



En empirisk analyse av leiepriser for kontoreiendom i Oslo 1998-2016

- *Med fokus på prognostisering av fremtidige leiepriser*

Henrik Hvesser

Henrik Lange Høvik

Veileder: Gernot Peter Doppelhofer

Masteroppgave i finansiell økonomi

NORGES HANDELSHØYSKOLE

Dette selvstendige arbeidet er gjennomført som ledd i masterstudiet i økonomi- og administrasjon ved Norges Handelshøyskole og godkjent som sådan. Godkjenningen innebærer ikke at Høyskolen eller sensorer inntår for de metoder som er anvendt, resultater som er fremkommet eller konklusjoner som er trukket i arbeidet.

Sammendrag

Denne utredningen presenterer en empirisk studie av den gjennomsnittlige leieprisen for kontoreiendommer i Oslo-området i perioden 1998-2016. Utredningen undersøker hvilke determinanter som driver leieprisene og gir videre en sammenlignende prognoseanalyse ved hjelp av tre økonometriske modeller; ARIMA, ARIMAX og VECM. Modellene blir sammenlignet med en naiv random-walk modell, og samtlige av de valgte modellene presterer bedre enn den naive modellen i prognoseperioden 2015q1-2016q4. ARIMAX er den mest treffsikre modellen for toårsperioden, etterfulgt av VECM. I ettårsprognosen for 2016 er også ARIMAX modellen den mest treffsikre og vi observerer at den i større grad fanger opp vendepunkter i leieprisen. Prognosen tar dog hensyn til virkelige verdier av de eksogene forklaringsvariablene og vil trolig prestere dårligere når verdien av variablene ikke er kjent ex-ante. Tar vi denne begrensningen i betraktning er VECM den foretrukne prognosemodellen for kontorleieprisen i Oslo-området. Vi observerer at modellen fanger opp langsiktige trender i leieprisen, men har problemer med å fange opp større svingninger kvartal til kvartal. VECM egner seg derfor til å predikere over lengre perioder.

Videre identifiserer vi lange renter og arbeidsledighet som de viktigste determinantene for leiepris i ARIMAX modellen. Ved hjelp av kointegrasjonsanalysen i VECM fant vi en langsiktig positiv sammenheng mellom forventet nybygging og leiepris. Resultatet indikerer at det er lønnsomt for utbyggerne å bygge nye kontorlokaler grunnet høy etterspørsel. Vi identifiserer også en langsiktig sammenheng mellom lange renter og leiepris. Basert på empiri er dette et forventet resultat, men likevel oppsiktsvekkende da det ceteris paribus viser at sentralbanken er forsiktig med å føre en motsyklisk pengepolitikk. Videre identifiserer vi en signifikant negativ sammenheng mellom arbeidsledighet og leiepris i ARIMAX modellen. Dette underbygges av VECM sin kointegrasjonsanalyse som identifiserer en kortsiktig kausal sammenheng. Basert på empiri var resultatet forventet da høy arbeidsledighet gir en indikasjon på lavere etterspørsel.

Forord

Denne masterutredningen er skrevet som en del av masterstudiet i økonomi og administrasjon ved Norges Handelshøyskole. Oppgaven er skrevet innenfor hovedprofilen finansiell økonomi.

Temaet ble valgt på bakgrunn av stor interesse for næringseiendomsmarkedet, samt et ønske om å bidra med beslutningsrelevant innsikt for aktører i det norske leie- og transaksjonsmarkedet. I tillegg var det motiverende å kunne bidra med å utvide eksisterende empiri for det norske næringseiendomsmarkedet som per i dag er svært mangelfull. Vi har tilegnet oss mye ny kunnskap om viktige verdidrivere for leiepriser og ikke minst innen økonometri og modellering av komplekse problemer.

Ferdigstillelsen av denne utredningen ville ikke vært mulig uten hjelp og støtte fra flere personer som fortjener en stor takk. Vi vil først takke vår veileder, professor Gernot Peter Doppelhofer fra Norges Handelshøyskole, for nyttige innspill og konstruktiv kritikk som har ledet oss på rett spor og økt vårt læringsutbytte.

Videre vil vi takke Sigmund Aas i Arealstatistikk, samt Sindre Bråtebæk og Ragnar Eggen i Akershus Eiendom for å hjelpe oss med datamateriale og relevant diskusjon. Vi vil også takke Hans M. Hilde, Bjørn Erik Nilsen og Henrik Sømme i Pareto Securities for innspill til tema, samt god introduksjon til markedet for næringseiendom.

Bergen, juni 2017



Henrik Hveser



Henrik Lange Høvik

Innholdsfortegnelse

SAMMENDRAG	3
FORORD.....	5
INNHALDSFORTEGNELSE	6
1. INNLEDNING	10
1.1 AVGRENSNINGER.....	11
1.2 PROBLEMSTILLING	12
2. DET NORSKE MARKEDET FOR NÆRINGSEIENDOM.....	13
2.1 MARKEDSSTRUKTUR.....	14
2.2 TILBUD/ETTERSPØRSEL LEIEPRIS	15
3. EMPIRI.....	17
3.1 FAKTORER SOM PÅVIRKER LEIEPRISER	17
3.1.1 Arealledighet.....	17
3.1.2 Arbeidsledighet.....	18
3.1.3 Lange renter.....	18
3.1.4 BNP	18
3.1.5 Forventet nybygg.....	19
3.1.6 Initial direkteavkastning (Cap-rate).....	19
3.2 ARIMA - AUTO-REGRESSIVE INTEGRATED MOVING AVERAGE	20
3.3 ARIMAX - AUTO-REGRESSIVE INTEGRATED MOVING AVERAGE WITH EXPLANATORY VARIABLES.....	21
3.4 VECM - VECTOR ERROR-CORRECTION MODELL.....	21
3.5 EVALUERING AV PROGNOSEMODELLENE	22
4. DATA	24
4.1 LEIEPRISER (LP).....	24
4.2 MAKROØKONOMISKE VARIABLER	27
4.2.1 Bruttonasjonalprodukt (BNP)	27
4.2.2 Arbeidsledighet (ARBL)	28
4.2.3 Lange renter (RRTI).....	29
4.3 MARKEDSSPESIFIKKE VARIABLER	30
4.3.1 Arealledighet (ARL).....	30
4.3.2 Forventet nybygg de neste 4 årene (FNYB).....	32
5. METODE OG TEORI.....	34
5.1 STASJONÆRE PROSESSER.....	34

5.2	INFORMASJONSKRITERIER	36
5.3	HVIT STØY OG LJUNG-BOX TEST	37
5.4	TEST FOR STRUKTURELLE BRUDD – WALD TEST	37
5.5	AR - AUTO REGRESSIV MODELL	38
5.6	ARIMA - AUTO REGRESSIVE INTEGRATED MOVING AVERAGE	38
5.6.1	<i>Box-Jenkins metode for ARIMA modellering</i>	39
5.6.2	<i>Diagnostiske tester</i>	39
5.7	ARIMAX	40
5.8	VECTOR ERROR-CORRECTION MODELLERING (VECM)	40
5.8.1	<i>Testing og estimering av kointegrerte systemer ved bruk av Johansen metode</i>	41
5.9	PROGNOSTISERING	42
5.10	METODE FOR EVALUERING AV PROGNOSEMODELLENE	43
6.	ANALYSE OG RESULTATER	45
6.1	STASJONÆRITETSANALYSE	45
6.2	NAIV MODELL (1)	49
6.3	ARIMA MODELL (2)	51
6.3.1	<i>Modellspesifikasjon</i>	51
6.3.2	<i>In-sample analyse ARIMA</i>	53
6.3.3	<i>Out-of-sample resultater ARIMA</i>	56
6.4	ARIMAX MODELL (3)	58
6.4.1	<i>Modellspesifikasjon</i>	58
6.4.2	<i>In-sample analyse ARIMAX</i>	59
6.4.3	<i>Out-of-sample resultater ARIMAX</i>	63
6.5	VECM (4)	66
6.5.1	<i>Lag struktur</i>	66
6.5.2	<i>Kointegrasjonstesting: Johansen metode</i>	67
6.5.3	<i>Modellestimering</i>	68
6.5.4	<i>Diagnostisk testing</i>	71
6.5.5	<i>In-sample prognose VECM</i>	73
6.5.6	<i>Out-of-sample resultater VECM</i>	74
7.	SAMMENLIGNING AV PROGNOSE	75
8.	KONKLUSJON	78
9.	LITTERATURLISTE	80

APPENDIKS	86
A. INTERPOLERING AV DATA.....	86
B. GRAFISK IN-SAMPLE PROGNOSE 1998Q1 – 2014Q4.....	87
C. OUT-OF-SAMPLE RESULTATER 2016Q1 – 2016Q4	88
D. ARIMAX, GRANSKENDE ANALYSE	90
E. STASJONÆRITETSTEST FOR ARIMAX MODELL (3B).....	91
F. VECM KOEFFISIENTMATRISE.....	92

Tabelloversikt:

Tabell 1: Oppsummerende statistikk.....	33
Tabell 2: ADF test for stasjonærhet	48
Tabell 3: Prognosefeil - naiv modell (1)	50
Tabell 4: Autokorrelasjon og delvis autokorrelasjon for leiepris og vekst i leiepris	52
Tabell 5: Sammenligning av informasjonskriteriene ARIMA modell (2).....	53
Tabell 6: Empiriske resultater ARIMA modell (2).....	54
Tabell 7: Prognosefeil - modell (2a) og (2b).....	57
Tabell 8: Sammenligning av informasjonskriteriene modell (3)	60
Tabell 9: Empiriske resultater ARIMAX modell (3).	61
Tabell 10: Prognosefeil - ARIMAX modell (3).....	64
Tabell 11: Informasjonskriterier VECM (4)	66
Tabell 12: Johansen test for kointegrasjon VECM(4).....	67
Tabell 13: Empiriske resultater VECM (4).....	69
Tabell 14: Lineærhetstest VECM (4).....	70
Tabell 15: Unit Moduli i VECM (4)	71
Tabell 16: Lagrange Multiplier test for autokorrelasjon i residualene VECM (4)	72
Tabell 17: Jarque-Bera normalitetstest VECM (4)	72
Tabell 18: Prognosefeil - VECM (4).....	74
Tabell 19: Sammenligning av prognosetreffsikkerhet out-of-sample.....	76
Tabell 20: Dynamisk out-of-sample prognose 2015q2 - 2017q1.....	77
Tabell 21: Granskende analyse ARIMAX modell (3)	90
Tabell 22: Eigenverdi stabilitetstilstand modell (3b).....	91

Figuroversikt:

Figur 1: Utvikling av leiepris relativt til hovedindeksen på Oslo Børs fra 1998q1 – 2016q4.....	13
Figur 2: Tilbud og etterspørselsfunksjon for leiepris	16
Figur 3: Indeks reelle leiepriser for kontormarkedet i Oslo fra 1998q1 – 2016q4.....	25
Figur 4: Utvikling av leiepris og vekst i leiepris i perioden 1998q1 - 2014q4.....	26
Figur 5: Utvikling av BNP og vekst i BNP i perioden 1998q1 til 2014q4.....	28
Figur 6: Utvikling i arbeidsledighet i Oslo-området i perioden 1998q1 – 2014q4	29
Figur 7: Utvikling i lange renter og endringer i lange renter i perioden 1998q1 – 2014q4.....	30
Figur 8: Utvikling i arealledighet og vekst i arealledighet i perioden 1998q1 – 2014q4.....	31
Figur 9: Utviklingen i forventet nybygg og vekst i forventet nybygg 1998q1 – 2014q4.....	32
Figur 10: Dynamisk prognose out-of-sample - naiv modell (1).....	49
Figur 11: In-sample prognose ARIMA modell (2a).....	55
Figur 12: Dynamisk prognose out-of-sample ARIMA modell (2).....	56
Figur 13: In-sample prognose ARIMAX modell (3a).....	63
Figur 14: Dynamisk out-of-sample prognose modell (3)	64
Figur 15: Grafisk fremstilling av unit moduli VECM (4)	71
Figur 16: In-sample prognose VECM (4).....	73
Figur 17: Dynamisk out-of-sample prognose modell (4)	74
Figur 18: Grafisk fremstilling av forskjell mellom linær- og cubic spline interpolasjon.....	86
Figur 19: In-sample prognose ARIMA modell (2b).....	87
Figur 20: In-sample prognose ARIMAX modell (3b).....	87
Figur 21: Dynamisk out-of-sample prognose Naiv modell (1), 2016q1 – 2016q4.	88
Figur 22: Dynamisk out-of-sample prognose ARIMA modell (2), 2016q1 - 2016q4	88
Figur 23: Dynamisk out-of-sample prognose ARIMAX modell (3), 2016q1 - 2016q4.....	89
Figur 24: Dynamisk out-of-sample prognose VECM (4), 2016q1 - 2016q4.	89

1. Innledning

Eiendom utgjør 60% av verdens aktiva og er med det den mest signifikante investeringsklassen regnet i verdi (Savills, 2016). Heterogenitet, illikviditet, fast lokasjon og høye enhetsverdier er alle karakteristikker som tradisjonelt har skilt eiendom fra andre investeringsklasser. Siden 1980-tallet har investeringsanalyse relatert til eiendom utviklet seg mye, og investorer har begynt å sammenligne eiendom med alternative investeringsklasser i en porteføljesammenheng. I dag kan næringseiendom kjøpes direkte ved overdragelse eller indirekte gjennom aksjeandeler i eiendomsselskap og andre verdipapirer. Dette åpner for kortere investeringshorisonter for investorer, og har engasjert teoretikere og praktikere i retningen av mer finanstøretiske verdivurderingsmodeller. Som et resultat av dette har man fått en gradvis økning i empirisk forskning på eiendomsmarkedet, hvor majoriteten av forskningen er konsentrert rundt leiepris av næringseiendommer (Karakozova, 2004).

Næringseiendom har tradisjonelt hatt lave transaksjonsvolumer, få krav til å oppgi transaksjonspriser og ingen sentralisert innsamling av data. Derfor er mangel på tilgjengelig data et vanlig problem i den empiriske forskningen av eiendomsmarkedet. En videre konsekvens av dette er at de fleste studiene som er gjennomført er basert på data fra større byer i USA, som har den beste tilgangen til data. Unntaket er til en viss grad det britiske markedet (Tsolacos et al., 1998; Brooks og Tsolacos, 2001; McGough og Tsolacos, 2001; Tsolacos, 2002; Hendershott et al., 2002). Eksisterende studier har brukt ulike økonometriske teknikker med forskjellige makroøkonomiske og finansielle variabler. Hvilke determinanter som driver leieprisen varierer mellom de forskjellige markedene grunnet nasjon-spesifikke karakteristikker. I noen av studiene blir modellene også brukt til å utarbeide prognoser for leiepriser, men ytterst få av disse har blitt evaluert for deres yteevne, eller blitt sammenlignet med alternative prognoseteknikker. Dette gir grunnlag for ytterligere empirisk arbeid som gjennomgår den norske utviklingen i leiepriser. Formålet med utredningen er derfor å bidra med beslutningsrelevant innsikt for aktører i leie- og transaksjonsmarkedet, og utvide det empiriske grunnlaget for valg av prognosemodeller.

Denne utredningen gir en sammenlignende prognoseanalyse for gjennomsnittlig leiepris i kontormarkedet i Oslo. Kontoreiendom er det segmentet som det norske analysemiljøet retter størst oppmerksomhet mot, noe som bedrer tilgjengeligheten på gode historiske data. Vi vil studere både tilbud- og etterspørselssiden for kontormarkedet i Oslo og implementere funnene våre i tre statistiske modeller; ARIMA, ARIMAX og VECM. Videre vil vi evaluere prognosemodellene mot hverandre for å finne den modellen som gjengir den faktiske utviklingen med størst treffsikkerhet i perioden 2015q1 til 2016q4. Out-of-sample prognosene for de ulike

modellene blir sammenlignet på grunnlag av ulike prognosefeilparametere. Modellen med lavest prognosefeil betraktes som den beste modellen.

Våre hovedfunn er at ARIMAX modellen er den mest treffsikre modellen for toårsperioden, etterfulgt av VECM. Siden ARIMAX modellen tar hensyn til virkelige verdier av eksogene forklaringsvariabler, som ikke vil være kjente *ex-ante*, vil VECM være den foretrukne modellen for fremtidig prognostisering av leieprisene for Oslo-området. Vi observerer at modellen klarer å fange opp langsiktige trender i leieprisen, men har tydelige problemer med å fange opp de største svingningene.

Videre identifiserer vi i kointegrasjonsanalysen en langsiktig positiv sammenheng mellom forventet nybygging og leiepris. Vi finner også en negativ langsiktig sammenheng mellom lange renter og leiepris. Basert på empiri er dette et forventet resultat, men likevel oppsiktsvekkende da det *ceteris paribus* viser at sentralbanken er forsiktig med å føre en motsyklisk pengepolitikk. Vi fant også signifikant negativ sammenheng mellom arbeidsledighet og leiepris i ARIMAX modellen. Dette underbygges av VECM sin kointegrasjonsanalyse som identifiserer en kortsiktig kausal sammenheng.

Resten av denne utredningen er strukturert på følgende måte: Kapittel 2 gir en introduksjon til leiemarkedet for kontoreiendommer i Oslo. Kapittel 3 gir en innføring i empiri knyttet til determinanter for leieprisutvikling, samt en empirisk forankring av de statistiske modellene som blir benyttet i oppgaven. Kapittel 4 gir en oversikt over datagrunnlaget som utredningen baserer seg på. I kapittel 5 blir metode og teori for de ulike modellteknikkene gjennomgått. Kapittel 6 inkluderer utarbeidelse av modellene samt analyse av disse. Yteevnen til prognosemodellene blir sammenlignet i kapittel 7, før utredningen konkluderes i kapittel 8.

1.1 Avgrensninger

Grunnet tilgang på data har vi valgt å avgrense utredningen til kun å omhandle kontoreiendommer i Oslo-området. Dette markedet er det mest likvide i Norge, og har derfor også størst datamateriale tilgjengelig. Videre har vi valgt å avgrense utredningen til å prognostisere aggregerte gjennomsnittspriser for hele Oslo-området under ett, da datamaterialet ville vært for lite til å kunne gjøre korrekte økonometriske analyser på mindre geografiske områder. Ved kategorisering i mindre geografiske områder ville det i tillegg vært vanskeligere å opprettholde anonymiteten til de ulike kontraktene, særlig i de tidligste årene grunnet svært få observasjoner per område. I tråd med den Europeiske Sentralbanken velger vi å avgrense definisjonen ”næringseiendom” til

bygninger, inkludert tomter, som er eiet for det eksplisitte formålet å generere inntekt (ESRB, 2015).

I prognoseanalysen avgrensner vi oss til tre økonometriske modeller; ARIMA, ARIMAX og VECM. Modellene er i hovedsak valgt på grunnlag av tidligere forskning på emnet. Wilson et al. (2000) konkluderte med at ARIMA var egnet til å predikere hovedtrender i sin analyse av eiendomsindekser, men har problemer med å predikere vendepunkter. For i større grad kunne predikere endringer i leieprisene ønsket vi å prognostisere modeller som inkluderer flere forklaringsvariabler for å undersøke om det bedret resultatet. Karakozova (2004) konkluderte med at ARIMAX var den foretrukne modellen til å predikere avkastning på kontoreiendommer i Helsinki. Hun benyttet også regresjonsmodeller og error-correction model (ECM). Sims (1980) og McNees (1986) konkluderte med at VAR/VEC modeller ofte er bedre til prognostisering enn tradisjonelle strukturelle modeller når det er vanskelig å definere hvilke variabler som skal behandles eksogene. Vi velger derfor også å inkludere VECM¹ som implisitt har færre føringer for hvilke variabler som må behandles som eksogene/endogene. Regresjonsmodeller ble valgt bort med hensyn på oppgavens omfang.

Utredningen avgrensner i tillegg out-of-sample prognoseperioden til maksimalt 2 år, da vi anser dette som en relevant tidshorisont for de som skal reforhandle kontrakter, samtidig som usikkerheten øker i takt med lengden på prognoseperioden.

1.2 Problemstilling

Med bakgrunn i avgrensningen har vi følgende problemstilling for utredningen:

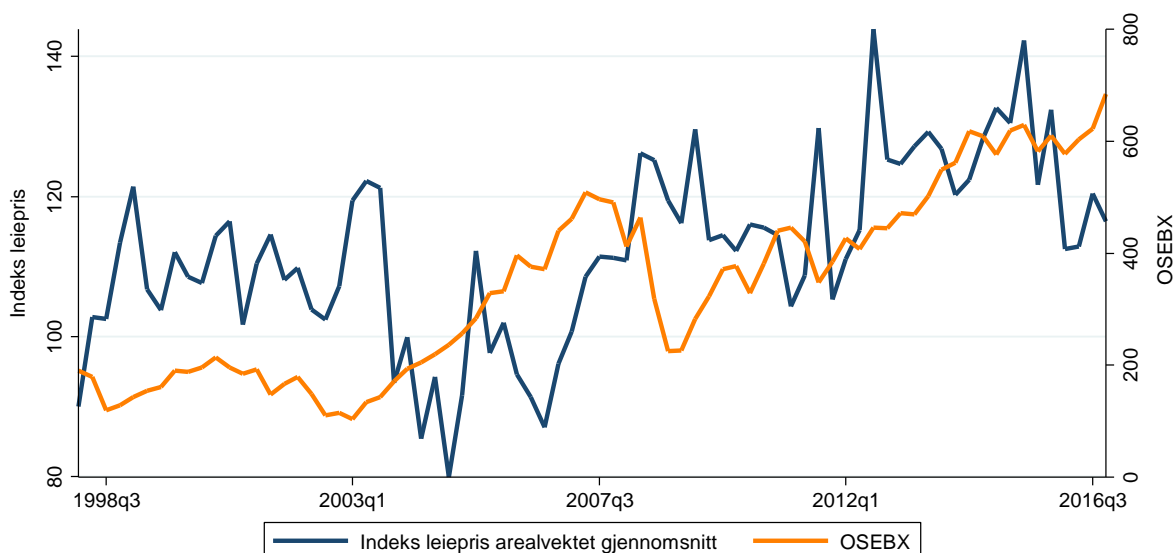
1. *Hvilke determinanter driver leieprisen for kontoreiendom i Oslo i perioden 1998q1 – 2016q4?*
2. *Hvordan kan determinantene anvendes til å utarbeide treffsikre prognosemodeller?*
3. *Hvilken prognosemodell gjengir den faktiske utviklingen med størst gjennomsnittlig treffsikkerhet?*

¹ VECM er en multivariat generalisering av ECM.

2. Det norske markedet for næringseiendom

Dette kapittelet gir først en introduksjon til historiske hendelser som har påvirket leieprisen for kontorbygg i Oslo de siste tiårene. Deretter gjennomgår vi hvordan det norske markedet er strukturert, før vi kort presenterer et eksempel på tilbud- og etterspørselsendringer for næringseiendom.

For å få et helhetlig bilde av årsaker til historiske hendelser som har påvirket leieprisen er det naturlig å nevne bankkrisen på slutten av 1980- og starten av 1990-tallet. Krisen kom etter en bølge med kredittliberalisering som resulterte i mangedobling av kredittvolumet i mange land. Det høye kredittvolumet førte igjen til bobletendenser i kreditt-, aksje- og boligmarkedet, før finanskrisen inntraff i perioden 1987-1993 (Grytten og Hunnes, 2016).



Figur 1: Utvikling av leiepris relativt til hovedindeksen på Oslo Børs fra 1998q1 – 2016q4

Venstre akse viser indekstert arealvektet reell gjennomsnittlig leiepris, høyre akse viser OSEBX indeks.

I løpet av 90-tallet forbedret leiemarkedet for kontoreiendom i Oslo seg gradvis. Arbeidsledigheten var avtagende, og nye toppnoteringer ble satt i leiemarkedet etter hvert som etterspørselen vokste fortere enn tilbudet. Særlig for de større selskapene ble situasjonen prekær, hvor gode lokaler på over $1000m^2$ var svært vanskelig å oppdrive. Dette førte til en viss divergens i prisforholdet mellom store og små kontorlokaler. Som en følge av «dot-com» krisen i perioden 2000 til 2002 raste børsene med nesten 50%, fulgt av en markant økning i arbeidsledighet og arealledighet. Som et resultat av dette ble leiemarkedet svært svakt frem til slutten av 2004 (Uniongruppen, 2002; 2004).

Med økt optimisme i markedene økte etterspørselen, og innen finanskrisen fikk grep i andre kvartal 2008 hadde leieprisene hentet inn nedgangen fra «dot-com» krisen og satt nye toppnoteringer. Som et resultat av finanskrisen falt dog leieprisen igjen gradvis etter hvert som eksisterende leiekontrakter utløp. Nedgangen fortsatte frem til midten av 2010, før høyere forventninger til fremtidig etterspørsel stabiliserte markedet. Som et resultat av dette økte leieprisen gradvis frem til rekordåret 2015, med unntak av mindre variasjoner. Også transaksjonsvolumet var rekordhøyt i 2015, med et samlet transaksjonsvolum på over 123 milliarder kroner (Brun, 2016). Som et resultat av oljeprisfallet sank etterspørselen etter kontoreiendom i 2015. Spesielt fra oljerelaterte foretak. Dette førte til økt arealledighet og redusert leiepris. I tråd med forventninger om at etterspørselsveksten ville være større enn tilbudssideveksten i årene fremover, styrket leieprisen seg noe i 2016 (Uniongruppen, 2016).

Vi observerer i Figur 1 at det ofte kan ta lang tid fra man har et fall på børsen til man får en korreksjon i leieprisen. Dette er sannsynligvis fordi store leietakere skriver kontrakter i lang tid før de flytter inn i lokalene, og at eksisterende leiekontrakter ikke blir reforhandlet umiddelbart etter økonomiske hendelser. Paradoksalt nok kan denne forsinkelsen gjøre at vi ser økende leiepriser mens børsen faller, slik vi for eksempel ser i slutten av 2002.

2.1 Markedsstruktur

I Norge er en stor andel av næringsbyggene eid av private selskaper. I følge tall fra Pangea Property Partners (Saltnes, 2017) har de private selskapene en eierandel på 76% av næringsbyggene, samtidig som andelen er henholdsvis 71%, 50% og 46% i Danmark, Finland og Sverige. Norge har historisk hatt en gunstig formuesbeskatning for eiendom kontra aksjer og andre investeringsklasser. I Sverige er det også et mer modent marked med flere og større institusjonelle aktører. Det er i tillegg en økende tendens til at flere internasjonale investorer ser til Norge for gode investeringer i næringsbygg, og i 2016 eide de omtrent 5% av næringsmassen (Saltnes, 2017).

Norske banker er direkte eksponert mot næringseiendom gjennom utlån til næringseiendomsforetak. I følge Norges Bank har disse eiendomsrelaterte foretakslåne vært den viktigste årsaken til banktap i de fleste økonomiske krisene de har undersøkt (Kragh-Sørensen og Solheim, 2014a). Bankene er også indirekte eksponert gjennom utlån til andre næringer, som gjerne inkluderer lån til bygg- og anlegg eller foretak i andre sektorer med pant i fast eiendom. Den direkte eksponeringen for bankene har siden 1990-tallet økt. På en annen side har den indirekte eksponeringen trolig blitt redusert grunnet at stadig flere foretak enten har solgt eller

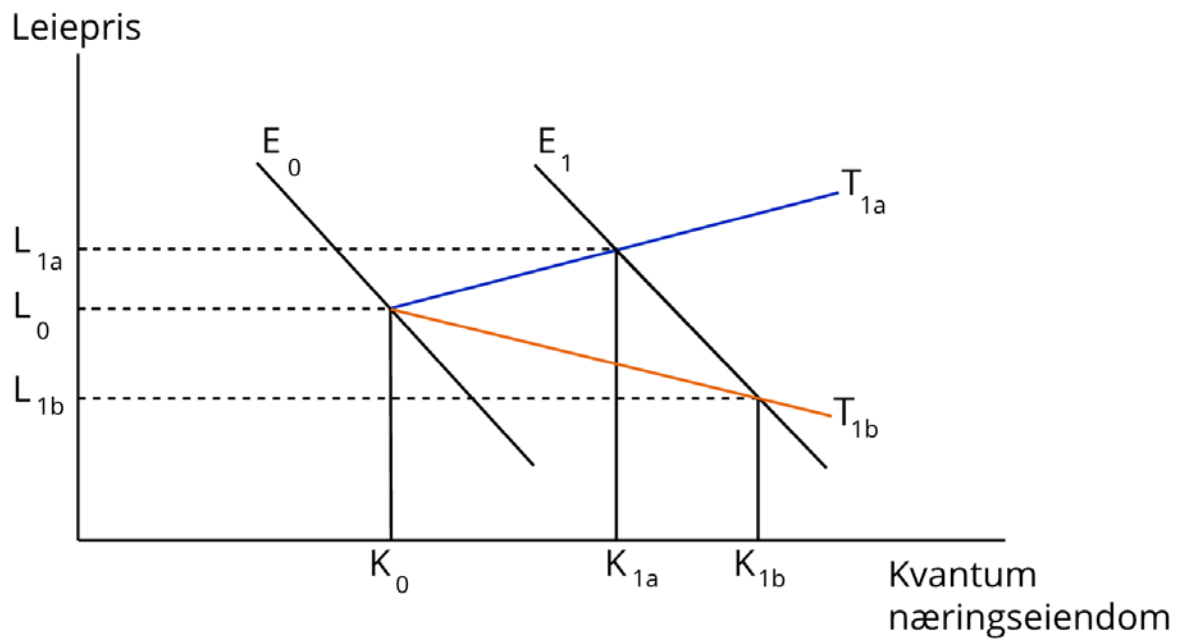
skilt ut eiendomsdrift i egne selskaper. Disse kategoriseres da som næringseiendom og inkluderes dermed i den direkte eksponeringen til banken (Kragh-Sørensen og Solheim, 2014b).

2.2 Tilbud/etterspørsel leiepris

På kort sikt vil endringer i leieprisen i hovedsak være drevet av endringer i etterspørselen. Siden det tar flere år før et bygg blir prosjektert til det er ferdigstilt, betraktes tilbudet av næringseiendom ofte som fast på kort sikt. Etterspørselen etter kontorlokaler påvirkes blant annet av endringer i sysselsettingen og forventet vekst. På lengre sikt vil også etterspørselen avhenge av langsiktige strukturelle faktorer som for eksempel endring i gjennomsnittlig kontorareal per ansatt (Hagen, 2016).

For å øke den langsiktige tilbudssiden av kontorareal har man to muligheter; bygge nytt, eller konvertere eksisterende bygg til kontoreiendom. Prisen for et nybygg kan igjen deles opp i byggekostnader og tomtepris. I følge Geltner et al. (2007) betraktes ofte byggekostnadene som relativt konstante, slik at prisvariasjonen på nybygg avhenger av tomteprisen. Tomteprisen er igjen avhengig av hvor høy arealknapphet det er innenfor det gitte geografiske området. Man får høyere marginalkostnad for nybygg/konvertering hvis arealknappheten øker, og vice versa. Siden det er mange fordeler med å leie lokaler som ligger sentralt plassert, er leieprisen nesten utelukkende høyere i bysentrum enn i mer perifere områder grunnet høyere arealknapphet.

