I (6.16) kalles  $\Gamma_i$  – matrisene de kortsiktige parameteren og  $\Pi \mathbf{y}_{t-1}$  den langsiktige delen.

La  $\text{rk}(\Pi) = r$  betegne rangen til matrisen  $\Pi$ , altså det største antallet lineært uavhengige kolonner i  $\Pi$ . Da kan  $\Pi$  skrives som et produkt av  $(n \times r)$  matrisene  $\alpha$  og  $\beta$  med  $\text{rk}(\alpha) = \text{rk}(\beta) = r$  på følgende måte:  $\Pi = \alpha \beta'$ . Om en premultipliserer en  $I(0)$  vektor med en vilkårlig matrise vil det igjen resultere i en  $I(0)$  prosess.  $\beta' \mathbf{y}_{t-1}$  må derfor være  $I(0)$ , siden elementet kan oppnås ved å premultiplisere  $\Pi \mathbf{y}_{t-1} = \alpha \beta' \mathbf{y}_{t-1}$  med  $(\alpha' \alpha)^{-1} \alpha'$ . Dermed kan en slutte at  $\beta' \mathbf{y}_{t-1}$  inneholder de kointegrerende relasjonene. Det følger fra dette at det er  $r = \text{rk}(\Pi)$  lineært uavhengige kointegrerende relasjoner mellom komponentene i  $\mathbf{y}_t$ . Det sies derfor at rangen til matrisen  $\Pi$  er den kointegrerende rangen til systemet og at  $\beta$  er kointegrasjonsmatrisen. Om det eksempelvis er tre variabler med to kointegrerende relasjoner ( $r = 2$ ), har man

$$(6.16) \quad \Pi y_{t-1} = \alpha \beta' y_{t-1} = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} \\ \beta_{12} & \beta_{22} & \beta_{32} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \end{pmatrix}$$

Som nevnt tidligere er ikke matrisene  $\alpha$  og  $\beta$  unike. For enhver  $(r \times r)$  matrise  $\mathbf{B}$ , der  $|\mathbf{B}| \neq 0$ , kan man få nye matriser  $\alpha\mathbf{B}$  og  $\beta\mathbf{B}^{-1}$ , som tilfredstiller  $\Pi = \alpha\mathbf{B}(\beta\mathbf{B}^{-1})'$ . Sånn sett ser en igjen at økonomisk innhold ikke kan trekkes direkte ut av de observerte seriene. Derfor er det bare det kointegrerende rommet og ikke de kointegrerende parameterene som estimeres korrekt. For å estimere  $\alpha$  og  $\beta$  korrekt er det nødvendig å innføre identifiserende restriksjoner. Uten slike vil bare en bare kunne estimere  $\alpha\beta' = \Pi$ .

## 6.6 Spesifisering

Før en kan estimere en VECM eller en VAR-modell er det nødvendig å fastslå hvor mange forsinkede ledd som må omfattes i modellen. For en VECM må også systemets kointegrerende rang konstateres. Under vil det kort greies ut om forskjellige metoder og teknikker for å finne ut av dette.

### 6.6.1 Modellens autoregressive orden

For å avgjøre en dynamisk modells autoregressive orden er det vanlig å benytte sekvensielle tester eller forskjellige utvelgelseskriterier. Det er ofte greit å starte med en VAR-modell som i (6.13) siden den kointegrerende rangen  $r$  til systemet i utgangspunktet ikke er kjent når det skal bestemmes hvor mange forsinkede verdier som skal tas med.

En mulig tilnærming er å begynne med en model med en forhåndsbestemt autoregressiv orden,  $p_{\max}$  og bruke sekvensielle tester for å finne en passende orden for modellen.

Eksempelvis kan en følgende sekvens med nullhypoteser testes intill testen forkastes:  $H_0 : \mathbf{A}_{p_{\max}} = 0$ ,  $H_0 : \mathbf{A}_{p_{\max}-1}$ , og videre. For denne tilnærmingen må det tas en beslutning om  $p_{\max}$ , for eksempel fra et teoretisk perspektiv. Om en vil inkludere forsinkede verdier for et år, må en ha 12 forsinkede verdier om man opererer med månedlige data, eller fire om en opererer med kvartalsvise data.

I stedet for sekvensielle tester kan en anvende mer generelle metoder gjennom forskjellige utvelgelseskriterier. Om en først tar utgangspunkt i en mer generell utgave av  $AR(1)$ -modellen i (6.4), en  $AR(p)$ -modell,

$$(6.17) \quad y_t = \rho_1 y_{t-1} + \dots + \rho_p y_{t-p} + \varepsilon_t,$$

vil et generelt utvelgelseskriterie være av formen:

$$(6.18) \quad Cr(p) = \ln \tilde{\sigma}_\varepsilon^2(p) + c_T \varphi(p),$$

der  $\tilde{\sigma}_\varepsilon^2 = T^{-1} \sum_{t=1}^T \hat{\varepsilon}(p)^2$  er estimatoren for variansen til feilledet, som baseres på OLS-

residualen  $\hat{\varepsilon}_t(p)$  fra  $AR(p)$ -modellen i (6.17). Det første leddet på høyresiden av (6.18) måler

hvordan en  $p$ -ordens modell passer til datasettet. Dette leddet vil avta for økende orden, siden det ikke korrigeres for frihetsgrader i variansestimatorene; flere forsinkede verdier øker modellens forklaringskraft. Videre betegner  $\ln$  den naturlige logaritmen,  $c_T$  er en sekvens som avhenger av størrelsen på utvalget og  $\varphi(p)$  er en funksjon som straffer store  $AR$ -ordener, slik at en ikke kan blåse opp modellens forklaringskraft med overdrevent høy orden. Den autoregressive ordenen som minimerer et utvelgelseskriterie velges som en estimator  $\hat{p}$  for den faktiske ordenen  $p$ .

En kan også utvide til et multivariat tilfelle der en vil tilpasse en  $VAR(p)$ -modell der  $p = 0, \dots, p_{\max}$ . Også i dette tilfellet vil en velge en estimator for  $p$  som minimerer et valgt kriterie. Nå vil forskjellige utvelgelseskriterier ha den generelle formen

$$(6.19) \quad Cr(p) = \ln \det(\tilde{\Sigma}_u(p)) + c_T \varphi(p),$$

der  $\det(\bullet)$  betegner determinanten og  $\tilde{\Sigma}_u(p) = T^{-1} \sum_{t=1}^T \hat{u}_t \hat{u}_t'$  er estimatoren for residualens kovariansmatrise i en  $p$ -ordens modell. De to andre leddene på høyre side av (6.19) er de samme som i (6.18), med samme funksjoner. Tre kjente kriterier er

$$(6.20) \quad AIC(p) = \ln \det(\tilde{\Sigma}_u(p)) + \frac{2}{T} pn^2,$$

$$(6.21) \quad HQ(p) = \ln \det(\tilde{\Sigma}_u(p)) + \frac{2 \ln \ln T}{T} pn^2$$

og

$$(6.22) \quad SC(p) = \ln \det(\tilde{\Sigma}_u(p)) + \frac{\ln T}{T} pn^2$$

der (6.20) er Akaikes informasjonskriterie, (6.21) er Hannan-Quinn- kriteriet og (6.22) er Schwarz-kriteriet.

Akaikes informasjonskriterie vil asymptotisk overestimere ordenen med positiv sannsynlighet, mens Hannan-Quinn- og Schwarz- kriteriet estimerer ordenen konsistent for ganske generelle betingelser hvis den faktiske datagenereringsprosessen har en endelig VAR orden og den maksimale ordenen  $p_{\max}$  er større en den faktiske ordenen. Disse resultatene vil da ikke bare gjelde for  $I(0)$  prosesser, men også kointegrerte  $I(1)$  prosesser. Om en betegner de forskjellige  $p$  som minimerer kriteriene med  $\hat{p}(AIC)$ ,  $\hat{p}(HQ)$  og  $\hat{p}(SC)$  vil

$$(6.23) \quad \hat{p}(SC) \leq \hat{p}(HQ) \leq \hat{p}(AIC),$$

gjelde, selv for små utvalg.

### 6.6.2 Kointegrerende rang

Som tidligere nevnt er matematikken som er nødvendig for å estimere kointegrerende rang noe utenfor denne oppgavens omfang. Likevel kan det være hensiktsmessig å redegjøre kort for hvordan rangen avgjøres.

Hvis enkelte av variablene i et system er  $I(1)$  kan en VECM være en passende modell, og kointegrerende rang  $r$  må velges i tillegg til autoregressiv orden. I utregningen av  $r$  anvendes det en sekvensiell testprosedyre der følgende rekkefølge med hypoteser testes:

$$H_0(0) : rk(\Pi) = 0 \quad \text{mot} \quad H_1(0) : rk(\Pi) > 0,$$

$$H_0(1) : rk(\Pi) = 1 \quad \text{mot} \quad H_1(1) : rk(\Pi) > 1,$$

⋮

$$H_0(n-1) : rk(\Pi) = n-1 \quad \text{mot} \quad H_1(n-1) : rk(\Pi) = n.$$

Testsekvensen avsluttes, og korresponderende kointegrerende rang velges når nullhypotesen ikke kan forkastes for første gang. Hvis den første nullhypotesen i denne sekvensen,  $H_0(0)$ , ikke kan forkastes betyr det at en VAR- modell med alle variablene i førstedifferanser bør anvendes. I det andre ytterpunktet, der alle nullhypotesene forkastes, inklusive  $H_0(n-1)$ , tyder det på at en VAR på nivåform bør benyttes til analyse av systemet.

## 7 Analyse

I denne delen av oppgaven vil all gjennomgått teori anvendes til å konstruere to forskjellige boligmodeller. Det som vil skille de to modellene er hvilke variabler som inngår i brukerkostnaden. Den første vil ta utgangspunkt i standard modellering, mens den andre vil forsøke å modellere mer rasjonelle aktører. Først i denne delen vil de anvendte variablene kort presenteres og testes for stasjonærhet før modellene spesifiseres, estimeres og kommenteres.

