

Forord

Masteroppgaven representerer slutten av masterstudiet i Finansiell Økonomi ved Institutt for Samfunnsøkonomi ved Norges Teknisk-Naturvitenskapelige Universitet (NTNU). Oppgaven utgjør 30 studiepoeng og tilsvarer ett semester.

Jeg vil takke min veileder, Snorre Lindset, for samarbeidet gjennom masteroppgaven.

Innholdet i denne oppgaven står for forfatterens regning.

Trondheim, Juni 2018

Veronica Birkeland Mo

Abstract

The primary aim of this master thesis is to create models that can predict future index returns of the Oslo Exchange Benchmark Index (OSEBX). The independent variables are OSEBXs index returns from previous years, and correlations between eight different sector returns and the OSEBX return. The data is collected from Thomson Reuters Datastream and the area of examination is the Norwegian stock market. Indices of eight sectors from Oslo Stock Exchange (OSE) are used as variables representing the performance of each sector to predict the return of OSEBX. The purpose of creating these models is to determine whether certain sectors are more suited for forecasting the return of OSEBX than the others.

This thesis focuses on four extensions of the EGARCH(1,1)-model, and their ability to explain and predict OSEBXs returns and conditional variance. The extension is represented by the correlation between sectors returns and OSEBX return in the preceding period.

The main finding is that two sectors are better at predicting future index returns than the original EGARCH(1,1)-model. The extended models that are more effective at predicting future index returns include the correlation between the consumer goods sector return and the OSEBX return, and the correlation between the finance sector and OSEBX. Among the total of 32 models estimated in this thesis, none yielded solely significant results, similar to the original EGARCH(1,1)-model. Still, two models are more effective at predicting future index return than the original EGARCH(1,1)-model. However, the models portrayed in this thesis does not offer perfect predictions as they are still fairly uncertain and yields some exclusively significant results.

Sammendrag

Hovedmålet med denne masteroppgaven er å lage modeller som kan forutse avkastningen til hovedindeksen på Oslo Børs (Oslo Exchange Benchmark Index, OSEBX). De uavhengige variablene er avkastningen til OSEBX og korrelasjoner mellom avkastningen til åtte forskjellige bransjeindekser og avkastningen til OSEBX. Dataene er hentet fra Thomson Reuters Datastream, og brukes til å undersøke det norske børsmarkedet. I oppgaven undersøkes det om korrelasjonen mellom avkastningene til 8 bransjeindekser ved Oslo børs og OSEBX kan brukes til å forutse avkastningen til OSEBX. Målet med å lage disse modellene er å vurdere om enkelte sektorer kan forutse avkastningen til OSEBX bedre enn de andre sektorene.

Denne oppgaven fokuserer på fire utvidelser av en EGARCH(1,1)-modell, og disses evne til å forklare og forutse avkastningen og variansen til OSEBX. Utvidelsen av disse EGARCH(1,1)-modellene utgjøres av korrelasjonen mellom avkastningene til de ulike bransjeindeksene og avkastningen til OSEBX i den forestående perioden.

Det viktigste funnet er at to sektorer forutser avkastningen til OSEBX bedre enn en opprinnelig EGARCH(1,1)-modell. De utvidede modellene som er bedre til å forutse framtidig avkastning inkluderer henholdsvis korrelasjonen mellom avkastningen til bransjeindeksen Forbruksvarer og avkastningen til OSEBX, og korrelasjonen mellom avkastningen til bransjeindeksen Finans og avkastningen til OSEBX. Blant de totalt 32 modellene som ble estimert i denne oppgaven, ga ingen av dem utelukkende signifikante resultater, i likhet med den opprinnelige EGARCH(1,1)-modellen. Likevel er de to modellene bedre til å forutse framtidig avkastning enn den opprinnelige EGARCH(1,)-modellen. Et forbehold er at modellene som presenteres i denne oppgaven ikke gir perfekte prognoser, da de fortsatt er usikre og gir ikke utelukkende signifikante resultater.

Innholdsfortegnelse

| | | |
|----------|---|-----------|
| 1 | Introduksjon | 1 |
| 1.1 | Tidligere forskning | 1 |
| 1.2 | Problemstilling | 4 |
| 2 | Metode | 5 |
| 2.1 | Avkastning | 5 |
| 2.2 | Korrelasjon | 5 |
| 2.3 | GARCH | 6 |
| 2.4 | TGARCH | 8 |
| 2.5 | EGARCH | 9 |
| 2.6 | Utvidet EGARCH(1,1) | 10 |
| 2.7 | Stasjonære prosesser | 12 |
| 2.8 | Autokorrelasjon | 13 |
| 2.9 | ARCH-effekter | 16 |
| 2.10 | Restleddets fordeling | 17 |
| 2.11 | Informasjonskriterier | 18 |
| 2.12 | Sammenligning av prediksjonsnøyaktighet | 20 |
| 2.12.1 | Root mean squared error (RMSE) | 20 |
| 2.13 | Mean absolute error (MAE) | 20 |
| 3 | Data | 21 |
| 3.1 | Beskrivelse av data | 21 |
| 3.1.1 | OSEBX | 21 |
| 3.1.2 | Bransjeindekser | 23 |
| 3.2 | Analyse av data | 25 |
| 3.2.1 | Historisk avkastning OSEBX | 25 |
| 3.2.2 | Deskriptiv statistikk | 27 |
| 3.3 | Korrelasjon | 28 |
| 4 | Resultater | 29 |
| 4.1 | Testverdier for Dickey-Fuller test | 30 |
| 4.2 | Testverdier for ARCH-effekt | 31 |
| 4.3 | Valg av orden i GARCH-modellene | 31 |
| 4.4 | Resultater av modellestimeringer | 32 |
| 4.5 | Test for autokorrelasjon i restleddene | 48 |
| 4.6 | Predikert avkastning OSEBX | 50 |
| 4.7 | Evaluering av prediksjonsmodellene | 53 |
| 5 | Diskusjon | 53 |

| | | |
|----------|--|-----------|
| 6 | Konklusjon..... | 58 |
| | Referanseliste | 59 |
| | Vedlegg/Appendiks | 64 |
| | <i>Vedlegg A: Sammensetningen av OSEBX.....</i> | <i>65</i> |
| | <i>Vedlegg B: Sektorsammensetning OSEBX.....</i> | <i>67</i> |
| | <i>Vedlegg C: Predikert avkastning OSEBX i tallverdi</i> | <i>67</i> |
| | <i>Vedlegg D: Begreper.....</i> | <i>69</i> |

1 Introduksjon

Denne masteroppgaven fokuserer på utvidelser av EGARCH(1,1)-modellen og deres evne til å predikere avkastningen til OSEBX. GARCH-modeller brukes i hovedsak til å predikere volatilitet, da dette er enklere å predikere enn avkastning. Likevel brukes modellene også mye til å predikere avkastning (Arowolo, 2013).

Masteroppgaven består av totalt seks deler. Del 1 inneholder gjennomgang av tidligere forskning og begrunnelse for valg av oppgave. Del 2 er en metodedel, som tar for seg hvilke mål og modeller som vil bli brukt gjennom oppgaven. Del 3 inneholder en kort presentasjon og analyse av dataene. Del 4 tar for seg resultatet av modellestimatene. Del 5 og 6 er henholdsvis diskusjon og konklusjon. Til slutt finnes referanseliste og vedlegg.

1.1 Tidligere forskning

Det har lenge vært interesse for å predikere avkastningen til aksjer og andre finansielle instrumenter ved bruk av økonometriske modeller. Ifølge Efficient market hypothesis (EMH), er det derimot ikke mulig å lage modeller som kan predikere avkastning, da denne sier at all relevant informasjon er tilgjengelig i markedet. Dette innebærer totalt fravær av arbitrasjemuligheter, og kan knyttes opp mot teorien om 'random walk'. Random walk forklares ytterligere senere, men innebærer at historiske priser ikke kan brukes til å predikere framtidige priser, da prosessene varierer over tid (Timmermann og Granger, 2004). Ingen har hittil klart å lage en prediksjonsmodell som fungerer over lengre tid, og selv om man skulle finne en prediksjonsmodell som klarer å predikere markedet en viss periode fram i tid er ikke dette i seg selv nok til å motbevise EMH, da teorien står sterkt i finansiell teori (Timmermann og Granger, 2004).

Det eksisterer tre sett av markedsformer i EMH: svak form, semi-sterk form og sterk form. Markedsformen avhenger av variablene som blir brukt i modelleringen man velger. Hvis det bare tas utgangspunkt i tidligere og nåværende aksjepriser i analysen, tester man EMH i sin svake form. Hvis analysen utvides til å inkludere all offentlig informasjon, testes EMH i sin semi-sterke form. Hvis privat informasjon også tillegges testes EMH i sin sterke form. Mesteparten av litteraturen på prediksjon av aksjemarkeder tar utgangspunkt i EMH i svak eller semi-sterk form (Tıtan, 2015), som det også vil tas utgangspunkt i ved denne oppgaven.

Til tross for at EMH står støtt i den finansielle verdenen, er det flere som forsøker å predikere avkastning gjennom å konstruere prediksjonsmodeller. Målet med dette er å finne den beste prediksjonsmodellen for et gitt finansielt instrument. Å finne en optimal prediksjonsmodell kalles ofte 'thin modeling'. Blant andre stiller Granger og Jeon (2004) seg kritisk til bruk av thin modeling som framgangsmåte. I stedet anbefaler forfatterne bruken av 'thick modeling', som innebærer at man tar utgangspunkt i flere modeller med statistisk like utfall (Granger og Jeon, 2004). Budskapet er at man ved thick modeling ikke kommer fram til en modell som er best, men tar i bruk flere av de beste modellene. Thick model-framgangsmåten er brukt i denne oppgaven, da vi ønsker å finne flere modeller som sammen kan predikere avkastningen til OSEBX.

Tidligere forsøk på å konstruere de beste prediksjonsmodellene for avkastning tar utgangspunkt i at volatiliteten til det finansielle instrumentet er konstant. Volatilitet er et viktig begrep i finans, og er et mål på usikkerhet i finansielle instrumentet. Den mest brukte metoden for prediksjon basert på antagelsen om konstant volatilitet er minste kvadraters metode (MKM). Derimot beviste Andersen og Bollerslev (1997) at volatiliteten ikke er konstant, men varierer med tid. For å fange opp effekten av varierende volatilitet, må vi ta utgangspunkt i en modell som tillater at volatiliteten varierer. De mest brukte modellene som oppfyller dette kriteriet er GARCH-modeller. GARCH står for generalized autoregressive conditional heteroscedasticity og er en utvidelse av autoregressive conditional heteroscedasticity (ARCH), som begge forklares i detalj senere. Hovedpoenget med ARCH og GARCH-modeller er at de estimerer avkastning og volatilitet samtidig, der varians brukes som mål på volatilitet.

På grunn av nettopp deres evne til å fange opp effekten av varierende volatilitet, har vi i denne oppgaven valgt å bruke GARCH-modeller til å predikere fremtidig avkastning ved OSEBX. Modellene er kjent for å fange opp både volatilitetsklumping og tykke haler i data, problemer som er vanlig i finansielle data (Niu og Wang, 2013). Volatilitetsklumping innebærer at volatiliteten har tendenser til å samle seg i perioder, som innebærer at perioder med lav volatilitet etterfølges med perioder med lav volatilitet og tilsvarende for perioder med høy volatilitet (Mandelbrot, 1963). Tykke haler er knyttet til fordelingen til restleddet i modellen som brukes, og innebærer at mer av restleddsfordelingen er å finne i halene, sammenlignet med hva man ville forventet av en normalfordeling (se figur 1 for eksempel på restleddsfordeling med fete haler). I de fleste økonometriske modeller går man ut ifra en

antagelse om normalfordelte restledd. Det viser seg derimot at restleddene i finansielle data ofte har tykke haler, som gjør at GARCH-modeller ofte er velegnet også av denne grunnen.