Siden kortsiktig tilbud av næringseiendom anses som gitt, er typisk langsiktige tilbudsfunksjoner fremstilt med en knekk i dagens likevektspunkt. Figur 2 illustrerer en tilbud/etterspørselsfunksjon for leiepris. I det opprinnelige likevektspunktet er leiepris, L_0 , gitt ved kvantum, $K_0 = E_0$. Knekket i tilbudsfunksjonen indikerer at elasticiteten endrer seg langsiktig. Dette betyr at en endring i etterspørsel vil ha en kortsiktig stor effekt. Som nevnt tidligere antar man byggekostnadene som konstante, samtidig som tomteprisen ofte følger en langsiktig trend. I Figur 2 har vi konstruert to scenarier for hvordan en langsiktig endring i tilbudet vil slå ut i leieprisen. Ved økende marginalkostnad får vi leiepris L_{1a} gitt $K_{1a} = E_1$. Økende marginalkostnad kan være grunnet økt arealknapphet og dyrere tomtepris. Leieprisen øker da fra utgangspunktet L_0 . Ved synkende marginalkostnad får vi leiepris L_{1b} gitt $K_{1b} = E_1$ hvor leieprisen synker fra utgangspunktet. En synkende marginalkostnad kan indikere at perifere områder har fått bedre tilgang til for eksempel kollektivtransport, som vil redusere arealknapphet og dermed også tomteprisen.



Figur 2: Tilbud og etterspørselsfunksjon for leiepris

Kortsiktig tilbud av næringsseiendom anses som gitt (K_0). På kort sikt vil endringer i leieprisen i være drevet av endringer i etterspørselen (E_0). Langsiktig tilbud drives av økende (T_{1a}) eller synkende (T_{1b}) marginalkost for nybygging/konvertering (medregnet tomtepris og byggekostnader). Etterspørsel på lang sikt er gitt ved (E_1). X-aksen beskriver kvantum næringsseiendom, Y-aksen beskriver leiepris.

Kilde: (Geltner et al., 2007)

3. Empiri

I denne delen gjennomgås relevant empiri og forskning relatert til problemstillingen. Empirien hjelper oss å forstå viktige sammenhenger, samt å forankre valg og begrensninger vi har satt. Utredningen begynner med å se på litteratur som omhandler hvilke faktorer som påvirker leiepris og avkastning på næringseiendom. Deretter gjennomgås tre statistiske modeller som vi bruker for å prognostisere leieprisen; ARIMA, ARIMAX og VECM. Til slutt blir litteratur om evaluering av yteevnene til prognosene presentert.

3.1 Faktorer som påvirker leiepriser

Det er produsert mye empirisk arbeid som har undersøkt hvilke faktorer som blir benyttet for å forklare leiepriser på næringseiendom. Disse faktorene spenner fra fysiske attributter som for eksempel nærhet til bysentrum og antall kvadratmeter, til makroøkonomiske faktorer som sysselsetting og bruttonasjonalprodukt (Sirmans og Benjamin, 1991).

Flere studier dokumenterer at leiepriser blir påvirket av endringer i økonomisk vekst og tilgang til utleieareal. Studiene viser til at endringer i leieprisene oppstår ved endringer i tilbud- og etterspørselen etter utleiearealet (se for eksempel (Giussani et al., 1992; Rosen, 1984; Hekman, 1985; Wheaton, 1987; Gardiner og Henneberry, 1989; Wheaton og Torto, 1988)). Næringseiendom er en heterogen aktivaklasse med stor variasjon i egenskapene til investeringsobjektene. Av denne grunn er det også ulikheter i hvilke faktorer som påvirker leieprisen. Vi velger å systematisk gjennomgå empiri relatert til hvilke determinanter som driver utviklingen av leieprisen.

3.1.1 Arealledighet

Frew og Jud (1988) estimerte en tverrsnitts-modell for kontorbygg i USA hvor arealledighet var en nøkkelvariabel til å forklare utviklingen i leiepris. Studien bruker simultanligningsprosedyrer for å estimere leiepris per kvadratfot. Studien konkluderer med at arealledighet må være inkludert i kalkulering av leiepris for kontorbygg. Liknende funn blir også identifisert i Wheaton og Torto (1988) hvor de fant at leieprisen falt med 2% for hver prosent arealledigheten økte ved å analysere nasjonale tidsseriedata for det amerikanske markedet. Sivitanides (1997) undersøkte hvordan arealledighet påvirker leieprisen ved å utvikle en modell som tar hensyn til at såkalt strukturell arealledighet kan variere. I følge Rosen og Smith (1983) kan strukturell arealledighet defineres som mengden arealledighet som er nødvendig for å tilfredsstille letebehovet til arealsøkende

leietakere. Sivitanides (1997) fant at den strukturelle arealledigheten varierer både gjennom tid og mellom ulike markeder. Han fant i tillegg at det er stor forskjell på hvor stor effekt arealledigheten har på leieprisen i ulike geografiske områder. Med tydelige indikasjoner på at arealledighet er en viktig determinant for leieprisen velger vi å inkludere variabelen i utredningen.

3.1.2 Arbeidsledighet

Empirien er tvetydig i hvilken grad arbeidsledighet påvirker leieprisen. Dobson og Goddard (1992) finner at effekten av endret arbeidsledighet for leieprisen er avhengig av hvilken næringseiendomstype som blir analysert. De fant at blant annet at arbeidsledighet signifikant påvirker leieprisen på industrieiendom, men ikke for kontorbygg. På den annen side fant De Wit og Van Dijk (2003) et signifikant negativt forhold mellom arbeidsledighet og leiepris for kontoreiendom. Forfatterne undersøkte ulike eiendomsspesifikke variabler og økonomisk vekst. De bekreftet hypotesen om at økonomisk vekst og endringer i tilbud og etterspørsel påvirker leieprisene. Studien utmerker seg i forhold til tidligere studier ved at den tar hensyn til problemet med glatting som man ofte finner i tidligere studier. Forfatterne brukte «generalized method of moments» til å estimere en dynamisk paneldata-modell som muliggjorde kombinasjonen av tverrsnitts- og tidsseriedimensjonen i de kvartalsvise dataene. Med et godt forankret empirisk grunnlag velger vi å inkludere arbeidsledighet i utredningen. Det er også interessant å undersøke om vi kan tilføye eksisterende empiri ny informasjon om forholdet mellom arbeidsledighet og leiepris for det norske kontormarkedet.

3.1.3 Lange renter

En studie av D'Arcy, McGough og Tsolacos (1997) fant en tydelig signifikant sammenheng mellom rentenivå og leiepris. Studien brukte tidsseriedata for 22 europeiske byer i perioden 1982-94. Studien testet også for om markedsstørrelse og endring i by-økonomien hadde en effekt på leieprisen, men fant ingen signifikant sammenheng. Vi velger å inkludere variabelen lange renter fordi den vil fange opp effekten som endring i pengepolitikk har for kontormarkedet. Dette er viktig fordi pengepolitikk gjennom pengetilbud og rentemanipulasjon påvirker økonomiske forhold som vil ha en effekt på etterspørselssiden for kontorbyggene.

3.1.4 BNP

Giussani, Hsia, og Tsolacos (1992) presenterte en empirisk undersøkelse av leiepriser for kontorbygg for noen av de største byene i Europa. De brukte årlige data for perioden 1983-91 til å teste forholdet mellom endringer i leiepris og variasjon i økonomisk aktivitet. Selv med store

databegrensninger fant de at BNP var en viktig determinant for leieprisen. D'Arcy, McGough og Tsolacos (1997) understøtter disse funnene, men peker på at det er geografiske forskjeller på når endringer i BNP slår ut i leieprisen. Også De Wit og Van Dijk (2003) fant en signifikant positiv sammenheng mellom BNP og leieprisen. Med et bredt empirisk grunnlag velger vi å inkludere BNP i utredningen.

3.1.5 Forventet nybygg

Karakozova (2004) brukte i sitt studie av kontormarkedet i Helsinki data på ferdigstilte nybygg for å undersøke hvilken effekt tilbudssiden hadde på kapitalverdien av kontoreiendom. Hun poengterer at andre variabler kunne vært bedre egnet til å kapre utviklingen i tilbudssiden, men mangel på data begrenser valg av variabel. Geltner (2007) argumenterer også for at nybygg er en sentral faktor for å forklare tilbudssiden som ble diskutert i kapittel 2.2. Vi velger å inkludere data på forventet nybygging som en proxy for tilbudssiden, på tross av at det foreløpig er relativt lite empirisk grunnlag for variabelen.

3.1.6 Initial direkteavkastning (Cap-rate)

I en studie av Plazzi et. (2010) ble det vist at man kan bruke "cap-rate" til å fange opp svingninger i forventet avkastning for utleieleiligheter, detaljhandelseiendommer og eiendommer knyttet til industri. Med cap-rate menes raten mellom leiepris og pris på eiendommen. Cap-rate omtales også som initial direkteavkastning på norsk. De fant imidlertid at cap-rate ikke kan forklare svingninger i forventet avkastning for kontoreiendom. Denne konklusjonen ble bygget på transaksjonsdata fra et tverrsnitt av sentrale byområder i USA. Siden vi i denne utredningen har avgrenset omfanget til kun å omhandle kontoreiendom har vi valgt å ikke bruke initial direkteavkastning som en forklarende faktor i prognostiseringen av leiepriser.

3.2 ARIMA - Auto-Regressive Integrated Moving Average

Box og Jenkins (1976) var de første til å utarbeide en systematisert Auto-Regressive Moving Average (ARMA) modell til å predikere finansielle variabler. I følge Brooks og Tsolacos (2010) brukes ARMA-modeller mest til kortsiktig prognostisering og som en benchmark for strukturelle modeller. En Auto Regressive Integrated Moving Average (ARIMA) modell er en generalisert versjon av ARMA-modeller. Disse modellene kan brukes der hvor dataene viser tegn til ikke-stasjonæritet, og hvor differensiering av variablene (som korresponderer med "integrated" delen av modellen) eliminerer ikke-stasjonæriteten.

Tse (1997) brukte ARIMA modeller for å prise kontor- og industrieiendomsindekser i Hong Kong. Utvalget bestod av kvartalsvise data fra 1. kvartal 1980 til 2. kvartal 1995. I følge Holden, Peel og Thompson (1990) er 50 observasjoner eller flere tilstrekkelig for å benytte seg av ARIMA-modellering. Basert på denne studien anså derfor Tse antall observasjoner som tilstrekkelig. Videre undersøkte Tse (1997) yteevnen over en tre-kvartalsperiode fra 3. kvartal 1995 til 1. kvartal 1996. Tse konkluderte med at ARIMA modellene indikerte et fall i kontor- og industrieiendommenes priser på henholdsvis 18,3% og 24,6%, som var veldig nært de faktiske endringene. I studien viste det seg at ARIMA modellene med spesifikasjonen (2.1.1) hadde best yteevne. I Brooks og Tsolacos (2010) sin gjennomgang av Tse (1997) sin studie hevder de at dette er fordi AR leddet indikerer at de sykliske effektene som er generert i den historiske informasjonen blir endogent overført til dagens priser. Vi vil dog i denne utredningen også teste for andre spesifikasjoner av ARIMA modellen.

Wilson *et al.* (2000) undersøkte evnen tidsseriemodeller hadde til å predikere vendepunkter i verdipapiriserte eiendomsindekser. Forfatterne brukte blant annet ARIMA modeller til å analysere kvartalsvise eiendomsindekser fra henholdsvis USA, Storbritannia og Australia. De genererte dynamiske "out-of-sample" prognoser for en åtte kvartalsperiode fra 4. kvartal 1996 til 3. kvartal 1998 for de amerikanske dataene, og 1. kvartal til 1997 til 4. kvartal 1998 for de britiske og australske dataene. Modellene klarte ikke å forutse den signifikante prisstigningen som inntraff i juni 1998, etterfulgt av fallet i prisene mot enden av 4. kvartal 1998. Likevel fant de at modellene traff godt i prediksjonene av indeksens verdi ved sluttidspunktet. Forfatterne konkluderte derfor med at modellene ikke var særlig gode til å predikere vendepunktene, men derimot gode til å predikere tilnærmet verdi av indeksen. Vi velger derfor å bruke ARIMA-modeller i denne utredningen, og merker oss at modellen kan ha problemer med å predikere vendepunkter.

3.3 ARIMAX - Auto-Regressive Integrated Moving Average with Explanatory Variables

Karakozova (2004) analyserer i sin studie avkastning på kontoreiendommer i Helsinki ved hjelp av en regresjonsmodell, en error-correction model (ECM) og en ARIMAX-modell. ARIMAX modellen bygger på ARIMA modellen, men inkluderer også en vektor av eksogene forklaringsvariabler. Hun fant at alle de valgte modellene gjorde det bedre enn en naiv modell, både på kort og lang sikt. ARIMAX modellen som inkluderte tidligere verdier av kapitalvekst, vekst i sysselsettingen for service-sektoren, og vekst i BNP klarte å fange opp ekstremvariasjoner i dataene. Modellen ble derfor det beste prognoseverktøyet for kontoreiendommer i Helsinki. Også i en empirisk studie av det kanadiske eiendomsmarkedet var ARIMAX modellen mest treffsikker i out-of-sample perioden. I studien ble modellen sammenlignet mot ARIMA og en ARIMA-GARCH modell (Auto-regressive conditional heteroskedastic model) (Vishwakarma, 2013). Vi velger derfor med basis i empirien å inkludere en ARIMAX modell i vår utredning.

3.4 VECM - Vector Error-Correction modell

I en artikkel av Sims (1980) ble tidligere strategier for å analysere økonometri analysert, og Vector Auto Regressive (VAR) modellen ble for første gang popularisert. Når variablene er integrert av første orden og det finnes en kointegrert sammenheng mellom variablene, bør man bruke en variant av VAR modellen som kalles Vector Error-Correction model (VECM). Vi utdyper mer om kointegrasjon og selve modellen i metode- og teorikapittelet.

I følge Brooks og Tsolacos (2010) er en av fordelene med VAR/VEC modellering at alle variablene er endogene. Dette er viktig fordi det i praksis ofte er vanskelig å definere hvilke variabler som skal behandles som eksogene. Videre hevder de at VAR/VEC er mer fleksible enn univariate AR modeller, fordi VAR modellene tillater verdien på en variabel å være avhengig av mer enn bare sine egne lags. Dette betyr at vi ikke bare kan se på de ulike variablene sin effekt på gjennomsnittsprisen, men også på variabelens univariate effekt på seg selv. Dette indikerer en mulighet til å fange opp flere egenskaper ved dataene, og at man kan bruke OLS på hver separate ligning. Forfatterne viser til studier av for eksempel Sims (1980) og McNees (1986) som konkluderer med at VAR/VEC modeller ofte er bedre til prognostisering enn tradisjonelle strukturelle modeller. De viser til at storskala strukturelle modeller ofte har lav treffsikkerhet på out-of-sample prognoser. McNees (1986) peker spesielt på at prognostisering av variabler som arbeidsledighet i USA, reelt BNP og enkelte andre variabler er VAR bedre egnet enn strukturelle

modeller. Brooks og Tsolacos (2003) peker dog på enkelte ulemper ved VAR/VEC-modellene. For det første er det utfordrende å velge optimal lag lengde. For det andre er VAR/VEC a-teoretiske, som betyr at de bruker lite teoretisk informasjon om forholdene mellom variablene for å guide spesifikasjonene til modellen. Til slutt blir mange frihetsgrader brukt opp i relativt små utvalgsstørrelser.

Brooks og Tsolacos (2003), og Ling og Naranjo (1997) gir en gjennomgang av studier som benytter VAR/VEC-modeller i eiendomssammenheng. Det blir særlig lagt vekt på at det ikke finnes noen universell enighet om hvilke variabler som best predikerer eiendomsmarkedet, og at VAR/VEC-modeller derfor må ses i sammenheng med de andre prognosemodellene vi benytter i denne utredningen.

3.5 Evaluering av prognosemodellene

Etter å ha valgt ulike prognosemodeller kan det være utfordrende å vite hvilken av modellene som yter best. Derfor er det viktig å ha noen evalueringsverktøy for å teste prognosemodellenes yteevne. I denne delen skal vi derfor studere hvordan yteevnen til prognosemodeller har blitt evaluert empirisk. Med yteevne mener vi hvor god modellen er til å prognostisere fremtidige verdier. Senere i utredningen sammenligner vi hvor godt våre egne modeller gjør det i forhold til hverandre. Med utgangspunkt i denne sammenligningen kan vi velge den modellen som er foretrukket til å prognostisere leiepriser for kontoreiendom i Oslo-området.

Det finnes lite forskning på hvilke prognosemodeller som fungerer best for det norske markedet. Vi vil derfor studere internasjonal litteratur for å gjennomgå tidligere evalueringer av prognosemodellene. I følge Brooks og Tsolacos (2010) er det flere økonometrikere som hevder at så lenge modellen produserer treffsikre prognoser er det mindre viktig om den inneholder ikke-signifikante verdier eller bryter med noen av forutsetningene for klassisk lineær regresjon (CLRM).

En studie av D'Arcy, McGough og Tsolacos (1999) analyserer data for kontorleiepriser i Dublin fra 1970-1997. Formålet med studien var todelt: å videreutvikle prognosemodellene brukt til prognostisering av leiepriser, og teste yteevnen til prognosemodellen i etterkant. Dette ble gjort ved å sammenligne prognosens "out-of-sample" leiepris med de virkelige leieprisene i out-of-sample perioden. I studien fant de at regresjonsmodellen utkonkurrerte to mye brukte prognosemetoder; dobbel-eksponentiell glatting og Holt-Winters metode. Regresjonsmodellen var særlig bedre når de forsøkte å prognostisere mer enn én periode frem i tid. Modellene ble

evaluert ved å sammenligne ”residual mean squared errors” (RMSE). Vi velger derfor med basis i empirien å inkludere prognosefeilparameteren RMSE til å sammenligne ulike prognosemodeller out-of-sample.

I det lignende studiet til Karakozova (2004) for avkastning på kontoreiendom i Helsinki ble ”mean absolute error” (MAE), ”mean absolute percentage error” (MAPE), og ”root mean squared error” (RMSE) sammenlignet for modellene out-of-sample. Vi velger derfor å anvende de samme prognosefeilparameterne i denne utledningen.

En annen metode for å sammenligne treffsikkerheten til prognosemodellene er konseptet om «forecast encompassing». Målet med konseptet er å komme frem til en modell som utkonkurrerer konkurrentene ved å kunne beskrive hvorfor prognosefeilene out-of-sample oppstår, og ved hjelp av den informasjonen kunne prognostisere bedre enn de andre modellene (se Hendry og Clements, 2000). Per i dag er det ikke noe praktisk metode å empirisk teste for dette konseptet og vil derfor ikke bli vurdert i denne utredningen.

Rossi (2014) argumenterer for at såkalte tetthetsprognoser gir viktig informasjon om hvor sannsynlig det er at bestemte hendelser inntreffer. I studiet evalueres treffsikkerheten ved ”Probability integral transform” (PIT) som måler sannsynligheten for å observere predikerte verdier lavere enn den faktiske verdien, hvor sannsynligheten er målt av tetthetsprognosen. Det nevnes også at det finnes flere andre metoder å evaluere tetthetsprognoser (se Corradi og Swanson, 2006). Av hensyn til oppgavens omfang vil vi i denne utledningen fokusere på gjennomsnittlige prognosefeilparameterne; MAD, RMSE, MAD og Theils U2. Parameterne vil bli forklart nærmere i metode- og teoridelen. I videre forskning vil det være interessant å studere om evaluering av tetthetsprognoser hadde gitt andre resultater.

4. Data

Utredningen inkluderer data fra flere forskjellige kilder til å belyse hvilke determinanter som driver leieprisutviklingen for kontorbygg i Oslo-området. I dette kapittelet presenteres dataene i detalj og hvor de er hentet fra. For å gjennomføre analysene og prognosene i henhold til problemstilling og metode har vi hentet data for følgende variabler: leiepriser, arbeidsledighet, BNP, lange renter, arealledighet og forventet nybygging. Alle variablene er forankret i empirien (se kapittel 3.1). På grunn av karakteristikene til næringseiendom vil lags av variablene implementeres i modellene. Dette valget underbygges av artikkelen Holland, Ott, og Riddiough (2000). Våre forventninger til variablene vil være basert på eksisterende empiriske funn.

4.1 Leiepriser (LP)

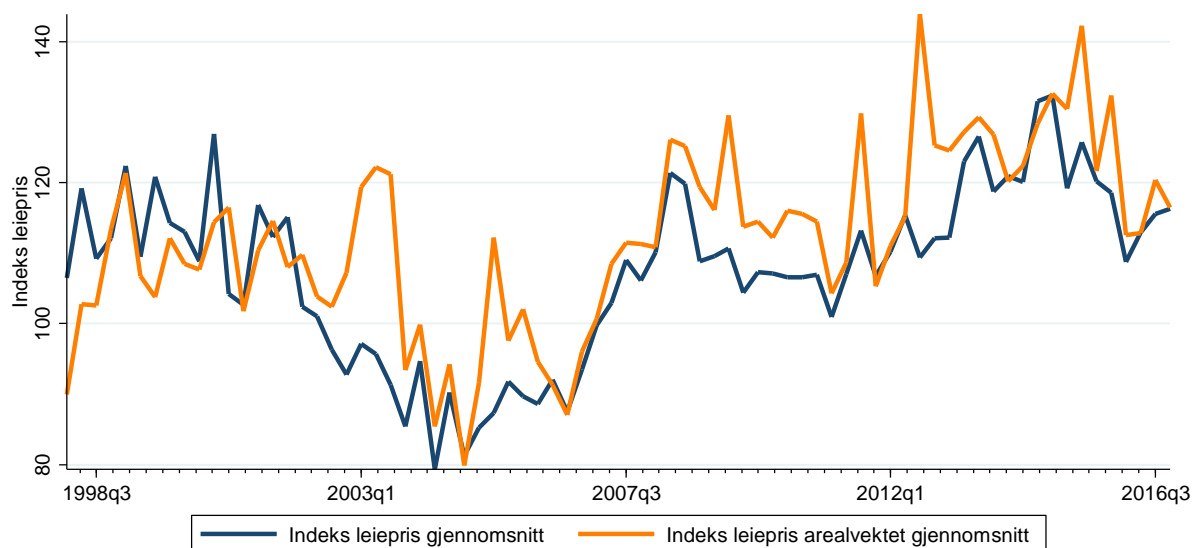
Leiepriser for kontorbygg i Oslo-området og underliggende informasjon om kvalitet, lengde på leiekontrakter og lokasjon har blitt hentet fra Arealstatistikk (AS). AS er et privateid uavhengig selskap som samler inn leiekontrakter fra de fleste eiendomsselskapene i Oslo. Selskapet utgir rapporter hvert kvartal hvor de presenterer utviklingen i markedet og kategoriserer kontraktene i åtte geografiske områder i og rundt Oslo. Videre deler selskapet dataen inn i fire nivåer basert på eiendommenes respektive attraktivitet. I følge selskapet har de en markedsdekning på rundt 90% av alle nye kontrakter som blir signert i dag.

Datagrunnlaget strekker seg fra første kvartal 1996 til fjerde kvartal 2016 som betyr totalt 84 observasjoner for leiepriser på kontorlokaler i Oslo-området. Utvalget i denne tidsperioden består av totalt 13256 kontrakter, men på grunn av konfidensialitet er vi nødt til å ta utgangspunkt i aggregerte data. I de tidligste årene var AS sin dekningsgrad lavere enn 90% av markedet, men denne øker gradvis gjennom perioden. Dette, i kombinasjon med at det var færre kontrakter i omløp, resulterer i at vi anser datagrunnlaget i de tidligste årene som utilstrekkelig til å kunne inkluderes i analysen. Vi begrenser derfor datagrunnlaget til perioden 1998q1-2016q4 da dataene blir gradvis bedre og anses som gode nok i denne perioden. Leieprisene er også kategorisert etter geografiske områder i Oslo tilbake til 2003, men på grunn av få kontrakter per område og et kortere tilgjengelig utvalg vil vi i denne utledningen ikke skille de geografiske områdene fra hverandre.

Leiekontraktene inngår i statistikken på den dagen leietaker flytter inn i lokalene og ikke på den dagen når kontrakten blir inngått. Dette er viktig å merke seg for den videre analysen fordi leietakere sjeldent flytter inn før fire til åtte måneder etter kontrakten er blitt signert. For større

kontrakter er denne perioden ofte lengre (Arealstatistikk, 2017). Andelen store kontrakter som blir signert varierer fra kvartal til kvartal. Dette kan føre til at leieprisen blir unormalt høy i enkelte kvartaler. Siden leieprisene er aggregerte data for hele Oslo-området har vi ikke mulighet å korrigere for dette, noe som er en svakhet i datasettet vårt.

I datasettet får vi presentert gjennomsnittlige nominelle leiepriser, arealvektede leiepriser og leiepriser for kategori A eiendommer, som er de dyreste eiendommene på markedet. For små- og mellomstore bedrifter er gjennomsnittsnivået av leieprisene mest relevant fordi disse oftest har behov for mindre lokaler. For fond, eiendomsselskaper eller større aktører med betydelige eiendomsporteføljer er kontrakter med store arealer typisk viktigere for kontantstrømmen enn kontrakter med små arealer. Derfor vil arealvektet gjennomsnitt gi et bedre bilde for leieprisutviklingen over tid for denne gruppen (Uniongruppen, 2016). Datagrunnlaget for kategori A eiendommer er såpass begrenset at vi velger å se bort fra denne kategorien. Denne utredningen forsøker å gi en metodisk tilnærming til prognostisering av leieprisene som videre kan brukes som et ledd i eiendomsinvesteringer. Vi velger med utgangspunkt i dette å ta utgangspunkt i arealvektede gjennomsnittspriser. Det er også verdt å merke seg at leieprisene har blitt justert for kvartalsvis KPI som er et vanlig mål på inflasjon (NorgesBank, 2017a).

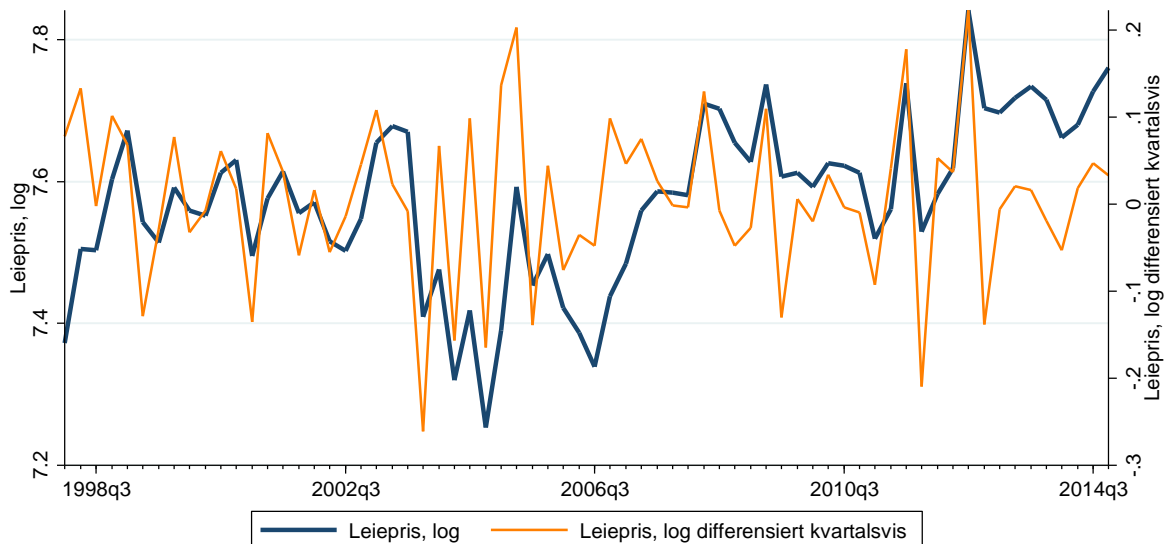


Figur 3: Indeks reelle leiepriser for kontormarkedet i Oslo fra 1998q1 – 2016q4

Vi ser i Figur 3 at vi finner en viss korrelasjon mellom gjennomsnittsleien og den arealvektede gjennomsnittsleien. Dette er et forventet resultat da den arealvektede gjennomsnitteleien er en justert versjon av gjennomsnittsleien. Likevel ser vi at det i perioder er relativt store avvik mellom de to størrelsene. Det betyr at konklusjoner som blir trukket som

resultat av analyser av den arealvektede gjennomsnittsleien ikke automatisk kan overføres til den ujusterte størrelsen.

Vi velger å log transformere serien slik at verdiene skal bli enklere å tolke mot andre variabler da de ikke nødvendigvis har en lineær sammenheng. Som det fremgår av Figur 4 tenderer utviklingen av log leiepriser svakt oppover med enkelte større svingvinger. Det er et større fall i leiepriser i perioden 2003-2005 før det er en positiv trend frem til 2014q4. Vi mistenker at serien ikke er stasjonær, noe som kan føre til spuriøse resultater i analysen (Enders, 2010). En mer formell test for stasjonærhet vil bli gjennomført i kapittel 6.1. Det er ingen tydelige sesongeffekter i serien og vi log differensierer derfor serien kvartalsvis. Vi observerer at den differensierte serien beveger seg relativt konstant rundt gjennomsnittet med et mer sporadisk mønster og vi forventer at serien er integrert av første orden. Videre har vi undersøkt datasettet med tanke på hvor mange kontrakter som er registrert per kvartal. Det er en tydelig trend at flere flytter inn i første kvartal hvert år og vil derfor undersøke sesongeffekter nærmere i kapittel 6.3 og 6.4 ved å inkludere sesongdummyer for hvert kvartal. Log leiepris blir heretter forkortet til «leiepris» og leiepris log differensiert kvartalsvis blir heretter referert til som «vekst i leiepris».



Figur 4: Utvikling av leiepris og vekst i leiepris i perioden 1998q1 - 2014q4.

Venstreaksen viser leieprisen på log nivå, mens høyreaksen viser leiepris log differensiert kvartalsvis.

4.2 Makroøkonomiske variabler

Underbygget av studien fra S.M Dobson og J.A. Goddard (1992) velger vi å inkludere makroøkonomiske variabler i analysen. Vi går i denne delen nærmere inn på de ulike makroøkonomiske variablene, samt kommenterer våre forventninger.