### 7.1 Datasett med kommentarer

For å gjennomføre analysen har det vært nødvendig med en rekke data. Siden det er bred enighet om at det var overvurderte boligpriser i Norge på slutten av 1980- tallet har det også vært et poeng å finne serier som strekker seg såpass langt tilbake i tid. Dataene er fra SSB og Norges Bank, og vil drøftes kort under. Hver variabel som inngår direkte i analysen vil også presenteres med en graf siden det blir hendig til Dickey- Fuller testene. Så langt det har vært mulig har det blitt brukt data av samme sort som i McCarthy og Peach (2002), der flere brukes i MODAG. Om ikke annet spesifiseres er priser faste.

#### *Boligpris.*

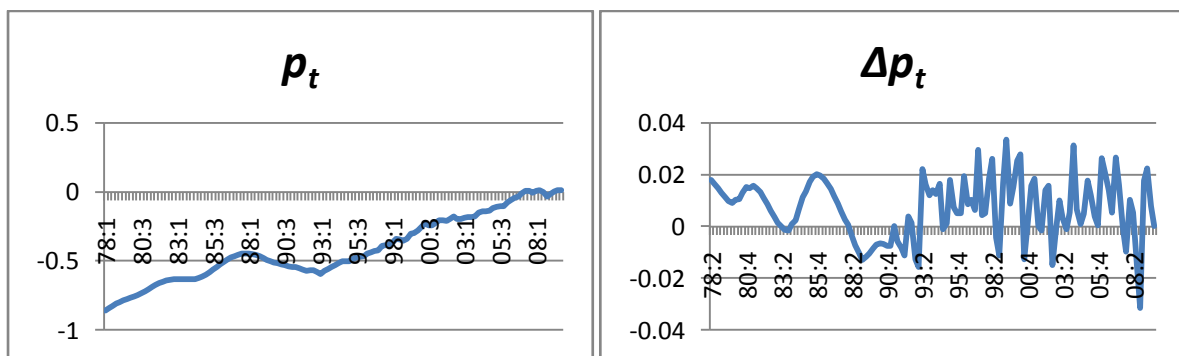
I denne oppgaven anvendes SSBs prisindeks for brukte selveierboliger *PBOL*. Tidligere var indeksen basert på data fra Tinglysningsregisteret, men har siden første kvartal 2003 vært beregnet på data fra FINN.no og Norges Boligbyggelags Landsforbund (NBBL). Indeksen ble tidligere satt sammen ved hedonisk metode. Dette betyr at man først finner sammenhenger mellom boligens kvalitetsegenskaper og pris, før man korrigerer for eventuelle kvalitetsendringer. Siden noe av datagrunnlaget fra NBBL består av gjennomsnittstall for

enkelte boligbyggerlag, kan ikke den hedoniske metoden benyttes på hele populasjonen. Disse omsetningene må derfor behandles annerledes enn selveieomsetningene. Data fra FINN.no er basert på enkeltobservasjoner og hedonisk metode kan derfor brukes.

Indeksen består av totalt 20 delindekser for selveide eneboliger, småhus og blokkleiligheter og småhus og blokkleiligheter i boligbyggerlag. I hovedindeksen vektes hver av disse ut ifra verdiandeler etter boligtype og geografisk sone.

$$(7.1) \quad p_t = \ln(PBOL_t)$$

**Figur 4:** Boligprisutvikling og boligprisvekst 1978:4 – 2009:4.



**Kilde:** SSB

### *Konsum.*

Som proxy for permanentinntekt benyttes konsum av ikke- varige goder i modellen, som modelleres slik i MODAG:

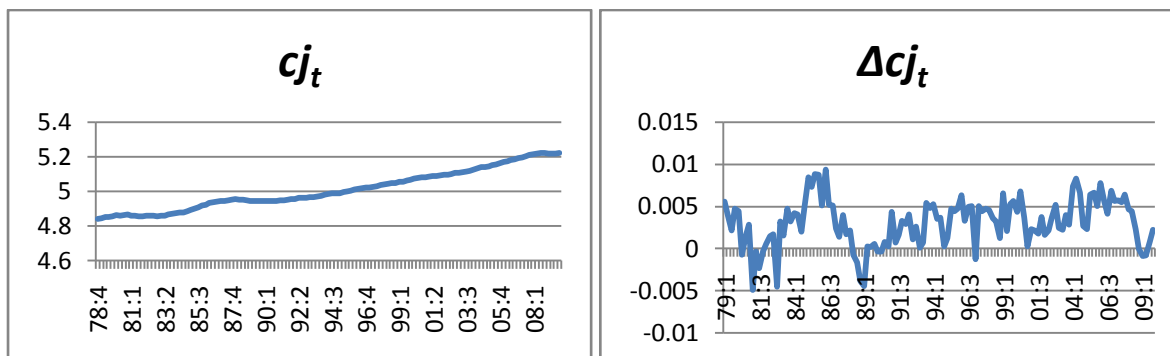
$$(7.2) \quad CP_{IV} = CPEB - CP_{30} - CP_{40},$$

der konsum av ikke- varige goder  $CP_{IV}$  er lik husholdningens konsum utenom bolig- og helsekonsum  $CPEB$  ekskludert kjøp av egne kjøretøyer  $CP_{30}$  og kjøp av andre varige konsumgoder  $CP_{40}$ . Siden konsum varierer sesongmessig vil gjennomsnittlig konsum over de fire siste kvartaler benyttes:

$$(7.3) \quad c_{j,t} = \ln[1/4(CP_{iv,t} CP_{iv,t-1} CP_{iv,t-2} CP_{iv,t-3})].$$

Data på  $CP_{IV}$  er tilgjengelig fra 1978:1 – 2009:4.

**Figur 5:** Justert konsum 1978:4 – 2009:4, og konsum på endringsform.



Kilde: SSB.

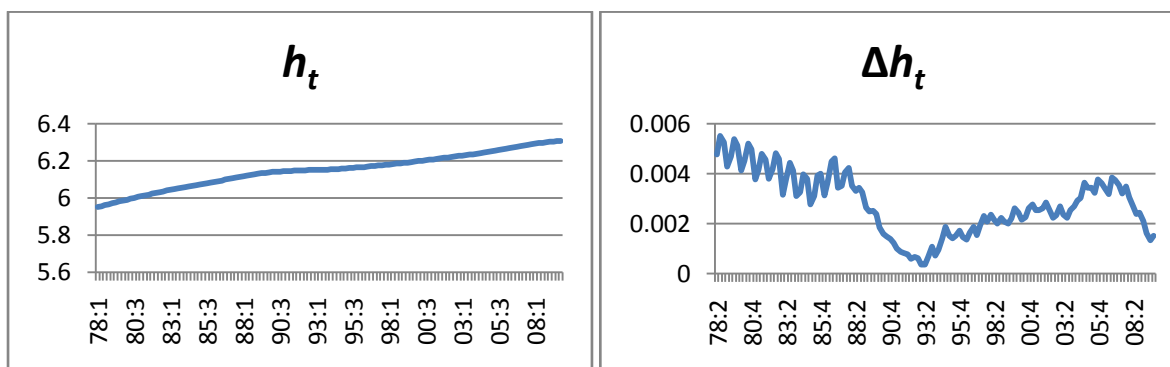
### *Boligbeholdning.*

Boligbeholdning, eller boligkapital, måles som produktet av en kvantitetsindeks og boligpriser, slik at  $K_{83}$  utgjør boligkapital målt i faste priser.

$$(7.4) \quad h_t = \ln(K_{83,t})$$

Data på  $K_{83}$  er tilgjengelig fra 1978:1 – 2009:4.

**Figur 6:** Boligmassen 1978:1 – 2009:4, og boligmassen på endringsform.



Kilde: SSB.

### *Brukerkostnader.*

I denne oppgaven anvendes brukerkostnader som stemmer overens med teorien i Poterba (1984), selv om det er noe forskjell mellom (7.5) og (7.8), som brukes i modelleringen i denne oppgaven, og (3.1). Mens det etter (3.1) i kapittel 3 skal brukes tall på marginal inntektsskatt og eiendomsskatt i utregning er det vanligere å bruke gjennomsnittlig marginalsatt på

kapitalinntekter<sup>7</sup>. Grunnen til dette er mangel på data for de to andre skattesatsene. Brukerkostnaden vil først regnes ut med flytende rente og bakoverskuende forventninger, basert på boligpriser, deretter med 10- års obligasjoner som proxy for fastrente og bakoverskuende forventninger basert på vekst i BNP. I utregningene benyttes data på styringsrente, obligasjoner med ti års løpetid, gjennomsnittlig marginalsatt på kapitalinntekter, KPI, bakoverskuende boligprisforventninger og bakoverskuende forventninger til vekst i BNP:

$$(7.5) \quad ufl_t = [RNOKj_{pa,t}(1 - \tau_t^K) - ir_t + \delta - E(\pi_t^B)],$$

der  $RNOKj_{pa,t}$  er gjennomsnittet av siste fire kvartalers årlige prosentvise styringsrente,  $\tau_t^K$  er gjennomsnittlig marginalsatt på kapitalinntekter,  $ir_t$  er årlig prosentvis inflasjonsrate:

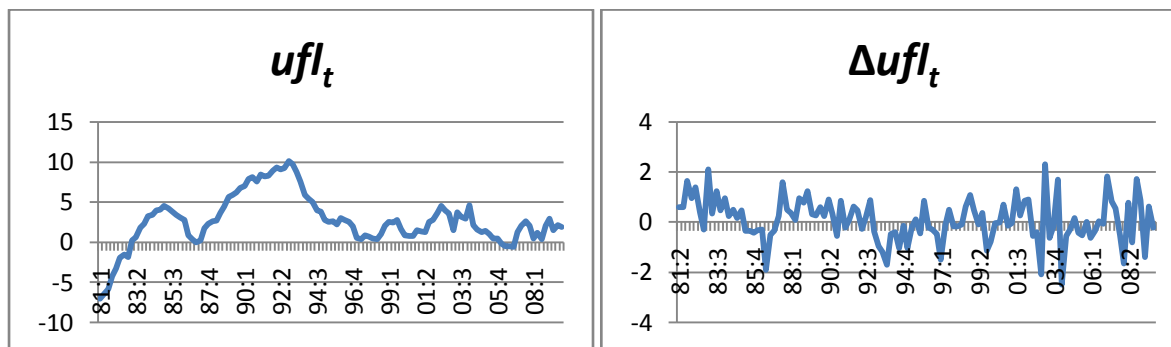
$$(7.6) \quad ir_t = [(KPI_t - KPI_{t-1})/KPI_{t-1}]400,$$

$\delta$  er kapitalslit og  $E(\pi_t^B)$  er forventet avkastning på boligkjøp der 12- kvartalers gjennomsnittlig prisvekst brukes som proxy:

$$(7.7) \quad E(\pi_t^B) = (\Delta p_t + \dots + \Delta p_{t-11})/12.$$

Det er verdt å merke seg at dette representerer bakoverskuende forventninger og blir en kilde til intertemporal prisavhengighet som nevnt i kapittel 4.