Tradisjonelle GARCH-modeller har derimot også svakheter. En svakhet ved de tradisjonelle GARCH-modellene er at de fungerer dårlig til å fange opp og reflektere asymmetri. Asymmetri innebærer at finansielle volatilitetsdata har en tendens til å la negative nyheter ha større effekt på fluktasjoner enn gode nyheter (Bouchaud, Matacz og Potters, 2001). Gode og dårlige nyheter vil således ha ulik effekt på volatiliteten, og denne asymmetriske effekten betegnes ofte også 'leverage effect'. Som nevnt tidligere estimerer GARCH-modeller avkastning og varians samtidig. Ved beregning av variansen vil tradisjonelle GARCH-modeller bare ta hensyn til størrelsen på avkastningen, men tar ikke høyde for om denne avkastningen er positiv eller negativ (Nelson, 1991). En annen svakhet ved de tradisjonelle GARCH-modellene er antagelsen om at alle koeffisientene skal være større enn null, da disse restriksjonene kan være vanskelig å oppnå i praksis. En ytterligere svakhet ved de tradisjonelle GARCH-modellene er at avkastning er en dårlig representant for volatilitetsnivået, noe som bidrar til at GARCH-modeller passer dårlig ved ekstreme hopp i volatilitet over kort tid (Hansen og Huang, 2016). Som følge av disse hoppene vil det ta flere perioder før volatiliteten justeres, og tradisjonelle GARCH-modeller vil derfor være for trege til å fange opp disse hoppene (Andersen *et al.*, 2003). Det finnes mer avanserte GARCH-modeller konstruert for å fjerne disse svakhetene. Blant annet finnes det GARCH-metoder som tar høyde for asymmetri, som exponential GARCH (EGARCH) (Nelson, 1991) og Threshold GARCH (TGARCH) (Glosten, Jagannathan og Runkle, 1993). EGARCH-modeller tillater også at noen, men ikke alle, koeffisienter kan være mindre enn null (Nelson, 1991). I denne oppgaven brukes derfor de mer avanserte GARCH-modellene EGARCH og TGARCH, i tillegg til en tradisjonell GARCH-modell, for å predikere avkastningen til OSEBX.

De fleste tidligere studier på GARCH-modeller lager prediksjonsmodeller på bakgrunn av tidligere avkastning og volatilitet til indeksen de undersøker. I denne oppgaven skal vi også inkludere effekten av korrelasjonen mellom indeksen som ønskes predikert og relevante sektorindekser i modellene. Dette er for å se om dette kan øke prediksjonsevnen til de ulike GARCH-modellene.

1.2 Problemstilling

Målet for oppgaven er å finne en god modell for forklaring av avkastningen til OSEBX, og nyttiggjøre seg av denne til å predikere den månedlige avkastningen til OSEBX.

Inspirasjon til oppgaven kom fra artikkelen «Ser farlige tendenser i børsroen» av Heljar Havnes, publisert i Dagens Næringsliv 15.mai 2017. Artikkelen bygger på DNB Markets analyse av resultatsesongen, altså tiden når aksjeselskaper publiserer resultater fra det forestående kvartalet. Hovedbudskapet i Dagens Næringsliv-artikkelen er at ved lav korrelasjonen mellom avkastning på OSEBX og sektorindekser, fokuseres det på enkeltelskaper i større grad enn på makroøkonomiske forhold. Videre hevdes det at når korrelasjonen mellom sektorer og OSEBX er lav øker det sannsynligheten for at markedet faller. Budskapet i artikkelen ga ideen om å undersøke om noen sektorer spiller større roller for utviklingen til OSEBX enn andre, og genererte et ønske om å lage en modell som kan forklare denne sammenhengen.

For å utvikle en slik modell er det tatt utgangspunkt i å utvide en EGARCH(1,1)-modell. GARCH-modeller kan brukes til å predikere avkastning og varians basert på tidligere års avkastning, varians og restledd. GARCH-modeller brukt tidligere har derimot ikke tatt hensyn til korrelasjonen mellom sektoravkastning og OSEBX sin avkastning. Dette vil undersøkes i denne oppgaven, da denne korrelasjonen kan bidra til å forklare OSEBX sin avkastning. Med bakgrunn i dette utvides en EGARCH(1,1)-modell ved å inkludere et ledd som tar høyde for korrelasjonen mellom sektoravkastning og OSEBX sin avkastning. Det er elleve sektorer, også kalt bransjer, på Oslo Børs, og i denne oppgaven vil det tas utgangspunkt i åtte av disse. Det lages derfor 32 utvidede EGARCH(1,1)-modeller, som hver inkluderer korrelasjonen mellom avkastningen til en enkelt sektor, og avkastningen til OSEBX. Målet med å lage slike modeller er å undersøke om noen av sektorene har større betydning for OSEBX sin avkastning enn andre, og følgelig kunne bruke noen av disse forbedrede modellene til å predikere avkastningen til OSEBX.

Endelig problemstilling:

«Kan korrelasjonen mellom hver bransjeindeks og hovedindeksen bidra til å forklare avkastningen i Oslo Børs sin hovedindeks, og kan denne brukes til å lage prognoser på OSEBX sin avkastning de nærmeste 12 månedene?»

2 Metode

Metodekapittelet tar for seg teorien til modellene som blir brukt i oppgaven, i tillegg til hvilke mål og tester som blir anvendt. Først forklares det hvordan avkastning og korrelasjon blir beregnet i denne oppgaven. Deretter presenteres de ulike GARCH-modellene som benyttes i estimeringene, nemlig GARCH, TGARCH og EGARCH. Det blir også presentert noen viktige begreper i finans som er relevante for oppgaven, som stasjonære prosesser og autokorrelasjon. Til slutt går det igjennom ulike kriterier som brukes til valg av modell, deriblant informasjonskriterier og root mean square error (RMSE).

2.1 Avkastning

I denne oppgaven beregnes nominelle logaritmiske avkastninger for hver indeks, altså OSEBX og de åtte andre bransjeindeksene. Vi har regnet ut både daglig og månedlig avkastning. Den daglige avkastningen er regnet ut for alle bransjeindeksene, og brukes til å beregne korrelasjonen mellom hver enkelt bransjeindeks og OSEBX. Den månedlige avkastningen er regnet ut for bare OSEBX, og brukes til å predikere avkastningen til OSEBX. Avkastningene som regnes ut er basert på totalt 5741 observasjoner av de daglige sluttprisene til indeksene. Vi har valgt å ta i bruk logaritmisk avkastning, fordi en logaritmisk skala kan vise informasjon som er vanskelig å oppdage ved bruk av lineær avkastning, og dermed kan belyse større nyanser i avkastningen. Avkastning blir generelt og i denne oppgaven regnet ut ved bruk av følgende formel:

$$R_t = \frac{\ln P_t}{\ln P_{t-1}} \quad (2.1.1)$$

R_t representerer avkastningen til indeksen i periode t . P_t og P_{t-1} representerer henholdsvis sluttprisen til indeksene i periode t og foregående periode ($t - 1$). Multipliseres denne formelen med 100 fås avkastningen i prosent, ofte kalt 'rate of return'.

2.2 Korrelasjon

Korrelasjonskoeffisienten Pearson's r ble introdusert av Karl Pearson i 1896, og er et mål på det lineære forholdet mellom to variabler. Pearson's r er en av de mest brukte korrelasjonskoeffisientene (Hauke og Kossowski, 2011), som er årsaken til at vi har valgt å bruke denne i oppgaven. Det er viktig å påpeke at korrelasjon ikke impliserer kausalitet, men kun at to hendelser tenderer til å forekomme sammen. I denne oppgaven forteller korrelasjon

noe om forholdet mellom avkastningene til OSEBX og de ulike bransjeindeksene. Som nevnt tidligere blir daglige avkastninger beregnet for å kunne beregne månedlige korrelasjoner mellom OSEBX og hver enkelt bransjeindeks. Korrelasjonskoeffisientene brukt i denne oppgaven beregnes ved bruk av Excel, på følgende måte:

$$\text{Pearsons } r = \text{Correl}(x, y) = \rho_{XY} = \frac{n \sum(xy) - (\sum x)(\sum y)}{\sqrt{[n \sum(x^2) - (\sum x)^2][n \sum(y^2) - (\sum y)^2]}} \quad (2.2.2)$$

I forbindelse med utregning av månedlig korrelasjon representerer x og y daglige avkastninger for henholdsvis OSEBX og hver enkelt bransjeindeks. Korrelasjonskoeffisienten (Pearsons r) viser et tall mellom -1.0 og 1.0 , der -1.0 er perfekt negativ korrelasjon og 1.0 perfekt positiv korrelasjon. En verdi på 0 betegner ingen korrelasjon. Fortegnet forteller altså hvilken retning de to indeksene beveger seg i forhold til hverandre. Ved positivt fortegn ses tendenser til at bransjeindeksen beveger seg i samme retning som OSEBX, mens ved negativt fortegn ses tendenser til at de beveger seg fra hverandre. Størrelsen på koeffisienten uttrykker styrken til korrelasjonen. En koeffisient nært 0 betegner svak korrelasjon, mens en koeffisient nært ± 1 indikerer sterk korrelasjon. Det finnes forskjellige retningslinjer på hva som betegnes som sterk og svak korrelasjon. Verdier mellom 0 og 0.3 indikerer svakt lineært forhold, mellom 0.3 og 0.7 indikerer moderat positivt lineært forhold, og mellom 0.7 og 1.0 indikerer sterkt positivt lineært forhold (Ratner, 2013).

2.3 GARCH

Autoregressive Conditional Heteroscedastic (ARCH)-modeller ble først introdusert av Engle (1982) og brukes til å estimere avkastningen og volatiliteten til finansielle instrumenter. Det spesielle ved ARCH-modeller er at de tillater volatiliteten å variere over tid og kan knyttes opp mot begrepet heteroskedastisitet. Heteroskedastisitet innebærer at volatiliteten til restleddene kan variere over tid (Engle, 1982), noe som passer bra ved bruk av finansielle data (Engle, 2001), som vist i underkapittel 1.1.

ARCH-modeller er forløperen til generaliserte ARCH-modeller (GARCH) som ble introdusert av Bollerslev (1986), og som vi skal ta utgangspunkt i ved denne oppgaven. GARCH ble utviklet som følge av at man måtte oppnå en høy ARCH-orden for å få fanget opp effekten av betinget varians. ARCH-orden er knyttet opp mot hvor mange restledd fra tidligere perioder som inkluderes i modellen, ofte kalt laggede restledd. Å ha krav om høy

ARCH-orden innebærer at man må inkludere flere variabler i modellen, som har en tendens til å gjøre modellene rotete og uoversiktlige. Dette problemet løser GARCH-modeller, da GARCH-modeller inneholder laggede variansledd i variansligningen som tar høyde for tidligere restledd. GARCH-modeller utvider altså ARCH-modeller med lagget varians, som gjør at man ikke trenger å inkludere mange restledd som i en tradisjonell ARCH-modell. På den måten får GARCH-modeller fram effekten av både laggede restledd og lagget varians.

GARCH-modeller fokuserer på prediksjon av volatilitet, da volatilitet er lettere å predikere enn avkastning, men kan også brukes til å predikere avkastning. Det er prediksjon av avkastning som er fokuset i denne oppgaven. GARCH(1,1)-modellen er variansligningen som vises i (2.3.3). ARCH-effekten i en GARCH(1,1)-modell er representert ved α_1 , som er koeffisienten til det laggede kvadrerte restleddet. GARCH-effekten vises ved β_1 , som er koeffisienten til den laggede variansen, som er et uttrykk for variansen i forrige periode.