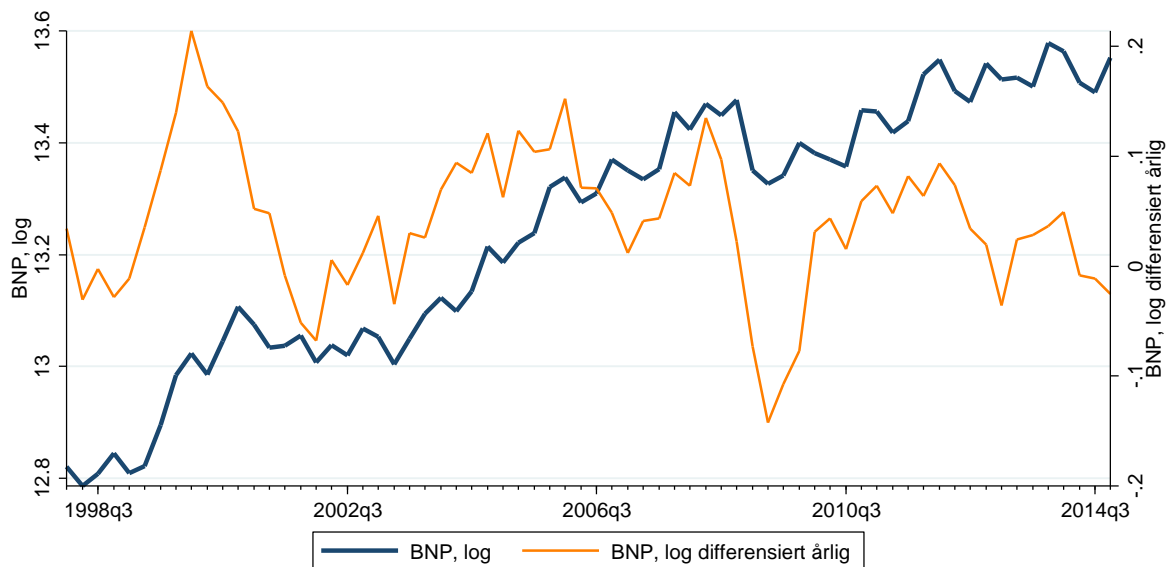
4.2.1 Bruttonasjonalprodukt (BNP)

Tall på bruttonasjonalprodukt (BNP) er hentet fra Statistisk Sentralbyrå (SSB) sin statistikkdatabase. BNP inkluderer all verdiskapning i Norge, altså både for fastland og offshorevirksomhet. Tallene blir presentert kvartalsvis og har blitt inflasjonsjustert til reelle tall med KPI². I følge SSB er det ofte betydelige sesongvariasjoner som kan vanskeliggjøre en direkte tolkning fra periode til periode. Dette kan være på grunn av klimatiske forhold, bevegelige helligdager, måneder med ferieavvikling etc. (SSB, 2016).

Vi har valgt å inkludere offshorevirksomheter i BNP tallene da disse selskapene er en viktig del av norsk økonomi. I tillegg preges dette markedet av store aktører som har behov for kontorarealer i Oslo-området. Dette forventer vi skal ha påvirkning på etterspørselssiden til kontorlokalene, og er derfor viktig for vår analyse. Vi forventer at BNP vil være en god indikator på økonomisk aktivitet, og dermed at en økning i BNP vil øke etterspørselssiden. Koeffisienten er forventet å være positiv og signifikant. Siden BNP måles i absolutte tall velger vi å log transformere serien. Figur 5 viser en tydelig positiv trend i dataserien for log BNP. Vi observerer også at det var et fall under finanskrisen i 2007 som hentet seg inn igjen i tiden etter 2010. Serien ser ut til å følge en deterministisk trend og vi mistenker at den ikke er stasjonær på nivåform.

For å hensynta betydelige sesongvariasjoner i dataene er log BNP differensiert årlig. Vi observerer ingen tydelige trender, og ser at serien beveger seg rundt gjennomsnittet. Vi forventer at log BNP er integrert av første orden I(1). Log reelt BNP blir heretter forkortet til «BNP», og reelt BNP log differensiert årlig blir heretter referert til som «vekst i BNP».

² BNP kan også inflasjonsjusteres ved bruk av andre metoder. BNP deflator er et vanlig mål på inflasjon og beregnes på grunnlag av all verdiskapning i landet, vektet med markedsverdien av totalt konsum for hver vare. Produsentprisindeks (PPI) er en annen metode som baserer seg på den faktiske utviklingen på varer i produsentleddet ved salg til det norske markedet og eksport (SSB, 2017).



Figur 5: Utvikling av BNP og vekst i BNP i perioden 1998q1 til 2014q4

Venstreaksen viser log BNP på nivå, mens høyreaksen viser log BNP differensiert kvartalsvis (vekst).

BNP er inflasjonsjustert til reelle tall med KPI.

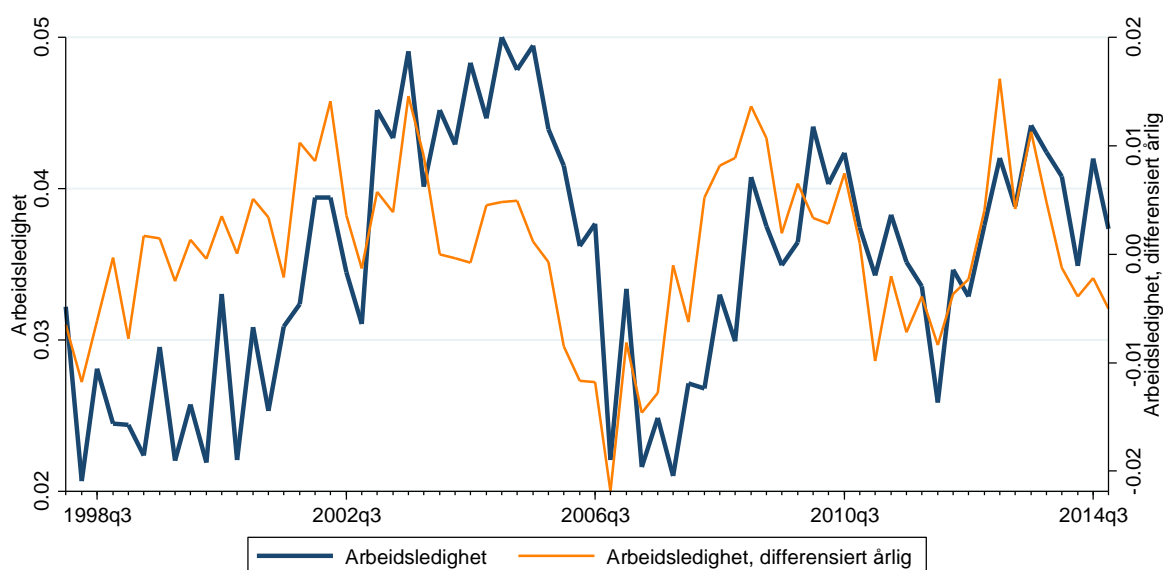
4.2.2 Arbeidsledighet (ARBL)

Tall på arbeidsledighet er hentet fra SSB sin statistikkdatabase. Tallene blir presentert kvartalsvis i absolutte tall. Siden vi i denne utredningen analyserer Oslo-området har vi skilt ut ledighetstallene for Oslo og Akershus fra landsoversikten. For å få arbeidsledighet som en rate har vi kalkulert forholdet mellom antall arbeidsledige og antall personer i arbeidsstokken for Oslo og Akershus. Vi står da igjen med en prosentandel av arbeidsstokken som er arbeidsledig som vi inkluderer i vår analyse. En økning i arbeidsledighet vil med bakgrunn i empirien redusere etterspørselssiden for kontorlokaler og vi forventer derfor at koeffisienten er negativ og signifikant.

Figur 6 viser tydelige trender for kortere perioder i dataserien for arbeidsledighet, men trendene er ikke entydige over hele perioden. Vi observerer også at det var en kraftig økning i arbeidsledigheten i perioden etter dot-com boblen i 2000 frem til 2005. Arbeidsledigheten øker igjen i tiden etter finanskrisen i 2007. Serien har altså noen sykler, men beveger seg ikke sporadisk rundt en konstant eller trend. Vi mistenker at serien kan inneholde enhetsrøtter.

I følge SSB (2016) kan også arbeidsledighet være preget av betydelige sesongvariasjoner som vanskeliggjør en direkte tolkning fra periode til periode. Vi velger derfor å differensiere arbeidsledighet årlig for å hensynta dette. Vi forventer at variabelen er integrert av første orden

I(1). Arbeidsledighet på nivåform blir forkortet til «arbeidsledighet» og arbeidsledighet differensiert årlig blir heretter referert til som «endring i arbeidsledighetsraten».



Figur 6: Utvikling i arbeidsledighet i Oslo-området i perioden 1998q1 – 2014q4

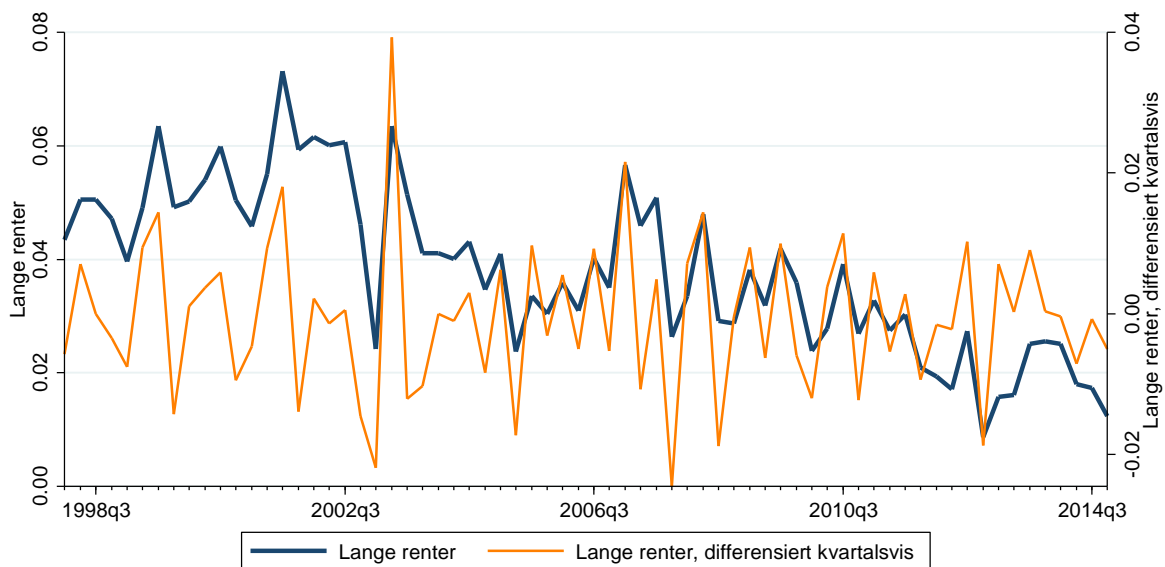
Venstreaksen viser raten mellom antall arbeidsledige og antall i arbeidsstokken i Oslo og Akershus
Høyreaksen viser arbeidsledigheten differensiert årlig (endring i arbeidsledighetsraten).

4.2.3 Lange renter (RRTI)

Vi har hentet data for 10årige statsobligasjoner fra Norges Bank. Dataene er basert på siste rentenotering per kvartal, før vi har inflasjonsjustert tallene med vekst i KPI. (NorgesBank, 2017b). Årsaken til at vi har valgt 10 års statsobligasjoner er at kontraktene i leiemarkedet som oftest går over lengre perioder fra 5 til 10 år. I tillegg kan vi se på variabelen som en proxy for risikofri rente som er en viktig parameter i beregning av verdien for kontoreiendom (Karakozova, 2004). I tråd med D'Arcy, McGough og Tsolacos (1997) forventer vi en negativ sammenheng mellom leiepris og lange renter. Vi har også observert en tydelig divergerende trend mellom disse variablene, noe som underbygger denne forventningen.

Figur 7 viser at vi totalt sett har hatt en nedadgående trend fra 1998-2015 for de lange rentene, med enkelte store variasjoner. Vi observerer at renten gikk en del opp under finanskrisen i 2007, og har falt i perioden etterpå. Basert på en tydelig nedadgående trend forventer vi at serien ikke er stasjonær på nivåform. Siden det ikke er indikasjoner på sesongeffekter i variabelen differensierer vi én gang kvartalsvis. I den differensierte serien observerer vi ingen tydelige trender og ser at serien beveger seg rundt gjennomsnittet med enkelte høye toppe og dype bunner. Vi forventer at serien er integrert av første orden I(1). 10årige statsobligasjoner på nivåform blir

heretter forkortet til «lange renter», og 10årige statsobligasjoner differensiert kvartalsvis blir heretter referert til som «endring i lange renter».



Figur 7: Utvikling i lange renter og endring i lange renter i perioden 1998q1 – 2014q4

Venstreaksen viser rentenivået på lange renter. Høyreaksen viser lange renter differensiert kvartalsvis.

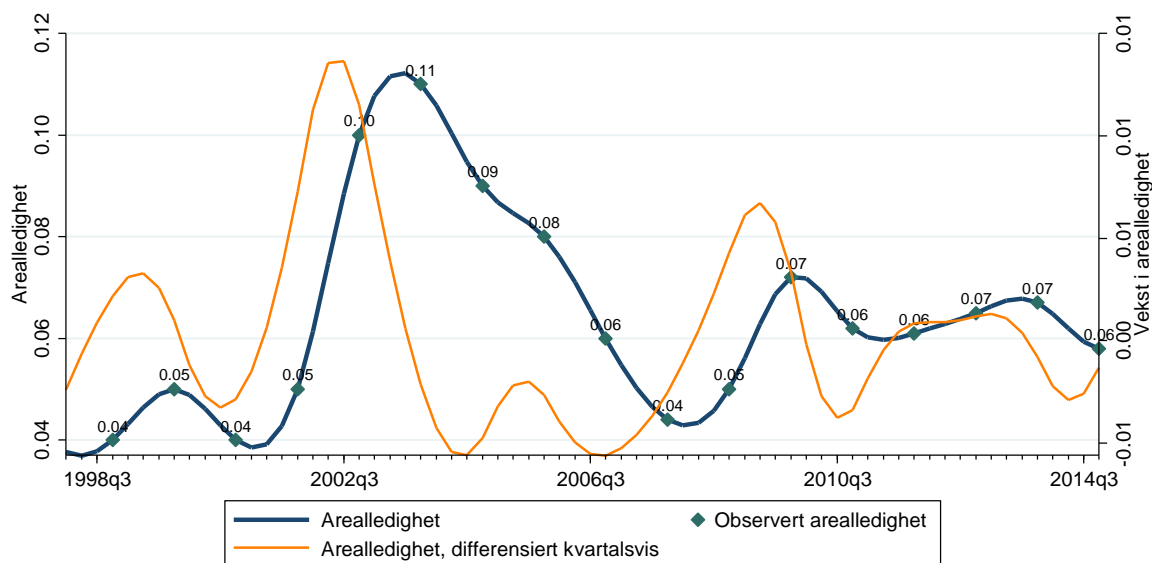
4.3 Markedsspesifikke variabler

Dataen som tilhører de markedsspesifikke variablene er hentet fra det fullintegreerte eiendomsselskapet Eiendomsspar sine årlige rapporter kalt Oslostudiet. Rapportene viser utviklingen i næringseiendomsmarkedet for kontor- og forretningsbygg i Oslo-området. Hovedvekten av studiet er på tilbudssiden, ved identifisering av ledige arealer og arealer under bygging/rehabilitering i forhold til eksisterende eiendomsmasse for de enkelte geografiske områdene i Oslo (Eiendomsspar, 2017).

4.3.1 Arealledighet (ARL)

Tallene på arealledighet (ARL) inkluderer både fremleiearealer og øvrige ledige arealer. Vi har valgt å bruke prosentvis ledighet i Oslo-området av totalt tilgjengelig utleieareal. Det har ikke vært mulig å oppdrive kvartalsvis data for arealledighet og vi har dermed valgt å interpolere dataen kvartalsvis ved hjelp av cubic splines. Fremgangsmåten og begrunnelse for dette blir forklart i appendiks A. En økning i arealledighet tyder på at økningen i tilbudet er større enn etterspørselen, noe som vil ha en negativ effekt på leieprisene. ARL er dermed en proxy for tilbudssiden for kontorlokaler i Oslo og vi forventer at koeffisienten er negativ og signifikant.

Figur 8 viser at arealledigheten for kontoreiendom i Oslo tok seg kraftig opp i perioden mellom 2001-2005 før den igjen falt dramatisk. I følge bransjeeksperter på næringsbygg i Oslo var en av hovedårsakene til dette ringvirkninger fra dot.com boblen i 2000. I årene 1999 – 2000 ble det prosjektert unormalt mange nye kontorlokaler og med høy arbeidsledighet i kontorareal-intensive yrker ble det følgelig et stort overskudd på kontorlokaler og arealledigheten steg (Uniongruppen, 2002). Vi observerer også at det ser ut som ledigheten er noe syklisk der den blant annet tar seg opp igjen etter finanskrisen i 2007. Vi forventer at serien ikke er stasjonær på nivåform.



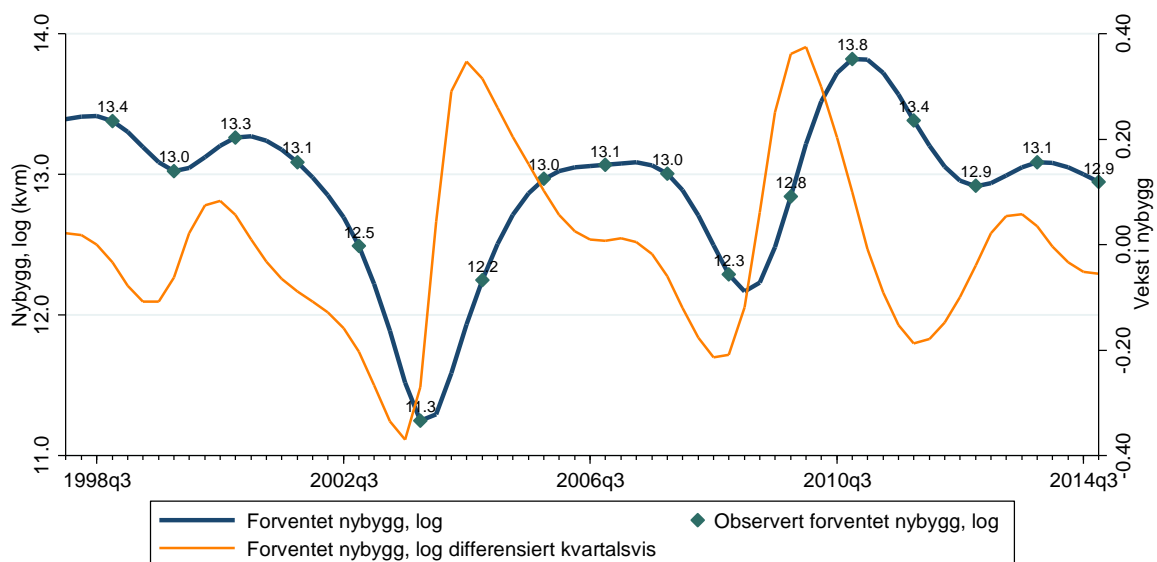
Figur 8: Utvikling i arealledighet og vekst i arealledighet i perioden 1998q1 – 2014q4

Venstreaksen viser raten mellom antall ledige kvadratmeter og totalt antall kvadratmeter utleieareal for kontorbygg. Høyreaksen viser arealledighet differensiert kvartalsvis. Punktene illustrerer de virkelige verdiene før serien ble interpolert kvartalsvis.

Første differensiering av dataen kvartalsvis ser ikke ut til å bedre stasjonærhetsproblemet i nevneverdig grad. Vi observerer en vekst i perioden 2000-2002, deretter er kraftig fall de neste to årene. Veksten i arealledigheten tar seg også opp igjen fra 2007 til slutten av 2008. Vi mistenker her ett eller flere strukturelle brudd i serien. Videre analyse av eventuelle strukturelle brudd vil bli presentert i kapittel 6.1. Implikasjonene ved å inkludere en variabel som ikke er stasjonær i tidsserieanalyse kan føre til spuriøse resultater (Enders, 2010). Arealledighet på nivåform blir heretter forkortet til arealledighet, og arealledighet differensiert kvartalsvis blir heretter referert til som «vekst i arealledighet».

4.3.2 Forventet nybygg de neste 4 årene (FNYB)

Data for forventet nybygging av kontorlokaler de neste 4 årene er hentet fra Eiendomsspar AS sine markedsrapporter og måles i antall kvadratmeter. Vi har basert oss på de byggene som selskapet mener er over 50% sannsynlig blir ferdigstilt i løpet av de neste 4 årene. Tallene blir presentert årlig, så vi har valgt å interpolere serien med samme metode som arealledighet. Interpoleringen blir forklart i detalj i appendiks A. Ceteris paribus vil nybygging av kontorlokaler øke tilbudssiden som vil redusere leiepris hvis ikke etterspørselen holder følge. Likevel er det vanskelig å forutse fortegnet til koeffisienten da forventet nybygging også kan være en indikator på økt etterspørsel i markedet. Som vi nevnte i introduksjonen blir også tilbudet av kontoreiendom ofte ansett som fast på kort sikt. Derfor er det mulig at variabelen har ulik effekt på leieprisen på kort og lang sikt. Basert på dette har vi ikke en klar forventning til fortegnet til koeffisienten. Det indikeres dog at effekten kan være positiv for leiepris på kort sikt (grunnet indikasjon på høyere etterspørsel), men at høy nybygging vil kunne resultere i for høyt tilbud på lang sikt. Resultatet av dette vil være negativt for utviklingen av leiepris på lang sikt. Serien blir log transformert siden den måles i absolutte kvadratmeter og for at det skal være enklere å tolke sammenhengen mot andre variabler senere i utredningen.



Figur 9: Utviklingen i forventet nybygg og vekst i forventet nybygg i perioden 1998q1 – 2014q4. Venstreaksen viser log av antall kvadratmeter forventet nybygg. Høyreaksen viser forventet nybygging log differensiert kvartalsvis. Punktene illustrerer de virkelige verdiene før serien ble interpolert kvartalsvis.

Figur 9 viser at det var en sterk nedgang i forventningene til nybygg i 2003. På samme tid var arealledigheten på historisk høye 11%. Etter dot.com boblen hadde arealabsorpsjonen i Oslo falt dramatisk som følge av lav sysselsettingsvekst i kontorintensive yrker. Allerede igangsatte prosjekter ble ferdigstilt i 2001 og 2002 noe som følgelig gjorde arealoverskuddet betydelig (Eiendomsspar, 2004). Forventningene tok seg opp igjen i løpet av de neste årene med unntak av en korreksjon under finanskrisen i 2007. Med nye prosjekter som «Barcode» i Oslo sentrum og satsning på Fornebu hvor store selskaper som Statoil og Aker Solutions relokerte seg økte nybyggingen betraktelig (Eiendomsspar, 2017). Serien har flere store svingninger uten noen veldig tydelig trend og vi forventer at serien ikke er stasjonær på log nivå.

Vi log differensierer variabelen kvartalsvis og ser at serien fortsatt har store svingninger, men varierer i større grad rundt gjennomsnittet. Spesielt er det stor vekst i første kvartal 2003 etter tre år med nedgang. Det er også stor vekst i perioden etter finanskrisen. Vi forventer at serien er integrert av første orden I(1). Log forventet nybygging vil heretter forkortes til «forventet nybygg», og forventet nybygging log differensiert kvartalsvis blir heretter referert til som «vekst i forventet nybygg».

Tabell 1: Oppsummerende statistikk

Tabellen presenterer gjennomsnittet, standardavviket (St.avvik), minimumsverdien (min) og maksimumsverdien (max) for variablene tidligere presentert i utledningen i perioden 1998q1 til 2014q4. Δ indikerer at variabelen er blitt differensiert en gang kvartalsvis. Δ_4 indikerer at variabelen er blitt differensiert en gang årlig. Ln viser at serien er log transformert, variablene uten ln er målt i rate på nivåform og er derfor ikke log transformert.

Variabel	Gjennomsnitt	St.avvik	Min	Max
lnLP	7,58	11,59 %	7,2536	7,8418
lnBNP	13,36	7,48 %	13,2041	13,4731
ARBL	3,51 %	0,81 %	0,0207	0,0500
RRTI	3,83 %	1,45 %	0,0087	0,0732
ARL	6,42 %	2,01 %	0,0370	0,1121
lnFNYB	12,90	55,35 %	11,2512	13,8213
Δ lnLP	0,68 %	9,22 %	-0,2607	0,2225
Δ_4 lnBNP	4,15 %	6,57 %	-0,1424	0,2140
Δ_4 ARBL	0,04 %	0,75 %	-0,0219	0,0162
Δ RRTI	-0,05 %	1,09 %	-0,0245	0,0393
Δ ARL	0,03 %	0,45 %	-0,0056	0,0136
Δ lnFNYB	-0,63 %	16,30 %	-0,3696	0,3747

5. Metode og teori

Hensikten med denne utredningen er å identifisere hvilke determinanter som driver leieprisene, og deretter utarbeide og evaluere prognoser basert på disse variablene. Vi skal som nevnt benytte oss av tre forskjellige statistiske modeller; ARIMA, ARIMAX, og VECM. Grunnet utilstrekkelig datagrunnlag for de første to årene baserer vi analysen på data fra 1.kvartal 1998 til 4. kvartal 2014. Dette tilsvarer 68 observasjoner. Deretter lager vi prognoser for ”out-of-sample” perioden fra 1.kvartal 2015 til 4. kvartal 2016, før vi sammenligner og evaluerer prognoseverdiene med de virkelige verdiene. Basert på denne sammenligningen vil vi kunne se hvilken prognosemodell som er best til å predikere out-of-sample verdiene, og vil være den modellen vi vil anbefale å bruke for videre prognostisering. Vi vil først presentere metode og teori knyttet til tidsserieanalyser. Deretter vil vi presentere metoden for de tre valgte statistiske modellene, før vi avslutter med metode for evaluering av modellene.

5.1 Stasjonære prosesser

En av de viktigste forutsetningene for analysen er stasjonære prosesser. En stokastisk variabel sies å være stasjonær hvis dens sannsynlighetsfordeling ikke endrer form over tid. I denne utredningen bruker vi betegnelsen stasjonær når en tidsserie er kovarians stasjonær, eller svakt stasjonær. Formelt må følgende betingelser være oppfylt for at en tidsserie skal være kovarians stasjonær:

$$E(y_t) = \mu$$

$$E(y_t - \mu)(y_t - \mu) = \sigma^2 < \infty$$

$$E(y_{t_1} - \mu)(y_{t_2} - \mu) = \gamma_{t_2-t_1} \quad \forall t_1, t_2$$

hvor μ , σ^2 og $\gamma_{t_2-t_1}$ er konstanter. En stasjonær serie kan defineres som en serie med konstant gjennomsnitt, konstant varians og konstant autokovarians. Dersom ikke-stasjonære tidsserier benyttes i univariate modeller kan det som nevnt tidligere føre til spuriøse regresjonsresultater. Dette kan medføre både uriktige betaestimer og ugyldige t-verdier. Spuriøse regresjoner kan også innebære at modellen indikerer en ikke-eksisterende sammenheng mellom to variabler (Enders, 2010). Hvis tidsserien ikke er stasjonær sier vi at den inneholder enhetsrøtter, og en vanlig løsning for å gjøre tidsserien stasjonær er differensiering av variablene (Brooks, 2014).

Vi velger å bruke Augmented Dickey-Fuller (ADF)-tester for å avgjøre om det finnes enhetsrøtter i tidsseriene. ADF-testen er en utvidelse av Dickey-Fuller (1979)-testen, men som fjerner alle de strukturelle effektene (autokorrelasjon) i tidsserien. Dickey-Fuller testen undersøker om $\phi = 1$ i

$$y_t = \phi y_{t-1} + \mu_t$$

mot det ensidige alternativet $\phi < 1$. Hypotesene blir da:

H_0 : serien inneholder en enhetsrot, mot

H_1 : serien er stasjonær.

I praksis blir dog følgende regresjon anvendt for å gjøre kalkulering og tolkning enklere:

$$\Delta y_t = \psi y_{t-1} + \mu_t$$

Dette medfører at en test om $\phi = 1$ er ekvivalent til en test om $\psi = 0$ (siden $\phi - 1 = \psi$). Det er viktig å merke seg er at Dickey-Fuller testen kun er gyldig når μ_t er hvit støy. Dette innebærer at μ_t ikke er autokorrelert, noe den vil være hvis det er autokorrelasjon i den avhengige variabelen i regresjonen (Δy_t). Hvis dette er tilfellet, vil testen bli overdimensjonert, noe som vil si at antall ganger nullhypotesen blir feilaktig forkastet er høyere enn den nominelle størrelsen som blir brukt (for eksempel 5%). Løsningen er å øke testens omfang ved å bruke p antall lags (på norsk: forsinkelse) av den avhengige variabelen (Brooks og Tsolacos, 2010). Modellen blir da som følger:

$$\Delta y_t = \psi y_{t-1} + \sum_{i=1}^p \alpha_i \Delta y_{t-i} + \mu_t$$

De laggede verdiene av Δy_t sikrer nå at μ_t ikke er autokorrelert. Hovedutfordringen som oppstår ved bruk av den utvidede Dickey-Fuller testen er valg av optimalt lag på den avhengige variabelen. Det er hovedsakelig to metoder som blir brukt for å velge antall lags i ADF-testen. Den første og enkleste metoden er å basere antall lags på frekvensen til dataen. Man bruker da normalt 12 lags på månedlige data, 4 lags på kvartalsvise osv. Den andre metoden går ut på å bruke såkalte informasjonskriterier, hvor man velger antall lags som minimerer verdien på ulike informasjonskriterier. Disse informasjonskriteriene blir diskutert i kapittel 5.2.

5.2 Informasjonskriterier

Informasjonskriteriene består av to faktorer: en funksjon av residualkvadratsummen (RSS), og et straffelegg for tapet av frihetsgrader ved å legge til ekstra parametere. Dette betyr at det å legge til en ny variabel eller en ekstra lag til en modell, vil ha to konkurrerende effekter på informasjonskriteriet: residualkvadratsummen vil falle, men verdien til straffelegget vil øke (Brooks og Tsolacos, 2010).

Formålet er å velge nummeret av parametere som minimerer verdien av informasjonskriteriet. I denne utledningen brukes informasjonskriteriene til å finne optimal lag lengde i ADF-tester og VECM. Kriteriene blir også brukt som et relativt mål mellom modeller for å bestemme antall parametere. Legger man til en ekstra parameter eller variabel vil kriteriene reduseres i verdi bare hvis fallet i RSS er større enn økningen i straffelegget.

I denne utredningen bruker vi Akaike's (AIC) (1974), Schwarz's Bayesianske (BIC) (1978) og Hannan-Quinns (HQIC) (1979) informasjonskriterier for å fastslå optimal lag lengde og den optimale modellspesifikasjonen. Informasjonskriterier er gitt ved:

$$AIC = \ln(\hat{\sigma}^2) + \frac{2k}{T}$$

$$BIC = \ln(\hat{\sigma}^2) + \frac{k}{T} \ln T$$

$$HQIC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \ln(\ln(T))$$

hvor $\hat{\sigma}^2 = \frac{RSS}{T}$, $k = p + q + 1$ er det totale antallet parametere og T er utvalgsstørrelse. BIC har et strengere straffelegg enn AIC, og HQIC plasserer seg noenlunde mellom de to nevnte. Det betyr at AIC gjennomsnittlig vil anbefale overtilpassede modeller, mens BIC anbefaler mer restriktive modeller. Det finnes også andre mulige kriterier, men de er i hovedsak varianter av de nevnte informasjonskriteriene og vil derfor ikke bli vurdert i denne utledningen (Brooks og Tsolacos, 2010).