**Figur 7:** Vanlige brukerkostnader 1981:1 – 2009:4, og brukerkostnader på endringsform. NB: Ikke logaritmisk form!



**Kilde:** SSB.

<sup>7</sup><http://www.norges-bank.no/upload/import/front/rapport/no/fs/2006-01/2006-01.pdf> og

<http://www.ssb.no/emner/09/90/sos111/sos111.pdf>



Siden det er 12- kvartalers gjennomsnitt av boligprisvekst som benyttes som proxy for forventet avkastning er brukerkostnad tilgjengelig fra 1981:1 – 2009:4.

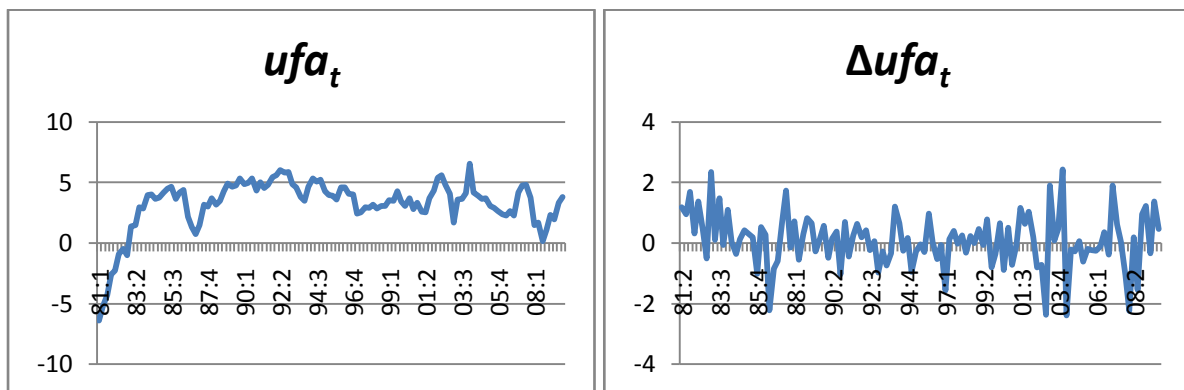
Om en vil benytte fastrente og BNP- vekst blir brukerkostnaden:

$$(7.8) \quad ufa_t = [FAST_t(1 - \tau_t^K) - ir_t + \delta - E(\pi_t^Q)],$$

der  $FAST_t$  er renten på obligasjoner med ti års løpetid og  $E(\pi_t^Q)$  er forventningen til BNP-vekst:

$$(7.9) \quad E(\pi_t^Q) = (\Delta q_t + \dots + \Delta q_{t-11})/12.$$

**Figur 8:** Alternative brukerkostnader 1981:1 – 2009:4, og alternative brukerkostnader på endringsform. NB: Ikke logaritmisk form!



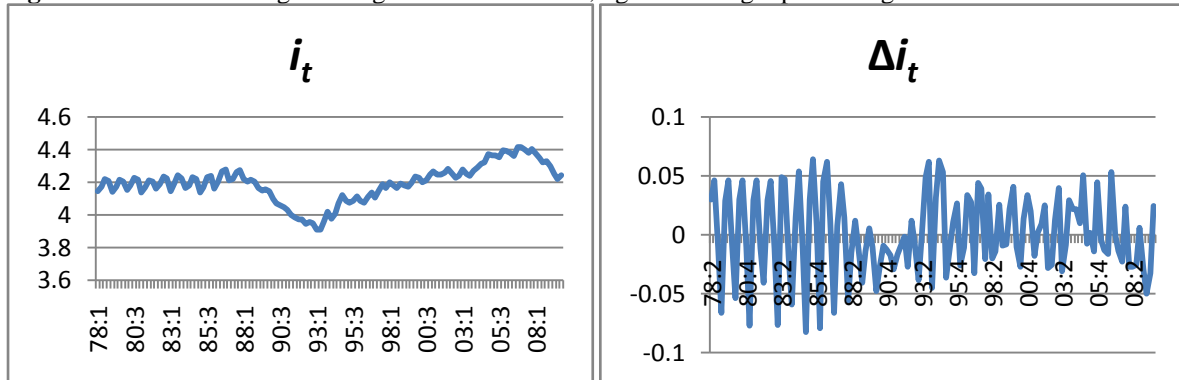
**Kilde:** Norges Bank, SSB.

### *Boliginvesteringer.*

$JKS_{83}$  fra MODAG representerer bruttoinvesteringer i boligmarkedet. Data er tilgjengelig 1978:1 – 2009:4.

$$(7.10) \quad i_t = \ln(JKS_{83,t})$$

**Figur 9:** Bruttoinvesteringer i boliger 1978:1 – 2009:4, og investeringer på endringsform.



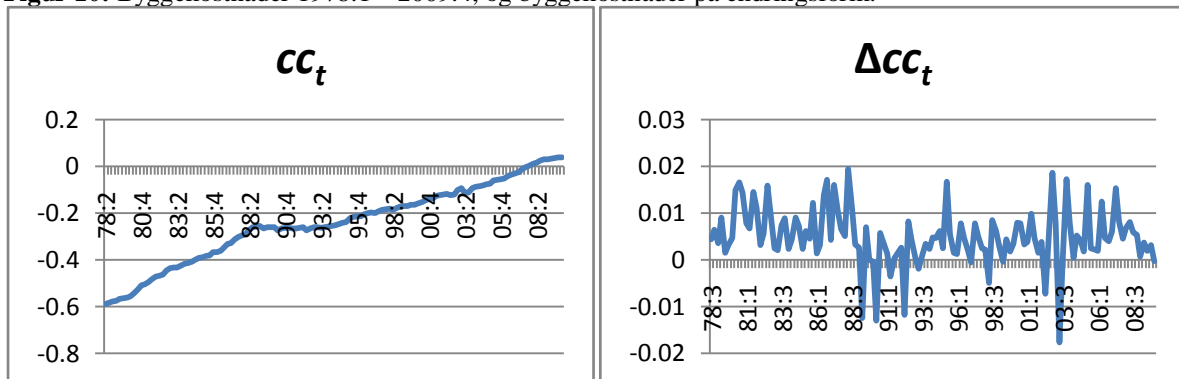
**Kilde:** SSB.

### Byggekostnader.

I modellen vil en indeks for pris på nye boliger eksklusive tomt  $PJKS_{83}$  fungere som en proxy for den teoretiske variabelen byggekostnader. Dette er inspirert av MODAG. Data på  $PJKS_{83}$  er tilgjengelig 1978:1 – 2009:4.

$$(7.11) \quad cc_t = \ln(PJKS_{83,t})$$

**Figur 10:** Byggekostnader 1978:1 – 2009:4, og byggekostnader på endringsform.



**Kilde:** SSB.

## 7.2 Dickey- Fuller testene

For å avgjøre om variablene ovenfor er stasjonære eller ikke brukes det her vanlig prosedyre med Dickey- Fuller tester. For å nøytralisere mulig autokorrelasjon bestemmes det hvor mange forsinkede verdier som skal inkluderes ved hjelp av tidligere nevnte seleksjonskriterier, før rett type Dickey- Fuller test velges og gjennomføres.

**Tabell 4:** Resultatene fra Dickey- Fuller testene.

Variabel	Lags	Modell	Test- verdi	Kritisk- verdi	Signifikans
$p_t$	6	Trend	-1.352	-3.147	10 %
$c_t$	3	Trend	-2.629	-3.147	10 %
$h_t$	6	Trend	-2.472	-3.147	10 %
$ufl_t^*$	6	Ingen	-2.699	-2.579	10 %
$ufa_t$	5	Ingen	-2.438	-2.579	10 %
$i_t$	5	Trend	-2.512	-3.147	10 %
$cc_t$	5	Trend	-3.141	-3.147	10 %
$\Delta p_t^*$	5	Ingen	-3.938	-3.503	1 %
$\Delta c_t^*$	2	Ingen	-3.183	-2.889	5 %
$\Delta h_t^*$	5	Drift	-2.213	-1.658	5 %
$\Delta ufa_t^*$	4	Ingen	-4.964	-3.503	1 %
$\Delta i_t^*$	4	Drift	-2.306	-1.658	5 %
$\Delta cc_t^*$	4	Ingen	-2.954	-2.889	5 %

**Kilde:** Egne utregninger, SSB, Norges Bank, variabler med markert med \* er stasjonære.

Som det kommer frem av tabell 4 er  $ufl_t$   $I(0)$ , og resten av variablene  $I(1)$ . Dermed kan alle variablene brukes i analysen, i tråd med kapittel 6.

### 7.3 Spesifisering

Helt sist før modellene kan estimeres, må det testes for hvor mange forsinkede verdier som skal inkluderes og hvor mange kointegrerende vektorer som eksisterer i systemet.