I denne oppgaven brukes den enkleste og mest brukte utgaven av GARCH-modeller, GARCH(1,1), i modellestimeringene. En GARCH(1,1)-modell er uttrykt matematisk som:

$$\text{Hovedligning: } y_t = a_0 + a_1 y_{t-1} + u_t \quad (2.3.1)$$

$$\text{Støtteligning: } u_t = v_t \sqrt{h_t} \quad (2.3.2)$$

$$\text{Variansligning: } h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \beta_1 h_{t-1} \quad (2.3.3)$$

GARCH(1,1)-modellen består av tre ligninger: (2.3.1) kalles ofte hovedligningen og representerer avkastningen i periode t , (2.3.2) er en støtteligning som viser hvordan restleddet er satt sammen, og (2.3.3) er variansligningen. Variablene i hovedligningen (2.3.1) forteller følgende: y_t representerer avkastningen i periode t , a_0 er et konstantledd, a_1 er koeffisienten til lagget avkastning, i denne oppgaven er dette OSEBX sin laggede avkastning, og u_t er restleddet. Støtteligningen (2.3.2), representert ved u_t , er satt sammen av to komponenter: 1) v_t som er det ubetingede leddet og representerer hvit støy i modellen, og 2) h_t som er det betingede leddet og representerer variansen i periode t . Variansligningen (2.3.3) viser hvilke komponenter variansen til det finansielle instrumentet er satt sammen av. α_0 representerer det finansielle instrumentets vektete gjennomsnittlige varians over en

bestemt tid. α_1 er koeffisienten til det kvadrerte laggede restleddet, og β_1 er koeffisienten til den laggede variansen.

En forutsetning for å kunne bruke en tradisjonell GARCH(1,1)-modell er at variansen til det finansielle instrumentet må være positiv. Dette innebærer at koeffisientene i variansligningen (2.3.3), α_0 , α_1 , og β_1 må være positive. En annen forutsetning for at GARCH(1,1)-modellen skal være gyldig er at det finansielle instrumentet er stasjonært. I en tradisjonell GARCH(1-1)-modell innebærer dette at $\alpha_1 + \beta_1 < 1$. Stasjonære prosesser beskrives nærmere i delkapittel 2.7.

Den største fordelen med GARCH-modeller er at de tar høyde for volatilitetsklumping, som nevnt innledningsvis. Det er derimot også noen ulemper med de tradisjonelle GARCH-modellene. Tradisjonelle GARCH-modeller kan ha problemer med å fange opp ekstreme observasjoner i data, noe som kan gjøre modellen upresis (Ross, 2013). Tradisjonelle GARCH-modeller tar heller ikke høyde for asymmetri i en modell (Nelson, 1991), som nevnt i delkapittel 1.1. Dette er et problem, da modeller som baseres på virkelige finansielle data ofte inneholder asymmetri. Det finnes heldigvis utvidete GARCH-modeller som kan ta høyde for denne asymmetrien. Eksempler på to slike asymmetriske GARCH-modeller, er Threshold GARCH (TGARCH) og Exponential GARCH (EGARCH), som vi vil presentere i detalj i de neste delkapitlene.

2.4 TGARCH

Glosten, Jagannathan og Runkle (1993) utviklet en GARCH-modell som tar høyde for asymmetri i modellen, kalt TGARCH. Modellen kalles også ofte GJR-GARCH, oppkalt etter forfatterens etternavn. TGARCH(1,1)-modellen som brukes i denne oppgaven, tar utgangspunkt i en tradisjonell GARCH(1,1)-modell, men legger til et ekstra ledd i variansligningen (2.4.1). Dette ekstra leddet tar høyde for asymmetri, altså effekten av at dårlige nyheter kan ha større innvirkning på volatilitet enn gode nyheter. Hovedligningen som representerer avkastning forblir uforandret og vil altså være den samme som (2.3.1) i en TGARCH-modell. Variansligningen i en TGARCH(1,1)-modell er uttrykt matematisk som:

$$\text{Variansligning: } h_t = \alpha_0 + \alpha_1 u_{t-1}^2 + \lambda_1 u_{t-1}^2 d_{t-1} + \beta_1 h_{t-1} \quad (2.4.1)$$

λ_1 representerer den asymmetriske effekten. For at λ_1 skal være nødvendig å inkludere i variansligningen, må $\lambda_1 > 0$. Når λ_1 er større enn null er det et bevis på det er asymmetri i modellen. Det er altså bare nødvendig å bruke TGARCH ved positiv λ_1 . u_{t-1}^2 er som tidligere beskrevet et kvadrert lagget restledd. d_{t-1} er en såkalt dummyvariabel, som betyr at den bare kan innta to verdier. d_{t-1} vil være lik 1 hvis det eksisterer asymmetri i modellen, og 0 hvis det ikke eksisterer asymmetri. Dette innebærer at $d_{t-1} = 1$ hvis $u_{t-1} < 0$ og at $d_{t-1} = 0$ hvis $u_{t-1} \geq 0$. α_1 viser dermed effekten av positive nyheter på volatiliteten, mens $\alpha_1 + \lambda_1$ viser effekten av negative nyheter.

Det er flere forutsetninger som må være oppfylt for at TGARCH-modeller skal kunne brukes. To av de viktigste forutsetningene er at koeffisienten til lagget kvadrert restledd vist ved $\alpha_1 \geq 0$, og som nevnt tidligere, at den asymmetriske effekten er positiv, altså at $\lambda_1 > 0$. Når disse forutsetningene er oppfylt, følger det at $\alpha_1 + \lambda_1 \geq 0$. Dette betyr at variansen er positiv, som er en ytterligere forutsetning for å kunne bruke en TGARCH-modell.

2.5 EGARCH

Exponential-GARCH (EGARCH) var den første asymmetriske GARCH-modellen og ble introdusert av Nelson (1991). Variansligningen i EGARCH(1,1)-modellen står i logaritmisk form, i motsetning til variansligningene i GARCH- og TGARCH. Det at EGARCH-modellene står i logaritmisk form, medfører at variansen aldri kan være negativ. Videre innebærer dette at enkelte av koeffisientene i EGARCH-modeller kan være negative. Dette er en styrke i forhold til GARCH- og TGARCH-modeller, hvor alle koeffisientene må være positive. Oftest er det koeffisienten til kvadrert lagget restledd, representert ved α_1 som er negativ, og som gjør at man å ta i bruk en EGARCH-modell, da denne kan være negativ i EGARCH-modeller. Variansligningen i en EGARCH(1,1)-modell er uttrykt matematisk som:

$$\text{Variansligning: } \ln h_t = \alpha_0 + \alpha_1 \left[\frac{u_{t-1}}{\sqrt{h_{t-1}}} \right] + \lambda_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln h_{t-1} \quad (2.5.1)$$

α_1 representerer den symmetriske effekten til modellen. Dette er den samme rollen som ARCH-effekten har i tradisjonelle GARCH- og TGARCH-modeller, forklart i delkapittel 2.3. λ_1 representerer den asymmetriske effekten, ofte kalt leverage effect. Det er dette leddet som representerer utvidelsen fra GARCH til EGARCH. Hvis λ_1 er 0 eksisterer det ingen asymmetrisk effekt, og det vil da være bedre å bruke en tradisjonell GARCH-modell. Dersom

λ_1 er forskjellig fra null, vil det eksistere en asymmetri i modellen, og det er bedre å bruke en asymmetrisk modell, som TGARCH eller EGARCH. Ved bruk av en EGARCH-modell, viser α_1 og λ_1 effekten av gode nyheter, mens $-\alpha_1$ og λ_1 viser effekten av dårlige nyheter. β_1 representerer effekten av lagget varians, og har samme rolle som i en tradisjonell GARCH-modell, men står i logaritmisk form i en EGARCH-modell i motsetning til i tradisjonelle GARCH- og TGARCH-modeller. Hvis β_1 er stor vil det si at variansene fra tidligere perioder er store, og henger igjen. Et eksempel på effekten av dette kan ses i tiden etter en markedskrise, hvor den laggede variansen vil være stor. Da vil β_1 være høy, og det vil ta lang tid før effekten av denne vil forsvinne. Ved en lav β_1 er det motsatt.

Hvis $\alpha_1 < 0$ burde EGARCH-modeller alltid velges som estimeringsmetode. Restleddene til EGARCH-modeller kan følge flere fordelinger, der normalfordeling og generalized error distribution (GED) er de mest vanlige. En styrke ved EGARCH-modeller er at de tar høyde for asymmetri, som nevnt tidligere. EGARCH-modeller har imidlertid også noen svakheter, og en av disse er at det ved bruk av EGARCH-modeller kan være vanskelig å lage prediksjoner på volatilitet. Dette er en svakhet generelt, da GARCH-modeller ofte brukes til å beregne volatilitet. I denne oppgaven brukes EGARCH-modellen derimot hovedsakelig til å estimere avkastning, og denne svakheten er dermed mindre relevant for denne oppgaven.

2.6 Utvidet EGARCH(1,1)

Et av målene med denne oppgaven var å utvide en variant av GARCH med et korrelasjonsledd i både hovedligningen (2.3.1) og i variansligningen (2.3.3). Fordi det ble funnet asymmetri i de finansielle instrumentene brukt i oppgaven vår, ble det valgt å bruke denne utvidelsen på en EGARCH(1,1)-modell. Motivasjonen for å inkludere denne utvidelsen i modellen er å undersøke om de ekstra forklaringsvariablene kan bidra til å bedre prediksjonskraften. I resultatdelen inkluderes fire hovedtabeller som viser estimeringene til fire ulike utvidede EGARCH(1,1)-modeller. Det som skiller disse fire utvidede EGARCH(1,1)-modellene er om lagget avkastning inkluderes i hovedligningen eller ikke, og om lagget korrelasjon mellom henholdsvis OSEBX og hver enkelt bransjeindeks inkluderes i variansligningen eller ikke. I det følgende presenteres de fire forskjellige utvidede EGARCH(1,1)-modeller:

$$\text{Hovedligning: } y_t = a_0 + a_1 y_{t-1} + a_2 \rho_{t-1} + u_t \quad (2.6.1)$$

$$\text{Variansligning: } \ln h_t = \alpha_0 + \alpha_1 \left[\frac{u_{t-1}}{\sqrt{h_{t-1}}} \right] + \lambda_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln h_{t-1} + \gamma \rho_{t-1} \quad (2.6.2)$$

De ulike koeffisientene har samme betydning som den opprinnelige EGARCH(1,1)-modellen diskutert i delkapittel 2.5. Det eneste forskjellen er utvidelsen av både hovedligning og variansligning med et korrelasjonsledd, der a_2 representerer effekten av lagget korrelasjon mellom OSEBX og bransjeindeks på avkastning, mens γ representerer effekten av lagget korrelasjon på volatilitet, her målt ved variansen.

$$\text{Hovedligning: } y_t = a_0 + a_2 \rho_{t-1} + u_t \quad (2.6.3)$$

$$\text{Variansligning: } \ln h_t = \alpha_0 + \alpha_1 \left[\frac{u_{t-1}}{\sqrt{h_{t-1}}} \right] + \lambda_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln h_{t-1} + \gamma \rho_{t-1} \quad (2.6.4)$$

Ligning (2.6.3) og (2.6.4) er den andre utvidede GARCH-modellen, legg merke til at ikke lagget avkastning er fjernet i denne modellen. Utviklingen til OSEBX sin avkastning vil således bare forklares av korrelasjonen mellom bransjeindeks og OSEBX.

$$\text{Hovedligning: } y_t = a_0 + a_1 y_{t-1} + a_2 \rho_{t-1} + u_t \quad (2.6.5)$$

$$\text{Variansligning: } \ln h_t = \alpha_0 + \alpha_1 \left[\frac{u_{t-1}}{\sqrt{h_{t-1}}} \right] + \lambda_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln h_{t-1} \quad (2.6.6)$$

Ligning (2.6.6) inneholder ikke lagget korrelasjon. Korrelasjonskoeffisienten i variansligningen fjernes for å se om dette bedrer forklaringskraften til modellen.

$$\text{Hovedligning: } y_t = a_0 + a_2 \rho_{t-1} + u_t \quad (2.6.7)$$

$$\text{Variansligning: } \ln h_t = \alpha_0 + \alpha_1 \left[\frac{u_{t-1}}{\sqrt{h_{t-1}}} \right] + \lambda_1 \left| \frac{u_{t-1}}{\sqrt{h_{t-1}}} \right| + \beta_1 \ln h_{t-1} \quad (2.6.8)$$

Ligning (2.6.7) og (2.6.8) viser den siste utvidede EGARCH-modellen. Her er avkastningen, representert ved y_t bare avhengig av korrelasjonen mellom bransjeindeks og OSEBX, i tillegg til et konstantledd og et restledd. Variansligningen er lik en vanlig EGARCH(1,1).