5.3 Hvit støy og Ljung-Box test

Vi har en hvit støy prosess når restleddet er ukorrelert med restleddet i tidligere perioder. En hvit støy prosess har konstant gjennomsnitt og varians, og ingen autokovarians bortsett fra ved lag 0 (Brooks, 2014). Vi velger å bruke Ljung-Box (1978)-test for å teste om vi har hvit støy prosesser. Testen kan uttrykkes ved:

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

hvor T er utvalgsstørrelse, \hat{t}_k er autokorrelasjon ved lag k , og m er antall lags som blir brukt. Under H_0 følger Q^* en χ^2 fordeling. Er det indikasjoner på autokorrelasjon i restleddet er det ifølge Granger og Newbold (1974) tre store konsekvenser for analysen: (1) estimatene av koeffisientene er spuriøse, (2) prognoser basert på modellen er ikke optimal, (3) signifikanstestene på koeffisientene er ikke gyldige. På bakgrunn av dette vil vi kun akseptere modeller i analysen hvor residualene følger en hvit-støy prosess i henhold til Ljung-Box Q^* test.

5.4 Test for strukturelle brudd – Wald test

Man antar ofte at parameterne (f.eks. β_1, β_2 og β_3) er konstante for hele utvalget, både for estimeringsperioden, og for alle etterfølgende perioder som blir brukt i konstruksjonen av prognosene. Man kan teste denne antakelsen ved å splitte datasettet i to underperioder og sammenligne koeffisientene eller standardfeilene mot hverandre. En klassisk tilnærming er å anvende Chow sin stabilitetstest for kjente bruddatoer. I senere tid har dog mer komplekse metoder blitt utviklet hvor man kan teste for ukjente bruddatoer, som for eksempel Wald stabilitetstest. Dette er spesielt nyttig i serier hvor bruddatoer ikke kan direkte linkes til en bestemt hendelse. Med bakgrunn i at det er vanskelig å vite nøyaktig når økonomiske hendelser slår ut i eksogene forklaringsvariabler for leiepris velger vi å bruke Wald sin stabilitetstest for ukjente bruddatoer som blir omtalt blant annet av Andrews (1993):

$$W(T_1) = n(\hat{\beta}_1 - \hat{\beta}_2)' \left(\hat{V}_1 \frac{n}{T_1} + \hat{V}_2 \frac{n}{n - T_1} \right)^{-1} (\hat{\beta}_1 - \hat{\beta}_2)$$

$$SupW = \sup_{T_1} W(T_1)$$

hvor \hat{V}_1 og \hat{V}_2 er standard asymmetrisk variansestimater for $\hat{\beta}_1 - \hat{\beta}_2$ i det splittede utvalget. SupW testen blir utført for alle bruddatoene i utvalget i regionen $[t_1, t_2]$ hvor $t_1 > 1$ og $t_2 < n$. Testen returnerer så den bruddatoen med høyest Wald verdi. Hvis bruddatoen kommer for nærme starten eller slutten av utvalget kan estimatene og testen være misledende. Vi benytter derfor anbefalt trimming på 15% av hver side av utvalget ($t_1 = [.15n], t_2 = [.85n]$). Er verdien over kritisk verdi forkaster vi nullhypotesen om ingen brudd i serien.

5.5 AR - Auto Regressiv modell

En autoregressiv modell er en modell hvor nåværende verdi av en variabel, y , kun avhenger av verdien variabelen tok i tidligere perioder pluss et feilledd (Brooks, 2014). En autoregressiv modell av orden p , betegnet som $AR(p)$, kan uttrykkes slik:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + u_t$$

hvor u_t er et hvit støy feilledd. Vi vil i denne utredningen bruke en random-walk modell til å utarbeide en naiv prognose for leieprisene hvor $AR(1)$ er lik 1. Modellen vil brukes som en benchmark for de mer sofistikerte modellene; ARIMA, ARIMAX og VECM. Fremgangsmåten for evaluering av prognosemodellene blir presentert senere i kapittel 5.10.

5.6 ARIMA - Auto Regressive Integrated Moving Average

ARIMA (p,d,q) modellen er som tidligere nevnt en univariat tidsseriemodell. Den skiller seg fra ARMA-modeller ved at variablene er differensiert d ganger. ARMA modellene er en kombinasjon av autoregressive ($AR(p)$) og glidende gjennomsnitts ($MA(q)$)-modeller. Modellen baserer seg på at verdien av en serie y er lineært avhengig av sin egen tidligere verdi, samt en kombinasjon av seriens nåværende- og tidligere verdi av feilleddet. Feilleddet er antatt å være hvit støy (Brooks og Tsolacos, 2010). ARMA modellen defineres:

$$y_t = \alpha + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \theta_2 u_{t-2} + \dots + \theta_q u_{t-q} + u_t$$

med

$$E(u_t) = 0; E(u_t^2) = \sigma^2; E(u_t u_s) = 0, t \neq s$$

5.6.1 Box-Jenkins metode for ARIMA modellering

For å spesifisere vår ARIMA modell bruker vi Box og Jenkins (1976) sin 3-steps systematiske fremgangsmåte:

Steg 1 er å bestemme hvilken orden modellen må være for å fange opp de dynamiske egenskapene til dataen. For å vurdere dette er det nyttig å introdusere autokorrelasjonsfunksjonen (ACF) som kan uttrykkes ved:

$$\tau_s = \frac{\gamma_s}{\gamma_0}, \quad s = 0, 1, 2, \dots$$

Autokovariansen blir normalisert ved å dele på variansen, og serien τ_s får standardegenskapene til korrelasjonskoeffisienter ved at verdiene er bundet til mellom -1 og +1 (Brooks og Tsolacos, 2010). Eksempelvis er autokorrelasjonen ved lag 0 ($s = 0$) lik 1, fordi man da finner korrelasjonen mellom y_t og y_t . Den partielle autokorrelasjonsfunksjonen (PACF) måler korrelasjonen mellom en observasjon k perioder tilbake og den nåværende observasjonen, etter å kontrollere for de mellomliggende laggene (alle lag $< k$). Ved å undersøke korrelogrammet til tidsserien finner vi ACF og PACF ved hver lag, og kan dermed undersøke autokorrelasjonsmønstre. Hvis tidsserien ikke har de hensiktsmessige egenskapene til å passe en ARMA modell transformerer vi serien ved første differensiering før vi på nytt undersøker korrelogrammet. For å fullføre steg 1 i Box og Jenkins (1976) metoden bruker vi informasjonskriteriene AIC og BIC sammen med korrelogrammet for å bestemme antall ARIMA komponenter.

Steg 2 er å estimere parameterne i del 1, noe vi utfører med maksimum sannsynlighetsteknikken integrert i Stata14 (StataCorp, 2013).

Steg 3 er en diagnostisk kontroll av modellen. Box og Jenkins (1978) foreslår å kontrollere residualene for lineær avhengighet. Vi velger her å undersøke ACF, PACF og Ljung-Box-test for hvit-støy.

5.6.2 Diagnostiske tester

For å kunne dra slutninger fra en ARIMA modell må den avhengige variabelen være kovariansstasjonær (Hamilton, 1994). Stasjonæritet av ARMA prosessen avhenger av den autoregressive AR komponenten. Hvis de inverterte røttene til AR-polynomtallet ligger innen i enhetssirkelen, er prosessen stasjonær, inverterbar, og har en uendelig rekkefølge av (MA) representasjoner. Motsatt kan MA delen av en ARMA prosess omskrives som en uendelig

rekkefølge av AR-prosesser forutsatt at MA-prosessen er inverterbar. Hamilton (1994) finner at det er mulig å bevise at dersom moduli av hver eigenverdi er strengt lavere enn 1 så er ARMA prosessen stasjonær. I utledningen vil vi benytte Stata14 for å teste om ARMA komponentene er stasjonær ved bruk av funksjonen «aroots» (StataCorp, 2013). Oppfyller ikke modellen kravet vil den ikke bli benyttet for prognoser out-of-sample.

Residualene er antatt å være normalfordelte: $\mu_t \sim N(0, \sigma^2)$ hvor μ_t er normalfordelt. I ARIMA og ARIMAX analysen gjennomfører vi en normalitetstest («sktest» (StataCorp, 2013)) utviklet av D'Agostino og Belanger (1990) som er lik den mer kjente Bera-Jarque (1981) testen, men som i tillegg korrigerer for utvalgsstørrelse.

5.7 ARIMAX

ARIMAX er en utvidet versjon av ARIMA modellen der man også inkluderer en vektor av forklaringsvariabler (X). Sagt på en annen måte er det en hybrid av ARIMA og regresjonsmodeller der regresjonsmodellen inkluderer korreksjoner for autokorrelerte feilledd ved å legge til ARMA komponenter. Forklaringsvariablene og den avhengige variabelen må være stasjonær for at modellen skal være gyldig (Karakozova, 2004). Modellen må også oppfylle kravene som beskrevet i teorien om ARIMA (se kapittel 5.6).

5.8 Vector Error-Correction modellering (VECM)

VAR og VEC modellene er systemregresjonsmodeller, altså at modellene har mer enn én avhengig variabel. Hver av verdiene til g variabler i systemet avhenger av k lags av de andre variablene og feilledd. Derfor omtales modellen ofte som en hybrid mellom univariate tidsseriemodeller og simultanlikningsmodeller (Brooks og Tsoacos 2010). Som nevnt i empiridelen blir i tillegg alle variabler behandlet som endogene i VAR/VEC-modeller, med de medfølgende fordelene dette har. VEC modellene er en variant av VAR modellen, men som tar hensyn til at variabler er integrert av samme orden og med kointegrerte forhold. Modellen er spesielt nyttig ved analyse av makroøkonomiske tidsserier. Dette er fordi man har muligheten til å analysere både kortsiktig og langsiktig informasjon i dataene ved å utnytte kointegrasjonsegenskapene. Vi har kointegrasjon når en lineær kombinasjon av $I(1)$ variabler er $I(0)$, altså stasjonære. Engle og Granger (1987) sin definisjon av kointegrasjon uttrykkes på følgende måte:

La w_t være en $k \times 1$ vektor av variabler; komponentene av w_t er integrert av orden (d, b) hvis:

- (1) alle komponentene i w_t er $I(d)$;
- (2) Det er minst én vektor av koeffisienter α slik at: $\alpha'w_t \sim (d - b)$

Mange tidsserier er ikke-stasjonære på nivåform, men beveger seg sammen over tid. Det betyr at variablene kan være bundet av et forhold på lang sikt. Et kointegrert forhold kan også ses på som et langsiktig likevekt fenomen, fordi det er mulig at de kointegrerte variablene kan avvike fra forholdet på kort sikt, men at deres sammenheng vil komme tilbake på lang sikt (Brooks, 2014).

5.8.1 Testing og estimering av kointegrerte systemer ved bruk av Johansen metode

Den vanligste testen for kointegrasjon er Johansen (1988) test. En fordel med testen er at den ikke er avhengig av en antakelse om at variablene er endogene eller eksogene i systemet. Johansen kointegrasjonstest kan uttrykkes på følgende måte:

$$y_t = v + A_1 y_{t-1} + \dots + A_p y_{t-p} + \mu_t$$

hvor y_t er en $k \times 1$ vektor av endogene variabler, v er en $k \times 1$ vektor av parametere, u_t er en $k \times 1$ vektor med normal- og uavhengig fordelte feilledd, og $A_1 - A_p$ er $k \times k$ matriser av parametere. Likningen kan skrives om til VECM form:

$$\Delta y_t = v + \Pi y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \mu_t$$

hvor $\Pi = \sum_{j=1}^{j=p} A_j - I_k$ og $\Gamma_i = -\sum_{j=i+1}^{j=p} A_j$. μ_t og v er identisk i begge likningene, og I_k er en $k \times k$ identitetsmatrise. Hvis variablene y_t er $I(1)$ har matrisen Π rang $0 \leq r < k$ hvor r er antallet lineært kointegrerte vektorer (Engle & Granger, 1987). Det langsiktige forholdet mellom y_t er bestemt av rangen til Π . Hvis vi antar at Π has redusert rang, $0 < r < k$ kan Π skrives som produktet av matrisene $\alpha\beta'$:

$$\Pi = \alpha\beta'$$

I et VECM system med fire endogene variabler og to kointegrerte vektorer vil α og β' være (4×2):

$$\Pi = \alpha\beta' = \begin{pmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \\ a_{14} & a_{24} \end{pmatrix} \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} & \beta_{14} \\ \beta_{21} & \beta_{22} & \beta_{23} & \beta_{24} \end{pmatrix}$$

Matrisen Π inneholder informasjon om de langsiktige forholdene i VECM systemet. Enders (2003) definerer β' som den langsiktige matrisen av kointegrerte parametere, og α kan tolkes som en koeffisient som viser justeringshastigheten mot likevekt.

Brooks og Tsolacos (2010) viser til at Johansen testen kan bli påvirket av lag-lengden som blir brukt i VECM, og understreker viktigheten av å forsøke å finne optimal lag lengde. Eksisterende eiendomsteori er tvetydig og lite konsistent i anbefaling av lag lengde for VECM. Forfatterne nevner to metoder som oftest blir foretrukket; krysslikningsrestriksjoner og informasjonskriterier. En test for krysslikningsrestriksjonene er sannsynlighetsrate-test (LR-test), men denne er kun gyldig når feilleddet fra hver likning er normalfordelt. Siden vi har relativt få observasjoner og kan ha problemer med normalitet velger vi å bruke informasjonskriteriene for valg av lags. Som tidligere velger vi lags som minimerer verdien på informasjonskriteriene.

5.9 Prognostisering

Vi bruker statistikkprogrammet Stata14 til å utarbeide prognosene basert på de valgte statistiske modellene. For ARIMA og ARIMAX utvikler vi statiske "ett-steg-frem" prognoser for in-sample dataene for å sjekke hvor godt modellen passer dataene. Siden statiske prognoser baserer seg på de virkelige dataene, kan de kun predikere én periode utover kjente datapunkt. Statisk ett-steg-frem prognose med én forklaringsvariabel kan defineres slik:

$$\hat{y}_f = \hat{\beta}_0 + \hat{\beta}_1 x_f + \hat{p} y_{f-1}$$

Vi utarbeider såkalte dynamiske prognoser i out-of-sample perioden for alle modellene for å kunne predikere lengre frem i tid enn et kvartal. Teknikken er også en ett-steg-frem prognose men bruker den predikerte verdien av y istedenfor den observerte verdien. Dynamisk prognose kan defineres slik:

$$\hat{y}_{f+1} = \hat{\beta}_0 + \hat{\beta}_1 x_{f+1} + \hat{p} \hat{y}_f$$

Når prognosen inkluderer en forklaringsvariabel x_t så kan vi kun prognostisere i perioder hvor vi har observasjoner for forklaringsvariabelen. Dette er tilfellet i ARIMAX modellen hvor det blir brukt eksogene forklaringsvariabler. I vanlige ARIMA modeller kan vi prognostisere så langt frem i tid som vi måtte ønske da de ikke er avhengige av eksogene variabler (StataCorp, 2013).

5.10 Metode for evaluering av prognosemodellene

Når man skal evaluere prognosenes yteevne er det vanlig å prognostisere hele out-of-sample perioden før man sammenligner prognosene med de virkelige verdiene. Videre aggregerer man opp prognosefeilene. Prognosen med minst prognosefeil anses som den beste. Eksempelvis er prognosefeilene for observasjon i definert som forskjellen mellom den virkelige verdien for observasjon i og prognosen til observasjonen. Siden prognosefeilene vil være positive (negative) hvis prognosen er for lav (høy), kan man ikke direkte summere prognosefeilene, fordi de positive og negative feilene vil trekke i motsatt retning. Derfor er det vanlig å kvadrere eller bruke absoluttverdi av feilene for å unngå dette (Brooks, 2014). Vanlige mål på prognosefeil er ”mean absolute deviation (MAD)”, ”mean absolute percentage error” (MAPE) og ”root mean squared error” (RMSE) og ”Theils U2”. Formelen for MAD er:

$$MAD = \frac{1}{n} \sum_{i=1}^n |A_i - P_i|$$

MAD gjengir gjennomsnittet av forskjellen mellom de faktiske (A_i) og de predikerte verdiene (P_i) i absolutte termer. Formelen for MAPE er:

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{A_i - P_i}{A_i} \right|$$

Siden MAD er skala-avhengig brukes MAPE som regel når man prognostiserer variabler på nivåform, mens MAD brukes hvis vi prognostiserer for eksempel vekst i leiepris.

MSE kan uttrykkes som:

$$MSE = \frac{1}{n} \sum_{i=1}^n (A_i - F_i)^2$$

Formelen for RMSE er:

$$RMSE = \sqrt{MSE}$$

MSE måler gjennomsnittet av kvadratene av prognosefeilene, altså forskjellen mellom estimatet og variabelen som blir estimert. I følge Brooks og Tsolacos er RMSE et godt mål på yteevnen til prognosemodellene når man sammenligner prognosemodeller basert på samme datasett, men må ikke brukes når man sammenligner datasett som bruker forskjellig skala (se Chatfield (1988) og Collopy og Armstrong (1992)). RMSE tillegger også større straff for store prognosefeil. Når den avhengige variabelen har store fluktuasjoner og vendepunkter er RMSE et bedre mål enn MSE og MAPE. Dette er fordi RMSE vil disproporsjonalt straffe store feil når prognosen ikke fanger opp fluktuasjonene. Siden vi i denne utredningen bruker data med store fluktuasjoner og flere vendepunkter velger vi å bruke RMSE som sammenligningsgrunnlag for prognosemodellene.

Den siste prognoseparameteren vi vil benytte i utledningen er Theils U2 koeffisient som sammenligner prognoseverdiene mot en naiv modell. Theils U2 er definert ved:

$$Theils\ U2 = \left(\frac{MSE}{MSE^{NAIV}} \right)^{0,5}$$

Theils U2 måler treffsikkerheten av prognosen ved det kvadratiske tapskriteriet. Hvis U2 er mindre enn 1 presterer modellen bedre enn den naive modellen, og vice versa (Brooks og Tsolacos, 2010).

6. Analyse og resultater

I dette kapittelet vil det bli tatt utgangspunkt i metodedelen med formål å konstruere økonomiske modeller for å predikere fremtidige leiepriser for kontorbygg i Oslo-området. Det vil bli tatt utgangspunkt i variablene vi tidligere har diskutert i teori og datadelen.

I den første delen vil vi gjennomføre en stasjonærhetsanalyse av variablene. Med utgangspunkt i kontorleiepriser for Oslo starter vi med en naiv random-walk modell for å predikere et utgangspunkt som vi kan måle de andre prognosemodellene opp mot. Deretter vil vi definere en ARIMA modell, utarbeide prognoser og evaluere yteevnen. Videre vil vi inkludere forklaringsvariabler i modellene ARIMAX og VECM, og deretter utarbeide prognoser for disse modellene. Vi vil også evaluere yteevnen relativt til den naive modellen.

6.1 Stasjonærhetsanalyse

Tidligere i oppgaven nevnte vi at stasjonærhet er viktig for den økonometriske analysen av tidsserier for å forhindre spuriøse resultater. I våre data har vi sjeldent stasjonærhet på nivåform og seriene har en tendens til å ha trendlignende utvikling eller inneholde sykler. I det følgende avsnittet vil vi presentere ulike ADF-tester for enhetsrøtter hvor vi tar utgangspunkt i et 5 prosent forkastningsnivå. Testene blir gjennomført for flere antall lags og vi presenterer lag 2, 4 og 6 i analysen. Det blir også testet med og uten trend og konstant. Testene er oppsummert i Tabell 2.

Leiepriser (lnLP)

Vi forventet at variabelen for leiepriser er integrert av første orden ved å undersøke grafen i datadelen. I følge ADF-testene kan vi ikke forkaste nullhypotesen om ikke-stasjonærhet på nivåform. I serien for vekst i leiepris viser ADF-testen ingen tegn til enhetsrøtter og vi forkaster nullhypotesen om ikke-stasjonærhet. Vi konkluderer med at leiepriser er integrert av første orden, $I(1)$.

BNP (lnBNP)

BNP har en tydelig deterministisk trend i vår analyseperiode og vi forventer at den er integrert av første orden. I følge ADF-testen kan vi ikke forkaste nullhypotesen om ikke-stasjonærhet på nivåform. I serien for vekst i BNP kan vi forkaste nullhypotesen om ikke-stasjonærhet på 5% nivå for majoriteten av lags hvor trend ikke er inkludert. Siden serien for vekst i BNP ikke har noen

tydelig trend, og det er ellers er sterke indikasjoner på at serien er stasjonær, konkluderer vi med at BNP er integrert av første orden $I(1)$.

Arbeidsledighet (ARBL)

I serien for arbeidsledighet mistenkte vi enhetsrøtter på nivåform ved å undersøke grafen. ADF-testen forkaster ikke nullhypotesen på 5% nivå. Serien for endring i arbeidsledighet viser ingen tegn til enhetsrøtter i seriene hvor trend ikke er inkludert. Siden serien ikke har noen tydelig trend, og det ellers er sterke indikasjoner på at serien er stasjonær, konkluderer vi med at arbeidsledighet er integrert av første orden $I(1)$.

Lange renter (Lange renter)

I serien for lange renter forventet vi at serien ikke er stasjonær på nivåform. Vi kan ikke forkaste nullhypotesen om ikke-stasjonærhet ved gjennomføring av ADF-testen. Differensierer vi serien en gang kvartalsvis viser ADF-testen ingen tegn til enhetsrøtter og vi forkaster nullhypotesen om ikke-stasjonærhet. Vi konkluderer med at lange renter er integrert av første orden, $I(1)$.

Arealledighet (ARL)

Arealledighet har en markant topp i 2003 før den reduseres de neste årene. Ellers ser serien ut til å følge sykliske bevegelser. Vi forventet at arealledighet ikke var stasjonær og det kan se ut som det er et strukturelt brudd i serien. I tillegg er variabelen interpolert som gjør at den inneholder glattningseffekter. Et strukturelt brudd er definert som en langsiktig endring i egenskapene for en tidsserie. En kritikk av standard ADF-tester for stasjonærhet er at de kan gi feilaktige konklusjoner dersom strukturelle brudd ikke tas hensyn til i estimeringen (Brooks, 2014). Det vil si at Dickey-Fuller testene kan konkludere med at serien inneholder enhetsrøtter til tross for serien er integrert på lavere orden før og etter brudd. ADF-testene for arealledighet viser at serien ikke er stasjonær på nivå eller i første orden og vi ønsker derfor å undersøke nærmere om variabelen arealledighet inneholder strukturelle brudd.

Et annet problem ved brudd i serien er at vi kan få feil i parameterne. Det kan være i konstantleddet, koeffisienten eller variansen. For å teste om serien inneholder strukturelle brudd gjennomfører vi en Wald test for ukjente brudd for arealledighet med utgangspunkt i en enkel dynamisk modell: $y_t = \alpha + \phi y_{t-1} + \epsilon_t$

Nullhypotesen for Wald testen er at det ikke er et skift i noen av parameterne, og blir beregnet for alle perioder i tidsserien. Perioden med høyest Wald verdi ($sWald$) blir plukket ut

og kan indikere et mulig strukturelt brudd. Dersom Wald parameteren overgår den kritiske verdien forkaster vi nullhypotesen om ingen strukturelle brudd. Vi aksepterer da alternativhypotesen om at det er et skifte i parameterne for det kvartalet som er identifisert av Wald testen. For at det skal være tilstrekkelig med observasjoner på hver side av testen trimmes utvalget. Vi benytter standard trimming fra Stata14 med 15% på hver side.

For arealledighet på nivåform forkaster vi nullhypotesen om ingen strukturelle brudd for bruddato 2003q3 (p-verdi = 0,000). For serien vekst i arealledighet forkaster vi nullhypotesen om ingen strukturelle brudd for bruddato 2002q4 (p-verdi = 0,000). I følge Figur 8 er dette toppunktene for arealledighet i de respektive seriene. Vi kan også lese av at serien for vekst i arealledighet treffer nullaksen i 2001q2, stiger til toppunkt i 2002q3, faller igjen og krysser nullaksen i 2003q4. For å sjekke om dette mulige bruddet påvirker parameterne til vekst i arealledighet inkluderer vi en dummyvariabel i perioden 2001q2 til og med 2003q4 som tar verdien 1 i denne perioden³. I analysen tar vi også med en interaktiv dummy for arealledighet i samme perioden ved å multiplisere vekst i arealledighet med dummyvariabelen. Empiriske resultater og videre tolkning vil bli presentert i kapittel 6.4.

Forventet nybygg (lnFNBYB)

Serien for forventet nybygging inneholder flere store sykler uten tydelig trend, og vi forventer at serien ikke er stasjonær på nivåform. Forventet nybygging er i likhet med arealledighet interpolert som fører til glattningseffekter i serien. ADF-testene forkaster ikke nullhypotesen om ikke-stasjonærhet på 5% nivå. Serien for vekst i forventet nybygging har fortsatt store sykler men varierer i større grad rundt gjennomsnittet. ADF testen konkluderer med at vi kan forkaste nullhypotesen om ikke-stasjonærhet på 5% nivå. Vi konkluderer med at log forventet nybygg er integrert av første orden I(1).

³ Se kapittel 4.3.1 for ytterligere info om arealledighet i perioden 2001q3 – 2003q4.

Tabell 2: ADF test for stasjonæritet

Tabellen gjengir resultatet for ADF-test for forklaringsvariablene i ulike lag strukturerer, med trend, konstant og uten konstant i perioden 1998q1 til 2014q4.

H_0 : serien inneholder en enhetsrot

H_A : serien er stasjonær

Forkastningsnivå: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$

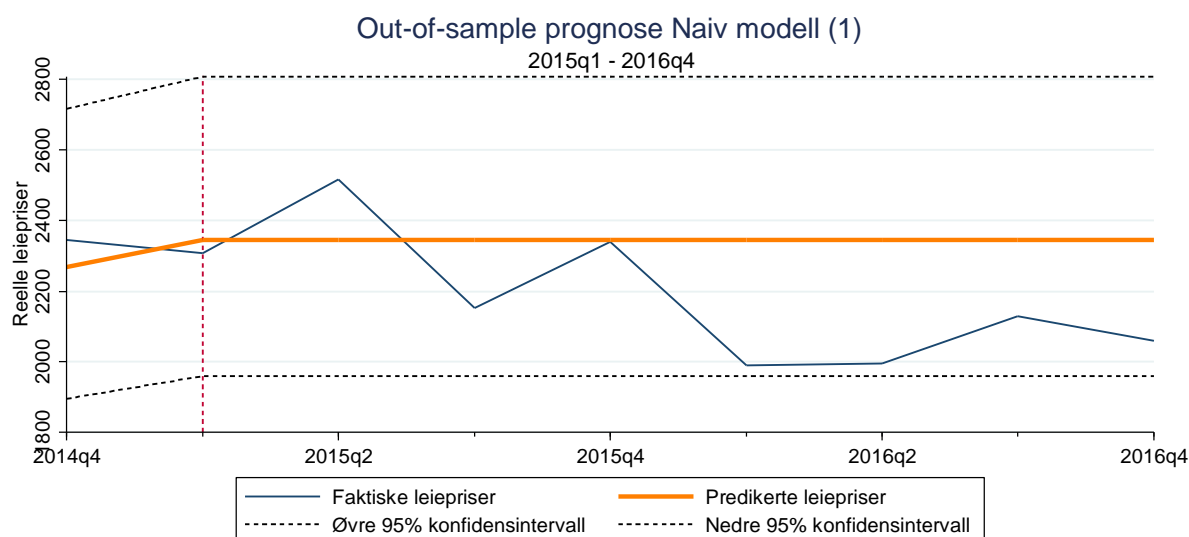
Δ indikerer at variabelen er blitt differensiert en gang kvartalsvis. Δ_4 indikerer at variabelen er blitt differensiert en gang årlig. Ln viser til at serien er log transformert, variablene uten ln er målt i rate på nivåform og er derfor ikke log transformert.

Variabel	Antall lags i ADF	Konstant, Trend	Konstant	Uten konstant	Variabel	Antall lags i ADF	Konstant, Trend	Konstant	Uten konstant
lnLP	2	-2,952	-2,522		Δ lnLP	2	-6,588***	-6,646***	-6,618***
	4	-2,663	-2,199			4	-5,063***	-5,107***	-5,020***
	6	-2,079	-1,706			6	-4,226***	-4,264***	-4,106***
lnBNP	2	-1,968	-0,977		Δ_4 lnBNP	2	-4,237***	-4,152***	-3,258***
	4	-2,873	-1,152			4	-2,965	-2,828*	-2,038**
	6	-1,863	-1,143			6	-3,385*	-3,123**	-1,961**
ARBL	2	-2,256	-2,160	-0,047	Δ_4 ARBL	2	-2,892	-3,021**	-3,035***
	4	-2,548	-2,480	-0,083		4	-2,813	-2,899*	-2,913***
	6	-3,400*	-3,332**	-0,156		6	-3,430*	-3,506**	-3,520***
RRTI	2	-3,479*	-1,225	-1,220	Δ RRTI	2	-8,566***	-8,599***	-8,489***
	4	-2,641	-0,663	-1,338		4	-4,468***	-4,479***	-4,326***
	6	-2,891	-0,640	-1,315		6	-4,497***	-4,511***	-4,216***
ARL	2	-1,185	-1,255	-0,185	Δ ARL	2	-2,174	-2,166	-2,161**
	4	-2,428	-2,477	-0,174		4	-2,898	-2,894*	-2,899***
	6	-2,284	-2,333	-0,113		6	-1,822	-1,799	-1,796*
lnFNYB	2	-1,863	-1,826	-0,349	Δ lnFNYB	2	-3,000	-2,996**	-3,012***
	4	-2,147	-2,131	-0,322		4	-3,368*	-3,357***	-3,376***
	6	-2,133	-2,118	-0,300		6	-3,275*	-3,266**	-3,286***
Augmented Dickey-Fuller test									
Kritisk verdi: *** $p < 0,01$ ** $p < 0,05$ * $p < 0,10$									

6.2 Naiv modell (1)

En naturlig start for å evaluere prognosemodeller er å predikere en naiv modell som kan benyttes som en benchmark for de mer sofistikerte modellene. Den mest naive modellen er en såkalt random-walk modell hvor man antar at leieprisen i siste periode forblir konstant i hele prognoseperioden.

Siden leiepriser ikke er stasjonær på nivåform differensierer vi serien en gang kvartalsvis. Den dynamiske prognosemodellen blir da en ARIMA(p,d,q=0.1.0) modell hvor d=1 indikerer at serien er differensiert av første orden. AR(0) betyr at den autoregressive komponenten er lik 1. Vi tar heller ikke hensyn til trender i in-sample perioden og utelukker derfor konstantleddet i prognosemodellen (Nau, 2017).