#### 7.3.1 Vanlige brukerkostnader

##### Seleksjonskriterier

Først testes det hvor mange lags som skal inkluderes for boligpris, boliginvesteringer, byggekostnader, vanlige brukerkostnader, konsum og boligbeholdning:

**Tabell 5:** Seleksjonskriterier for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufl_t$ ,  $c_t$  og  $h_t$ .

Selection-order criteria  
 Sample: **11 - 100**

Number of obs = **90**

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	<b>908.424</b>				<b>7.9e-17</b>	<b>-20.0539</b>	<b>-19.9867</b>	<b>-19.8872</b>
1	<b>1879.69</b>	<b>1942.5</b>	<b>36</b>	<b>0.000</b>	<b>7.4e-26</b>	<b>-40.8377</b>	<b>-40.3672</b>	<b>-39.6711</b>
2	<b>1974.92</b>	<b>190.45</b>	<b>36</b>	<b>0.000</b>	<b>2.0e-26</b>	<b>-42.1537</b>	<b>-41.2801*</b>	<b>-39.9872*</b>
3	<b>2012.17</b>	<b>74.493</b>	<b>36</b>	<b>0.000</b>	<b>2.0e-26</b>	<b>-42.1815</b>	<b>-40.9046</b>	<b>-39.015</b>
4	<b>2051.98</b>	<b>79.621</b>	<b>36</b>	<b>0.000</b>	<b>1.9e-26</b>	<b>-42.2661</b>	<b>-40.586</b>	<b>-38.0998</b>
5	<b>2099.28</b>	<b>94.616</b>	<b>36</b>	<b>0.000</b>	<b>1.6e-26</b>	<b>-42.5174</b>	<b>-40.4341</b>	<b>-37.3511</b>
6	<b>2149.97</b>	<b>101.37</b>	<b>36</b>	<b>0.000</b>	<b>1.4e-26</b>	<b>-42.8437</b>	<b>-40.3572</b>	<b>-36.6775</b>
7	<b>2207.42</b>	<b>114.9</b>	<b>36</b>	<b>0.000</b>	<b>1.0e-26</b>	<b>-43.3204</b>	<b>-40.4306</b>	<b>-36.1543</b>
8	<b>2263.75</b>	<b>112.67</b>	<b>36</b>	<b>0.000</b>	<b>8.7e-27</b>	<b>-43.7723</b>	<b>-40.4793</b>	<b>-35.6063</b>
9	<b>2321.52</b>	<b>115.53</b>	<b>36</b>	<b>0.000</b>	<b>8.0e-27*</b>	<b>-44.256</b>	<b>-40.5597</b>	<b>-35.09</b>
10	<b>2381.95</b>	<b>120.86*</b>	<b>36</b>	<b>0.000</b>	<b>8.2e-27</b>	<b>-44.7989*</b>	<b>-40.6994</b>	<b>-34.633</b>

Endogenous: `lnp lni lncc ufl lncj lnh`  
 Exogenous: `_cons`

**Kilde:** Egne utregninger, SSB.

Fra tabell 5 finner en at  $HQ(p) = SC(p) = 2$ , mens  $AIC(p) = 10$ . Det er ganske typisk for disse testene at  $HQ(p) = SC(p)$ , mens  $AIC(p)$  er større<sup>8</sup>. Begge ordnene vil testes.

### Systemets kointegrerende rang

$HQ(p) = SC(p) = 2$ : Test for å fastslå  $rk(\Pi)$ .

**Tabell 6:** Johansens Trace Statistic for kointegrasjon i systemet,  $HQ(p) = SC(p) = 2$ .

Johansen tests for cointegration						
Trend: none				Number of obs =		<b>98</b>
Sample: <b>3 - 100</b>				Lags =		<b>2</b>
maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value	
<b>0</b>	<b>36</b>	<b>2033.2533</b>	<b>.</b>	<b>116.7533</b>	<b>90.45</b>	
<b>1</b>	<b>47</b>	<b>2054.999</b>	<b>0.35840</b>	<b>73.2618</b>	<b>66.52</b>	
<b>2</b>	<b>56</b>	<b>2070.0206</b>	<b>0.26403</b>	<b>43.2185*</b>	<b>45.58</b>	
<b>3</b>	<b>63</b>	<b>2083.0439</b>	<b>0.23339</b>	<b>17.1719</b>	<b>29.75</b>	
<b>4</b>	<b>68</b>	<b>2088.0773</b>	<b>0.09762</b>	<b>7.1053</b>	<b>16.31</b>	
<b>5</b>	<b>71</b>	<b>2091.6237</b>	<b>0.06982</b>	<b>0.0124</b>	<b>6.51</b>	
<b>6</b>	<b>72</b>	<b>2091.6299</b>	<b>0.00013</b>			

**Kilde:** Egne utregninger, SSB.

<sup>8</sup> Lütkepohl og Krätzig (2004).

I tabell 6 ser en at  $H_0(2) : rk(\Pi) = 2$  ikke kan forkastes. Dermed kan man konkludere med at det finnes to kointegrerende vektorer for systemet når  $HQ(p) = SC(p) = 2$  benyttes. Det gjøres oppmerksom på at testen er gjennomført med spesifiseringen om at konstantledd og trend er lik null, men som det vises i vedlegg A.1, vil  $rk(\Pi) = 2$  være tilfellet også om konstant eller trend inkluderes.

$AIC(p) = 10$ : Test for å fastslå  $rk(\Pi)$ .

**Tabell 7:** Johansens Trace Statistic for kointegrasjon i systemet,  $AIC(p) = 10$ .

Johansen tests for cointegration					
Trend: none				Number of obs =	<b>90</b>
Sample: 11 - 100				Lags =	<b>10</b>
maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value
<b>0</b>	<b>324</b>	<b>2230.5549</b>	<b>.</b>	<b>236.7584</b>	<b>90.45</b>
<b>1</b>	<b>335</b>	<b>2273.9675</b>	<b>0.61891</b>	<b>149.9332</b>	<b>66.52</b>
<b>2</b>	<b>344</b>	<b>2310.0975</b>	<b>0.55197</b>	<b>77.6732</b>	<b>45.58</b>
<b>3</b>	<b>351</b>	<b>2328.0333</b>	<b>0.32872</b>	<b>41.8015</b>	<b>29.75</b>
<b>4</b>	<b>356</b>	<b>2343.051</b>	<b>0.28375</b>	<b>11.7662*</b>	<b>16.31</b>
<b>5</b>	<b>359</b>	<b>2348.8239</b>	<b>0.12040</b>	<b>0.2203</b>	<b>6.51</b>
<b>6</b>	<b>360</b>	<b>2348.9341</b>	<b>0.00244</b>		

**Kilde:** Egne utregninger, SSB.

I tabell 7 finner man at  $H_0(4) : rk(\Pi) = 4$  ikke kan forkastes, og impliserer at det eksisterer 4 kointegrerende vektorer for systemet når  $AIC(p) = 10$  anvendes. Som over er det testet med konstantledd og trend spesifisert lik null.  $rk(\Pi) = 5$  og 0 når henholdsvis konstant og trend inkluderes.

Når  $AIC(p) = 10$ , benyttes fører det til varierende resultater, som ikke stemmer overens med stock-flow teori. Derfor vil videre  $HQ(p) = SC(p) = 2$  og  $rk(\Pi) = 2$  benyttes i analysen med vanlig brukerkostnader. I den estimerte modellen i McCarthy og Peach (2002) er forøvrig  $p = 4$ . Dermed velges også den ordenen som er nærmest deres resultater.

### 7.3.2 Alternative brukerkostnader

#### Seleksjonskriterier

Her testes det hvor mange ordener en modell med boligpris, boliginvesteringer, byggekostnader, alternative brukerkostnader, konsum og boligbeholdning har:

**Tabell 8:** Seleksjonskriterier for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufa_t$ ,  $c_j$  og  $h_t$ .

```

Selection-order criteria
Sample: 11 - 100
Number of obs = 90

```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	927.444				5.2e-17	-20.4765	-20.4093	-20.3099
1	1868.55	1882.2	36	0.000	9.5e-26	-40.59	-40.1196	-39.4234
2	1967.28	197.45	36	0.000	2.4e-26	-41.9839	-41.1103*	-39.8174*
3	2012.05	89.536	36	0.000	2.0e-26	-42.1788	-40.9019	-39.0124
4	2047.22	70.345	36	0.001	2.2e-26	-42.1604	-40.4803	-37.994
5	2111.9	129.37	36	0.000	1.2e-26	-42.7978	-40.7144	-37.6315
6	2165.62	107.44	36	0.000	9.6e-27	-43.1916	-40.705	-37.0254
7	2210.82	90.407	36	0.000	9.5e-27	-43.3961	-40.5063	-36.23
8	2264.74	107.83	36	0.000	8.5e-27	-43.7942	-40.5012	-35.6282
9	2340.11	150.73	36	0.000	5.3e-27*	-44.669	-40.9728	-35.503
10	2398.34	116.46*	36	0.000	5.7e-27	-45.163*	-41.0635	-34.9971

Endogenous: lnp lni lnci ufa2 lncj lnh  
Exogenous: \_cons

**Kilde:** Egne utregninger, SSB, Norges Bank.

Som en ser av tabell over er  $HQ(p) = SC(p) = 2$ , mens  $AIC(p) = 10$ . Igjen testes begge de estimerte ordnene.

### Systemets kointegrerende rang

$HQ(p) = SC(p) = 2$ : Test for å fastslå  $rk(\Pi)$ .

**Tabell 9:** Johansens Trace Statistic for kointegrasjon i systemet,  $HQ(p) = SC(p) = 2$ .

```

Johansen tests for cointegration
Trend: constant
Sample: 3 - 100
Number of obs = 98
Lags = 2

```

maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value
0	42	2030.5072	.	141.1766	103.18
1	53	2058.6449	0.43687	84.9011	76.07
2	62	2073.8868	0.26733	54.4173*	54.46
3	69	2086.5822	0.22825	29.0265	35.65
4	74	2096.1181	0.17684	9.9547	20.04
5	77	2100.3312	0.08239	1.5285	6.65
6	78	2101.0955	0.01548		

**Kilde:** Egne utregninger, SSB, Norges Bank.

Igjen er  $H_0(2) : rk(\Pi) = 2$  første nullhypotese som forkastes, slik at systemets kointegrerende rang er 2. Det gjøres her oppmerksom på at dette kun vil gjelde når et konstantledd inkluderes i testen. For de to andre tilfellene vil systemets rang være lik 3.