2.7 Stasjonære prosesser

Om en prosess er stasjonær eller ikke har konsekvenser for den økonometriske analysen og om GARCH-modeller kan brukes eller ikke. Ikke-stasjonære prosesser er uforutsigbare, og kan derfor ikke brukes til å lage prognoser. Ikke-stasjonære prosesser kalles ofte random walk. En stasjonær prosess på sin side er definert som en prosess hvor sannsynlighetsfordelingen til dataene ikke endres over tid, og kan derfor brukes til å lage estimater og prognoser. Følgelig er det viktig at prosessen, i dette tilfellet avkastningen til OSEBX, er stasjonær for at den skal kunne estimeres og brukes til prediksjon av framtidig avkastning. Av stasjonære prosesser skilles det hovedsakelig mellom strenge stasjonære prosesser og svake stasjonære prosesser. Det er bare svake stasjonære prosesser som er relevant for denne oppgaven, og derfor er det bare disse som vil bli presentert. En prosess er svak stasjonær hvis forventningsverdien, variansen og autokovariansen er konstant over tid. Autokovarians refererer til kovariansen mellom en variabel på et gitt tidspunkt og verdien av denne variabelen på et tidligere tidspunkt, også kalt laggede variabler. I denne oppgaven er altså autokovarians kovariansen mellom variabelen avkastning, representert ved y_t , og dens laggede variabler, representert ved y_{t-1} .

For å finne ut om en prosess er stasjonær eller ikke, må det gjennomføres en test. En Dickey-Fuller test kan brukes til dette formålet. For å gjennomføre en Dickey-Fuller test tas det først utgangspunkt i en autoregressiv (AR) modell, som betyr at den endogene variabelen er avhengig av dens laggede variabler. Prosessen som testes skrives så om til endringsform. Deretter testes det om koeffisienten til den laggede variabelen er lik 0 eller mindre enn 0. Dersom den er lik null, er prosessen ikke-stasjonær, altså ikke predikerbar. Dersom koeffisienten til den laggede variabelen derimot er mindre enn 0, vil prosessen være stasjonær, altså predikerbar.

Det er dermed viktig å finne ut om prosessen av interesse i denne oppgaven, OSEBX sin avkastning, er stasjonær eller ikke-stasjonær. For å komme fram til om denne variabelen, y_t , er stasjonær eller ikke tas det utgangspunkt i følgende AR-modell med et lag (AR(1)):

$$y_t = a_0 + \beta y_{t-1} + \varepsilon_t \quad (2.7.1)$$

Hvor a_0 er et konstantledd, β er koeffisienten til lagget avkastning og ε_t representerer hvit støy, $\varepsilon_t \sim iid(0, \sigma^2)$. Hvis denne prosessen er ikke-stasjonær, vil β være lik 1, mens hvis prosessen er stasjonær vil β være mindre enn 1. (2.7.2) skrives derfor på endringsform for å kunne gjennomføre testen. y_t skrives på endringsform ved å trekke fra y_{t-1} på hver side av likhetstegnet:

$$y_t - y_{t-1} = a_0 + \beta y_{t-1} - y_{t-1} + \varepsilon_t$$

$$\Delta y_t = a_0 + (\beta - 1)y_{t-1} + \varepsilon_t$$

$$\Delta y_t = a_0 + \mu y_{t-1} + \varepsilon_t \quad (2.7.2)$$

Hypotesene blir da:

$$H_0: \mu = 0 \text{ (ikke - stasjonær prosess)}$$

$$H_1: \mu < 0 \text{ (stasjonær prosess)}$$

Bruker en ensidig t-test for å finne ut om nullhypotesen kan forkastes:

$$T = \frac{\hat{\mu} - \mu}{se(\hat{\mu})} \sim T_{kritisk\ verdi} \quad (2.7.3)$$

Der $\hat{\mu}$ er koeffisienten til lagget avkastning, μ er hypoteseverdien som testet og $se(\hat{\mu})$ er standardavviket til koeffisienten til lagget avkastning. Kritiske verdier ved Dickey-Fuller test varierer fra vanlige t-verdiene og rapporteres automatisk i OxMetrics. Hvis $\mu < 0$ og vi kan forkaste nullhypotesen, vil prosessen konvergere mot å være stasjonær (Dickey og Fuller, 1979). Tester for hvorvidt prosessene som brukes i denne oppgaven er stasjonære eller ikke finnes i delkapittel 4.1.

2.8 Autokorrelasjon

Siden volatiliteten i GARCH(1,1)-modellen avhenger av forrige periodes volatilitet fører dette til at det er en mulighet for at volatiliteten i en periode t og den laggede volatiliteten blir tilnærmet like. Graden av likhet mellom volatiliteten og den laggede volatiliteten uttrykkes

gjennom begrepet autokorrelasjon. Autokorrelasjon uttrykker altså i hvilken grad en variabel er korrelert med laggede verdier av variabelen. Grunnen til at man tester for autokorrelasjon i en GARCH-modell er altså for å finne ut hvor mye støy som henger igjen fra forrige periode, representert ved restleddet.

Man kan også regne ut hvor lang tid autokorrelasjonen mellom volatilitet i en periode og den laggede volatiliteten bruker på å avta og gå mot null. Dettes bestemmes av størrelsen på $\alpha_1 + \beta_1$, altså summen av laggede kvadrerte restledd og lagget varians fra variansligningen. Dette innebærer altså at jo nærmere $\alpha_1 + \beta_1$ er 1, jo lengre tid vil det ta før autokorrelasjonen går mot null. $\alpha_1 + \beta_1$ vil ofte være nær 1 ved bruk av GARCH(1,1) (Bollerslev, Chou og Kroner, 1992).

Dersom det inntreffer en plutselig endring i volatilitet, vil det ta lang tid før volatiliteten får stabilisert seg, dersom volatiliteten og den laggede volatiliteten har høy grad av autokorrelasjon. At effekten av en plutselig endring i volatilitet forsvinner sakte, altså at det kan gå flere perioder før volatiliteten får stabilisert seg, kan da føre til at man kan få problemer pga. autokorrelasjon, ved at for mye støy fra restleddet fra forrige periode henger igjen i modellen.

Autokorrelasjonen er definert ved følgende matematiske uttrykk:

$$\text{Corr}(u_t, u_{t-i}) = \frac{\text{Cov}(u_t, u_{t-i})}{\sqrt{\text{Var}(u_t)\text{Var}(u_{t-i})}} = \frac{\text{Cov}(v_t\sqrt{h_t}, v_{t-1}\sqrt{h_{t-1}})}{\sqrt{\text{Var}(u_t)\text{Var}(u_{t-i})}} \quad (2.8.1)$$

Hvor u_t er restleddet i periode t , v_t er hvit støy i periode t og h_t er variansledd i periode t , som representerer volatiliteten. u_{t-1} , v_{t-1} og h_{t-1} er henholdsvis disse variablene i forrige periode. Av formelen for autokorrelasjon ser vi at restleddene er en funksjon av hvit støy og volatilitet. Dette innebærer at når man tester om det er en sammenheng mellom restleddet i periode t og lagget restledd, tester man også for om det er en sammenheng mellom volatiliteten i periode t og lagget volatilitet.

Det er vanlig å få problemer med autokorrelasjon ved bruk av finansielle data. Derfor er det viktig å teste for autokorrelasjon og evt. ta høyde for det i modellen. Box og Pierce (1970)

introduserte en test for å vurdere autokorrelasjon mellom restledd og laggede restledd. I denne oppgaven tas det utgangspunkt i en variant av denne testen kalt Ljung-Box (Ljung og Box, 1978). Ljung-Box testen gjennomføres på de standardiserte restleddene, u_t/σ_t^2 , og er gitt ved:

$$Q_k = T(T + 2) \sum_{i=1}^k \frac{\tau_k^2}{T - k} \sim \tau_k^2 \quad (2.8.2)$$

Der T representerer utvalgsstørrelsen, k er antall lag og τ_k er autokorrelasjonen på lag k. Vi bruker denne testen for å undersøke om det er autokorrelasjon tilstede i modellene våre, og vi bruker følgende hypotese:

H₀: Ingen autokorrelasjon i restleddene

H₁: Autokorrelasjon i restleddene

Resultatene av Ljung-Box testes vises i resultatdelen av oppgaven, i delkapittel 4.5. Dersom autokorrelasjon eksisterer mellom restleddene må man bruke en modell som tar høyde for denne korrelasjonen. I følge Hull og Basu (2016) skal en GARCH-modell som fanger fordelingen godt kunne ta høyde for effekter av autokorrelasjon. Hvor god en GARCH-modell er til dette bestemmes av ordenen til GARCH-modellen. Ordenen til en GARCH-modell bestemmes av hvor mange laggede kvadrerte restledd, og laggede volatilitetsledd som inkluderes.

Autokorrelasjonsfunksjonen (ACF) og den partielle autokorrelasjonsfunksjonen (PACF) kan også brukes til å si noe om korrelasjonen mellom restleddene og de laggede restleddene i en modell. Korrelasjonen mellom restleddene innebærer at det eksisterer en systematisk sammenheng mellom restleddene, slik at restleddet i periode t inneholder informasjon om restleddet i foregående periode, $(t - 1)$.

Autoregressive-moving-average (ARMA)-modeller kan brukes til å finne ut ordenen i en GARCH-modell (Enders, 2015, s. 129). Følgelig presenteres en ARMA(1,1)-prosess, som det vil tas utgangspunkt i ved valg av GARCH-orden i delkapittel 4.3.

$$y_t = a_0 + a_1 y_{t-1} + \beta_1 \varepsilon_{t-1} \quad (2.8.3)$$

a_1 representerer det autoregressive leddet, som er koeffisienten til lagget avkastning, mens β_1 viser til moving-average leddet, som innebærer at avkastningen avhenger av hvit støy fra foregående periode. I delkapittel 4.3 i resultatet vil autokorrelasjonsfunksjonen (ACF) og den partielle autokorrelasjonsfunksjonen (PACF) til de standardiserte restleddene i en ARMA(1,1)-prosess illustreres.

I denne oppgaven brukes GARCH-modeller av ordenen (1,1). Dersom autokorrelasjonen er veldig høy, må man likevel vurdere å bruke en GARCH-modell av annen orden. Av testene for autokorrelasjon i delkapittel 4.5 virker det ikke å være mye autokorrelasjon i modellen vår, da vi ikke kan forkaste nullhypotesen om autokorrelasjon i restleddene. GARCH-modellene våre av ordenen (1,1) kan dermed brukes videre til predikering av avkastning og volatilitet.