Figur 10: Dynamisk prognose out-of-sample - naiv modell (1)

Prognose for vekst i leiepriser konvertert til reelle leiepriser. Prognoseperiode fra 2015q1-2016q4.

Følgende modell er estimert: $\Delta \ln LP_t = p \Delta \ln LP_{t-1} + \varepsilon_t$; $p = 1$ (1)

Vi ønsker å predikere utviklingen i fremtidige leiepriser og følgelig konverterer vi prognosene fra vekst i leiepriser til reelle leiepriser på nivå for at resultatet skal bli mer informativt for leseren. I Figur 10 er prognosen fremstilt grafisk sammenlignet med den faktiske leieprisutviklingen i perioden fra 2015q1 til 2016q4. Prisutviklingen er konstant på 2364,- gjennom hele toårsperioden. Den treffer godt i 1. og 4. kvartal i 2015, men bommer systematisk i de andre periodene.

For at vi skal kunne analysere resultatene fra prognosemodellene for kortere perioder prognostiserte vi også to ettårsperioder for 2015 og 2016. I Tabell 3 presenterer vi prognosefeilparameterne som ble beskrevet i teorikapittelet. Generelt er prognosene for de første fire kvartalene bedre enn de siste 4 kvartalene. Toårsperioden gir også bedre prognoser enn ettårs prognosen for 2016. Den naive modellen vil utgjøre utgangspunktet som vi måler de mer sofistikerte modellene opp mot. En komplett sammenligning av alle modellene vil bli presentert i kapittel 7 og den grafiske fremstillingen for ettårsprognosen i 2016 er presentert i appendiks C.

Tabell 3: Prognosefeil- naiv modell (1)

Denne tabellen rapporterer prognosefeilparametere Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE) og Mean Absolute Percentage Error (MAPE) for den naive modellen (1).

Naiv modell	Fra	Til	MAD	RMSE	MAPE
	2015q1	2015q4	102,220	130,615	4,428
<i>Dynamisk</i>	2016q1	2016q4	296,526	301,859	14,594
	2015q1	2016q4	202,150	236,118	9,647

6.3 ARIMA modell (2)

I dette kapittelet bruker vi den noe mer sofistikerte ARIMA-modellen for å prognostisere fremtidige leiepriser. Det vil bli tatt utgangspunkt i Box-Jenkins (1976) fremgangsmåte for å definere den optimale modellen. AR prosessen gjenspeiler tidligere leiepriser og MA prosessen fanger opp tidligere bevegelser i residualene. For å kunne sammenligne resultatene med prognosene i den naive modellen benytter vi den samme estimeringsmetoden og prognoseperioden. Det vil si en estimeringsperiode mellom 1998q1 til 2014q4 og prognoseperiode mellom 2015q1 til 2016q4.

6.3.1 Modellspesifikasjon

Tabell 4 viser korrelegrammet for leiepriser og vekst i leiepriser hvor det er inkludert 10 lags. For leiepriser synker autokorrelasjonen gradvis fra første ordens autokorrelasjon på 0,64 til og med lag 10. Ljung Box Q* test med 10 lags gir en verdi mellom 29 og 95 (p-verdi = 0,0000) som er høyst signifikant, det bekrefter et sterkt autokorrelasjonsmønster i serien og vi forkaster nullhypotesen om hvit støy i feilledet. Delvis autokorrelasjon viser en høy topp på lag 1 fulgt av en sterk reduksjon i lagene etterpå. Dette indikerer en svært vedvarende autoregressiv struktur i serien.

For serien vekst i leiepriser er førsteordens autokorrelasjon negativ med -0,28 og varierer rundt null i de påfølgende lag periodene. Det tyder på at transformasjonen har gjort serien mindre autokorrelert sammenlignet med serien på nivå form. Ljung Box Q* test med 10 lags gir en verdi mellom 6 og 14 (p-verdi = 0,017 tom 0,159) som kun er signifikant på første lag på 5% nivå. Det indikerer at vi har fått bort noe av problemet med autokorrelasjonsmønstre i residualene og at det kun er indikasjoner på autokorrelasjon i første lag. Vi konkluderer med at vi må inkludere minst én AR eller MA prosess for at modellen ikke skal inneholde autokorrelasjon i residualene. Vi observerer også at det ikke er tydelige tegn til sesongvariasjoner i Tabell 4 for autokorrelasjon og delvis autokorrelasjon.

Tabell 4: Autokorrelasjon og delvis autokorrelasjon for leiepris og vekst i leiepris

Tabellen rapporterer autokorrelasjon (AC) og delvis autokorrelasjon (PAC) for leiepriser (lnLP) og vekst i leiepriser (Δ lnLP) i perioden 1998q1 – 2014q4. Q er Ljung-Box test for hvit støy i feilleddene.

Lag	lnLP				Δ lnLP			
	AC	PAC	Q	Prob>Q	AC	PAC	Q	Prob>Q
1	0,643	0,670	29,374	0,000	-0,284	-0,285	5,735	0,017
2	0,512	0,179	48,242	0,000	-0,015	-0,112	5,752	0,056
3	0,397	0,033	59,797	0,000	-0,097	-0,153	6,435	0,092
4	0,378	0,121	70,428	0,000	0,126	0,040	7,613	0,107
5	0,288	-0,061	76,712	0,000	-0,145	-0,133	9,210	0,101
6	0,246	0,058	81,358	0,000	-0,109	-0,223	10,126	0,119
7	0,265	0,182	86,855	0,000	0,115	-0,003	11,155	0,132
8	0,239	-0,006	91,376	0,000	-0,042	-0,094	11,293	0,186
9	0,226	0,064	95,493	0,000	0,184	0,174	14,024	0,122
10	0,063	-0,225	95,820	0,000	-0,061	0,083	14,325	0,159
Periode	1998Q1 - 2014Q4							

I henhold til analysene i forrige avsnitt tar vi utgangspunkt i vekst i leiepris. Det vil si at utgangspunktet er en ARIMA(p.d.q=0.1.0) hvor d=1 refererer til første ordens differensiering kvartalsvis for leiepriser. I kapittel 4.1 observerte vi at vekst i leiepriser varierer rundt null og i Tabell 1 rapporterte den oppsummerende statistikken at gjennomsnittet for vekst i leiepris er 0,68%, med et standardavvik på 9,22%. For å ta hensyn til dette inkluderer vi derfor ikke et konstantledd i ARIMA modellen. For å velge antall AR og MA prosesser som optimerer modellen så tester vi for flere mulige strukturer og benytter oss av informasjonskriteriene BIC og AIC til å velge den optimale modellen in-sample. De beste modellene bruker vi videre til å prognostisere out-of-sample. Se Tabell 5 for resultatene fra de ulike ARIMA strukturene.

BIC kriteriet foreslår ARIMA(p.d.q=0.1.1) som den beste modellen og ARIMA(1.1.0) som det nest beste modellen. Negativ autokorrelasjon i første lag kan tyde på at serien er svakt overdifferensiert og ifølge Nau (2017) er da MA(1) det foretrukne valget. AIC kriteriet foreslår en større struktur, noe som ikke er overraskende da AIC har en tendens til å overtilpasse modellen. AIC foretrekker ARIMA(2.1.3) og det nest beste valget er ARIMA(0.1.1) som er den samme strukturen som BIC foreslår.

Tabell 5: Sammenligning av informasjonskriteriene ARIMA modell (2)

Tabellen oppsummerer informasjonskriteriene Akaike informasjonskriterie (AIC) og Schwarz Bayesian informasjonskriterie (BIC) for vekst i leiepriser ved flere ulike ARIMA strukturer for perioden 1998q1 til 2014q4. Markert er den modellen som er foretrukket av informasjonskriteriet. Følgende modell ble estimert:

$$\Delta \ln LP_t = \phi_1 \Delta \ln LP_{t-1} + \dots + \phi_p \Delta \ln LP_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

BIC	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)
AR(0)	-127,6	-130,2	-126,6	-123,4	-119,3
AR(1)	-128,8	-127,4	-123,7	-119,8	-115,7
AR(2)	-125,4	-123,7	-119,7	-122,5	-120,7
AR(3)	-122,7	-118,7	-115,5	-118,6	-119,3
AR(4)	-118,6	-114,5	-111,8	-117,1	-117,1
AIC					
AR(0)	-129,8	-134,7	-133,3	-132,2	-130,4
AR(1)	-133,3	-134,1	-132,6	-130,9	-129,1
AR(2)	-132,1	-132,6	-130,8	-135,8	-134
AR(3)	-131,6	-129,8	-128,8	-134,2	-134,8
AR(4)	-129,7	-127,8	-127,3	-132,6	-134,8

6.3.2 In-sample analyse ARIMA

I Tabell 6 er de empiriske resultatene for modell (2) presentert. Modell (2a) rapporterer at MA(1) er signifikant på 1% nivå med negativ verdi på -0.364. En ARIMA(0.1.1) modell uten konstant er en enkel eksponentiell glattings modell. Den bruker gjennomsnittet av de siste observasjonene til å filtrere ut støy og øke treffsikkerheten til det estimerte gjennomsnittet (Nau, 2017). I modell (2b) er AR(1) signifikant på 5% nivå med negativ verdi på -0,278. Standardavviket i residualene (sigma) øker noe, men det er ingen indikasjon på autokorrelasjon og enhetsrøtter i residualene ifølge Box-Ljung Q* test og ADF testen. Totalt sett ser modellene ut til være robuste.

I modell (2c) inkluderes flere AR og MA prosesser. AR koeffisientene er signifikante på 1% nivå men koeffisientene er nå mer negative. MA koeffisientene er ikke signifikante og positive. Det fører videre til at AR og MA prosessene trekker i motsatt retning og hvis vi summerer koeffisientene blir verdien omtrent én. Dette kan være tegn på at modellen er ustabil (Nau 2017). Vi observerer også at standardavviket til residualene (sigma) ikke er signifikant. Test for hvit støy, normalitet og enhetsrøtter i residualene blir godkjente for modellen. På bakgrunn av at MA prosessene ikke er signifikante, at det er tegn til ustabilitet, og at modellen har en mer kompleks struktur velger vi å ikke ta med modell (2c) videre i analysen.

Tabell 6: Empiriske resultater ARIMA modell (2)

Denne tabellen rapporterer resultatene fra estimering av ARMA modell (2) for vekst i leiepriser i Oslo-området ved bruk av kvartalsvis tidsseriedata fra 1998q1 til 2014q4. Følgende ARMA modell er estimert:

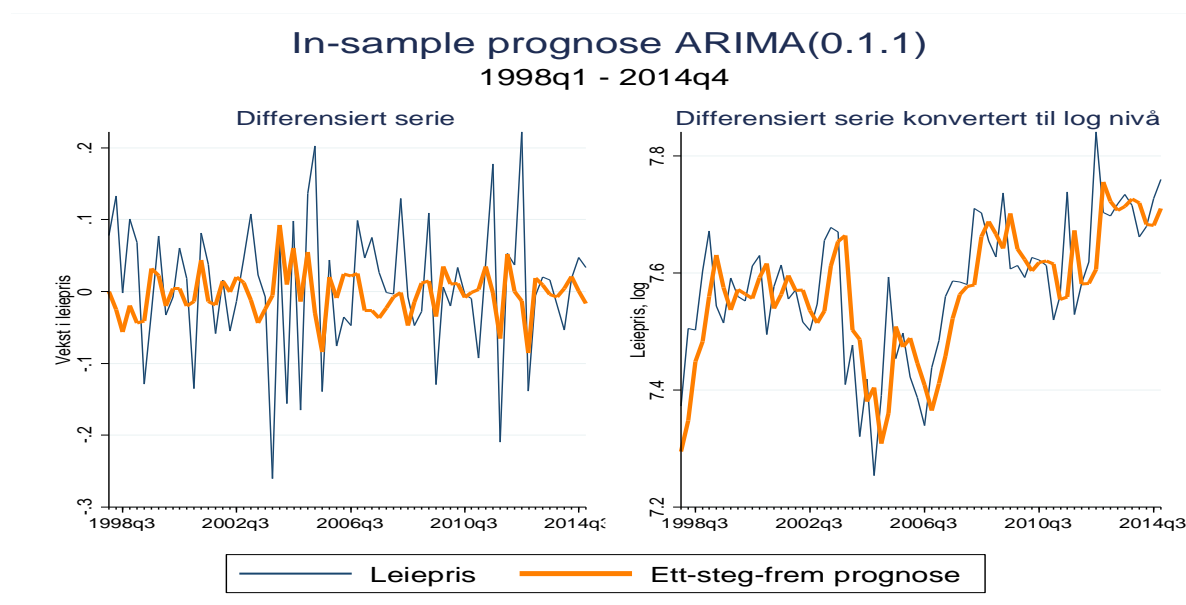
$$\Delta \ln LP_t = \phi_1 \Delta \ln LP_{t-1} + \dots + \phi_p \Delta \ln LP_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (2)$$

Δ betyr at variabelen er differensiert av første orden kvartalsvis ($d=1$), p er antall AR prosesser og q er antall MA prosesser «ARIMA(p,d,q)». Sigma gjengir det estimerte standardavviket for hvit støy i feilledet ε . AIC og BIC er mål på relativ kvalitet for modellene. ll er log likelihood, LR-test er likelihood ratio test for å sammenligne modellene. Box-Ljung Q^* er portmanteau test for hvit støy i residualene. Sktest er en versjon av Jarque-Bera sin test for normalitet i residualene og ADF 4.lag er test for enhetsrøtter i residualene.

Modell	(2a)	(2b)	(2c)	(2d)
ARIMA	(0.1.1)	(1.1.0)	(2.1.3)	(0.1.1) D(q)
AR(1)		-0.278* (-2.21)	-0.854*** (-4.67)	
AR(2)			-0.717** (-3.17)	
MA(1)	-0.364** (-2.76)		0.623 (0.00)	-0.366** (-2.82)
MA(2)			0.387 (0.00)	
MA(3)			-0.531 (-0.00)	
q1				0.00590 (0.22)
q2				0.0307 (1.49)
q3				0.0179 (0.79)
q4				-0.0311 (-1.35)
sigma	0.0872*** (14.10)	0.0881*** (13.77)	0.0791 (0.00)	0.0840*** (11.29)
AIC	-134.7	-133.3	-135.8	-131.7
BIC	-130.2	-128.8	-122.5	-118.4
ll	69.33	68.64	73.92	71.85
LR-test (χ^2)	NA	NA	9.18(2a)	5.04(2a)
Justert R^2	0.10	0.08	0.22	0.16
Box-Ljung Q^*	35.31	38.75	27.37	33.15
Sktest (χ^2)	3.46	3.28	5.90	1.46
ADF 4.lag	-4.863***	-5.115***	-4.676***	-5.003***
z-stat i parentes: * $p < 0,05$, ** $p < 0,01$, *** $p < 0,001$				

I modell (2d) inkluderer vi dummyvariabler for hvert kvartal i en enkel ARIMA(0.1.1) modell. Ingen av dummyvariablene blir signifikante, noe som bekrefter våre forventninger fra kapittel 4.1 om at det ikke er noen tydelige tegn til sesongeffekter i serien.

I Figur 11 er de predikerte in-sample verdiene for modell (2) presentert sammen med den faktiske utviklingen. Modell (2a) ser ut til å følge den faktiske utviklingen for vekst i leiepriser relativt dårlig. Den får ikke med seg de høye toppene eller de dype dalene i serien. Enkelte perioder treffer den heller ikke helt på retningen. Den differensierte serien forklarer omtrent 10% av bevegelsene i leieprisene. Konverterer vi leieprisene til log nivå ser modellen ut til å treffe bedre både på retning, topper og bunner. Forklaringskraften er omtrent 48%. Metoden for in-sample prediksjonene er såkalt ett-steg-frem prognoser. Det betyr at for hvert kvartal som blir predikert blir forrige periodes virkelig verdi hensyntatt i neste periodes prognose. Grafisk in-sample prognose for modell (2b) er presentert i appendiks B.



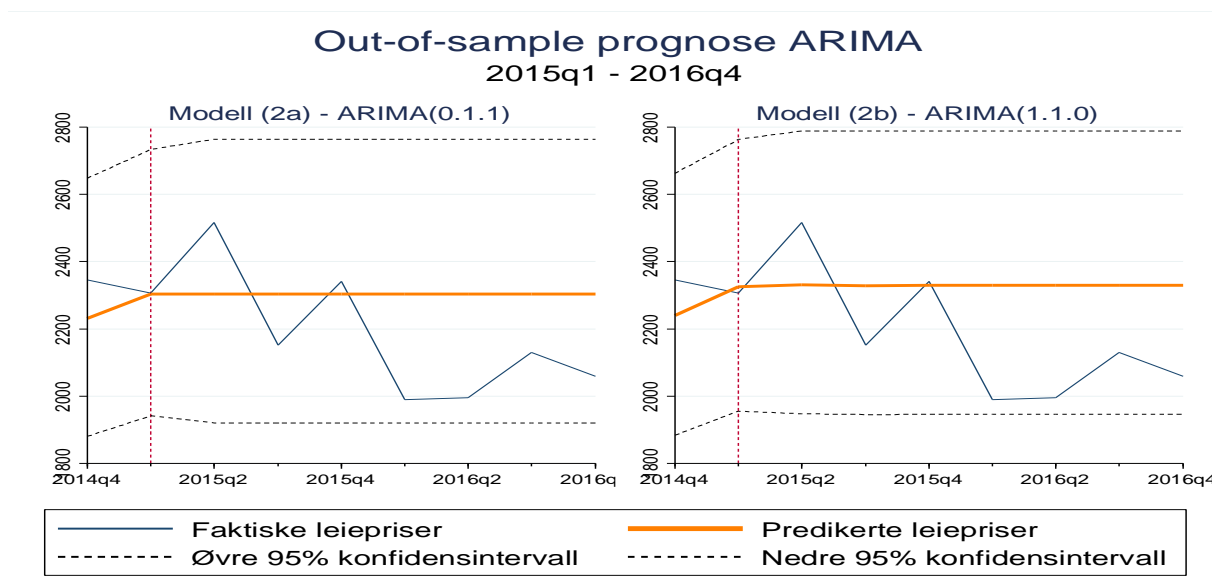
Figur 11: In-sample prognose ARIMA modell (2a)

Denne figuren viser in-sample prognose for vekst i leiepriser og leiepriser for kontoreiendom i Oslo-området sammenlignet med den faktiske utviklingen i perioden 1998q1 – 2014q4. Venstre graf viser utviklingen i vekst i leiepris. Høyre graf viser utviklingen i leiepris transformert tilbake til log-nivå.

6.3.3 Out-of-sample resultater ARIMA

I prediksjon av leiepriser out-of-sample benytter vi såkalte dynamiske prognoser. Empirien sier at ARMA modeller kan være gode verktøy i korte perioder (Brooks og Tsolacos 2010). Derfor lager vi tre prognoseperioder. En toårsperiode 2015q1-2016q4 og to ettårsperioder for 2015 og 2016. For sistnevnte modellerer vi frem til 2015q4 og deretter utarbeider vi prognoser for de neste 4 kvartalene.

For modell (2a) og (2b) predikerer vi verdien for vekst i leiepris og konverterer verdiene tilbake til nivåform for reelle leiepriser. I Figur 12 vises dynamisk prognose med konfidensintervall på to standardavvik. Modell (2a) treffer veldig godt i første kvartal 2015 og holder den samme verdien gjennom hele perioden. Det gjør at den ikke får med seg fluktusjonene fra kvartal til kvartal. At prognosen er flat gjennom hele perioden er forventet da en enkel eksponentiell glattings modell (arima(0.1.1) ikke tar hensyn til trender, sesongvariasjoner eller andre mønstre i serien. Sagt på en annen måte tar modellen utgangspunkt i siste kvartals vekst i leiepris, og en justeringsfaktor fra prognosens feilledd fra forrige periode. Siden de virkelige verdiene ikke er kjent ex-ante vil det ikke være noe feilledd som prognosen kan justere seg mot og flater dermed ut i de neste periodene (Nau, 2017).



Figur 12: Dynamisk prognose out-of-sample ARIMA modell (2)

Figuren viser toårig prognosene for reelle leiepriser på nivåform for modell (2a) og (2b). Prognosen er gjennomført på log differensierte leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2015q1-2016q4.

Modell (2b) er også tilnærmet lineær gjennom hele estimeringsperioden. Den treffer ikke like godt som (2a) i første kvartal 2015 og predikerer en litt høyere pris gjennom hele perioden. Heller ikke denne modellen egner seg til å predikere endringer fra periode til periode. I begge modellene holder den faktiske utviklingen seg innenfor 2 standardavvik av prognosen. Den grafiske fremstillingen for ettårsprognoser i 2016 er presentert i appendiks C.

For å evaluere prognosene måler vi modellene opp mot den naive tabellen (se Tabell 7). Vi observerer at modell (2a) presterer bedre enn den naive modellen for toårsperioden og etterårsperioden i 2016. For ettårsprognosen i 2015 scorer den kun bedre på MAD og MAPE. Studerer vi RMSE scorer den omtrent likt som den naive modellen.

Modell (2b) presterer bedre enn den naive modellen i alle periodene, men som med modell (2a) er det en relativt begrenset forbedring. Dette gjenspeiles også i Theils U2 der en verdi under 1 betyr at prognosen er bedre enn den naive modellen, og vice versa. Måler vi modell (2b) opp mot modell (2a) er den bedre i den ettårige prognosen for 2016, men presterer dårligere i 2015 og den toårige prognosen. Totalt sett ser ikke ARIMA modellen til å være veldig godt egnet til å predikere fremtidige leiepriser da den ikke klarer å fange opp endringene fra kvartal til kvartal.

Tabell 7: Prognosefeil - modell (2a) og (2b)

Tabellen rapporterer prognosefeilene relativt til den naive modellen. Negative verdier (grønn farge) betyr at modellen har prestert bedre enn den naive modellen. Positive verdier (rød farge) betyr at den naive modellen har prestert best. Prognosefeilparameterne er Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) og Theils U2.

ARIMA - Naiv modell	Fra	Til	MAD	RMSE	MAPE	Theils U2
(2a): (0.1.1)						
Dynamisk	2015q1	2015q4	-1,063	1,315	-0,122	1,010
	2016q1	2016q4	-33,009	-32,355	-1,616	0,893
	2015q1	2016q4	-21,679	-26,256	-1,096	0,889
(2b): (1.1.0)						
Dynamisk	2015q1	2015q4	-4,376	-1,889	-0,218	0,986
	2016q1	2016q4	-48,694	-48,008	-2,387	0,841
	2015q1	2016q4	-10,601	-11,250	-0,521	0,952

6.4 ARIMAX modell (3)

I denne delen ønsker vi å utvide ARIMA modellen ved å inkludere eksogene forklaringsvariabler for å undersøke om det bedrer prognosene. AR komponenten fanger opp tidligere effekter fra vekst i leiepris og MA komponenten fanger opp tidligere tilfeldige bevegelser i residualene. X-komponenten inkluderer en vektor av forklaringsvariabler som vil påvirke leieprisen. Variablene vi inkluderer er blitt presentert under datadelen og stasjonærhetsanalysen. Som tidligere tar vi utgangspunkt i estimeringsperiode fra 1998q1 til 2014q4 og prognoseperiode fra 2015q1 til 2016q4.

6.4.1 Modellspeifikasjon

For å tilpasse modellen til å forklare leieprisutviklingen velger vi en såkalt generell til spesifikk fremgangsmåte (Karakozova, 2004). I kapittel 6.3 konkluderte vi med at ARIMA spesifikasjon (0.1.1) for vekst i leiepriser er den foretrukne modellen. I denne delen blir denne modellen benyttet som en benchmark for ARIMAX modellene, hvor vi sammenligner informasjonskriteriene ved å utvide modellen med eksogene forklaringsvariabler. Det vil også bli tatt høyde for signifikansnivå og robustheten til de utvidete modellene. Signifikante variabler som forbedrer AIC og BIC vil bli inkludert i ARIMAX modellen. Er variablene ikke-signifikante og/eller bidrar til at modellen blir ustabil forkaster vi variabelen. Vi tester den optimale modellen for ulike ARIMA strukturer for å sjekke om det forbedrer informasjonskriteriene. Avslutningsvis prognostiseres out-of-sample perioden og prognosen sammenlignes mot den naive modellen.

En av forutsetningene for ARIMAX modellen er at alle variablene må være stasjonære enten på nivå eller av første orden $I(1)$ og at residualene følger en hvit støy prosess. Siden arealledighet ikke er stasjonær og inneholder mulige strukturelle brudd i perioden 2001q2 til 2003q4 inkluderer vi en dummyvariabel i bruddperioden. Dette blir presisert nærmere i neste delkapittel. Dersom koeffisienten til vekst i arealledighet blir signifikant og modellen blir mer robust, inkluderer vi variabelen i videre analyse.

For å tilpasse modellen til å forklare utviklingen i leiepriser er det kritisk å ta hensyn til antall lags i de eksogene forklaringsvariablene. Alle forklaringsvariablene er lagget to ganger for å ta høyde for at leiekontraktene gjennomsnittlig trer i kraft ca. 4-8 måneder etter de er signert. Leieprisene i datasettet blir som nevnt tidligere registrert når leietaker flytter inn, og ikke når kontrakten blir signert. Dette er også i tråd med tidligere forskning på leiepriser i kontormarkedet (Holland et al., 2000). I tillegg er alle variablene testet opp til fem lags for å sjekke om det forbedrer modellen ved å sammenligne signifikansnivå og informasjonskriteriene AIC og BIC.

Det vil bli gjennomført flere diagnostikktester som må oppfylles før vi kan konkludere med at modellene er tilstrekkelig robuste til å prognostisere fremtidige leiepriser. Hvis AR eller MA komponentene summerer til $|1|$ kan det indikere at komponentene inneholder enhetsrøtter. Vi sjekker dette med Stata14 sin stabilitetstest «aroots» (StataCorp, 2013). Blir eigenverdien til AR eller MA prosessen $|1|$ konkluderer vi med at modellen er ustabil og forkaster modellen. Box-Ljung's Q^* test vil bli benyttet for å sjekke om residualene følger en hvit støy prosess. Forkastes nullhypotesen er det indikasjon på autokorrelasjon i residualene og modellen vil ikke være gyldig. Residualene blir også sjekket for normalitet med Stata14 sin «sktest» basert på D'Agostino et.al (1990), og stasjonærhet blir testet ved hjelp av en ADF test. Videre gjennomfører vi en LR-test (likelihood-ratio test) mellom modellene for å få en indikasjon på om den ene modellen utkonkurrerer den andre.

I Tabell 4 viste korrelegrammet for vekst i leiepriser at vi har problemer med autokorrelasjon i første lag. For å unngå spuriøse resultater inkluderer vi en moving-average prosess ($q=1$). På grunnlag av at vi nå inkluderer en vektor med eksogene forklaringsvariabler inkluderer vi også et konstantledd i ARIMAX modellen for å ta hensyn til lineære sammenhenger.

6.4.2 In-sample analyse ARIMAX

I modell (3) er vekst i leiepriser for kontorbygg i Oslo-området forklart av forrige periodes støy i residualene (MA(1)), endring i arbeidsledighetsraten ($\Delta_4 ARBL_{t-2}$), vekst i BNP ($\Delta_4 \ln BNP_{t-2}$), endring i lange renter ($\Delta RRTI_{t-2}$), vekst i nybygg ($\Delta \ln FNYB_{t-2}$), og vekst i arealledighet (ΔARL_{t-2}). De eksogene forklaringsvariablene er alle lagget to kvartaler, slik notasjonen bemerker. Serien for vekst i arealledighet er ikke stasjonær og har antydninger til strukturelle brudd i perioden 2001q2 til 2003q4 og vi inkluderer derfor en dummyvariabel for bruddperioden. Vi inkluderer også en interaktiv dummy som multipliseres med arealledighet.

Resultatene fra modell (3) viste at endring i arbeidsledighetsraten lagget to kvartaler var høyst signifikant og bedret informasjonskriteriene betraktelig i forhold til benchmark modellen. Vi beslutter derfor å inkludere endring i arbeidsledighetsraten i modellen. Endring i lange renter bedret informasjonskriteriene ytterligere og er høyst signifikant lagget to kvartaler. Den blir derfor også inkludert i modellen. Modell (3a) vil fungere som en ny benchmark for videre analyse.

Modell (3a) ble så utvidet med vekst i BNP, vekst i arealledighet og vekst i nybygg. Ingen av variablene ble signifikante da vi testet for andre til femte lag. Vi observerte også at modellen ble mindre stabil ved å inkludere vekst i BNP og vekst i arealledighet der MA komponenten ble strengt negativ. Dette kan være indikasjoner på enhetsrøtter i MA prosessen, og vi testet derfor

med ulike AR og MA strukturer. Ingen av resultatene utkonkurrerte modell (3a). Ved inkludering av AR komponenter økte Box-Ljung's Q^* som indikerer at residualene blir mer autokorrelert. For at modellene skulle bli gyldige måtte to AR komponenter inkluderes. I appendiks D er resultatene for de optimaliserte modellene som inkluderer vekst i BNP, arealledighet og nybygg. Modellene er optimalisert for den ARMA spesifikasjonen der modellen er stabil og residualene følger en hvitstøy prosess. For alle modellene er både AIC og BIC relativt dårligere enn modell (3a) og koeffisientene er ikke signifikante på 5% nivå. Vi konkluderer med at endring i arbeidsledighetsraten og lange renter er viktige determinanter for vekst i leiepris og blir inkludert som eksogene forklaringsvariabler i ARIMAX modellen.

Videre testet vi for ulike ARMA strukturer og sammenlignet informasjonskriteriene AIC og BIC for å finne den foretrukne modellen (se Tabell 8). BIC foretrekker (p.d.q=0.1.1) og både BIC og AIC foretrekker (p.d.q=0.1.2) i in-sample perioden 1998q1 til 2014q4. Resultatene fra ARIMAX modell (3) er presentert i Tabell 9.