$AIC(p) = 10$ : Test for å fastslå  $rk(\Pi)$ .

**Tabell 10:** Johansens Trace Statistic for kointegrasjon i systemet,  $AIC(p) = 10$ .

Johansen tests for cointegration					
Trend: constant				Number of obs =	<b>90</b>
Sample: 11 - 100				Lags =	<b>10</b>
maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value
<b>0</b>	<b>330</b>	<b>2245.3887</b>	<b>.</b>	<b>305.8929</b>	<b>103.18</b>
<b>1</b>	<b>341</b>	<b>2300.3068</b>	<b>0.70489</b>	<b>196.0568</b>	<b>76.07</b>
<b>2</b>	<b>350</b>	<b>2341.7786</b>	<b>0.60212</b>	<b>113.1132</b>	<b>54.46</b>
<b>3</b>	<b>357</b>	<b>2366.1031</b>	<b>0.41757</b>	<b>64.4642</b>	<b>35.65</b>
<b>4</b>	<b>362</b>	<b>2386.4261</b>	<b>0.36341</b>	<b>23.8182</b>	<b>20.04</b>
<b>5</b>	<b>365</b>	<b>2396.1188</b>	<b>0.19378</b>	<b>4.4328*</b>	<b>6.65</b>
<b>6</b>	<b>366</b>	<b>2398.3352</b>	<b>0.04806</b>		

**Kilde:** Egne utregninger, SSB, Norges Bank.

$H_0(5) : rk(\Pi) = 5$  er første nullhypotese som forkastes når  $AIC(p) = 10$  og et konstantledd inkluderes. Om det ikke inkluderes noe konstantledd eller trend blir  $rk(\Pi) = 4$ . Om trend inkluderes blir  $rk(\Pi) = 0$ .

Seleksjonskriteriene valgte de samme ordenene når man bruker alternative brukerkostnader i stedet. Dessverre ble Johansen Trace testene vanskelige å tolke for alternative brukerkostnader. Likevel er det vist ovenfor at  $rk(\Pi) = 2$  om det inkluderes et konstantledd og ordenen som spesifiseres av både Hannan-Quinn- og Schwarz-kriteriene brukes. Derfor vil modellen også estimeres for dette tilfellet under.

## 7.4 Estimering

Før matrisen  $\beta$ , bestående av de kointegrerende vektorene, estimeres må en som nevnt i kapittel 6.4 og 6.5 legge identifiserende restriksjoner på  $\beta$ . Uten restriksjoner ser modellen slik ut:

$$(7.12) \quad ce_{1t} = \beta_{11}p_t^* + \beta_{12}c_t + \beta_{13}h_t + \beta_{14}u_t + \beta_{15}i_t + \beta_{16}cc_t,$$

$$(7.13) \quad ce_{2t} = \beta_{21}p_t^* + \beta_{22}c_t + \beta_{23}h_t + \beta_{24}u_t + \beta_{25}i_t + \beta_{26}cc_t,$$

der  $ce_{jt}$  er en av de  $j$  kointegrerende vektorene.  $j = 1, 2$ .

For å komme frem til modellen (4.5) og (4.6) vil syv identifiserende restriksjoner inkluderes:

1. Koeffisientene til investeringer og byggekostnader skal være lik null i (7.12).
2. Koeffisientene til konsum og brukerkostnader skal være lik null i (7.13).
3. Summen av koeffisientene til investering og boligmassen skal være null i (7.13).
4. Det normaliseres på priser i begge ligningene.

**Tabell 11:** Identifiserende restriksjoner på den kointegrerende matrisen.

Vektor	Variabel	Restriksjon
1	$p_t^*$	$\beta_{11} = 1$
1	$i_t$	$\beta_{15} = 0$
1	$cc_t$	$\beta_{16} = 0$
2	$p_t^*$	$\beta_{21} = 1$
2	$c_t$	$\beta_{22} = 0$
2	$u_t$	$\beta_{24} = 0$
2	$h_t$ og $i_t$	$\beta_{23} + \beta_{25} = 0$

Med disse restriksjonene kan (7.12) og (7.13) skrives som:

$$(7.14) \quad ce_{1t} = p_t^* + \beta_{12}c_t + \beta_{13}h_t + \beta_{14}u_t,$$

$$(7.15) \quad ce_{2t} = p_t^* + \beta_{23}(i_t - h_t) + \beta_{26}cc_t.$$

(7.14) og (7.15) estimeres i STATA med Johansen maximum likelihood estimation, med alle beskrankningene som spesifiseres i tabell 11, først for standard, deretter alternative brukerkostnader i perioden 1985:1 – 2009:4<sup>9</sup>:

### Standard brukerkostnader:

<sup>9</sup> I vedlegg A.2 fremstilles estimering med vanlige brukerkostnader over perioden 1981:1 – 2009:4, og ellers samme betingelser som for (7.16) og (7.17). Modellen blir svært lik, med unntak av endret fortegn for brukerkostnader. Derfor anvendes i stedet modellen for 1985:1 – 2009:4, som fortsatt omfatter boligboblen på slutten av 80- tallet. Se videre diskusjon i A.2.



$$(7.16) \quad ce_{1t} = p_t^* - 4,16c_t + 3,45 h_t + 0,001ufl_t,$$

$$(7.17) \quad ce_{2t} = p_t^* - 0,08(i_t - h_t) - 1,67cc_t,$$

der (7.16) og (7.17) er systemets to kointegrerende vektorer.

**Tabell 12:** Estimer for de to kointegrerende vektorene, standard brukerkostnader.

Variabel	Etterspørsel (7.18)		Tilbud (7.19)	
	Koeffisient	Standardfeil	Koeffisient	Standardfeil
Konsum	4.16	0.169	—	—
Boligbeholdning	-3.45	0.139	0.08	0.008
Brukerkostnad	-0.001	0.004	—	—
Investeringer	—	—	0.08	0.008
Byggekostnader	—	—	1.67	0.047

Test på de overidentifiserende restriksjonene:  
 $\chi^2(3) = 20.84$

**Kilde:** Egne utregninger, SSB. Se vedlegg A.3 for flere detaljer.

Den endelige modellen blir:

$$(7.18) \quad p_t^{d*} = 4,16c_t - 3,45h_t - 0,001ufl_t,$$

$$(7.19) \quad p_t^{s*} = 0,08(i_t - h_t) + 1,67cc_t.$$

### Alternative brukerkostnader:

$$(7.20) \quad ce_{1t} = p_t^* - 1,33c_t - 0,95h_t - 0,003ufa_t + 13,09,$$

$$(7.21) \quad ce_{2t} = p_t^* + 0,68(i_t - h_t) - 1,73cc_t + 1,74,$$

der (7.20) og (7.21) er systemets to kointegrerende vektorer.

**Tabell 13:** Estimer for de to kointegrerende vektorene, alternative brukerkostnader.

Variabel	Etterspørsel (7.22)		Tilbud (7.23)	
	Koeffisient	Standard feil	Koeffisient	Standard feil

Konsum	1.33	0.183	—	—
Boligbeholdning	0.95	0.277	0.68	0.142
Brukerkostnad	0.003	0.001	—	—
Investeringer	—	—	-0.68	0.142
Byggekostnader	—	—	1.74	0.051
Konstantledd	-13.09	—	-1.74	—

Test på de overidentifiserende restriksjonene:

$$\chi^2(3) = 3.421$$

**Kilde:** Egne utregninger, SSB, Norges Bank. Se vedlegg A.4 for flere detaljer.

Den endelige modellen blir:

$$(7.22) p_t^{d*} = 1,33c_t + 0,95h_t + 0,003ufa_t - 13,09,$$

$$(7.23) p_t^{s*} = -0,68(i_t - h_t) + 1,74cc_t - 1,74.$$

### **Kommentarer:**

(7.18) og (7.19) ser temmelig bra ut fra et teoretisk ståsted: Alle koeffisientene utenom brukerkostnader er signifikante, og alle har riktig fortegn. Faktisk er brukerkostnader også innsignifikante i McCarthy og Peach (2002). Utenom investeringsrate og brukerkostnader er koeffisientene også av rimelig størrelse i forhold til andre lignende modeller<sup>10</sup>. De identifiserende restriksjonene forkastes dessverre basert på vanlig  $\chi^2$ -fordeling siden  $\chi^2(3) = 20.84$ , men som det vises i Johansen (2000) vil ofte sanne hypoteser feilaktig forkastes ved få observasjoner (type 1 feil). Basert på det vil (7.18) og (7.19) brukes til analysen i neste del. (7.22) og (7.23) utgjør derimot ikke en like god modell. Det var ventet at koeffisienter skulle

<sup>10</sup> Se blant andre Jacobsen og Naug (2004) eller McCarthy og Peach (2002).

bli av omtrent samme størrelse, og naturligvis ha samme fortegn, som i (7.18) og (7.19). Det er dessverre ikke tilfelle og (7.22) og (7.23) kan ikke anvendes i analysen.