2.9 ARCH-effekter

For at det skal være nødvendig å bruke en GARCH-modell, må man først sjekke om restleddene er heteroskedastiske. For å sjekke om restleddene er heteroskedastiske anvendes en ARCH-test. Dette gjennomføres ved å teste om variansen til restleddene inneholder ARCH-effekter ved bruk av Engle sin ARCH-test (Engle, 1982). Hvis restleddene inneholder ARCH-effekter betyr det at variansen til restleddene ikke er konstant, som er definisjonen på heteroskedastisitet. ARCH-testen gjennomføres ved å kjøre en lineær regresjon på avkastningen, med laggede avkastninger som forklaringsvariabel. Tester for ingen ARCH-ledd med hypotesene:

$$H_0: \text{Ingen ARCH - effekt}$$

$$H_1: \text{ARCH - effekt}$$

Hvis nullhypotesen om ingen ARCH-effekter i de standardiserte restleddene kan forkastes, bør en ikke-lineær modell bør benyttes i estimeringene. Dette innebærer da at man må bruke en modell som inkluderer ARCH-ledd. En ARCH- eller GARCH-modell inneholder et ARCH-ledd og passer derfor bra dersom nullhypotesen forkastes.

2.10 Restleddets fordeling

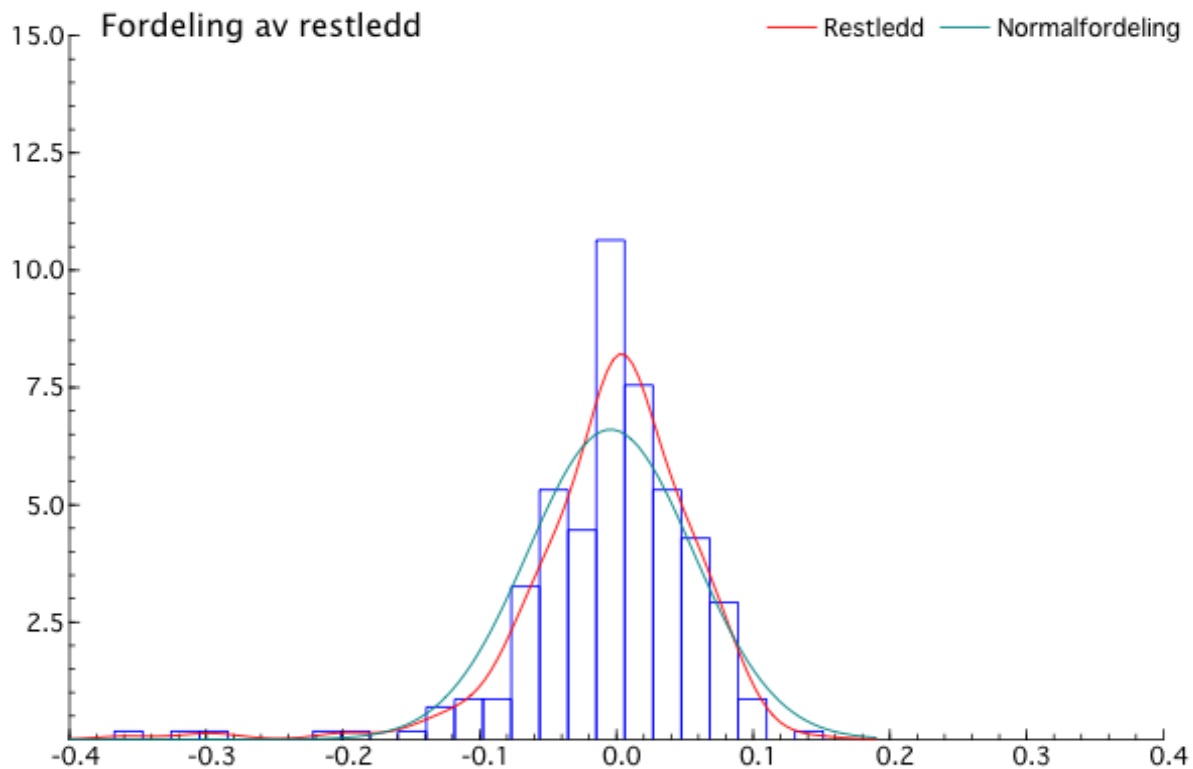
Ved estimering av GARCH-modeller i OxMetrics brukes estimeringsmetoden log-likelihood. Denne estimeringsmetoden innebærer å finne de parameterverdiene som er mest sannsynlige for datasettet. En forutsetning som må være oppfylt ved bruk av denne estimeringsmetoden er at de standardiserte restleddene er normalfordelte og uavhengige over tid. Dette kan vi teste ved å teste for normalitet, som gir oss følgende hypoteser:

$$H_0: \text{Normalfordelte restledd}$$

$$H_1: \text{Ikke – normalfordelte restledd}$$

Av testene for normalfordeling, kommer vi fram til at vi kan forkaste nullhypotesen om normalfordelte restledd ved 1% signifikansnivå. Dette tyder på at fordelingen vår ikke er normalfordelt og vi må derfor benytte oss av en annen fordeling for å fortsatt kunne bruke log-likelihood som estimeringsmetode. En student t-fordeling passer godt til dette formålet, ved estimering av GARCH(1,1) og TGARCH(1,1). Ved estimering av EGARCH(1,1)-modellene vil likevel normalfordelingen bli brukt av praktiske hensyn.

Grunnen til at spesielt student t-fordelingen passer bedre til vår data, skyldes de tykkere halene. Dette er illustrert i et søylediagram i figur 1, som tar for seg en grafisk analyse av de standardiserte restleddene for å illustrere fordelingen. Søylediagrammet viser fordelingen av restleddene, og hvordan denne fordelingen ser ut i forhold til en normalfordeling. Man ser i figuren at de standardiserte restleddene fra våre data har litt tykkere haler enn normalfordelingen, noe som tyder på at student t-fordelingen er mer passende å bruke.



Figur 1: Figuren viser fordelingen til restleddene i modellen. Den røde linjen viser hvordan restleddene i modellen fordeler seg, mens den grønne linjen viser normalfordelingen. Av figuren kan man se at denne fordelingen har noe tykkere haler enn normalfordelingen. Dette tyder på at en student t-fordeling vil være mer passende for dataene i denne oppgaven.

2.11 Informasjonskriterier

Informasjonskriterier brukes til å evaluere hvor god en modell er i forhold til andre modeller. Informasjonskriterier sier altså ikke noe om hvor bra en modell er i seg selv, men hvor god modellen er relativt til andre modeller. Verdien til informasjonskriteriene går mot minus uendelig, der det laveste tallet indikerer den beste modellen. Ved bruk av informasjonskriterier til å bestemme hvilken modell som er best må alle modellene være innenfor samme utvalgsperiode. I tillegg må det også være like antall observasjoner i modellene som sammenlignes.

De tre mest brukte informasjonskriteriene er Akaike's information criterion (AIC), Schwarz's Bayesian information criterion (SBC) og Hannan-Quinn criterion (HQC) (Hannan og Quinn, 1979). De har forskjellige styrker og svakheter, men ingen av dem er generelt sett overlegne de andre (Brooks, 2008, s. 233). Informasjonskriteriene fungerer ved teste hvilken modell som er best, ved å straffe modeller som inkluderer flere variabler. Her er SBC strengest på

denne straffen, AIC er «snillest» og HQC utgjør en middelvei mellom de to (Brooks, 2008, s. 233). Alle disse informasjonskriteriene vil tendere til å foretrekke modeller med færrest variabler, men SBC vil nesten alltid peke ut modellen med færrest variabler som den beste. AIC var det første informasjonskriteriet som ble innført (Akaike, 1974). Ved bruk av AIC til å velge modell kan valget fortsatt falle på modeller med flest variabler, da det er kriteriene som straffer i minst grad på bakgrunn av antall variabler. Fordi AIC kunne velge modeller med flest variabler, ble SBC derfor introdusert av Schwarz (1978) som et strengere alternativ til AIC. SBC er konsekvent, men mindre effektiv, mens AIC er lite konsekvent, men mer effektiv (Brooks, 2008, s. 233). AIC kan fungere bedre enn SBC når man håndterer mindre utvalg, men da vi i denne oppgaven ser på en relativ stor utvalgsperiode, 22 år, er AIC mindre relevant. Vi trenger derfor ikke i like stor grad å ta høyde for dette. Det er vanlig å bruke alle tre informasjonskriteriene når man vurderer hvilken modell som er best, og det er også gjort i denne oppgaven. De ulike informasjonskriteriene er uttrykt matematisk som:

$$AIC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \quad (2.11.1)$$

$$SBC = \ln(\hat{\sigma}^2) + \frac{k}{T} \ln T \quad (2.11.2)$$

$$HQC = \ln(\hat{\sigma}^2) + \frac{2k}{T} \ln(\ln T) \quad (2.11.3)$$

Hvor $\hat{\sigma}^2$ representerer variansen til restleddene, $k = p + q + 1$ er antall variabler som estimeres i en valgt modell og T er utvalgsstørrelsen.

Utrykkene for informasjonskriteriene inneholder to ledd, hvor det første leddet er en funksjon av de kvadrerte restleddene, som også ofte omtales som residual sum of squares (RSS). Det andre leddet representerer straffen ved å inkludere flere variabler i modellen. Å legge til flere variabler i modellen vil derfor ha to motvirkende effekter, der det første leddet reduserer restleddsvariansen, mens det andre leddet øker som følge av tapet av frihetsgrader ved inkludering av ekstra variabler. Formlene viser tydelig hvordan SBC straffer mer enn HQC og AIC og HQC. Resultater av informasjonskriteriene når de er brukt til å vurdere modellene i denne oppgaven, kan finnes i resultatdelen, i delkapittel 4.4.

2.12 Sammenligning av prediksjonsnøyaktighet

For å finne ut hvilken modell som gir mest presise prediksjoner, dvs. at prediksjonsverdiene til avkastning sammenfaller med virkelige avkastninger, brukes forskjellige mål på prediksjonsnøyaktighet. I denne oppgaven brukes det to slike mål: root mean squared error (RMSE) og mean absolute error (MAE). Begge disse målene sammenligner forventet avkastning med virkelig avkastning, og måler hvor presise forskjellige prediksjoner er i forhold til hverandre (Akay, 2012). Informasjonskriterier kan også brukes til å sammenligne prediksjonsnøyaktighet til modellene, og vil derfor også inkluderes i analysen av prediksjonsnøyaktighet.

2.12.1 Root mean squared error (RMSE)

Root mean squared error (RMSE) er et mål på prediksjonsnøyaktighet, og kan derfor brukes til å finne hvilken modell som lager de beste prediksjonene. Det er ønskelig at RMSE er så nært null som mulig. Jo lavere verdi RMSE gir, jo bedre er modellen, dvs. jo bedre er prediksjonene modellen gjør. RMSE er det mest brukte målet i prediksjonslitteratur (Akay, 2012), som er en av grunnene til at vi har valgt å bruke den til å vurdere prediksjonsnøyaktigheten til modellene i oppgaven. RMSE er uttrykt matematisk som:

$$RMSE_{e_{t+h,t}} = \sqrt{\frac{1}{T} \sum_{t=0}^T e_{t+h,t}^2} \quad (2.12.1.1)$$

Hvor t står for tid, T er antall observasjoner og $e_{t+h,t}$ er prediksjonsfeil.