Tabell 8: Sammenligning av informasjonskriteriene modell (3)

Tabellen oppsummerer informasjonskriteriene Akaike information criterion (AIC) og Bayesian information criterion (BIC) for vekst i leiepriser ved flere ulike ARMA strukturer for perioden 1998q1 til 2014q4. Markert er den modellen som er foretrukket av informasjonskriteriet. Følgende modell ble estimert:

$$\Delta \ln LP_t = \alpha_0 + \beta_1 \Delta ARBL_{t-2} + \beta_2 \Delta RRTI_{t-2} + \phi_1 \Delta \ln LP_{t-1} + \dots + \phi_p \Delta \ln LP_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad (3)$$

BIC	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)
AR(0)	-127.8	-141.3	-141.3	NA	-138.9
AR(1)	-133.2	NA	NA	NA	NA
AR(2)	-131.3	-138.5	-139.0	NA	NA
AR(3)	-131.2	-134.6	-135.0	-136.6	-137.2
AIC					
AR(0)	-136.7	-152.4	-154.6	NA	-154.5
AR(1)	-144.3	NA	NA	NA	NA
AR(2)	-144.6	-154.1	-154.5	NA	NA
AR(3)	-146.7	-152.3	NA	NA	NA
NA: Ikke stabil modellspefisikasjon (minst en eigenverdi lik 1)					

Tabell 9: Empiriske resultater ARIMAX modell (3)

Denne tabellen rapporterer resultatene fra ARIMAX modellen for vekst i leiepriser for kontormarkedet i Oslo-området ved bruk av kvartalsvis tidsseriedata fra 1998q1 til 2014q4. Følgende modell er estimert:

$$\Delta \ln LP_t = \alpha_0 + \beta_1 \Delta ARBL_{t-2} + \beta_2 \Delta RRTI_{t-2} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} \quad (3)$$

Δ betyr at variabelen er differensiert av første orden kvartalsvis ($d=1$), p er antall AR prosesser og q er antall MA prosesser «ARIMAX(p,d,q)». Forklaringsvariablene er lagget to ganger for å ta høyde for at leiekontraktene gjennomsnittlig trer i kraft ca. 4-8 måneder etter de er signert.

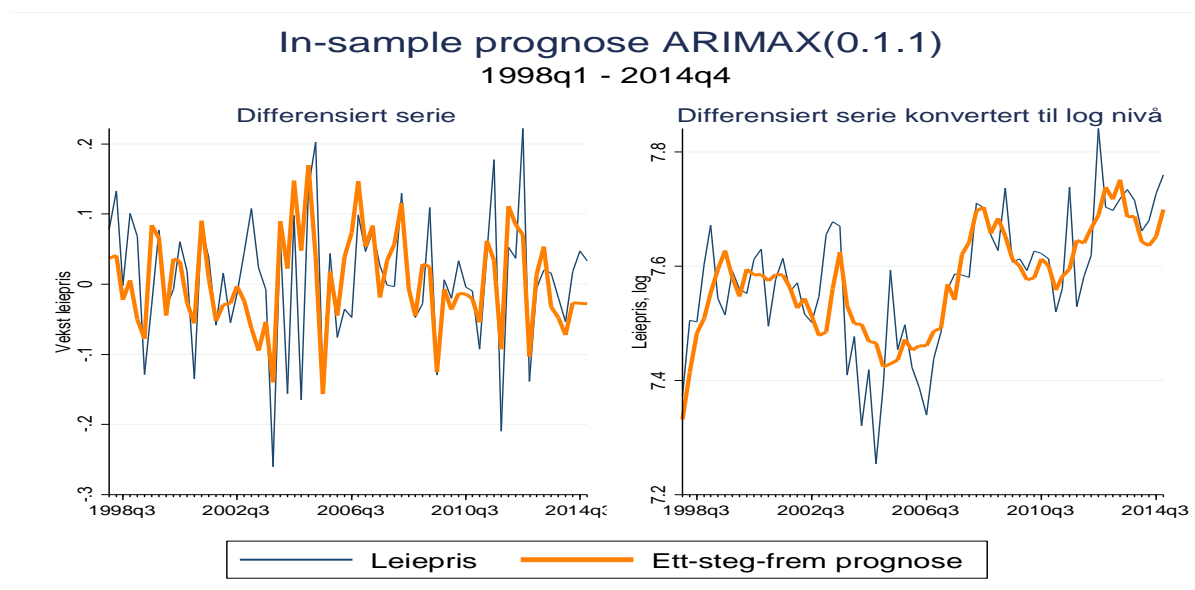
Sigma gjengir det estimerte standardavviket for hvit støy i feilleddet ε . AIC og BIC er mål på relativ kvalitet for modellene. ll er log likelihood, LR-test er likelihood-ratio test for å sammenligne modellene. Box-Ljung Q^* er portmanteau test for hvit støy i residualene. Sktest er en versjon av Jarque-Bera sin test for normalitet i residualene og ADF 4.lag er test for enhetsrøtter i residualene. Standardfeilene er kalkulert robust for å ta høyde for mulig heteroskedastisitet.

Modell	(3a)	(3b)
ARIMAX	(0.1.1)	(0.1.2)
konstant	0.00396 (1.91)	0.00319*** (4.83)
$\Delta_4 ARBL_{t-2}$	-2.815*** (-6.44)	-2.839*** (-6.29)
$\Delta RRTI_{t-2}$	-2.447* (-2.53)	-2.262** (-2.60)
ARIMA		
MA(1)	-0.806** (-2.94)	-0.730*** (-8.79)
MA(2)		-0.270** (-3.25)
<i>sigma</i>	0.0727*** (9.87)	0.0691*** (10.54)
AIC	-152.4	-154.6
BIC	-141.3	-141.3
ll	81.20	83.29
LR-test (χ^2)		4.18*(3a)
Justert R^2	0.38	0.40
Box-Ljung Q^*	33.50	31.11
Sktest (χ^2)	2.41	1.48
ADF 4.lag	-4.606***	-3.609***
z statistikk i parentes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$		

Resultatene fra modell (3a) indikerer at endring i arbeidsledighetsraten, lange renter og forrige periodes støy i residualene er signifikant på 5% nivå. Som forventet har endring i arbeidsledighetsraten negativ påvirkning på leieprisen. Endringer i lange renter er også negativ som samsvarer med funnene til D'Arcy, McGough og Tsolacos (1997). Residualene i modell (3a) følger en hvit støy prosess, er normalfordelte, og det er ingen tegn til enhetsrøtter hverken i residualene eller i MA komponenten.

I modell (3b) merker vi oss at det er lagt til to MA komponenter. Summerer vi koeffisientene sammen er de lik -1 som kan være en indikasjon på enhetsrøtter i MA leddet som kan gi ustabilitet i modellen. Vi undersøker dette nærmere i Stata14 ved bruk av funksjonen «aroots». Den konkluderer med at alle MA komponentene ligger innenfor enhetssirkelen og vi beholder derfor modellen (se appendiks E). Vi merker oss at standardavviket til residualene har blitt redusert fra 0.0727 til 0.0691 og at LR-testen forkaster modell (3a) til fordel for (3b). Residualene følger en hvit-støy prosess, er normalfordelte, og har ikke tegn til enhetsrøtter ifølge de diagnostiske testene. Vi testet også begge modellene for sesongeffekter ved å inkludere dummyvariabler for første til tredje kvartal. Dummyvariablene ble ikke signifikant forskjellig fra null. Vi konkluderer med at modell (3a) og (3b) er tilstrekkelig robuste for prognoser out-of-sample.

I Figur 13 er de predikerte in-sample verdiene for modell (3a) presentert sammen med den faktiske utviklingen. Modell (3a) ser ut til å følge den faktiske utviklingen for vekst i leiepriser relativt dårlig der den ikke får med seg flere av de høye toppene og dype dalene i serien. Den differensierte serien forklarer omtrent 38% av bevegelsene i leieprisene som er betraktelig bedre enn ARIMA modell (2a). Konverterer vi leieprisene til log nivå følger modellen perioder med vekst relativt bra, men når leieprisen synker er prognosen veldig forsinket. Forklaringskraften er omtrent 60%, som er bedre enn ARIMA og den naive modellen som begge har en forklaringskraft på omtrent 48%. Metoden for prediksjonene er på samme måte som tidligere en ett-steg-frem prognose. In-sample prognosen for modell (3b) er presentert grafisk i appendiks B.



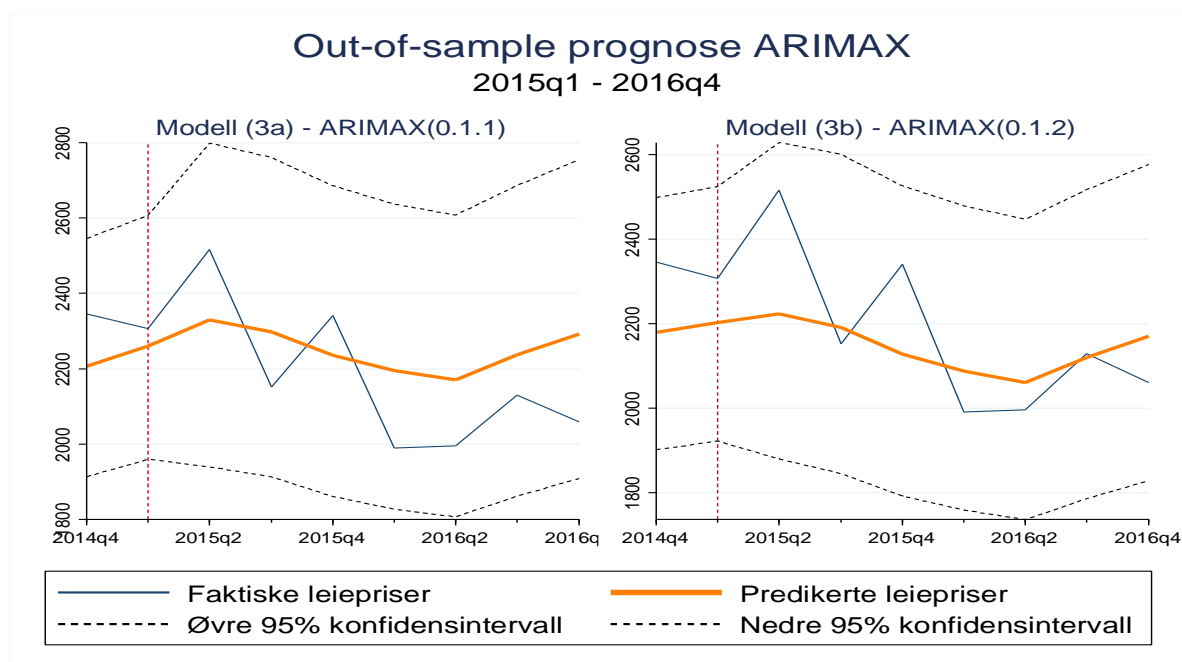
Figur 13: In-sample prognose ARIMAX modell (3a)

Denne figuren viser in-sample prognose for vekst i leiepriser og leiepriser for kontoreiendom i Oslo-området sammenlignet med den faktiske utviklingen i perioden 1998q1 – 2014q4. Venstre graf viser utviklingen i vekst i leiepris. Høyre graf viser utviklingen i leiepris transformert tilbake til log-nivå.

6.4.3 Out-of-sample resultater ARIMAX

I prognose av leiepriser out-of-sample benytter vi dynamiske prognoser. Det er dog viktig å merke seg at ARIMAX prognosen tar hensyn til virkelige verdier av de eksogene forklaringsvariablene. Dermed vil resultatene trolig være bedre enn det man kan forvente hvis man kun har estimater på forklaringsvariablene. Prognoseperiodene er som tidligere, en toårsperiode fra 2015q1 til 2016q4 og to ettårsperioder i 2015 og 2016. For sistnevnte modellerer vi frem til 2015q4 og utarbeider deretter en prognose for de neste fire kvartalene.

Prognosene blir utarbeidet på serien vekst i leiepris og konvertert tilbake til nivåform for reelle leiepriser. I Figur 14 presenteres den dynamiske toårsprognosen med konfidensintervall på to standardavvik. Den grafiske fremstillingen av den ettårige prognosen i 2016 blir presentert i appendiks C. Modell (3a) treffer på retningen de to første kvartalene og den nedadgående trenden etter 2015q2. Den får ikke med seg økningen i 2015q4 og overpredikerer i hele 2016. Vi observerer også at den får med seg retningsendringen i 2016q2.



Figur 14: Dynamisk out-of-sample prognose modell (3)

Figuren viser toårig prognosene for reelle leiepriser på nivåform for modell (3a) og (3b). Prognosen er gjennomført på log differensierte leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2015q1-2016q4.

Modell (3b) har en litt lavere prognose gjennom hele perioden som gjør at den treffer dårlig i 2015 der den er langt fra toppene i 2015q2 og 2015q4. I 2016 ser den ut til å prestere bedre enn modell (3a) og vi observerer at også modell (3b) får med seg retningsendringen i 2016q2.

Tabell 10: Prognosefeil - ARIMAX modell (3)

Tabellen viser prognosefeilene i forhold til den naive modellen. Negative verdier (grønn farge) betyr at modellen har prestert bedre enn den naive modellen. Positive verdier (rød farge) betyr at den naive modellen har prestert best. Prognosefeilparameterne er Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) og Theils U2.

ARIMAX - Naiv modell	Fra	Til	MAD	RMSE	MAPE	Theils U2
(3a): (0.1.1)						
	2015q1	2015q4	18,895	1,076	0,752	1,008
Dynamisk	2016q1	2016q4	-88,214	-88,302	-4,358	0,707
	2015q1	2016q4	-51,629	-75,020	-2,633	0,682
(3b): (0.1.2)						
	2015q1	2015q4	60,123	58,965	2,342	1,451
Dynamisk	2016q1	2016q4	-109,182	-107,495	-5,374	0,644
	2015q1	2016q4	-85,726	-90,449	-4,519	0,617

For å evaluere prognosene måler vi de opp mot den naive modellen (se Tabell 10). Ingen av ARIMAX modellene klarer å utkonkurrere den naive modellen i ettårsprognosen for 2015. For kortere horisonter er ikke dette et uvanlig resultat, men over fire kvartaler hadde vi ventet at modellen skulle score bedre enn den naive modellen. Noe av forklaringen kan være den nedadgående trenden i 2015 som inkluderer to store topper. Toppene jevner ut gjennomsnittet i perioden og den naive modellen treffer derfor godt. ARIMAX får med seg den negative trenden, men greier ikke å justere seg tilstrekkelig for de store svingningene mellom kvartalene. Vi observerer derfor at modellen er bedre posisjonert for lengre prognoser enn den naive modellen, men blir utkonkurrert over firekvartalsprognosen i 2015 (se Figur 14).

Modell (3a) er den som scorer relativt best i den første perioden. I ettårsprognosen for 2016 klarer begge modellene å utkonkurrere den naive modellen. Modell (2b) scorer bedre på alle parameterne både i 2016 og toårsperiodene. Den virker derfor å være den foretrukne ARIMAX modellen for å predikere leiepriser i Oslo-området over lengre perioder.

6.5 VECM (4)

Målet med denne delen er å finne ut hvilke innbyrdes forhold det er mellom variablene, og hvordan de påvirker leieprisen. Videre vil vi prognostisere leiepriser med basis i analysen. Som vi nevnte i teoridelen er VAR/VEC modeller spesielt nyttige ved analyse av makroøkonomiske tidsserier. Dette er fordi man har muligheten til å analysere både kortsiktig og langsiktig informasjon i dataene ved å utnytte kointegrasjonsegenskapene. Siden VECM er a-teoretisk velger vi å bruke variablene; $\ln LP$, $\ln FNYB$, $ARBL$ og $RRTI$. Som vi har observert tidligere er variablene integrert av første orden, $I(1)$. Siden alle variablene er integrert av samme orden velger vi derfor å utarbeide en VEC modell i denne delen.

6.5.1 Lag struktur

Vi begynner med å finne ut hvilken lag struktur vi skal bruke i modellen. Som tidligere baserer vi optimal lag lengde ved å minimere informasjonskriteriene AIC, BIC og HQIC. Rapporten for valg av laglengde er presentert i Tabell 11. AIC indikerer 10 lags som optimalt, mens HQIC og BIC indikerer 4. Med få observasjoner i datasettet er vi påpasselige med å ikke overtilpasse modellen, og velger derfor 4 lags som utgangspunkt. Vi er dog spesielt oppmerksomme når vi undersøker autokorrelasjon i residualene etter modellestimeringen. Autokorrelasjon i residualene vil indikere at vi har en feilspesifisert modell, og at det kan være nødvendig å inkludere flere lags.

Tabell 11: Informasjonskriterier VECM (4)

AIC er Akaikes informasjonskriterie, HQIC er Hannan-Quinn informasjonskriterie, og BIC er Schwarz Bayesianske informasjonskriterie estimert i perioden 1998q1 – 2014q4.

Endogene variabler: $\ln LP$, $\ln FNYB$, $ARBL$, $RRTI$. Maks lag: 10

Lag	AIC	HQIC	BIC
0	-12.676	-12.624	-12.543
1	-16.830	-16.568	-16.167
2	-19.344	-18.872	-18.150
3	-20.099	-19.417	-18.373
4	-20.727	-19.835*	-18.471*
5	-20.493	-19.392	-17.706
6	-20.546	-19.235	-17.229
7	-20.583	-19.062	-16.734
8	-20.528	-18.798	-16.149
9	-21.053	-19.113	-16.143
10	-21.207*	-19.057	-15.766

6.5.2 Kointegrasjonstesting: Johansen metode

Siden vi tidligere i utredningen har gjort de innledende testene for variablene og sikret at de oppfyller kravene vi har stilt underveis, kan vi nå estimere om dataene inneholder en (flere) kointegrerende vektor(er) eller ikke. Kointegrasjonstesten vi benytter oss av, vecrank, er integrert i stata14 og er basert på Johansen sin metode gjennomgått i teoridelen (StataCorp, 2013). Tabell 12 presenterer teststatistikkene og deres kritiske verdier til nullhypotesen om ingen kointegrasjon (i rad 1), og en eller færre kointegrerte likninger i rad 2, osv. Eigenverdien blir brukt til å beregne «trace statistic» i raden over. Johansens testprosedyre starter med å teste for ingen kointegrerte likninger (maksimal rang 0) og så akseptere den første nullhypotesen som ikke blir forkastet.

Tabell 12: Johansen test for kointegrasjon VECM (4)

r er antall kointegrerte likninger, parametere er antall parametere i testen, LL er log-likelihood.

* er antall kointegrerte likninger testen velger. Se kapittel 5.8.1 for definisjon av test statistikk.

H_0 : r eller færre kointegrerte likninger, H_A : flere enn r kointegrerte likninger.

Maksimum rang	Parametere	LL	Eigenverdi	Trace statistikk	5% kritisk verdi
$r = 0$	52	743.523		62.397	47.210
$r = 1$	59	758.500	0.353	32.746	29.680
$r = 2$	64	769.364	0.277	10.717*	15.410
$r = 3$	67	774.427	0.138	0.589	3.760
$r = 4$	68	774.722	0.009		
Lags = 4					
Utvalgsperiode 1998q1-2014q4					

H_0 er at det er r eller færre kointegrerte likninger. Alternativhypotesen er at det er flere enn r kointegrerte likninger. Nullhypotesen blir forkastet hvis «trace statistic» er større enn den kritiske verdien. Når testen ikke blir forkastet, blir verdien av r estimatet på antall kointegrerte forhold (Kuzmenko et al., 2017). Som vi observerer i Tabell 12 kan vi tydelig forkaste nullhypotesen om ingen kointegrasjon ($H_0 r = 0$ er forkastet på 5% nivå siden $62,397 > 47,210$). Vi kan også forkaste nullhypotesen med maksimalt én kointegrert likning ($H_0 r = 1$ er forkastet på 5% nivå siden $32,746 > 29,680$). Vi kan ikke forkaste $H_0 r = 2$ og aksepterer derfor nullhypotesen om 2 kointegrerte likninger.

I tråd med Lütkepohl (2004) dobbeltsjekker vi resultatet ved å undersøke bivariat kointegrasjon og finner totalt to kointegrerte sammenhenger; (lnLP & lnFNYB), og (lnFNYB & RRTI). Dette understøtter konklusjonen med to kointegrerte sammenhenger vi fant da vi testet hele systemet.

6.5.3 Modellestimering

Nå som vi har bestemt antallet kointegrerte likninger kan vi estimere VECM parameterne. Resultatet av modellen er presentert i Tabell 13. Her finner vi informasjon om estimeringen av de kointegrerte vektorene i modellen, sammen med deres standardfeil, z-statistikker og konfidensintervaller. Koeffisientmatrisene er presentert i appendiks F av plasshensyn.

Vi kan uttrykke VECM relatert til leiepris på følgende måte:

$$\begin{aligned} \Delta \ln LP_t = & \pi_0 + \sum_{i=1}^4 \pi_1 \Delta \ln LP_{t-i} + \sum_{i=1}^4 \pi_2 \Delta_4 ARBL_{t-i} + \sum_{i=1}^4 \pi_3 \Delta \ln FNYB_{t-i} \\ & + \sum_{i=1}^4 \pi_4 \Delta RRTI_{t-i} + \delta ECT_{t-1} + \epsilon_{1t} \end{aligned}$$

hvor koeffisienten δ til ECT_{t-1} representerer lnLP sin justeringshastighet mot sin langsiktige likevekt, π_1, \dots, π_4 er koeffisientene til de laggede kortsiktige sammenhengene for de ulike variablene mot $\Delta \ln LP_t$. Parameterne i feilkorreksjonsleddet (ECT) stammer fra den kointegrerte vektoren:

$$\delta ECT_{t-1} = \delta(\beta_1 \ln LP_t + \beta_2 ARBL_t + \beta_3 \ln FNYB_t + \beta_4 RRTI_t)$$

Vi ser at feilkorreksjonsleddet i VECM er negativt og signifikant på 5% nivå ($\Delta \ln LP_{ce1L1} = -0,346$, $p = 0,034$). Dette indikerer at det finnes en justering mot likevekt i systemet, og at det eksisterer en langsiktig likevekt mellom lnLP og de andre variablene.

Når vi normaliserer med hensyn til lnLP finner vi følgende kointegrerte forhold:

$$\ln LP_t = 0,115 \ln FNYB_t - 4,628 RRTI_t + 6,290$$

Koeffisienten til forventet nybygg er positiv og signifikant på 1% nivå, koeffisienten til lange renter er negativ og signifikant på 1% nivå. Arbeidsledighet er ikke signifikant til å forklare en langsiktig sammenheng. Kointegrasjonsresultatet viser at det finnes en langsiktig positiv sammenheng mellom forventet nybygg og leieprisen. Det er gjort lite forskning på hvilken effekt tilbudssidevariabler har på leieprisen, og vi har derfor lite tilgjengelig empiri for å forankre dette funnet. Likevel er resultatet logisk, da økt forventet nybygging indikerer at det er lønnsomt for utbyggere å bygge nye arealer. Høy lønnsomhet for utbyggere indikerer igjen høy arealknapphet og høy leiepris.

Tabell 13: Empiriske resultater VECM (4)

Denne tabellen rapporterer resultatene til VECM (4), med de endogene variablene $\ln LP$ (leiepris), $ARBL$ (arbeidsledighetsrate), $\ln FNYB$ (forventet nybygg), og $RRTI$ (lange renter). Δ indikerer at variabelen er differensiert av første orden. Øverste tabellen presenterer tilpasningen til hver ligning, hvor R^2 forklaringskraften og χ^2 er khi-kvadratstatistikken for ligningen. I tabellen under finner vi de estimerte parameterne til de kointegrerte vektorene i modellen, sammen med deres standardfeil, z-statistikker og konfidensintervaller. Koeffisientmatrisen er presentert i appendiks F.

Likning	Parametere	RMSE	R^2	χ^2	$P - verdi > \chi^2$	
$\Delta \ln LP$	15	0.075	0.487	49.448	0.000	
$\Delta ARBL$	15	0.004	0.623	85.906	0.000	
$\Delta \ln FNYB$	15	0.024	0.983	2934.254	0.000	
$\Delta RRTI$	15	0.009	0.419	37.472	0.001	
Kointegrasjonslikninger						
$_{ce1}$	2			35.335	0	
$_{ce2}$	2			62.246	0	
Beta	Koeffisient	Standardfeil	z	p-verdi > z	95% konfidensintervall	
$_{ce1}$						
$\ln LP$	1
$ARBL$	0
$\ln FNYB$	-0.115	0.032	-3.61	0.000	-0.177	-0.052
$RRTI$	4.628	1.054	4.39	0.000	2.563	6.694
konstant	-6.290
$_{ce2}$						
$\ln LP$	0
$ARBL$	1
$\ln FNYB$	0.018	0.003	6.810	0.000	0.013	0.024
$RRTI$	0.406	0.089	4.560	0.000	0.232	0.058
konstant	-0.287
Alpha						
$\Delta \ln LP$						
$_{ce1L1}$	-0.346	0.163	-2.12	0.034	-0.067	-0.026
$_{ce2L1}$	-2.581	1.769	-1.46	0.145	-6.049	0.886
$\Delta ARBL$						
$_{ce1L1}$	0.028	0.009	3.01	0.003	0.010	0.047
$_{ce2L1}$	-0.147	0.102	-1.44	0.151	-0.346	0.053
$\Delta \ln FNYB$						
$_{ce1L1}$	-0.080	0.053	-1.52	0.130	-0.184	0.024
$_{ce2L1}$	-2.683	0.575	-4.66	0.000	-3.810	-1.556
$\Delta RRTI$						
$_{ce1L1}$	-0.023	0.020	-1.12	0.261	-0.063	0.017
$_{ce2L1}$	-0.148	0.222	-0.67	0.506	-0.058	0.029

Kointegrasjonsresultatet viser også at vi har en signifikant negativ langsiktig sammenheng mellom lange renter og leieprisen. Dette er i samsvar med eksisterende empiri, hvor blant annet D'Arcy, McGough og Tsolacos (1997) fant en negativ sammenheng mellom leieprisen og rentenivå. Det er også i tråd med vår tidligere observasjon om divergerende trender i rentenivå og leiepris. Den negative koeffisienten vi finner på de lange rentene er derfor forventet. På den annen side er resultatet oppsiktsvekkende, da man normalt antar at sentralbanken skal «lene seg mot vinden» ved å drive motsyklisk pengepolitikk. Den langsiktige negative sammenhengen mellom rentene og leieprisen kan være et bevis på det motsatte.

Videre tester vi for om det finnes kortsiktig kausalitet mellom variablene. I Tabell 14 er H_0 ingen kortsiktig kausalitet mellom de analyserte variablene. I praksis tester vi om koeffisientene til alle tre lags av hver variabel = 0 samtidig. Hvis vi forkaster H_0 indikerer dette at vi har kortsiktig kausalitet mellom variablene.

Tabell 14: Lineæritetstest VECM (4)

Δ indikerer at variabelen er differensiert av første orden. L# angir hvilken lag av variabelen som er valgt, og ln indikerer den naturlige logaritmen. χ^2 er khi-kvadratfordelingen for lineæritetstesten.

H_0 : ingen kortsiktig kausalitet

Lineæritetstest	χ^2	p-verdi $> \chi^2$
$[\Delta \ln LP]L\Delta ARBL = 0$ $[\Delta \ln LP]L2\Delta ARBL = 0$ $[\Delta \ln LP]L3\Delta ARBL = 0$	10.83	0.013
$[\Delta \ln LP]L\Delta \ln FNYB = 0$ $[\Delta \ln LP]L2\Delta \ln FNYB = 0$ $[\Delta \ln LP]L3\Delta \ln FNYB = 0$	3.89	0.2741
$[\Delta \ln LP]L\Delta RRTI = 0$ $[\Delta \ln LP]L2\Delta RRTI = 0$ $[\Delta \ln LP]L3\Delta RRTI = 0$	6.36	0.0952

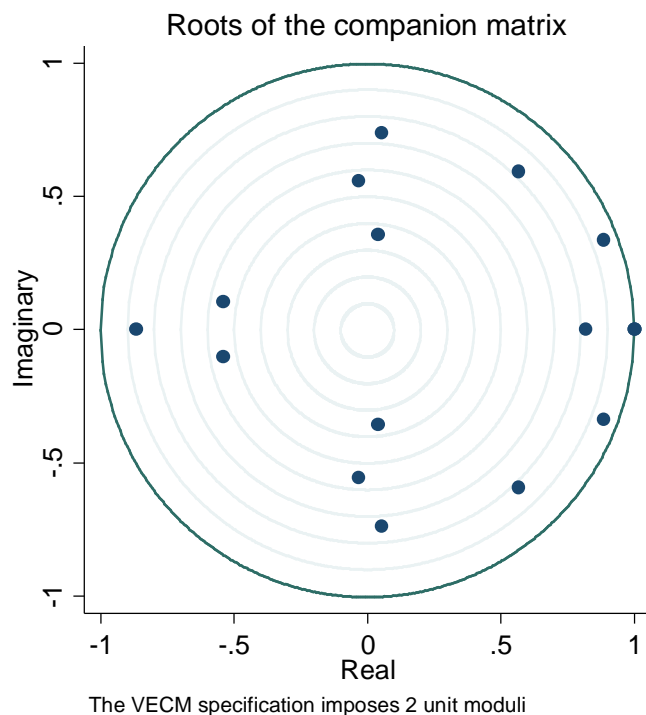
Basert på resultatene i tabellen kan vi si at det kun finnes kortsiktig kausalitet mellom ARBL og lnLP da dette er den eneste testen med p-verdi under 5% nivå ($p=0,013$). Vi kan derfor forkaste nullhypotesen om ingen kortsiktig kausalitet mellom de valgte variablene. Vi finner dog ikke kortsiktig kausalitet mellom lnFNYB til lnLP, eller RRTI til lnLP. At vi finner en kortsiktig sammenheng mellom arbeidsledighet og leieprisen er et forventet resultat. Dette er i tråd med funnene fra ARIMAX modellen som fant en signifikant negativ sammenheng og understøttes av studien til De Wit og Van Dijk (2003). Vi tolker funnet som at høy arbeidsledighet gir en indikasjon på lavere etterspørsel, som vil kunne slå negativt ut i leieprisen relativt raskt.

6.5.4 Diagnostisk testing

6.5.4.1 Stabilitet

Når vi nå har estimert VECM parameterne sjekker vi stabilitetsbetingelsen til estimatene. Hvis en VECM har K endogene variabler, og r kointegrerte vektorer, vil det være $K - r$ unit moduli i følgematriksen. For vår modell med 4 endogene variabler og 2 kointegrerte vektorer, vil følgematriksen ha $(4 - 2 = 2)$ unit eigenverdier. Hvis modellen skal være stabil må moduli av de resterende r eigenverdiene være strengt lavere enn 1. Resultatet fra stabilitetstesten er presentert i Tabell 15 og Figur 15.

Eigenverdi	Moduli
1	1
1	1
0.886 + .336i	0.948
0.886 - .335i	0.948
-0.867	0.867
0.566 + .593i	0.820
0.566 - .593i	0.820
0.818	0.818
0.054 + .737i	0.739
0.054 - .737i	0.739
-0.033 + .556i	0.557
-0.033 - .556i	0.557
0.540 - .103i	0.550
0.540 + .103i	0.550
0.039 + .357i	0.359
0.039 - .357i	0.359



Tabell 15: Unit Moduli i VECM (4)

Figur 15: Grafisk fremstilling av unit moduli VECM (4)
Figuren viser hvor nært røttene er til enhetssirkelen

Tabell 15 rapporterer eigenverdiene til følgematriksen og deres assosierte moduli. Vi observerer at to av røttene er lik 1, som er forventet siden $(4 - 2 = 2)$. Hvis noen av de resterende moduli er lik 1 indikerer dette enten at de kointegrerte likningene ikke er stasjonære, eller at det er en annen trend i dataene. Det finnes ingen distribusjonsteori for å måle hvor nært røttene er til 1, men problematikken er blant annet diskutert i Johansen (1995). Vi ser at modellen har en rot med verdi 0,948, som på bakgrunn av diskusjonen i Johansen kan indikere noen problemer med stabilitet, dog kan vi hverken bekrefte eller avkrefte dette grunnet mangel på distribusjonsteori. Vi velger å akseptere modellen, og Leser bes dermed være oppmerksom på at en eventuell ustabilitet i modellen vil kunne ha implikasjoner for prognosene, men at temaet ikke blir problematisert ytterligere i utredningen.