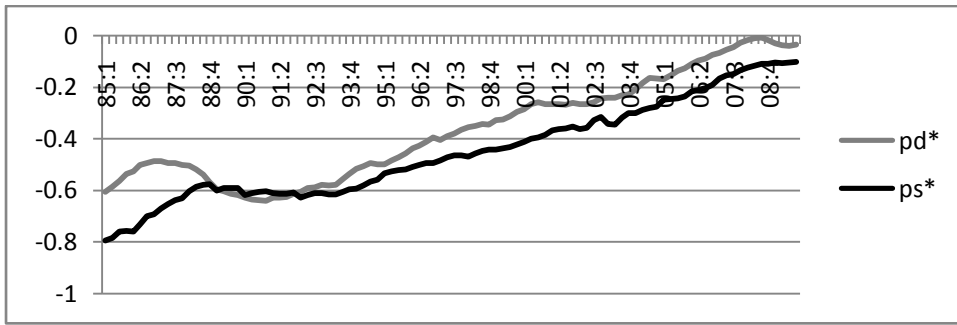
## 8 Resultatbasert vurdering av boligprisnivået

I kapittel 7 ble datasettet presentert og testet for å se om det kunne brukes til å konstruere en modell av boligmarkedet. Det viser seg at det eksisterer to kointegrerende vektorer for tidsseriene når en anvender vanlige brukerkostnader under ganske generelle betingelser, mens det er mer diffust når en anvender alternative brukerkostnader. Som vist sist i kapittel 7 kan dessverre bare modellen med vanlige brukerkostnader anvendes:

$$(7.18) \quad p_t^{d*} = 4,16c_t - 3,45h_t - 0,001ufl_t,$$

$$(7.19) \quad p_t^{s*} = 0,08(i_t - h_t) + 1,67cc_t.$$

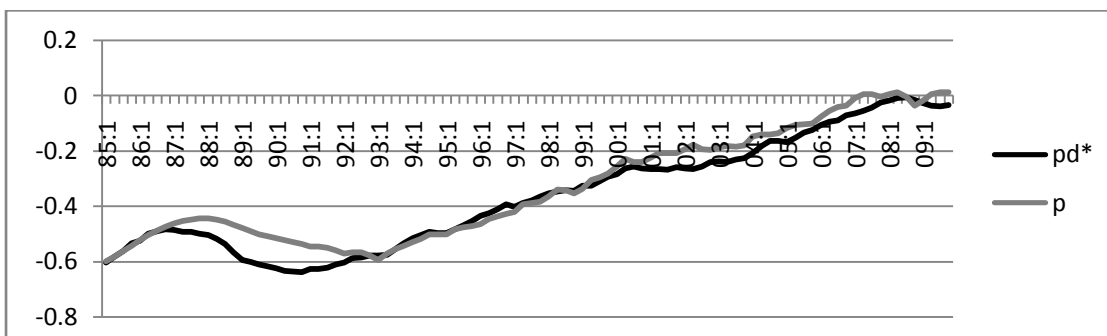
**Figur 11:** Langsiktig tilbud og etterspørsel 1985:1 – 2009:4.

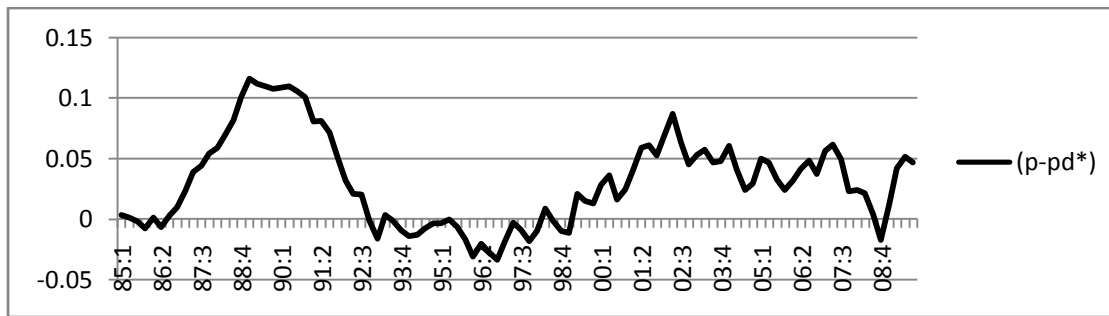


Kilde: Egne utregninger, SSB.

I slutten av kapittel 4 ble det argumentert for at det er (7.18) som avgjør boligprisnivået for en gitt boligbeholdning, slik at  $(p_t - p_t^{d*})$  kan benyttes til å vurdere observerte boligpriser opp mot det boligprisnivået fundamentale faktorer støtter opp under. I figur 12 er faktiske priser og estimerte priser tegnet opp, samt avviket mellom disse.

Figur 12: Utvikling i observerte og fundamentale priser, og  $(p_t - p_t^{d*})$  i perioden 1985:1 – 2009:4.





**Kilde:** Egne utregninger, SSB.

Fra figuren kan en se hvordan estimerte og faktiske boligpriser samvarierer over perioden 1985:1 – 2009:4. Modellen fanger opp de sterkt overvurderte boligprisene på slutten av 80-tallet, og indikerer at priser først stemte overens med fundamentale faktorer i løpet av 1992. Boliger ser ut til å være korrekt priset over størsteparten av 90- tallet, før en priskile igjen kan observeres fra begynnelsen av 2000- tallet omtrent til høsten 2008. I løpet av 2009 ble en ny priskile dannet. Siden brukerkostnader er innsignifikante i den endelige modellen, forklares det aller meste av fundamentale priser ut ifra befolkningens permanentinntekt og boligbeholdningen. Dermed er det ikke overraskende at modellen tenderer til å hevde boliger er høyt priset under og like etter høykonjunkturer, og mer fornuftig priset under lavkonjunkturer siden permanentinntekt i liten grad påvirkes av kortsiktige svingninger i inntekt. Denne tendensen stemmer også godt overens med treg prisjustering, som impliseres av nyere stock-flow teori.

Ifølge modellen har altså boligpriser vært tydelig overvurdert størsteparten av de siste ti årene. I andre kvartal 2002 oppnådde faktisk priskilen en verdi på 75 % av priskilens høyeste verdi under 80- tallets boligboble. Selv om dette åpenbart impliserer bobletendenser i Norges boligmarked på 2000- tallet er det verdt å peke på en vesentlig forskjell mellom de to periodene. På slutten av 80- tallet vedvarte priskilen nær toppnivå fra første kvartal 1989 helt til andre kvartal 1990, før den falt hurtig mot null fjerde kvartal 1992. Til sammenligning nådde priskilen et kortvarig toppnivå andre kvartal 2002 med en lavere toppverdi enn i 1989, og falt gradvis mot null tredje kvartal 2008. Riktignok med enkelte avvik.

Mens liberalisering av kredittmarkedet tilskrives mye av skylden for 80- tallets bobledannelser, er det trolig historisk lav rente som er årsaken på 2000- tallet. Selv om brukerkostnader ender opp innsignifikante er den negative sammenhengen mellom

pengepolitikk og boligpriser fastslått empirisk flere steder<sup>11</sup>. Ut ifra figur 12 ser boligpriser eksempelvis ut til å falle mot et fundamentalt støttet nivå i løpet av lavkonjunkturen rundt 2003, men får ny oppsving i første kvartal 2004 som tydelig sammenfaller med kraftig reduserte brukerkostnader (se figur 7).

Mens en større boble sprakk og førte til kollaps i bolig og finansmarked på slutten av 80-tallet, tyder modellen i denne oppgaven på at en mindre boble ble dannet i begynnelsen av 2000-tallet, som luften gradvis har seget ut av frem mot høsten 2008. Ifølge modellen er priser igjen noe overvurderte ved utgangen av 2009, slik at det blir svært interessant å følge markedsutviklingen fremover.

## **9 Avslutning**

Denne oppgaven begynte med en problemstilling om norske boligpriser er overvurderte. Med overvurderte priser menes det at observerte priser er på et nivå som ikke kan forklares eller støttes opp av verdier på langsiktige fundamentale faktorer i markedet, og dermed ikke er opprettholdbare. I løpet av oppgaven har det blitt vist hvordan man med VAR- metode kan estimere den langsiktige delen av en stock-flow boligmarkedsmodell. Den estimerte modellen

---

<sup>11</sup> Se for eksempel Kahn (1989), Pozdena (1990) eller Bernanke og Gertler (1995).

kan igjen anvendes til å betrakte forholdet mellom faktiske og fundamentalt forankrede boligpriser.

I overensstemmelse med konsensus påviser den estimerte modellen tydelig en boligboble på slutten av 80- tallet til begynnelsen av 90- tallet. Videre impliserer modellen at boligpriser har vært overvurdert fra omkring årtusenskiftet, men ble korrigert ned til fundamentalnivå i løpet av 2008. På ny kan dannelse av en positiv priskile observeres i begynnelsen av 2009, dog med negativ vekst i fjerde kvartal 2009 og av langt mindre verdi enn for omtrent tyve år siden. Trolig skyldes denne siste priskilen en kombinasjon av historisk lav rente og stimuli fra de omfattende internasjonale hjelpepakkene, ment å begrense omfanget av den internasjonale finanskrisen.

I skrivende stund har Norges Bank nettopp justert opp styringsrenten, Finanstilsynet, med flere, har lenge advart mot boligbobler og sentralbanksjef Svein Gjedrem har flere ganger oppfordret regjeringen til å fase inn mindre gunstig boligbeskatning. Økt rente og gjennomføringen av sentralbanksjefens oppfordringer vil begge føre til økte brukerkostnader, og således lavere boligpriser. Alle tre punkter vil også trolig senke markedsaktørens forventninger til fremtidig prisvekst, selv om dette ikke inngår direkte i modellen slik prisforventning er modellert.

Innledningsvis ble det lagt noen nødvendige begrensninger på oppgaven. Disse begrensningene har vært nødvendig delvis på grunn av tilgjengelig tid, og delvis på grunn av tilgjengelige data. Det ville vært svært interessant å utvide modellen med flere relevante variabler, eller analysere forskjellige delmarkeder.

Innsignifikante brukerkostnader er bemerkelsesverdig i denne oppgaven og gitt mer tid ville det blitt eksperimentert med forskjellige metoder for å inkludere disse. Det ville også vært interessant å modellere forventninger på en annen måte, slik at de også inkorporerer fremtidsutsikter. Disse og andre eventuelle utvidelser må nok dessverre vente til eventuelt videre arbeid.

## **Litteraturliste**

Alberts W. W. (1962): *Residential Construction Cycles, and the Mortgage Market*, The Journal of Political Economy, Vol. 70, No. 3, s. 263-281.



Allen F., Gale D. (2000): *Bubbles and Crises*, The Economic Journal, Vol. 110 No. 460, s. 236-255.

Bernanke B. S., Gertler M. (1995): *The Credit Channel of Monetary Policy Transmission*, The Journal of Economic Perspectives, Vol 9., No. 4, s. 27-48.

Case K. E., Quigley J. M., Shiller R. J. (2001): *Comparing Wealth Effects: The Stock Market versus the Housing Market*, NBER Working Paper No. 8606.