2.13 Mean absolute error (MAE)

MAE er mindre vanlig enn RMSE, men MAE straffer i mindre grad modellenes prediksjonsverdi på bakgrunn av store feilledd enn RMSE (Akay, 2012). Av denne grunnen, er MAE også inkludert som mål på prediksjonsnøyaktigheten til modellene i oppgaven. MAE er uttrykt matematisk som:

$$MAE_{e_{t+h,t}} = \frac{1}{T} \sum_{t=0}^T e_{t+h,t} \quad (2.12.2.1)$$

Hvor t er tid, T er antall observasjoner, og $e_{t+h,t}$ er prediksjonsfeil.

3 Data

Dette kapitlet tar for seg presentasjon og analyse av dataene som brukes til å estimere GARCH-modellene i denne oppgaven. Først beskrives OSEBX og de åtte bransjeindeksene som vil brukes i modellestimeringene. Her analyseres utviklingen til hver enkelt bransjeindeks i perioden 1996-2017. Deretter presenteres en analyse av dataene, som blant annet inneholder OSEBX sin historiske månedlige avkastning og deskriptiv statistikk. Til slutt presenteres og diskuteres korrelasjonen mellom OSEBX og hver enkelt bransjeindeks.

3.1 Beskrivelse av data

Dataene brukt i denne oppgaven er hentet fra Thomson Reuters Datastream. Daglige sluttpriser for hver bransjeindeks og hovedindeksen på Oslo Børs, OSEBX, er hentet ut fra perioden 29.12.1996-31.12.2017. Begrunnelsen for valg av tidsperiode er at indeksene bare er tilbakeregnet til og med 1996. Ved arbeid med finansielle instrumenter er det vanligst å se på prisendringer, altså avkastningen. Dette skyldes blant annet at avkastninger er enklere å håndtere på grunn av egenskapene de besitter. I denne oppgaven modelleres derfor avkastning. Det finnes flere definisjoner av avkastning som kan brukes, men i denne oppgaven brukes logaritmisk avkastning. Dataprogrammene som er brukt til behandling av data Excel og OxMetrics.

3.1.1 OSEBX

OSEBX er hovedindeksen på Oslo Børs og inneholder et representativt utvalg av aksjene som er tilgjengelige på Oslo Børs (*Hovedindeksen*, 2017). Indeksen ble opprettet i 2001, men er i senere tid blitt tilbakeregnet til 1996. Den historiske utviklingen til OSEBX er illustrert i figur 2. OSEBX er utbyttejustert og friflytjustert, som innebærer at aksjer som ikke er tilgjengelige i markedet fjernes (*Hovedindeksen*, 2017). Dette skjer når indeksen revideres, som er to ganger i året. Indeksens sammensetning avhenger av hvilke selskaper som er børsnotert på revideringstidspunktene og verdien av disse selskapene.

I internasjonal sammenheng er Oslo børs en liten børs. En konsekvens av dette er at store selskaper kan utgjøre store deler av OSEBX. Per 31.12.2017 består hovedindeksen av totalt 67 selskaper, der indeksen er vektet i forhold til aksjeverdien til hvert selskap. Følgelig vil en endring i aksjekursen til en av OSEBX sine største selskaper påvirke hovedindeksen i større grad enn et mindre selskap. Equinor er eksempelvis vektet med 28.63% av hovedindeksen og de fire største selskapene i OSEBX (Equinor, DNB, Telenor, Aker BP) utgjør over 50% av

indeksens verdi (se vedlegg A). Det faktum at enkeltelskaper utgjør så store deler av hovedindeksen kan gjøre utslag i korrelasjonsberegningene, da store selskaper kan utgjøre en stor andel av både sin bransjeindeks og hovedindeksen. Dette kan medføre sterkere korrelasjon mellom bransjeindeksen og hovedindeksen.



Figur 2: Grafen viser utviklingen til hovedindeksen på Oslo Børs i perioden 1996-2017. Kilde: Thomson Reuters Datastream. Det framgår at OSEBX har en overveiende stigende kurve, da den begynte på 100 den 29.12.1995, og har steget til 814,45 den 29.12.2017. Likevel er det verdt å merke seg perioder med negativ utvikling. En av disse er den lave indeksverdien i perioden rundt dotcombobla i 2001, der høyt verdsatte internettaksjer mistet mye av sin verdi. Den negative utviklingen i 2008 sammenfaller tidsmessig med finanskrisen, hvor investeringsbanken Lehman Brothers kollapset som følge av stort antall misligholdte boliglån.

3.1.2 Bransjeindekser

Ved Oslo Børs eksisterer det elleve ulike bransjeindekser. I denne oppgaven ekskluderes bransjeindeksene for eiendom (OSE60GI) (*Eiendom*, 2017), forsyning (OSE50GI) (*Forsyning*, 2017) og telekom (OSE55GI) (*Telekom*, 2017), da disse indeksene ga dårlige estimater og i noen tilfeller ikke noen estimater i det hele tatt ved modellering.

Bransjeindeksene for eiendom og forsyning har til felles at de begge utgjør en liten del av OSEBX, jf. vedlegg B vedrørende sektorsammensetningen til OSEBX. Bransjeindeksen til telekom utgjør en vesentlig større del av OSEBX, nemlig 10%. Procentsatsen på 10% i OSEBX utgjøres av et selskap, Telenor. Dette kan forklare at korrelasjonene mellom telekom og OSEBX fungerer dårlig som forklaringsvariabel i GARCH-modeller. Det er også andre sektorer som utgjør en liten del av hovedindeksen, eksempelvis helsevern, IT og forbruksvarer. Disse inkluderes likevel i analysen, da de ga opphav til brukbare modeller. I estimeringene av GARCH-modeller vil det tas utgangspunkt i de resterende åtte bransjeindeksene ved Oslo Børs. Utviklingen av de åtte bransjeindeksene som inkluderes i denne oppgaven i perioden 1996-2017 illustreres i figur 3.

- OSE10GI – Energi (*Energi*, 2017)

Bransjeindeksen energi inneholder selskaper som opererer innenfor olje, gass, industriell transport og supporttjenester. Per 31.12.2017 består bransjeindeksen av totalt 49 selskaper, blant andre Equinor, Aker BP og Subsea 7.

- OSE15GI – Materialer (*Materialer*, 2017)

Bransjeindeksen materialer inneholder selskaper som driver innenfor industrielle metaller, gruvedrift, kjemikalier og industrielle ingeniørfag. Ved utgangen av 2017 består bransjeindeksen av åtte selskaper. Blant de mest sentrale her er Norsk Hydro, Yara International og Elkem.

- OSE20GI – Industri (*Industri*, 2017)

Bransjeindeksen industri består av selskaper som driver innenfor industrielle ingeniørfag, reiseliv, fritid, industriell transport, konstruksjon, materialer, generell industri, supporttjenester, alternativ energi, eiendomsinvestering, eiendomstjenester og finansielle tjenester innenfor industrisektoren. Industrisektoren består per 31.12.2017 av totalt 35 selskaper. Eksempler på sentrale selskaper innenfor denne sektoren er Tomra Systems, Kongsberg Gruppen og Wallenius Wilhelmsen Logistics.

- OSE25GI – Forbruksvarer (*Forbruksvarer*, 2017)

Bransjeindeksen forbruksvarer inneholder medieselskaper og generelle forhandlere, i tillegg til selskaper innenfor husholdningsvarer, biler og bildeler. Bransjeindeksen består av totalt ni selskaper, blant andre Schibsted ser. B, Schibsted ser. A og XXL.

- OSE30GI – Konsumvarer (*Konsumvarer*, 2017)

Bransjeindeksen konsumvarer består i hovedsak av selskaper som produserer mat og drikke. Per 31.12.2017 utgjøres denne bransjeindeksen av 11 selskaper, deriblant Marine Harvest, Orkla og SalMar.

- OSE35GI – Helsevern (*Helsevern*, 2017)

Bransjeindeksen helsevern inneholder selskaper som arbeider med legemidler, bioteknologi, supporttjenester, utstyr til helse og tjenester innenfor helsevern.

Bransjeindeksen består av ni selskaper, eksempelvis BerGenBio, Nordic Nanovectora og Medistim.

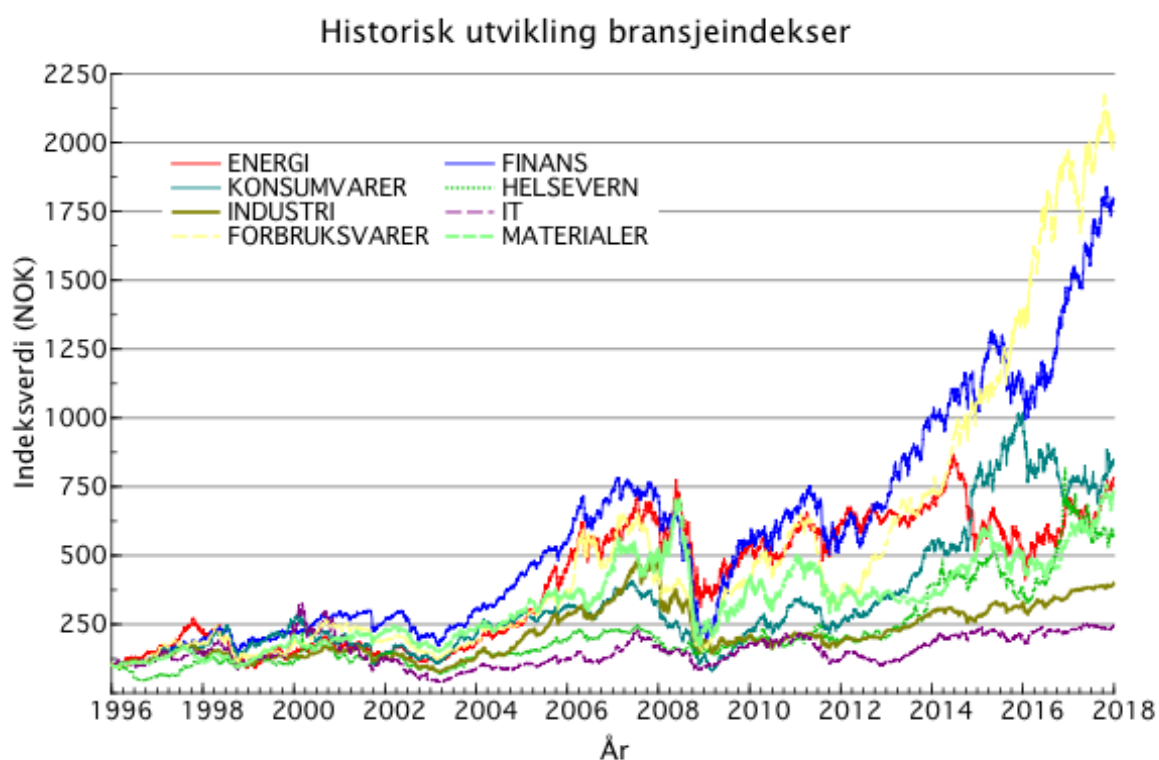
- OSE40GI – Finans (*Finans*, 2017)

Bransjeindeksen finans består av selskaper som driver innenfor bankvirksomhet, forsikring (ikke livsforsikring), og finansielle tjenester innenfor finanssektoren.

Bransjeindeksen består av 16 selskaper, blant andre DNB, Gjensidige Forsikring, og Aker.

- OSE45GI – Informasjonsteknologi (IT) (*IT*, 2017)

Bransjeindeksen informasjonsteknologi inneholder selskaper som arbeider med programvare, datatjenester, maskinvare, utstyr for teknologi, elektronisk utstyr, elektrisk utstyr, mobil telekommunikasjon og alternativ energi. Bransjeindeksen består av totalt 25 selskaper, hvor de mest sentrale er Atea, EVRY og Nordic Semiconductor.