6.5.4.2 Autokorrelasjon

Den neste diagnostiske testen vi utfører er å teste residualene for autokorrelasjon. Vi tester for dette med Lagrange Multiplier test implementert i stata14 (StataCorp, 2013). Resultatet fra testen er presentert i Tabell 16.

Tabell 16: Lagrange Multiplier test for autokorrelasjon i residualene VECM (4)

χ^2 er khi-kvadratfordelingen for Lagrange Multiplier testen og df er antall frihetsgrader

H_0 : ingen autokorrelasjon

Lag	χ^2	df	p-verdi > χ^2
1	12.470	16	0.711
2	18.352	16	0.304
3	17.262	16	0.369
4	15.509	16	0.488
H_0 : Ingen autokorrelasjon ved lag nummer #			

H_0 i testen er ingen autokorrelasjon. En p-verdi lavere enn 0,05 vil derfor indikere autokorrelasjon i residualene. Tabell 16 presenterer ingen p-verdier under 0,05 og vi kan dermed ikke forkaste H_0 om ingen autokorrelasjon i residualene. Autokorrelasjonstesten indikerer derfor ingen bevis på at VECM er feilspesifisert.

6.5.4.3 Normalitet

Videre tester vi for normalitet. Vi benytter den implementerte kommandoen «vecnorm» i Stata14 hvor H_0 er at feilleddene i VECM er normalfordelte (StataCorp, 2013). Resultatet fra normalitetstesten er presentert i Tabell 17.

Tabell 17: Jarque-Bera normalitetstest VECM (4)

χ^2 er khi-kvadratfordelingen for Jarque-Bera testen og df er antall frihetsgrader

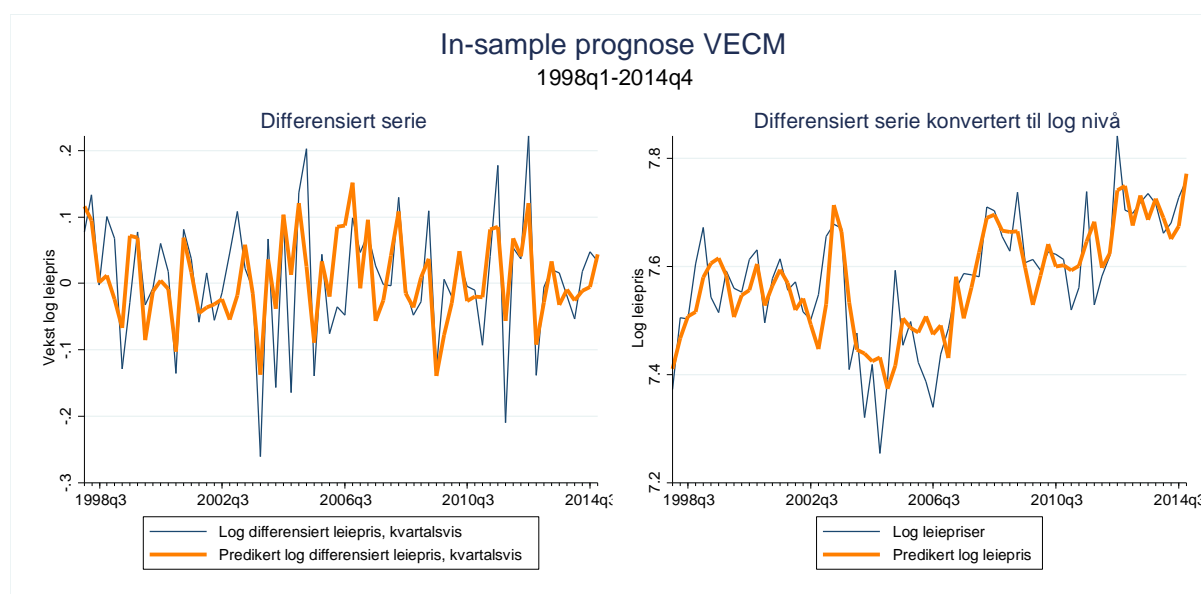
H_0 : feilleddene i VECM er normalfordelte

Likning	χ^2	df	p-verdi > χ^2
$\Delta \ln LP$	1.667	2	0.435
$\Delta ARBL$	0.455	2	0.797
$\Delta \ln FNYB$	18.288	2	0.000
$\Delta RRTI$	1.521	2	0.467
ALL	21.931	8	0.005

Som vi observerer i Jarque-Bera testen finner vi en veldig lav p-verdi på 0,000 på $\Delta \ln FNYB$. Vi kan derfor forkaste H_0 om at feilleddene er normalfordelte. At Jarque-Bera testen feiler kan være en indikasjon på at man ikke har inkludert nok antall lags i modellen. Vi tester for

dette, og finner at det ikke hjelper nevneverdig å inkludere flere lags. En sannsynlig årsak til dette er at vi analyserer et relativt lite utvalg. I tillegg er nybygg interpolert fra årlig data til kvartalsvise data som bidrar til glatting i serien. Hvis feilleddene ikke er normalfordelte vil standard inferens bare være korrekt i store utvalg. I mindre utvalg vil dermed standardfeil og teststatistikker være tilnærmet korrekt. Siden vi i denne utledningen i hovedsak bruker modellene til out-of-sample prognoser vil dette mest sannsynlig ikke påvirke våre resultater nevneverdig (Sukati, 2013).

6.5.5 In-sample prognose VECM



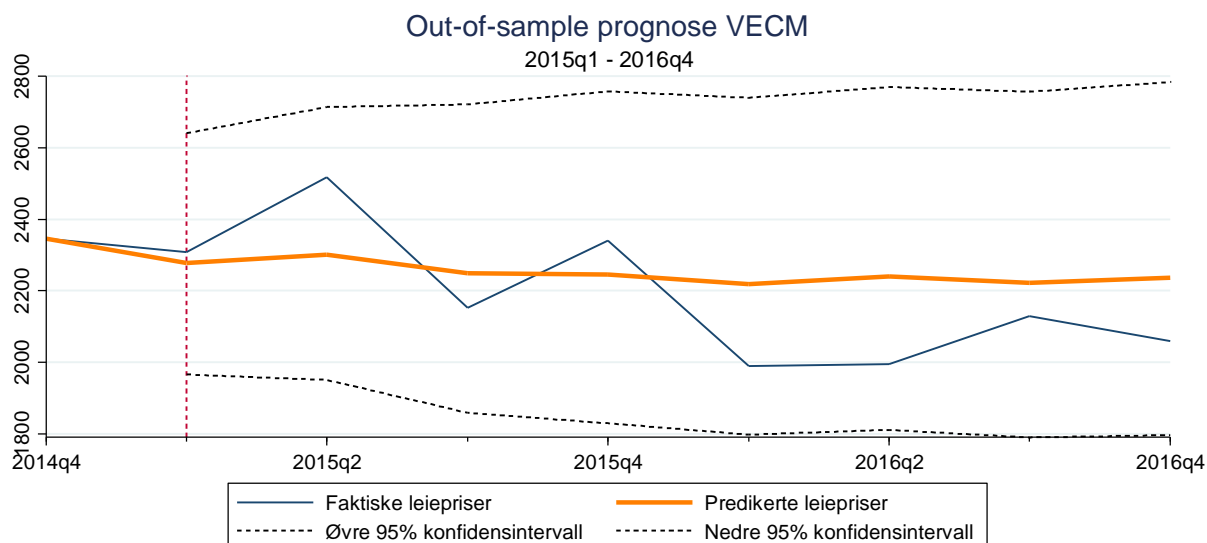
Figur 16: In-sample prognose VECM (4)

Denne figuren viser in-sample prognose for vekst i leiepriser og leiepriser for kontoreiendom i Oslo-området sammenlignet med den faktiske utviklingen i perioden 1998q1 – 2014q4. Venstre graf viser utviklingen i vekst i leiepris. Høyre graf viser utviklingen i leiepris transformert tilbake til log-nivå.

I Figur 16 er de predikerte in-sample verdiene for modell (4) presentert sammen med den faktiske utviklingen. VECM ser ut til å følge den faktiske utviklingen for log differensierte leiepriser i stor grad gjennom estimeringsperioden. I de første årene ser modellen ut til å treffe bedre en ARIMA og ARIMAX modellen. Det er enkelte store svingninger den ikke får med seg og avviker spesielt i år 2002 og 2006. Forklaringskraften for den differensierte serien er omtrent 49% som er bedre enn ARIMAX modell (3a) på 38%.

Konverterer vi leieprisene til log nivå ser modellen ut til å treffe godt med retningen og får med seg store deler av svingningene med unntak nedgangen i 2006 og 2011. Forklaringskraften øker til 67% som er bedre enn ARIMAX modell (3a) på 60%.

6.5.6 Out-of-sample resultater VECM



Figur 17: Dynamisk out-of-sample prognose modell (4)

Figuren viser toårig prognosene for reelle leiepriser på nivåform for VECM (4). Prognosen er gjennomført på log leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2015q1-2016q4.

I likhet med de andre modellene utarbeider vi dynamiske prognoser for out-of-sample perioden. Figur 17 presenterer den toårige prognoseperioden 2015q1-2016q4. Den ettårige prognosen for 2016 er presentert i appendiks C. Vi observerer at VECM prognosen treffer relativt godt på retning og fanger opp en svakt nedadgående trend i leieprisen. VECM får derimot ikke med seg svingningene i den faktiske leieprisene fra kvartal til kvartal.

For å evaluere prognosene måler vi de opp mot den naive modellen (se Tabell 18). Theils U2 er lavere enn 1 for alle prognoseperiodene, som indikerer at VECM er bedre enn den naive modellen målt med RMSE. Den utkonkurrerer den naive modellen spesielt i toårsprognosen, men er ikke bedre enn den naive modellen målt etter MAD og MAPE i 2015.

Tabell 18: Prognosefeil - VECM (4)

Tabellen rapporterer prognosefeilene i forhold til den naive modellen. Negative verdier (grønn farge) betyr at modellen har prestert bedre enn den naive modellen. Positive verdier (rød farge) betyr at den naive modellen har prestert best. Prognosefeilparameterne er Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) og Theils U2.

VECM - Naiv modell	Fra	Til	MAD	RMSE	MAPE	Theils - U2
	2015q1	2015q4	6,669	-2,581	0,159	0,980
Dynamisk	2016q1	2016q4	-50,026	-48,905	-2,467	0,838
	2015q1	2016q4	-55,028	-71,362	-2,776	0,698

7. Sammenligning av prognoser

Dette kapitlet sammenligner de ulike prognosemodellene for kvartalsvise kontorleiepriser i Oslo-området i perioden 2015q1 til 2016q4. Prognosene er generert dynamisk, noe som vil si at det tas hensyn til forrige periodes predikerte leiepris før neste periode predikeres. Slik fortsetter det til alle åtte kvartalene er prognostisert. Utredningen sammenligner også ettårige prognoser for perioden 2015 og 2016 for å undersøke hvilken modell som predikerer best i kortere perioder. I tillegg vil prognosemodellene bli sammenlignet mot en naiv random-walk modell som sier at fremtidige leiepris er den samme som leieprisen i siste periode i all fremtid.

Out-of-sample prognosene for de ulike modellene er evaluert og sammenlignet på grunnlag av prognosefeilparameterne Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE) og Theils U2. Den beste modellen er definert som den med lavest MAD, RMSE, MAPE og Theils U2 verdier, da det indikerer at modellen har de mest treffsikre gjennomsnittlige prognosene. Tabell 19 presenterer sammenligningen av resultatene.

I toårsprognosen fra 2015q1 – 2016q4 scorer ARIMAX modell (3b) klart best på alle parameterne. VECM presterer ikke mye dårligere enn modell (3a) i sammenligning av RMSE, men scorer betraktelig dårligere på MAD og MAPE. Alle modellene presterer bedre enn den naive modellen for toårsprognosen.

I ettårsprognosen fra 2015q1 – 2015q4 scorer ARIMA modell (2b) best på MAD og MAPE. VECM modellen scorer marginalt bedre på RMSE og Theils U2, men ingen av modellene greier å utkonkurrere den naive modellen (1) i nevneverdig grad. ARIMAX modell (3b), som var den beste på toårsprognosen, gjør det klart svakest i ettårsprognosen for 2015.

I ettårsprognosen fra 2016q1 – 2016q4 scorer ARIMAX modell (3b) klart best på alle parameterne etterfulgt av ARIMAX modell (3a). Totalt sett er prognosefeilene større i denne perioden sammenlignet med de to andre prognoseperiodene. Det tyder på at modellene har vanskeligheter med å fange opp den negative veksten fra 2015q4 til 2016q1 og dermed overprognostiserer leieprisen konsekvent gjennom hele perioden.

Tabell 19: Sammenligning av prognosetreffsikkerhet out-of-sample

Tabellen rapporterer out-of-sample prognosetreffsikkerhet for fire alternative modeller; ARIMA, ARIMAX, VECM, og random-walk modellen som er en naiv prognose teknikk. Prognosene for modellene er evaluert og sammenlignet ved prognose estimatene Mean Absolute Deviation (MAD), Root Mean Square Error (RMSE) og Mean Absolute Percentage Error (MAPE). Theils U2 er et mål på prognosetmodellen relativt til den naive modellen.

$$\text{MAD} = \left(\sum_{t=1}^N |A_t - P_t| \right) / N, \quad \text{RMSE} = \sqrt{\text{MSE}}, \quad \text{MAPE} = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{A_i - F_i}{A_i} \right|, \quad \text{Theils U2} = \sqrt{\frac{\text{MSE}}{\text{MSE}^{\text{NAIVE}}}}, \quad \text{hvor MSE er definert: } \text{MSE} =$$

$$\frac{1}{N} \sum_{t=1}^N (A_t - P_t), \quad A_t \text{ er faktiske leiepriser og } P_t \text{ er predikerte leiepriser for periode } t.$$

Estimert modell	Ett år: 2015q1 - 2015q4				Ett år: 2016q1 - 2016q4				To år: 2015q1 - 2016q4			
	MAD	RMSE	MAPE	Theils U2	MAD	RMSE	MAPE	Theils U2	MAD	RMSE	MAPE	Theils U2
ARIMA												
Modell (2a): (0.1.1)	101,157	131,930	4,306	1,010	263,517	269,504	12,978	0,893	180,471	209,863	8,550	0,889
Modell (2b): (1.1.0)	97,844	128,726	4,210	0,986	247,832	253,851	12,207	0,841	191,550	224,868	9,126	0,952
ARIMAX												
Modell (3a): (0.1.1)	121,115	131,691	5,180	1,008	208,312	213,557	10,235	0,707	150,521	161,098	7,013	0,682
Modell (3b): (0.1.2)	162,343	189,581	6,770	1,451	187,344	194,364	9,219	0,644	116,424	145,670	5,128	0,617
VECM												
Modell (4)	108,889	128,034	4,587	0,980	246,500	252,954	12,127	0,838	147,122	164,757	6,870	0,698
Naiv random-walk												
Modell (1)	102,220	130,615	4,428	1,000	296,526	301,859	14,594	1,000	202,150	236,118	9,647	1,000

Totalt sett presterer ARIMAX og VECM bedre enn ARIMA modellen. ARIMA treffer greit med den generelle retningen men får ikke med seg vendepunktene i stor grad. Dette samsvarer med funnene til Wilson et. al (2000). ARIMAX modellen greier å fange opp vendepunktene i større grad, som er en viktig egenskap for investorer slik at de har mulighet til å justere sine investeringsstrategier. Utfordringen med modellen er at den tar hensyn til virkelige verdier av de eksogene forklaringsvariablene. Disse verdiene vil ikke være kjente ex-ante og for lengre prognoseperioder vil man derfor måtte prognostisere forklaringsvariablene. Ved bruk av estimater vil usikkerheten øke og trolig redusere treffsikkerheten til modellen. Siden arbeidsledighet og lange renter er lagget to kvartaler er det dog mulig å prognostisere fremtidige leiepriser to kvartaler frem med dagens informasjon.

VECM treffer godt med den generelle retningen gjennom hele perioden men får ikke med seg de store vendepunktene. Tar vi i betraktning at ARIMAX modellen egentlig kun kan predikere to kvartaler frem er VECM det foretrukne valget for lengre prognoser.

I løpet av arbeidet med utredningen ble data for første kvartal 2017 tilgjengelig. I Tabell 20 er resultatene presentert når vi utvider prognosen med ett kvartal. Både VECM (4) og ARIMAX modell (3a) treffer godt og feilpredikerer leieprisen med omtrent 50kr. ARIMAX modell (3b) utkonkurrerer de andre modellene og feilpredikerer med kun 24kr. Dette resultatet underbygger funnene i utredningen om at VECM og ARIMAX modellene egner seg til å predikere leiepriser over lengre perioder.

Tabell 20: Dynamisk out-of-sample prognose 2015q2 - 2017q1

Tabellen rapporterer den reelle leieprisen ARIMA, ARIMAX og VECM prognostisere i 2017q1 ved en toårig prognose i perioden 2015q2 – 2017q1 for kontoreiendom i Oslo-området.

	2017q1	Differanse
ARIMA		
Modell (2a): (0.1.1)	2306	106
Modell (2b): (1.1.0)	2316	116
ARIMAX		
Modell (3a): (0.1.1)	2253	53
Modell (3b): (0.1.2)	2176	-24
VECM		
Modell (4)	2246	46
Faktisk leiepris	2200	
<i>Prognoseperiode: 2015q2 - 2017q1 (8 perioder)</i>		

8. Konklusjon

Vi har i denne utredningen forsøkt å svare på problemstillingene:

1. *Hvilke determinanter driver leieprisen for kontoreiendom i Oslo i perioden 1998q1-2016q4?*
2. *Hvordan kan determinantene anvendes til å utarbeide treffsikre prognosemodeller?*
3. *Hvilken prognosemodell gjengir den faktiske utviklingen med størst gjennomsnittlig treffsikkerhet?*

Fremgangsmåten har vært å analysere historiske kontorleiepriser i Oslo-området for perioden 1996q1-2016q4 med ulike statistiske verktøy. Grunnet svært få observasjoner i de tidligste årene valgte vi å begrense analyseperioden til 1998q1-2016q4. Videre utarbeidet vi tre prognosemodeller; ARIMA, ARIMAX og VECM. Ulike spesifikasjoner av modellene ble sammenlignet med en naiv modell for å finne den beste prognosemodellen out-of-sample.

Gitt utredningens forbehold har våre analyser gitt flere interessante funn. Vi har funnet at alle våre prognosemodeller er mer treffsikre enn den naive modellen for toårsperioden 2015q1-2016q4. ARIMAX er den mest treffsikre modellen for toårsperioden, etterfulgt av VECM. I ettårsprognosen for 2016 er også ARIMAX modellen den mest treffsikre og vi observerer at den i større grad fanger opp vendepunkter i leieprisen. Dette samsvarer med den lignende studien til Karakozova (2004). ARIMAX prognosen tar dog hensyn til virkelige verdier av de eksogene forklaringsvariablene og vil trolig prestere dårligere når verdien av variablene ikke er kjent ex-ante. Tar vi denne begrensningen i betraktning er VECM den foretrukne prognosemodellen for kontorleieprisen i Oslo-området. Vi observerer at modellen fanger opp langsiktige trender i leieprisen, men har problemer med å fange opp større svingninger kvartal til kvartal. VECM egner seg derfor best til å predikere over lengre perioder.

Som et resultat av kointegrasjonsanalysen fikk vi bekreftet en positiv statistisk signifikant langsiktig sammenheng mellom forventet nybygging og leiepris. Resultatet indikerer at det er lønnsomt for utbyggerne å bygge nye kontorlokaler og at etterspørselen vil overgå tilbudet i fremtiden. Vi fant også en negativ langsiktig sammenheng mellom lange renter og leiepris. Basert på empiri fra blant annet D'Arcy, McGough og Tsolacos (1997) var dette et forventet resultat. Resultatet er likevel oppsiktsvekkende da det ceteris paribus viser at sentralbanken er forsiktig med å føre en motsyklisk pengepolitikk. I tillegg fant vi en kortsiktig kausal sammenheng mellom

arbeidsledighet og leiepris. Dette samsvarer med ARIMAX modellen som fant en signifikant negativ sammenheng. Resultatet underbygger studien til De Wit og Van Dijk (2003).

Analysen som inkluderer variablene forventet nybygging og arealledighet er de mest usikre da vi her kun hadde årlige data tilgjengelig. Resultatet av dette er at dataene ble interpolert med cubic splines. Siden interpolert data er estimater og ikke virkelige verdier blir analysene mer usikre. Det er også mulig at andre interpoleringsteknikker ville gitt andre resultater, for eksempel bruk av Kalman filter.

I empirisk forskning vil tilgjengelighet og kvalitet på dataene til en viss grad diktere hvilke økonomiske metoder og tilnærminger som kan anvendes. Siden det fortsatt er lite eiendomsdata tilgjengelig begrenser dette hvilke metoder som kan brukes for å modellere eiendomsmarkedet. I videre forskning er det særlig tre problemstillinger som vil være spesielt interessante. Det første er å undersøke hvor robuste resultatene i utredningen er. Når mer datamateriale er tilgjengelig vil man kunne gjøre liknende analyse og dermed underbygge (eventuelt tilbakevise) resultatene som er funnet i denne utredningen. Videre ville det vært interessant å undersøke metoder som tar mer hensyn til at eiendom er en særdeles heterogen aktivaklasse, som vanskeliggjør estimeringen av tilbud og etterspørsel. For eksempel ville det vært interessant å kategorisere Oslo-området i mindre geografiske områder ved bruk av paneldata. Dette vil kunne bidra til å gjøre prognosene mer beslutningsrelevante for leietakere som ønsker å leie eiendom innenfor et gitt geografisk område. Til slutt ville det vært interessant å analysere markedet ved bruk av såkalte hedoniske modeller. Disse modellene vil ta mer hensyn til de ulike karakteristikkene ved de forskjellige eiendommene, og kan dermed blitt brukt til å utarbeide en prisindeks eller vurdere verdien til kontoreiendommene. Dette er særlig interessant i markedet for næringseiendom siden lite transaksjonsdata er tilgjengelig.

9. Litteraturliste

- Andrews, D. W. K. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, 821-856.
- Arealstatistikk. *Om oss*. Tilgjengelig fra: <http://www.arealstatistikk.no/om-oss/>.
- Arealstatistikk. 2017. Leieprisoversikten
- Armstrong, J. S. & Collopy, F. 1992. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8(1), s 69-80.
- Bera, A. K. & Jarque, C. M. 1981. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 7(4), s 313-318.
- Box, G. E. P. & Jenkins, G. M. 1976. *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.
- Brooks, C. 2014. *Introductory Econometrics for Finance*: Cambridge University Press.
- Brooks, C. & Tsolacos, S. 2001. Forecasting real estate returns using financial spreads. *Journal of Property Research*, 18(3), s 235-248.
- Brooks, C. & Tsolacos, S. 2003. International evidence on the predictability of returns to securitized real estate assets: econometric models versus neural networks. *Journal of Property Research*, 20(2), s 133-155.
- Brooks, C. & Tsolacos, S. 2010. *Real Estate Modelling and Forecasting*, University Printing House, Cambridge CB2 8BS, United Kingdom: Cambridge University Press.
- Brun, T. A. 2016. Rekordhøyt transaksjonsvolum i Norden. *Estate nyheter*. <http://www.estatenyheter.no/2016/12/22/rekordhoyt-transaksjonsvolum-i-norden/>
- Chatfield, C. & Yar, M. 1988. Holt-Winters Forecasting: Some Practical Issues. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 37(2), s 129-140.
- Corradi, V. & Swanson, N. R. 2006. Predictive density evaluation. *Handbook of economic forecasting*, 1(197-284).

-
- D'Agostino, R. B. & Belanger, A. 1990. A Suggestion for Using Powerful and Informative Tests of Normality. *The American Statistician*, 44(4), s 316-321.
- D'Arcy, E., McGough, T. & Tsolacos, S. 1997. National economic trends, market size and city growth effects on European office rents. *Journal of Property Research*, 14(4), s 297-308.
- D'Arcy, E., McGough, T. & Tsolacos, S. 1999. An econometric analysis and forecasts of the office rental cycle in the Dublin area. *Journal of Property Research*, 16(4), s 309-321.
- De Wit, I. & Van Dijk, R. 2003. The Global Determinants of Direct Office Real Estate Returns. *Journal of Real Estate Finance and Economics*, 26(1), s 27-45.
- Dickey, D. A. & Fuller, W. A. 1979. Distributions of the Estimators for Autoregressive Time Series With a Unit Root. *Journal of the American Statistical Association*, 74(366), s 427-431.
- Dobson, S. M. & Goddard, J. A. 1992. The Determinants of Commercial Property Prices and Rents. *Bulletin of Economic Research*, 44(4), s.
- Eiendomsspar. 2004. *Oslostudiet*. Tilgjengelig fra: <https://www.eiendomsspar.no/wp-content/uploads/2016/05/Oslostudiet-2004.pdf>.
- Eiendomsspar. 2017. *OsloStudiet*. Tilgjengelig fra: https://www.eiendomsspar.no/wp-content/uploads/2017/02/OsloStudiet_2017.pdf.
- Enders, C. K. 2010. *Applied missing data analysis*: Guilford Press.
- Engle, R. F. & Granger, C. W. J. 1987. Co-integration and error correction: representation, estimation, and testing. *Econometrica: journal of the Econometric Society*, 251-276.
- ESRB. 2015. Report on commercial real estate and financial stability in the EU
- Frew, J. & Jud, D. 1988. The vacancy rate and rent levels in the commercial office market. *Journal of Real Estate Research*.
- Gardiner, C. & Henneberry, J. 1989. The development of a simple regional office rent prediction model. *Journal of Valuation*, 7(1), s 36-52.

- Geltner, D. M., Miller, N. G., Clayton, J. & Eichholtz, P. 2007. Commercial real estate. *Thomson Higher Education, Mason*.
- Giussani, B., Hsia, M. & Tsolacos, S. 1992. A Comparative Analysis of the Major Determinants of Office Rental Values in Europe. *Journal of Property Valuation and Investment*, 11(2), s 157-173.
- Granger, C. W. J. & Newbold, P. 1974. Spurious regressions in econometrics. *Journal of econometrics*, 2(2), s 111-120.
- Grytten, O. H. & Hunnes, A. 2016. *Krakk og Kriser i historisk perspektiv*: Cappelen Damm Akademiske.
- H., A. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), s 716-723.
- Hagen, M. 2016. Næringseiendom i Norge,
- Hamilton, J. D. 1994. *Time series analysis*: Princeton university press Princeton.
- Hekman, J. S. 1985. Rental Price Adjustment and Investment in the Office Market. *Real Estate Economics*, 13(1), s 32-47.
- Hendershott, P., MacGregor, B. & White, M. 2002. Explaining real commercial rents using an error correction model with panel data. *The Journal of Real Estate Finance and Economics*, 24(1-2), s 59-87.
- Hendry, D. F. & Clements, M. P. 2000. Economic forecasting in the face of structural breaks. *Econometric Modelling: Techniques and Applications*, 3(37).
- Holden, K., Peel, D. A. & Thompson, J. L. 1990. *Economic Forecasting: an introduction*, The Pitt Building, Trumpington Street, Cambridge CB2 1RP: Cambridge University Press.
- Holland, A. S., Ott, S. H. & Riddiough, T. J. 2000. The role of uncertainty in investment: An examination of competing investment models using commercial real estate data. *Real Estate Economics*, 28(1), s 33-64.
- Johansen, S. 1995. *Likelihood-based inference in cointegrated vector autoregressive models*: Oxford University Press on Demand.

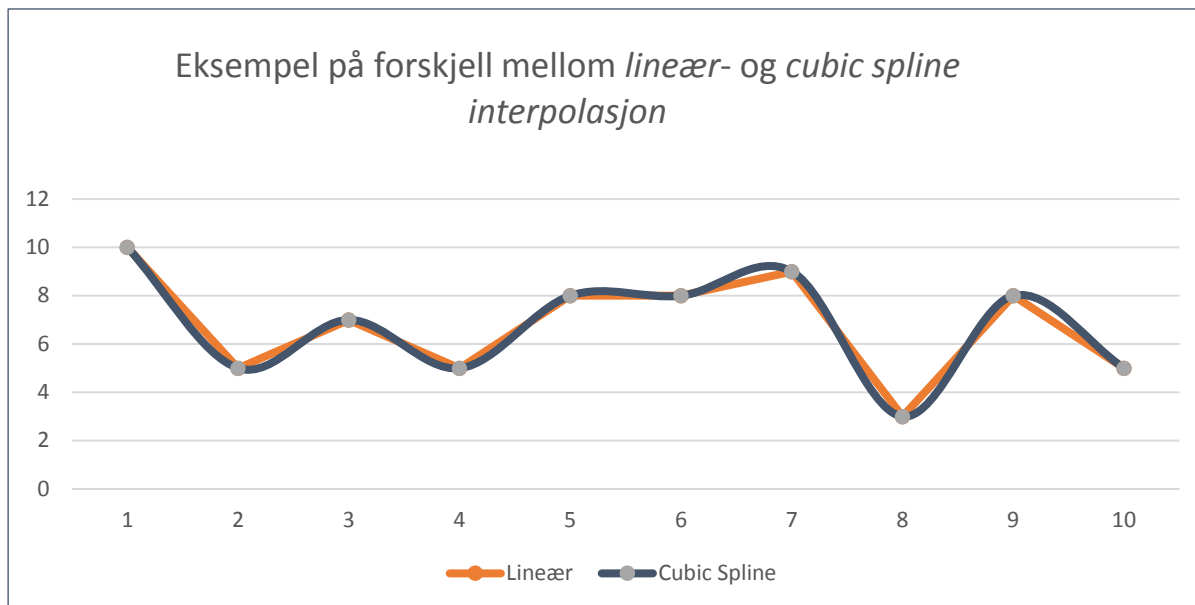
-
- Karakozova, O. 2004. Modelling and forecasting office returns in the Helsinki area. *Journal of Property Research*, 21(1), s 51-73.
- Kragh-Sørensen, K. & Solheim, H. 2014a. Hva taper bankene penger på under kriser?, NorgesBank.
- Kragh-Sørensen, K. & Solheim, H. 2014b. Kanaler fra høy gjeld i husholdningene til tap i bankene, NorgesBank.
- Kuzmenko, E., Smutka, L., Pankov, M. & Efimova, N. 2017. The Success of Economic Policies in Russia: Dependence on Crude Oil vs. Export Diversification. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 65(1), s 299-310.
- Ljung, G. M. & Box, G. E. P. 1978. On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65(2), s 297-303.
- McGough, T. & Tsohacos, S. 2001. Do yields reflect property market fundamentals. *Real Estate and Finance Investment Research Paper*.
- McKinley, S. & Levine, M. 1998. Cubic spline interpolation. *College of the Redwoods*, 45(1), s 1049-1060.
- McNees, S. K. 1986. Forecasting accuracy of alternative techniques: A comparison of US macroeconomic forecasts. *Journal of Business and Economic Statistics*, 4(5-23).
- Naranjo, A. & Ling, D. C. 1997. Economic Risk Factors and Commercial Real Estate Returns. *The Journal of Real Estate Finance and Economics*, 14(3), s 283-307.
- Nau, R. 2017. *ARIMA models for time series forecasting*. Fuqua School of Business. Duke University. Tilgjengelig fra: <https://people.duke.edu/~rnau/411home.htm>.
- NorgesBank. 2017a. Pengepolitisk rapport 1/17, NorgesBank (www.norges-bank.no).
- NorgesBank. 2017b. *Statsobligasjoner, Månedsgjennomsnitt*. Oslo Børs. Tilgjengelig fra: <http://www.norges-bank.no/Statistikk/Rentestatistikk/Statsobligasjoner-Rente-Manedsgjennomsnitt-av-daglige-noteringer/>.
- Plazzi, A., Torous, W. & Valkanov, R. 2010. Expected Returns and Expected Growth in Rents of Commercial Real Estate. *The Review of Financial Studies*, 23(9), s 3469-3519.