Case K. E., Shiller R. J. (1989): *Efficiency of the Market for Single-Family Homes*, The American Economic Review, Vol. 79, No. 1, s. 125-137.

Case K. E., Shiller R. J. (2003): *Is There a Bubble in the Housing Market?*, Brooking Papers on Economic Activity, Vol. 34, issue 2003-2, s. 299-362

Chen O., (2009): *Kan ikke overse boligprisene*, Kommentar på e24.no, <http://e24.no/kommentar/spaltister/chen/article3203306.ece>  
[Lesedato: 01.09.2009]

Christensen, T. H. (2003): *Boligprisindeksen: Datagrunnlag og beregningsmetode*, Notater 2003/83, Statistisk Sentralbyrå.

Christiano L. J., Eichenbaum M., Evans C. (1996): *Policy Shocks: Evidence from the Flow of Funds*, The Review of Economics and Statistics, Vol. 78, No. 1, s. 16-34.

Davidson R, MacKinnon J. G. (1993): *Estimation and Inference in Econometrics*, Oxford University Press, New York.

DiPasquale D., Wheaton W. C. (1994): *Housing Market Dynamics and the Future of Housing Prices*, The Journal of Urban Economics Vol. 35, No. 1, s. 1-27.

Eitrheim Ø., Erlandsen S. K. (2003): *House price indices for Norway 1819-2003*, (In: *Historical Monetary Statistics for Norway 1819-2003*, Eds: Eitrheim, Ø, Klovland J. T. og Qvigstad J.F., Norges Bank Occasional Papers No. 35, Oslo 2004, chapter 9, s. 349-376.)

Engle R. F., Granger C. W. J. (1987): *Cointegration and Error Correction: Representation, Estimation and Testing*, Econometrica Vol.55, No. 2 (Mar., 1987) s. 251-276

Fair R. C. (1972): *Disequilibrium in Housing Models*, The Journal of Finance 27, No. 2, s. 207-21.

Favero C. A. (2001): *Applied Macroeconometrics*, Oxford University Press, New York.

Fredriksen, H. (2007): *En kritisk gjennomgang av Jacobsen og Naug sin modell for hva som driver boligprisene*, Masterutredning, Norges Handelshøyskole

Gordon M., Shapiro E. (1956): *Capital Equipment Analysis: The Required Rate of Profit*, Management Science, 3, s. 102-110.

Granger C. W. J. (1981): *Some Properties of Time Series Data and their use in Econometric Model Specification*, The Journal of Econometrics Vol. 16, No.1, s. 121-130.

Grytten O. H. (2004): *A Consumer Price Index for Norway 1516-2003*, (In: *Historical Monetary Statistics for Norway 1819-2003*, Eds: Eitrheim, Ø, Klovland J. T. og Qvigstad J.F., Norges Bank Occasional Papers No. 35, Oslo 2004, chapter 3, s. 47-98.)

Grytten O. H. (2009): *Boligkrakk og Finanskriser i historisk perspektiv*, Samfunnsøkonomen, nr. 4, 2009.

Hamilton J. D. (1994): *Time Series Analysis*, Princeton University Press, New Jersey.

Hendershott P., Slemrod J. (1983): *Taxes and the User Cost of Capital for Owner-Occupied Housing*, The Journal of the American Real Estate and Urban Economics Association, Vol. 10, No. 4, s. 375-393.

Hill R. C., Griffiths W. E., Lim G. C. (2008): *Principles of Econometrics*, John Wiley and Sons, Inc. Hoboken.

Himmelberg C., Mayer C. J., Sinai T. (2005): *Assessing high house prices: Bubbles, fundamentals, and misperceptions*, Staff Report No. 218, Federal Reserve Bank of New York.

Jacobsen D. H., Naug B. E. (2004): *Hva driver boligprisene?*, Penger og Kreditt nr. 4/2004.

Johansen S. (2000): *A Bartlett Correction Factor for Tests on the Cointegrating Relations*, Econometric Theory, Vol 16, No. 5, s. 740-778.

Kahn, G. A. (1989): *The Changing Interest Sensitivity of the U.S. Economy*, Federal Reserve Bank of Kansas City Economic Review, s. 13-34.

Kearl J. R. (1979): *Mortgage and Housing*, The Journal of Political Economy, Vol. 87, No. 5, s. 1115-1138.

Leamer E. E. (2002): *Bubble Trouble? Your House has a P/E-ratio too*, UCLA Anderson Forecast, Nation-1.1.

Lütkepohl H., et. al. (2004): *Applied Time Series Econometrics*, Red: Lütkepohl H., Krätzig M., Cambridge University Press, Cambridge.

Mankiw N. G., Weil (1989): *The Baby Boom, the Baby Bust and the Housing Market*, Regional Science and Urban Economics, Vol. 19, No. 2, s. 235-258.

Mayer, C. J., and Somerville C. T (2000): *Residential Construction: Using the Urban Growth Model to Estimate Housing Supply*, Journal of Urban Economics Vol. 48, No. 1, s. 85-109.

McCarthy J., Peach R. M. (2004): *Are Home Prices The Next "Bubble"?*, Federal Reserve Bank of New York, Economic Policy Review, Vol. 10, No. 3.

McCarthy J., Peach R. M. (2002): *Monetary Policy Transmission To Residential Investment*, Federal Reserve Bank of New York, Economic Policy Review, Vol. 8, No. 1.

Miller M. H., Modigliani M. (1961): *Dividend Policy, Growth, and the valuation of Shares*, The Journal Of Business, Vol. 34, No. 4, s. 411-433.

Norges Bank (2006): *Finansiell Stabilitet*, Norges Banks rapportserie, Nr.2 – 2006.

Norges Eiendomsmeglerforbund, Eiendomsmeglerforetakenes Forening, ECON, FINN.no:  
Eiendomsmeglerbransjens boligprisstatistikk, mars 2010. Internettadresse:

<http://www.nef.no/index.gan?id=1647&subid=0>

[Lesedato: 01.03.2010]

OECD (2006): (iii). *Recent House Price developments: The Role of Fundamentals*, OECD ECONOMIC OUTLOOK 2006.

Poterba, J. M. (1984): *Tax Subsidies to Owner-Occupied Housing: An Asset-Market Approach*, The Quarterly Journal of Economics Vol. 98, No. 4, s. 731-752.

Poterba, J. M. (1992): *Taxation and Housing: Old Questions, New Answers*, The American Economic Review, Vol. 82, No. 2 (May, 1992) s. 237-242.

Pozdena, R. J. (1990): *Do Interest Rates Still Affect Housing?*, Federal Reserve Bank of San Francisco Economic Review, s. 3-14.

Reite, E. (2008): *Er vi i en boligboble?*, Masterutredning, Norges Handelshøyskole

Romer D. (2006): *Advanced Macroeconomics*, McGraw-Hill/Irwin, New York.

Rosen K. T., Smith L. B. (1983): *The Price-Adjustment Process for Rental Housing and the Natural Vacancy Rate*, The American Economic Review Vol. 73, No. 4, s. 779-786.

Sheffrin S. (1983): *Rational Expectations*, Cambridge University Press, Cambridge.

Sjølie Ø. (2009): *Gjedrem må stanse boligboblen*, Kommentar på e24.no,

<http://e24.no/kommentar/e24-kommentar/sjolie/article3203194.ece>

[Lesedato: 01.09.2009]

Skinner J. S. (1996): *Is Housing Wealth a Sideshow?* i *Advances in the Economics of Aging*, Red: Wise D. A., University of Chicago Press. Chicago s. 241-68.

Smith L. B. (1969): *A Model of the Canadian Housing and Mortgage Markets*, The Journal of Political Economy, Vol. 77, No.5, s. 795-816.

SSB (1999): *Kristianiakrakket 1899, 1999*.

Internettadresse: <http://www.ssb.no/emner/00/00/sa39/del8.pdf>

[Lesedato: 18.02.2010]

SSB (2008): *MODAG- En Makroøkonomisk Modell for Norsk Økonomi*, August 2008. Eds: Boug P., Yngvar D.

Internettadresse: <http://www.ssb.no/emner/09/90/sos111/sos111.pdf>

[Lesedato: 05.03.2010]

Stiglitz, J. E. (1990): *Symposium on Bubbles*, The Journal of Economic Perspectives, Vol. 4, No. 2, s: 13-18.

Stock J. H., Watson M. W. (2001): *Vector Autoregressions*, The Journal of Economic Perspectives, Vol. 15, No. 4, s. 101-115.

Sydsæter K., Hammond P. (2008): *Essential Mathematics for Economic Analysis*, Pearson Education Limited, Essex.

Sydsæter K., Hammond P., Seierstad A., Strøm A. (2008): *Further Mathematics for Economic Analysis*, Pearson Education Limited, Essex.

Topel R., Rosen S. (1988): *Housing Investment in the United States*, The Journal of Political Economy, Vol. 96, No. 4, s. 718-740.

## **Vedlegg:**

### **A.1**

**Tabell 14:** Johansens Trace Statistic for kointegrasjon i systemet,  $HQ(p) = SC(p) = 2$ , med hhv. konstant og trend variabel.

Johansen tests for cointegration						Number of obs =	98
Trend: constant						Lags =	2
Sample: 3 - 100							
maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value		
0	42	2037.4791	.	148.0866	103.18		
1	53	2065.8268	0.43927	91.3914	76.07		
2	62	2087.3674	0.35571	48.3101*	54.46		
3	69	2100.3185	0.23226	22.4079	35.65		
4	74	2107.0505	0.12837	8.9440	20.04		
5	77	2110.9094	0.07573	1.2260	6.65		
6	78	2111.5224	0.01243				

Johansen tests for cointegration						Number of obs =	98
Trend: trend						Lags =	2
Sample: 3 - 100							
maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value		
0	48	2039.1698	.	159.7303	114.36		
1	59	2067.7289	0.44169	102.6121	85.78		
2	68	2089.9478	0.36457	58.1744*	61.21		
3	75	2105.7962	0.27634	26.4776	40.49		
4	80	2112.4042	0.12616	13.2616	23.46		
5	83	2116.3585	0.07753	5.3530	6.40		
6	84	2119.035	0.05316				

Kilde: Egne utregninger, SSB.