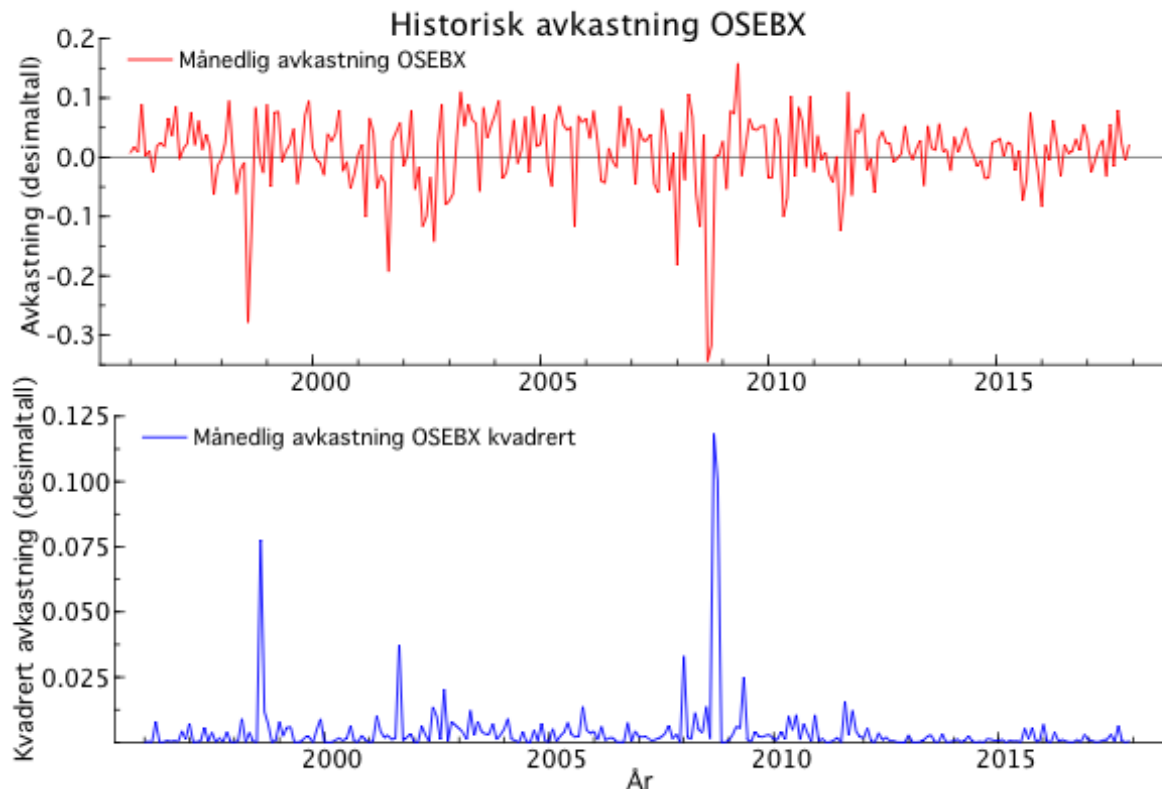
Figur 3: Grafen viser utviklingen til de åtte bransjeindeksene på Oslo Børs som inkluderes i analysen, for perioden 1996-2017. Kilde: Thomson Reuters Datastream. I likhet med OSEBX starter også bransjeindeksene på 100 poeng per 29.12.1995. Av figuren ser man at bransjeindeksene for forbruksvarer og finans har hatt den høyeste stigningen i indeksverdi de siste årene. Det er også verdt å legge merke til bransjeindeksene for industri og IT, som har lav volatilitet, som gjenspeiles i små endringer av indeksverdi.

3.2 Analyse av data

3.2.1 Historisk avkastning OSEBX

Figur 4 illustrerer den månedlige avkastningen til OSEBX. Det er denne avkastningen vi ønsker å predikere gjennom modellene som estimeres i denne oppgaven. Figuren viser tydelige tegn til volatilitetsklumping (Mandelbrot, 1963), som innebærer at perioder med lav volatilitet etterfølges av perioder med lav volatilitet, mens perioder med høy volatilitet etterfølges av perioder med høy volatilitet. Volatilitet sier noe om usikkerhet og risiko knyttet til et finansielt instrument, og måles gjennom standardavviket til avkastningen. Volatilitet er et mål på variasjon, endring, spredning og risiko, og uttrykkes ofte matematisk som en funksjon av variansen. Tilstedeværelse av volatilitetsklumping i finansielle data tyder på at det eksisterer ARCH-effekter i dataene, noe som kan ses grafisk ved at volatiliteten kommer i perioder, og illustreres i figur 4. Dette er en indikasjon på at restleddet er heteroskedastisk,

som innebærer at variansen til restleddet varierer over tid (Engle, 1982). ARCH/GARCH modeller beskriver volatilitetsklumping bedre enn andre ikke-lineære modeller, og ser derfor ut til å passe bra til våre data. Volatilitetsklumping ses tydeligere ved høyere oppløsning i dataene. Effekten av volatilitetsklumping er altså mer synlig ved bruk av data som er samlet inn med kortere intervall. Eksempelvis vil data som er samlet inn daglig eller ukentlig tendere til å ha mer volatilitetsklumping enn data som er samlet inn månedlig eller årlig.



Figur 4: Den røde grafen viser den månedlige logaritmiske avkastningen til hovedindeksen på Oslo Børs i perioden 1996-2017. Mens den blå grafen viser månedlig avkastning til OSEBX kvadrert. Verdiene vises som desimaltall. Kilde: Thomson Reuters Datastream. Man kan tydelig se flere ekstremverdier i avkastningene, men særlig påfallende er to ekstremverdier ved henholdsvis 1998-1999 og 2008-2009, altså samtidig som finanskrisen.

Flere faktorer kan altså påvirke utviklingen til OSEBX. Figur 4 viser den tydelige effekten økonomiske hendelser i USA kan ha på OSEBX. Det er tidligere skrevet oppgaver som viser at S&P500, en indeks som er representativ for aksjemarkedet i USA, har en klar påvirkning på utviklingen til OSEBX. Vår masteroppgave belyser om norske bransjeindekser kan ha en tilsvarende effekt på OSEBX, gjennom bruk av korrelasjonen mellom hver enkelt sektors avkastning og OSEBX sin avkastning.

3.2.2 Deskriptiv statistikk

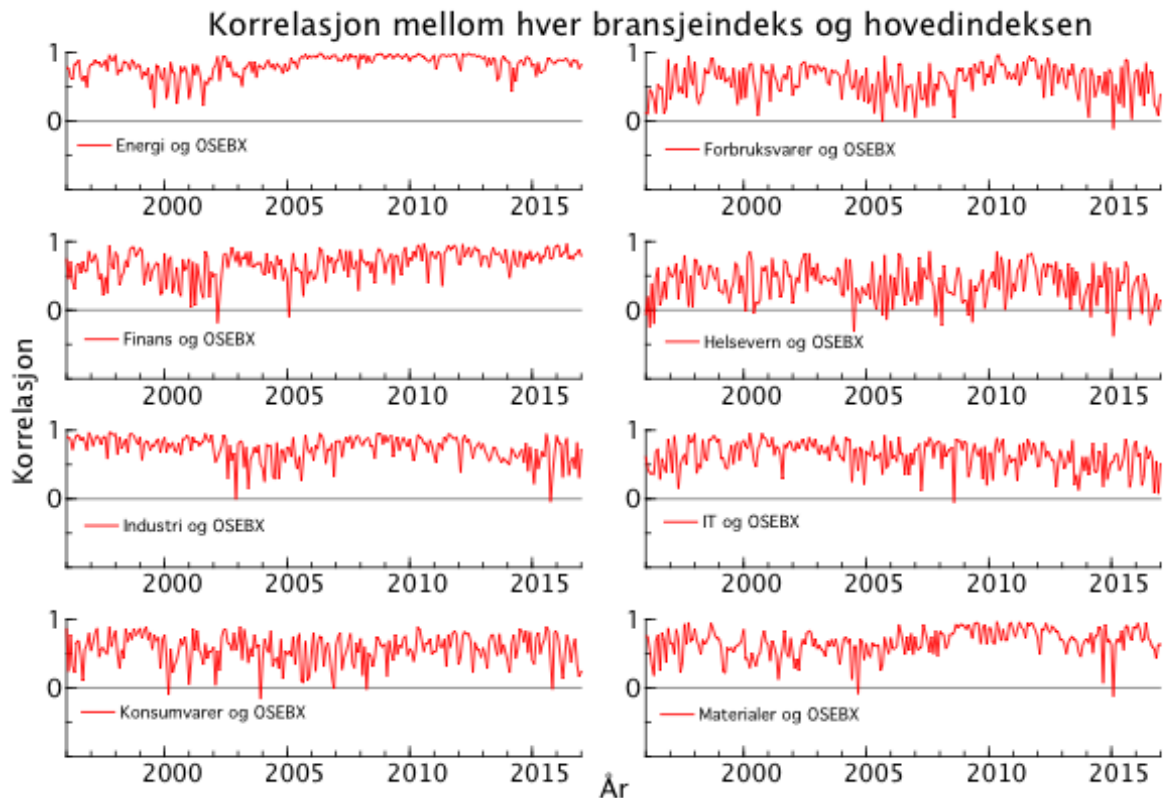
Tabell 1: Deskriptiv statistikk for bransjeindekser

| | Gj.snitt | St.avvik | Skjevhet | Kurtose | Min. | Maks. |
|------------------------------|----------|----------|----------|---------|---------|---------|
| R_{OSEBX} | 0.0079 | 0.0629 | -1.7926 | 6.9384 | -0.3439 | 0.15 |
| $\rho_{OSEBX,energi}$ | 0.8235 | 0.1430 | -1.5993 | 3.2306 | 0.1919 | 0.9904 |
| $\rho_{OSEBX,forbruksvarer}$ | 0.5811 | 0.2238 | -0.5587 | -0.3437 | -0.1107 | -0.9608 |
| $\rho_{OSEBX,finans}$ | 0.6815 | 0.2018 | -1.2627 | 1.8196 | -0.1856 | 0.9655 |
| $\rho_{OSEBX,helsevern}$ | 0.3748 | 0.2588 | -0.3089 | -0.4385 | -0.3818 | 0.8619 |
| $\rho_{OSEBX,industri}$ | 0.7303 | 0.1839 | -1.3515 | 1.9825 | -0.0434 | 0.9709 |
| $\rho_{OSEBX,IT}$ | 0.6278 | 0.2033 | -0.7166 | -0.0229 | -0.0502 | 0.9555 |
| $\rho_{OSEBX,konsumvarer}$ | 0.5595 | 0.2175 | -0.7559 | 0.1896 | -0.1649 | 0.8892 |
| $\rho_{OSEBX,materialer}$ | 0.6653 | 0.1917 | -0.9887 | 1.5746 | -0.1234 | 0.9612 |

Tabell 1: I tabellen vises deskriptiv statistikk for OSEBX og de 8 bransjeindeksene brukt i denne oppgaven. For OSEBX vises logaritmisk avkastning, mens det for de resterende indeksene vises deskriptiv statistikk for korrelasjonen mellom hver indeks og OSEBX. Se vedlegg D for forklaring av begrepene skjevhet og kurtose.

3.3 Korrelasjon

Denne oppgaven tar utgangspunkt i korrelasjonen mellom hver enkelt bransjeindeks og OSEBX som forklaringsvariabel. Som nevnt tidligere forteller korrelasjonen oss om styrken til det lineære forholdet mellom to variabler. Det brukes daglige avkastninger til å regne ut månedlige korrelasjon. Dette utgjør totalt 264 ulike korrelasjonsverdier, som illustreres grafisk i figur 5.