- Rosen, K. T. 1984. Toward a Model of the Office Building Sector. *Real Estate Economics*, 12(3), s 261-269.
- Rosen, K. T. & Smith, L. B. 1983. The price-adjustment process for rental housing and the natural vacancy rate. *The American Economic Review*, 73(4), s 779-786.
- Rossi, B. 2014. Density Forecasts in Economics, Forecasting and Policymaking.
- Saltnes, D.-J. 2017. Høyest privat eierandel i Norden. *Estate Magasin*. Estate.
- Savills. 2016. Around the world in dollars and cents,
- Schwarz, G. 1978. Estimating the Dimension of a Model. 461-464.
- Sims, C. A. 1980. Macroeconomics and reality. *American Economic Review*, 62(4), s 540-552.
- Sirmans, S. G. & Benjamin, J. D. 1991. Determinants of Market Rent. *Journal of Real Estate Research*, 6(3), s 357-380.
- Sivitanides, P. 1997. The rent adjustment process and the structural vacancy rate in the commercial real estate market. *Journal of Real Estate Research*, 13(2), s 195-209.
- SSB. 2016. *Kvartalsvis Nasjonalregnskap*. Tilgjengelig fra: <https://ssb.no/nasjonalregnskap-og-konjunkturer/statistikker/knr/kvartal/2017-02-09#content> [Lastet ned 9. februar 2017].
- SSB. 2017. *Produsentprisindeksen*. SSB, Statistisk Sentralbyrå. Tilgjengelig fra: <https://www.ssb.no/priser-og-prisindekser/statistikker/ppi>.
- StataCorp. 2013. *Stata Time-Series Reference Manual*. Stata Press.
- Sukati, M. A. 2013. Cointegration Analysis of Oil Prices and Consumer Price Index in South Africa using STATA Software.
- Tse, R. Y. C. 1997. An application of the ARIMA model to real-estate prices in Hong Kong. *Journal of Property Finance*, 8(2), s 152-163.
- Tsolacos, S. A Symmetric Responses of UK Real Estates Returns to the Business Cycle. 2002.

-
- Tsolacos, S., Keogh, G. & McGough, T. 1998. Modelling use, investment, and development in the British office market. *Environment and Planning A*, 30(8), s 1409-1427.
- Uniongruppen. 2002. *Markedsrapport næringsseiendom 04/02* [Online]. Tilgjengelig fra: http://www.union.no/wp-content/uploads/marketreport_2002_q4.pdf.
- Uniongruppen. 2004. *Markedsrapport Næringsseiendom* [Online]. Tilgjengelig fra: http://www.union.no/wp-content/uploads/marketreport_2004_q4.pdf.
- Uniongruppen. 2016. *Markedsrapport Høst* [Online]. Tilgjengelig fra: <http://www.union.no/wp-content/uploads/markedsrapport-hst-20161.pdf>.
- Vishwakarma, V. K. 2013. Forecasting real estate business: empirical evidence from the canadian market.
- Wheaton, W. C. 1987. The Cyclic Behavior Of The National Office Market. *AREUEA Journal*, 15(4), s.
- Wheaton, W. C. & Torto, R. G. 1988. Vacancy Rates and the Future of Office Rents. *Real Estate Economics*, 16(4), s 430-436.
- Wilson, P., Okunev, J., Ellis, C. & Higgins, D. 2000. Comparing univariate forecasting techniques in property markets. *Journal of Real Estate Portfolio Management*, 6(3), s 283-306.

Appendiks

A. Interpolering av data

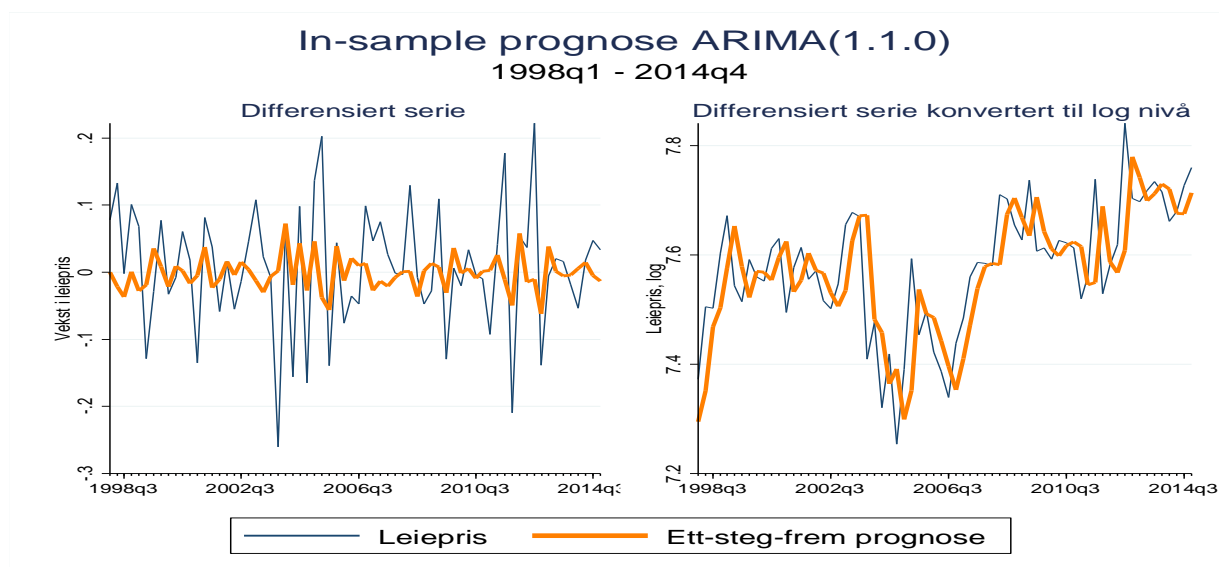


Figur 18: Grafisk fremstilling av forskjell mellom lineær- og cubic spline interpolasjon

En av hovedutfordringene ved analyser av eiendomsmarkedet er mangel på data. Siden markedet er relativt illikvid er det få transaksjoner og det er lav frekvens på oppdateringen av dataene. Denne utfordringen møtte vi særlig under de markedsspesifikke dataene, hvor det kun fantes årlige data. Vi måtte dermed ta en avgjørelse om vi skulle bruke årlige eller kvartalsvise data, hvor begge utfall hadde både positive og negative sider. Ved å benytte oss av årlige data ville vi miste et stort antall observasjoner, som ville kunne gjøre analysen svakere. I tillegg var det kun de markedsspesifikke variablene vi manglet kvartalsvise data på, så vi valgte derfor å interpolere disse variablene for å finne et estimat på kvartalsvise data. Det finnes flere metoder for å interpolere data som gir noe forskjellige estimater på verdiene. Den vanligste er lineær interpolasjon, men vi ønsket å benytte oss av en interpoleringsmetode som jevner ut observasjonene bedre (McKinley og Levine, 1998). Vi brukte derfor en metode som kalles Cubic Spline Interpolasjon.

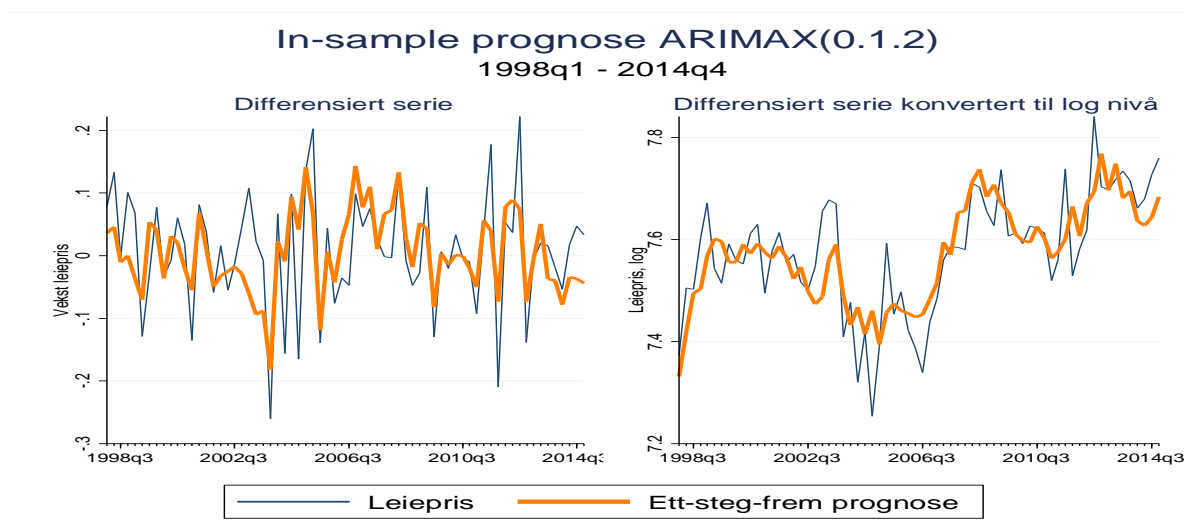
Som man ser i Figur 18 gir Cubic Spline interpolasjonen en jevnere linje mellom de kjente datapunktene. Vi anser dette som mer realistiske estimater på de manglene verdiene, og valgte derfor å bruke denne interpolasjonsmetoden for å estimere kvartalsvise markedsspesifikke data. Selv med realistiske estimater på de kvartalsvise dataene vil aldri interpolasjon gi de faktiske verdiene, og vi må derfor ta forbehold om at de markedsspesifikke dataene er estimater, og ikke faktiske verdier.

B. Grafisk in-sample prognose 1998q1 – 2014q4



Figur 19: In-sample prognose ARIMA modell (2b)

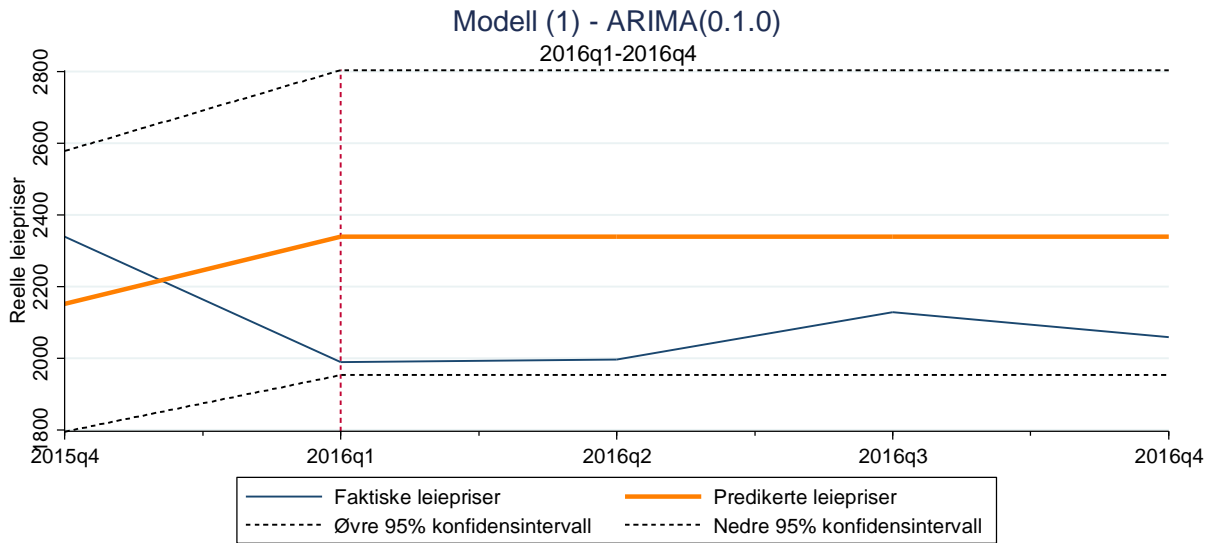
Denne figuren viser in-sample prognose for vekst i leiepriser og leiepriser for kontoreiendom i Oslo-området sammenlignet med den faktiske utviklingen i perioden 1998q1 – 2014q4. Venstre graf viser utviklingen i vekst i leiepris. Høyre graf viser utviklingen i leiepris transformert tilbake til log-nivå.



Figur 20: In-sample prognose ARIMAX modell (3b)

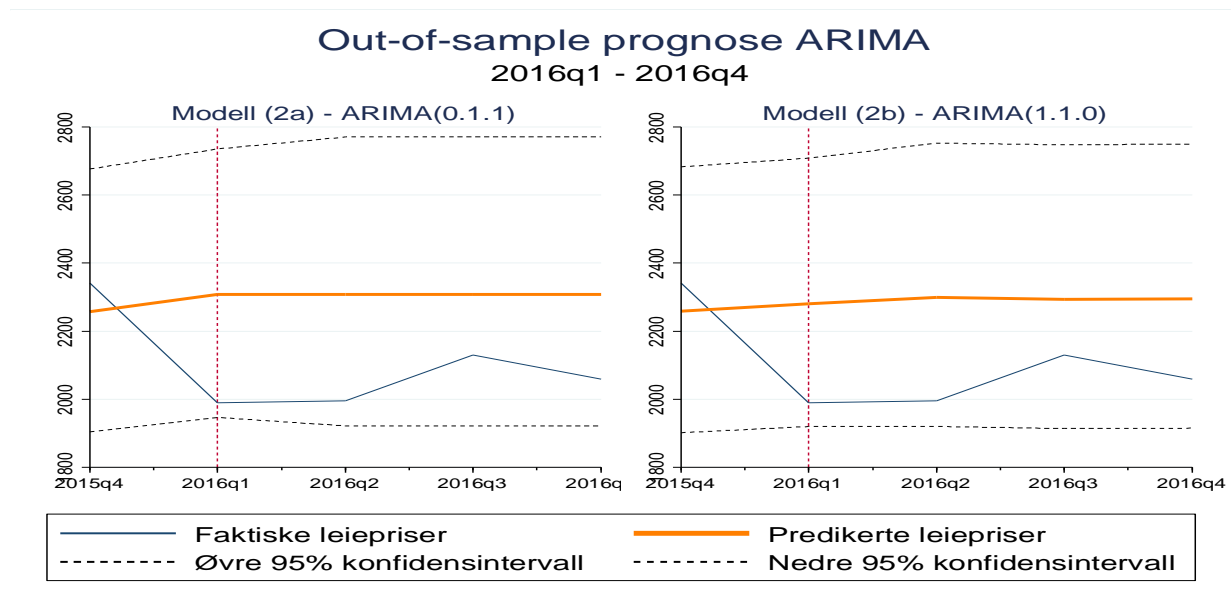
Denne figuren viser in-sample prognose for vekst i leiepriser og leiepriser for kontoreiendom i Oslo-området sammenlignet med den faktiske utviklingen i perioden 1998q1 – 2014q4. Venstre graf viser utviklingen i vekst i leiepris. Høyre graf viser utviklingen i leiepris transformert tilbake til log-nivå.

C. Out-of-sample resultater 2016q1 – 2016q4



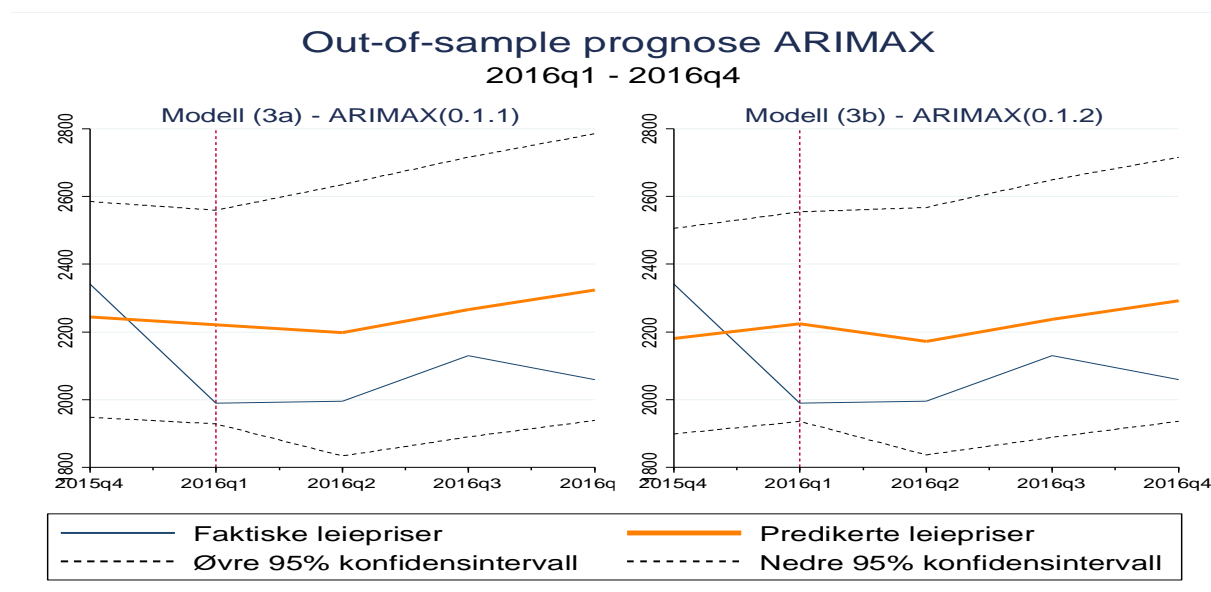
Figur 21: Dynamisk out-of-sample prognose Naiv modell (1), 2016q1 – 2016q4

Figuren viser ettårig prognose for reelle leiepriser på nivåform for modell (1). Prognosen er gjennomført på log differensierte leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2016q1-2016q4.



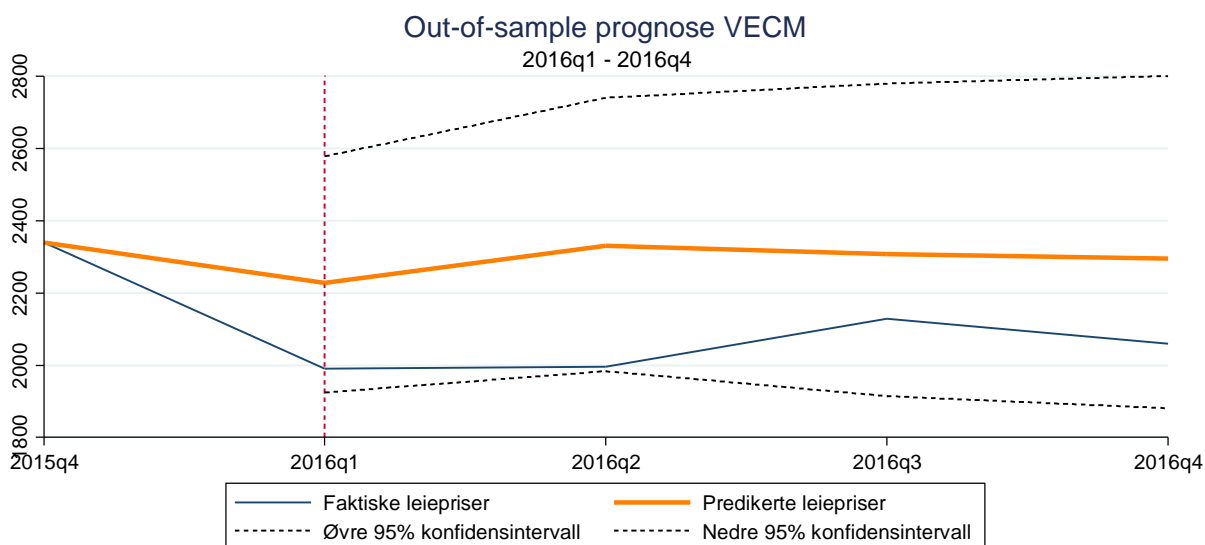
Figur 22: Dynamisk out-of-sample prognose ARIMA modell (2), 2016q1 - 2016q4

Figuren viser ettårig prognose for reelle leiepriser på nivåform for modell (2a) og (2b). Prognosen er gjennomført på log differensierte leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2016q1-2016q4.



Figur 23: Dynamisk out-of-sample prognose ARIMAX modell (3), 2016q1 - 2016q4

Figuren viser ettårig prognose for reelle leiepriser på nivåform for modell (3a) og (3b). Prognosen er gjennomført på log differensierte leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2016q1-2016q4.



Figur 24: Dynamisk out-of-sample prognose VECM (4), 2016q1 - 2016q4

Figuren viser ettårig prognose for reelle leiepriser på nivåform for modell (4). Prognosen er gjennomført på log leiepriser og konvertert tilbake til leiepriser på nivå i perioden 2016q1-2016q4.

D. ARIMAX, granskende analyse

Tabell 21: Granskende analyse ARIMAX modell (3)

Denne tabellen rapporterer resultatene fra ARIMAX modellen for vekst i leiepriser for kontormarkedet i Oslo-området ved bruk av kvartalsvis tidsseriedata fra 1998q1 til 2014q4. Følgende modell er estimert:

$$\Delta \ln LP_t = \alpha_0 + \phi_p \Delta \ln LP_{t-i} + \beta_1 \Delta \ln BNP_{t-2} + \beta_2 \Delta ARBL_{t-2} + \beta_3 \Delta RRTI_{t-2} + \beta_4 \Delta \ln FNYB_{t-2} + \beta_5 \Delta ARL_{t-2} + \alpha_1 D_{t-2}(2001q1 - 2003q4) + \delta_1 D_{t-2}(2001q1 - 2003q4) \Delta ARL_{t-2} + \varepsilon_t + \theta_q \varepsilon_{t-q} \quad (3)$$

Standardfeil er estimert robust. Se Tabell 9 for mer informasjon om benevninger og diagnostiske tester.

Modell	(3a)				
$\Delta \ln LP$	(0.1.1)	(2.1.0)	(2.1.0)	(0.1.1)	(2.1.0)
konstant	0.00396 (1.91)	0.00595 (0.89)	0.00648 (1.16)	0.00417 (1.68)	0.00539 (0.84)
$\Delta_4 ARBL_{t-2}$	-2.815*** (-6.44)	-3.187** (-3.98)	-2.616* (-2.34)	-2.801*** (-5.59)	-3.070* (-2.69)
$\Delta RRTI_{t-2}$	-2.447* (-2.53)	-2.726** (-2.66)	-2.325 (-2.13)	-2.575* (-2.68)	-2.357* (-2.21)
$\Delta_4 \ln BNP_{t-2}$		-0.0155 (-0.17)			0.0101 (0.10)
ΔARL_{t-2}			-1.366 (-0.49)		0.00248 (0.00)
$D_{t-2}(2001q1 - 2003q4)$			-0.0466 (-1.85)		-0.0349 (-1.43)
$D_{t-2}(2001q1 - 2003q4) \Delta ARL_{t-2}$			7.168 (1.76)		5.711 (1.43)
$\Delta \ln FNYB_{t-2}$				0.0185 (0.54)	0.0419 (0.98)
ARIMA					
AR(1)		-0.433*** (-3.63)	-0.470*** (-4.13)		-0.480*** (-4.26)
AR(2)		-0.181 (-1.53)	-0.246* (-2.03)		-0.250 (-2.11)
MA(1)	-0.806** (-2.94)			-0.761* (-2.18)	
<i>sigma</i>	0.0727*** (9.87)	0.0764*** (9.78)	0.0737*** (10.12)	0.0725*** (10.13)	0.0732*** (11.44)
AIC	-152.4	-142.6	-143.4	-151.0	-140.2
BIC	-141.3	-127.1	-123.5	-137.7	-115.8
ll	81.20	78.30	80.72	81.50	81.12
LR-test (X^2)		-5.80(3a)	-0.96(3a)	0.60(3a)	-0.17(3a)
Justert R^2	0.38	0.30	0.35	0.38	0.36
Box-Ljung Q	33.50	38.70	38.61	34.77	40.62
sktest(X^2)	2.41	2.83	2.14	1.91	2.17
ADF 4.lag	-4.606***	-5.509***	-5.061***	-4.573***	-4.934***
z statistikk i parentes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$					

E. Stasjonærhetstest for ARIMAX modell (3b)

Tabell 22: *Eigenverdi stabilitetstilstand modell (3b)*

Eigenverdi	Moduli
.9999987	.999999
.2698804	.26988

*Alle eigenverdiene ligger innenfor enhets sirkelen.
MA parameterne tilfredsstiller inverter barhets tilstand*

F. VECM koeffizientmatrise

	Coef.	Std. Err.	z	P> z	[95% Conf. Intervall]	
D_lnLP						
_ce1						
L1.	-.3460265	.1632086	-2.12	0.034	-.6659096	-.0261434
_ce2						
L1.	-2.581314	1.769052	-1.46	0.145	-6.048592	.8859641
lnLP						
LD.	-.2767645	.1431906	-1.93	0.053	-.557413	.0038839
L2D.	-.137995	.1267392	-1.09	0.276	-.3863992	.1104093
L3D.	-.1338247	.1118395	-1.20	0.231	-.3530261	.0853766
ARBL						
LD.	5.916721	2.540986	2.33	0.020	.9364813	10.89696
L2D.	-.9926752	3.074589	-0.32	0.747	-7.018759	5.033409
L3D.	-4.328427	2.464851	-1.76	0.079	-9.159447	.5025928
lnFNYB						
LD.	-.3907028	.2796525	-1.40	0.162	-.9388117	.1574061
L2D.	.4380869	.4756804	0.92	0.357	-.4942295	1.370403
L3D.	-.1135814	.2897481	-0.39	0.695	-.6814772	.4543144
RRTI						
LD.	2.244691	1.296245	1.73	0.083	-.2959029	4.785286
L2D.	-.3135754	1.280545	-0.24	0.807	-2.823397	2.196246
L3D.	-.0357227	1.09312	-0.03	0.974	-2.178199	2.106753
_cons	.0002251	.01019	0.02	0.982	-.0197469	.0201972
D_ARBL						
_ce1						
L1.	.0283039	.0094101	3.01	0.003	.0098605	.0467472
_ce2						
L1.	-.1465467	.1019975	-1.44	0.151	-.3464581	.0533647

lnLP							
LD.	-.0147834	.0082559	-1.79	0.073	-.0309647	.0013978	
L2D.	-.0032729	.0073073	-0.45	0.654	-.017595	.0110493	
L3D.	-.0034335	.0064483	-0.53	0.594	-.0160719	.0092049	
ARBL							
LD.	-.6891355	.1465045	-4.70	0.000	-.9762792	-.4019919	
L2D.	-.2914041	.1772703	-1.64	0.100	-.6388475	.0560393	
L3D.	-.3217657	.1421149	-2.26	0.024	-.6003058	-.0432256	
lnFNYB							
LD.	-.0246778	.0161238	-1.53	0.126	-.0562799	.0069243	
L2D.	.0484512	.0274261	1.77	0.077	-.005303	.1022053	
L3D.	-.0173355	.0167059	-1.04	0.299	-.0500785	.0154074	
RRTI							
LD.	-.1194411	.0747371	-1.60	0.110	-.2659231	.0270409	
L2D.	-.1754192	.0738318	-2.38	0.018	-.320127	-.0307115	
L3D.	-.1180079	.0630256	-1.87	0.061	-.2415357	.00552	
_cons	.0008473	.0005875	1.44	0.149	-.0003042	.0019989	
D_lnFNYB							
_ce1							
L1.	-.080437	.0530613	-1.52	0.130	-.1844352	.0235612	
_ce2							
L1.	-2.68294	.5751423	-4.66	0.000	-3.810198	-1.555681	
lnLP							
LD.	-.0639453	.0465532	-1.37	0.170	-.1551879	.0272972	
L2D.	-.0759136	.0412046	-1.84	0.065	-.1566731	.0048459	
L3D.	-.0174886	.0363605	-0.48	0.631	-.0887539	.0537767	
ARBL							
LD.	2.511678	.8261081	3.04	0.002	.8925362	4.130821	
L2D.	3.638173	.9995897	3.64	0.000	1.679013	5.597332	
L3D.	2.186413	.8013559	2.73	0.006	.6157847	3.757042	
lnFNYB							
LD.	2.065709	.0909188	22.72	0.000	1.887512	2.243907	
L2D.	-1.739154	.15465	-11.25	0.000	-2.042263	-1.436046	
L3D.	.6450216	.0942009	6.85	0.000	.4603911	.829652	

RRTI							
LD.	1.399167	.4214266	3.32	0.001	.5731858	2.225148	
L2D.	1.744372	.4163221	4.19	0.000	.928396	2.560348	
L3D.	1.404976	.3553878	3.95	0.000	.708429	2.101524	
_cons	-.0001659	.0033129	-0.05	0.960	-.0066591	.0063273	
<hr/>							
D_RRTI							
_ce1							
L1.	-.0230185	.0204687	-1.12	0.261	-.0631364	.0170995	
_ce2							
L1.	-.1476086	.2218646	-0.67	0.506	-.5824553	.2872381	
lnLP							
LD.	-.0012025	.0179582	-0.07	0.947	-.0363998	.0339949	
L2D.	.0069103	.0158949	0.43	0.664	-.0242432	.0380638	
L3D.	-.0023115	.0140263	-0.16	0.869	-.0298026	.0251795	
ARBL							
LD.	.469903	.3186763	1.47	0.140	-.154691	1.094497	
L2D.	.5423777	.3855978	1.41	0.160	-.2133801	1.298136	
L3D.	-.1527917	.3091279	-0.49	0.621	-.7586713	.4530879	
lnFNYB							
LD.	.0222865	.0350725	0.64	0.525	-.0464543	.0910273	
L2D.	-.0494015	.0596572	-0.83	0.408	-.1663275	.0675244	
L3D.	.0308843	.0363386	0.85	0.395	-.040338	.1021066	
RRTI							
LD.	-.5061736	.1625679	-3.11	0.002	-.8248008	-.1875464	
L2D.	-.3958645	.1605988	-2.46	0.014	-.7106323	-.0810967	
L3D.	-.2667591	.137093	-1.95	0.052	-.5354565	.0019383	
_cons	-.0017628	.001278	-1.38	0.168	-.0042676	.000742	