## A.2

**Tabell 15a:** Seleksjonskriterier for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufl_t$ ,  $cj_t$  og  $h_t$ .



Selection-order criteria  
 Sample: **11 - 116** Number of obs = **106**

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	<b>999.338</b>				2.9e-16	-18.7422	-18.6811	-18.5915
1	<b>2167.36</b>	<b>2336.1</b>	<b>36</b>	<b>0.000</b>	1.5e-25	-40.1012	-39.6735	-39.0459
2	<b>2263.6</b>	<b>192.47</b>	<b>36</b>	<b>0.000</b>	5.0e-26	-41.2377	-40.4434*	-39.2778*
3	<b>2315.15</b>	<b>103.11</b>	<b>36</b>	<b>0.000</b>	3.8e-26	-41.5312	-40.3702	-38.6667
4	<b>2361.13</b>	<b>91.956</b>	<b>36</b>	<b>0.000</b>	3.2e-26	-41.7194	-40.1918	-37.9504
5	<b>2427.3</b>	<b>132.34</b>	<b>36</b>	<b>0.000</b>	1.9e-26	-42.2887	-40.3945	-37.6151
6	<b>2466.75</b>	<b>78.891</b>	<b>36</b>	<b>0.000</b>	2.0e-26	-42.3537	-40.0929	-36.7756
7	<b>2516.89</b>	<b>100.29</b>	<b>36</b>	<b>0.000</b>	1.7e-26	-42.6206	-39.9931	-36.1379
8	<b>2565.59</b>	<b>97.398</b>	<b>36</b>	<b>0.000</b>	1.5e-26	-42.8602	-39.8661	-35.4729
9	<b>2614.88</b>	<b>98.571</b>	<b>36</b>	<b>0.000</b>	1.5e-26*	-43.1109*	-39.7502	-34.819
10	<b>2648.3</b>	<b>66.848*</b>	<b>36</b>	<b>0.001</b>	2.1e-26	-43.0623	-39.3349	-33.8659

Endogenous: `lnp lni lnci ufl lncj lnk`  
 Exogenous: `_cons`

**Kilde:** Egne utregninger, SSB.

**Tabell 15b:** Johansens Trace Statistic for kointegrasjon i systemet,  $HQ(p) = SC(p) = 2$ , med hhv. konstant og trend variabel.

Johansen tests for cointegration

Trend: none Number of obs = **114**  
 Sample: **3 - 116** Lags = **2**

maximum rank	parms	LL	eigenvalue	trace statistic	1% critical value
<b>0</b>	<b>36</b>	<b>2337.5298</b>	.	<b>126.3022</b>	<b>90.45</b>
<b>1</b>	<b>47</b>	<b>2365.4895</b>	<b>0.38769</b>	<b>70.3826</b>	<b>66.52</b>
<b>2</b>	<b>56</b>	<b>2382.5055</b>	<b>0.25809</b>	<b>36.3507*</b>	<b>45.58</b>
<b>3</b>	<b>63</b>	<b>2394.8509</b>	<b>0.19474</b>	<b>11.6600</b>	<b>29.75</b>
<b>4</b>	<b>68</b>	<b>2398.9166</b>	<b>0.06884</b>	<b>3.5285</b>	<b>16.31</b>
<b>5</b>	<b>71</b>	<b>2400.6737</b>	<b>0.03036</b>	<b>0.0143</b>	<b>6.51</b>
<b>6</b>	<b>72</b>	<b>2400.6809</b>	<b>0.00013</b>		

**Kilde:** Egne utregninger, SSB.

**Tabell 15c:** Den estimerte modellen for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufl_t$ ,  $cj_t$  og  $h_t$  1981:1 – 2009:4.

```

Vector error-correction model

Sample: 3 - 116                                No. of obs   =    114
                                                AIC          = -40.72173
Log likelihood = 2374.139                       HQIC         = -40.20546
Det(Sigma_ml) = 3.28e-26                       SBIC         = -39.44964

Cointegrating equations

Equation      Parms    chi2      P>chi2
-----
_ce1          3      300.4636  0.0000
_ce2          2      2719.018  0.0000

Identification:  beta is overidentified

( 1)  [_ce1]lnp = 1
( 2)  [_ce1]lni = 0
( 3)  [_ce1]lncc = 0
( 4)  [_ce2]lnp = 1
( 5)  [_ce2]lncj = 0
( 6)  [_ce2]uf1 = 0
( 7)  [_ce2]lni + [_ce2]lnh = 0

+-----+-----+-----+-----+-----+-----+-----+
      beta |      Coef.  Std. Err.      z    P>|z|    [95% Conf. Interval]
+-----+-----+-----+-----+-----+-----+-----+
_ce1      |
      lnp      |           1           .           .           .           .           .
      lni      |      (omitted)
      lncc      |      (omitted)
      uf1      |  -.0503639   .0065421   -7.70   0.000   -0.0631861   -0.0375416
      lncj      |  -5.513203   .358896   -15.36   0.000   -6.216626   -4.80978
      lnh      |   4.587407   .294738   15.56   0.000    4.009731    5.165083
+-----+-----+-----+-----+-----+-----+
_ce2      |
      lnp      |           1           .           .           .           .           .
      lni      |  -.0988497   .0073639  -13.42   0.000   -0.1132826   -0.0844167
      lncc      |  -1.607058   .0408045  -39.38   0.000   -1.687033   -1.527082
      uf1      |      (omitted)
      lncj      |      (omitted)
      lnh      |   .0988497   .0073639   13.42   0.000    .0844167    .1132826
+-----+-----+-----+-----+-----+-----+
LR test of identifying restrictions:  chi2( 3) =  16.73  Prob > chi2 = 0.001

```

**Kilde:** Egne utregninger, SSB.

*Kommentarer:* Trolig blir fortegn feil på brukerkostnadskoeffisienten i denne modellen på grunn av at brukerkostnader stiger i perioden 1981:1 – 1984:2, en periode der boligpriser også stiger. Gitt mer tid ville en modell som tillot strukturelle skift eller lignende blitt estimert, som også ville vært egnet til å forklare boligpris utvikling i perioder med svært høy inflasjon.

### A.3

**Tabell 16:** Den estimerte modellen for  $p_t, i_t, cc_t, ufl_t, c_j$ , og  $h_t$  1985:1 – 2009:4.

```

Vector error-correction model

Sample: 3 - 100
Log likelihood = 2059.599
Det(Sigma_ml) = 2.24e-26
No. of obs = 98
AIC = -40.95099
HQIC = -40.38553
SBIC = -39.553

Cointegrating equations

Equation      Parns    chi2      P>chi2
-----
_ce1           3      2188.52   0.0000
_ce2           2      2972.315 0.0000

Identification: beta is overidentified

( 1)  [_ce1]lnp = 1
( 2)  [_ce1]lni = 0
( 3)  [_ce1]lncc = 0
( 4)  [_ce2]lnp = 1
( 5)  [_ce2]lncj = 0
( 6)  [_ce2]ufl = 0
( 7)  [_ce2]lni + [_ce2]lnh = 0

```

	beta	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
<b>_ce1</b>	lnp	1	.	.	.	.	.
	lni	(omitted)					
	lncc	(omitted)					
	ufl	.0009663	.0035375	0.27	0.785	-.0059671	.0078997
	lncj	-4.164303	.1692336	-24.61	0.000	-4.495995	-3.832612
	lnh	3.452129	.139376	24.77	0.000	3.178957	3.725301
<b>_ce2</b>	lnp	1	.	.	.	.	.
	lni	-.0800488	.007841	-10.21	0.000	-.0954168	-.0646807
	lncc	-1.670477	.047343	-35.28	0.000	-1.763268	-1.577687
	ufl	(omitted)					
	lncj	(omitted)					
	lnh	.0800488	.007841	10.21	0.000	.0646807	.0954168

LR test of identifying restrictions: chi2( 3) = 20.84 Prob > chi2 = 0.000

**Kilde:** Egne utregninger, SSB.

**Tabell 17:** Den estimerte modellen for  $p_t$ ,  $i_t$ ,  $cc_t$ ,  $ufa_t$ ,  $cj_t$  og  $h_t$  1985:1 – 2009:4.

```

Vector error-correction model

Sample: 3 - 100
Log likelihood = 2072.176
Det(Sigma_ml) = 1.73e-26
No. of obs = 98
AIC = -41.08523
HQIC = -40.45575
SBIC = -39.52897

Cointegrating equations

Equation      Parms    chi2      P>chi2
-----
_ce1           3      2319.8    0.0000
_ce2           2      1210.812  0.0000

Identification:  beta is overidentified

( 1)  [_ce1]lnp = 1
( 2)  [_ce1]lni = 0
( 3)  [_ce1]lncc = 0
( 4)  [_ce2]lnp = 1
( 5)  [_ce2]lncj = 0
( 6)  [_ce2]ufa2 = 0
( 7)  [_ce2]lni + [_ce2]lnh = 0

beta
-----
      Coef.  Std. Err.  z    P>|z|    [95% Conf. Interval]
-----
_ce1
      lnp      1          .          .          .          .          .
      lni  (omitted)
      lncc  (omitted)
      ufa2  -.003017   .0016452   -1.83   0.067    -.0062416   .0002076
      lncj  -1.329461   .1828829   -7.27   0.000    -1.687905   -.9710168
      lnh   -.9486825   .2771862   -3.42   0.001    -1.491957   -.4054076
      _cons  13.09371   .          .          .          .          .
-----
_ce2
      lnp      1          .          .          .          .          .
      lni  -.6790304   .1423933    4.77   0.000    -.3999447   .9581161
      lncc  -1.736292   .0506967   -34.25  0.000    -1.835655   -1.636928
      ufa2  (omitted)
      lncj  (omitted)
      lnh   -.6790304   .1423933    4.77   0.000    -.9581161   -.3999447
      _cons  1.744652   .          .          .          .          .
-----
LR test of identifying restrictions:  chi2( 3) = 3.421  Prob > chi2 = 0.331

```

**Kilde:** Egne utregninger, SSB, Norges Bank.