Figur 5: Figuren viser korrelasjonen mellom hver bransjeindeks og hovedindeksen på Oslo Børs i perioden 1996-2017. Y-aksen viser korrelasjonen og X-aksen viser årstall. Figuren viser at energisektoren har positiv korrelasjon med hovedindeksen gjennom hele observasjonsperioden. Store deler av perioden har korrelasjonen mellom energiindeksen og hovedindeksen et sterkt positivt lineært forhold, som indikerer at korrelasjonen ligger mellom 0.7 og 1.0 (Ratner, 2013). Denne korrelasjonen har også holdt seg på et nokså høyt stabilt nivå i perioden 2005-2013. Ut fra grafene samlet sett ser det ut til at energisektorens korrelasjon med OSEBX endres minst av alle korrelasjonene i utvalgsperioden, da den holder seg på et nokså høyt stabilt nivå. Korrelasjonen mellom de resterende syv bransjeindeksene og hovedindeksen varierer mer i perioden og har alle perioder med negativ korrelasjon. Det er også verdt å legge merke til at korrelasjonen mellom OSEBX og bransjeindeksene for forbruksvarer, finans, helsevern, IT og konsumvarer ser ut til å bevege seg et lignende mønster.

4 Resultater

Dette kapitlet presenterer resultatene av de mest vesentlige estimeringene som er gjort i forbindelse med oppgaven. Først presenteres testverdiene av Dickey-Fuller testene, som må gjennomføres for å sikre at de finansielle instrumentene som estimeres er stasjonære. Se delkapittel 3.7 for beskrivelse av stasjonære prosesser. Deretter presenteres testverdiene for om det finnes ARCH-effekter i dataene. Disse testene gjennomføres for å teste om bruk av GARCH-modeller og utvidelser av denne er passende for dataene i oppgaven. Etter dette brukes autokorrelasjon for å vurdere hvilken orden som bør brukes i GARCH-modellene som estimeres. Etter dette presenteres en vurdering av modellene på bakgrunn av informasjonskriteriene, med mål om å finne de beste prediksjonsmodellene. Deretter følger resultater av testene for graden av autokorrelasjon mellom restledd og laggede restledd i de valgte prediksjonsmodellene. Etter dette følger resultater fra predikerte avkastninger for OSEBX, med bakgrunn i prediksjonsmodellene som er estimert i denne oppgaven. Til slutt inkluderes mål på prediksjonsnøyaktighet for å se hvor nøyaktige våre prediksjonsmodeller er.

Som nevnt tidligere er estimeringene gjennomført ved bruk av månedlige avkastnings- og korrelasjonsdata i perioden 1996-2017. Forklaringsvariablene som er brukt er forrige måneds avkastning av OSEBX og forrige måneds korrelasjon mellom bransjeindeks og OSEBX. Det blir kjørt ulike estimeringsmodeller der det rulleres på hvilke forklaringsvariabler som blir inkludert.

4.1 Testverdier for Dickey-Fuller test

Dickey-Fuller (DF) test ble gjennomført for å finne ut om de finansielle dataene brukt i denne oppgaven var stasjonære eller ikke. I denne oppgaven er variablene som testes for stasjonaritet avkastningen til OSEBX og korrelasjonene mellom avkastningene til de 8 bransjeindeksene og OSEBX. Resultater fra Dickey-Fuller testene vises i tabell 2.

Tabell 2: Resultater fra Dickey-Fuller test

| | Konstant | | Konstant og trend | |
|------------------------------|----------|----------------|-------------------|----------------|
| | μ | \hat{t}_{DF} | μ | \hat{t}_{DF} |
| R_{OSEBX} | -0.8264 | -6.4618 | -0.8245 | -10.31 |
| $\rho_{OSEBX,energi}$ | -0.2735 | -4.7018 | -0.3138 | -4.7243 |
| $\rho_{OSEBX,forbruksvarer}$ | -0.4532 | -4.7818 | -0.4469 | -4.6965 |
| $\rho_{OSEBX,finans}$ | -0.3840 | -4.6250 | -0.6411 | -5.7947 |
| $\rho_{OSEBX,helsevern}$ | -0.5419 | -4.7267 | -0.5764 | -4.9752 |
| $\rho_{OSEBX,industri}$ | -0.4428 | -4.8752 | -0.5611 | -5.8342 |
| $\rho_{OSEBX,IT}$ | -0.4981 | -4.9084 | -0.6274 | -5.8468 |
| $\rho_{OSEBX,konsumvarer}$ | -0.6517 | -5.4006 | -0.6529 | -5.4060 |
| $\rho_{OSEBX,materialer}$ | -0.4004 | -4.7764 | -0.4948 | -5.1538 |

Tabell 2: Tabellen viser testverdiene for Dickey-Fuller (DF) testene. De kritiske verdiene ved DF er ulike ved hver estimering og blir oppgitt i Oxmetrics.

Kritisk DF-verdi ved beregning av konstant er -2.86 ved 5% signifikansnivå, mens den kritiske DF-verdien ved inkludering av både konstant og trend er -3.41 ved 5% signifikansnivå. Dette innebærer at vi på bakgrunn av DF test kan forkaste nullhypotesen om at prosessene er ikke-stasjonære. Dette betyr at prosessene brukt i oppgaven kan ses på som stasjonære. En konsekvens av dette er at avkastning og korrelasjon kan brukes til predikere både avkastning og volatilitet, som er ønskelig.

4.2 Testverdier for ARCH-effekt

En ARCH-test ble gjennomført for å undersøke om det ARCH-effekter er til stede i de finansielle dataene. Resultater fra ARCH-testen vises i tabell 3.

Tabell 3: Resultater fra ARCH-test

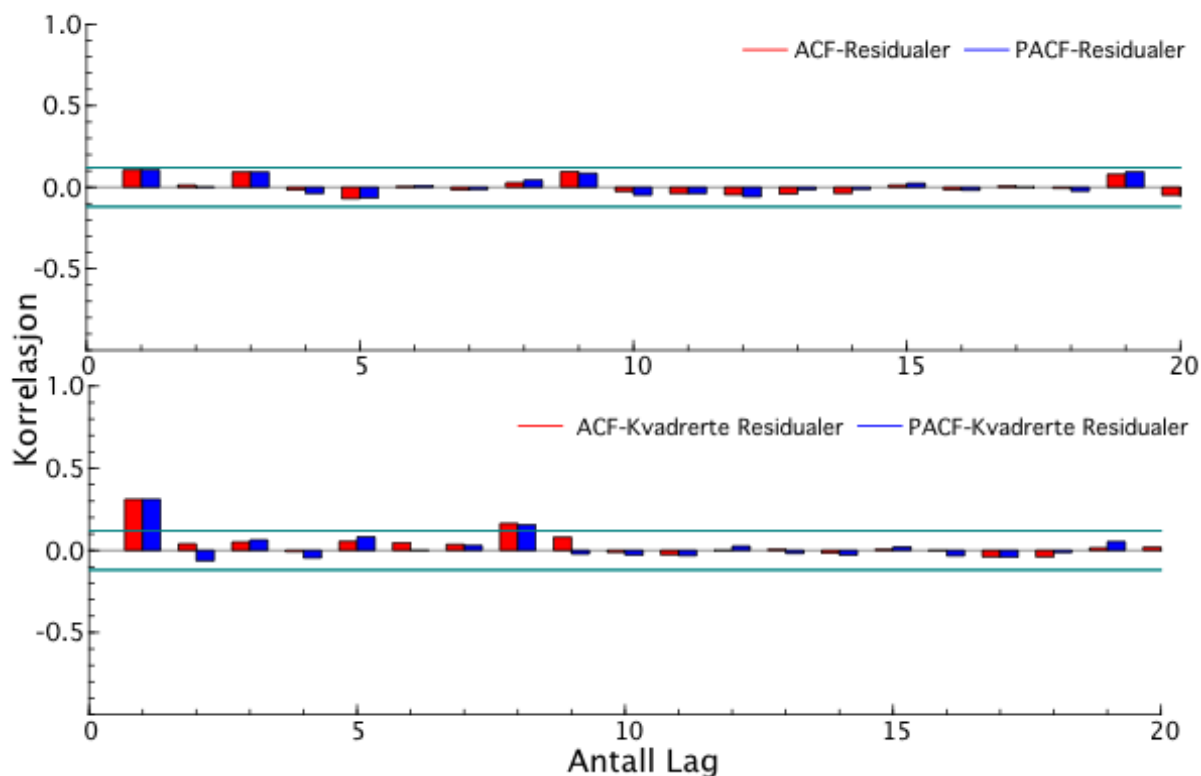
| | Testverdi | p-verdi |
|----------------|-----------|----------|
| ARCH 1-10 test | 2.9546 | 0.0016** |

*Tabell 3: Tabellen viser testverdien av en ARCH-test på en lineær regresjonsmodell som estimerer avkastning i periode t . Forklaringsvariabelen er lagget avkastning i periode $(t-1)$. ** indikerer at nullhypotesen om ingen ARCH-effekter kan forkastes ved 1% signifikansnivå.*

Resultatet fra ARCH-testen er signifikant ved 1 % signifikansnivå, som betyr at vi kan forkaste nullhypotesen om ingen ARCH-effekter. Dette betyr at en ARCH/GARCH-modell vil passe bra for å predikere avkastning ut fra dataene som brukes i oppgaven.

4.3 Valg av orden i GARCH-modellene

Autokorrelasjon kan brukes til å finne ut hvor mange lag som bør inkluderes i modellen, der antall lag forteller hvor mange perioder bak i tid som bør inkluderes. Ifølge Enders (2015) kan autokorrelasjonsfunksjonen til de kvadrerte restleddene hjelpe til med å finne ordenen til GARCH-modellen (Enders, 2015, s. 129). Autokorrelasjonsfunksjonen (ACF) og den partielle autokorrelasjonsfunksjonen (PACF) til de standardiserte restleddene i en autoregressive-moving-average (ARMA) (1,1)-prosess er illustrert i figur 6.



Figur 6: Figuren viser autokorrelasjonsfunksjonen (ACF) og den partielle autokorrelasjonsfunksjonen (PACF) mellom restleddene og de laggede restleddene i modellene som estimeres. Visuell inspeksjon av figur 6 viser tendenser til en svakt avtakende trend i korrelasjonsfunksjonen, som innebærer at prosessen kan være stasjonær. Dette tyder på at en GARCH-modell kan brukes. Ved studering av ACF og PACF til de kvadrerte restleddene ser vi en svak avtakende trend, og modellen ser derfor ut til å følge et GARCH-mønster. Av PACF ser vi en tydelig signifikant søyle på lag 1, som kan tyde på at modellen følger en AR(1)-prosess. Et annet kjennetegn ved GARCH-prosesser er geometrisk avtakende ACF, som vi ser svake tendenser til i denne figuren. ACF har også en signifikant søyle på 1.lag, som tyder på at det er ARCH-effekter i restleddene. At modellen bør inneholde ARCH-ledd støttes også opp av geometrisk avtakende PACF. Siden både ACF og PACF har signifikante søyler ved 1.lag, indikerer det at prosessen er av ordenen (1,1). En GARCH-modell av orden (1,1) ser derfor ut til å passe våre data bra. Kilde: Thomson Reuters Datastream

4.4 Resultater av modellestimeringer

Først vil resultatene av GARCH(1,1)-, TGARCH(1,1)- og EGARCH(1,1)-modeller med lagget avkastning som eneste forklaringsvariabel i hovedligningen estimeres. GARCH(1,1) og TGARCH(1,1) har begge restledd som følger student t-fordelingen, mens EGARCH estimeringene følger en normalfordeling. Tabell 4 viser estimater fra en tradisjonell GARCH(1,1)-modell med student t-fordelte restledd.

Tabell 4: Resultater fra tradisjonell GARCH(1,1)-modell med student t-fordelte restledd

| Hovedligning | | | | Variansligning | | | |
|--------------|--------|---------|---------|----------------|--------|---------|---------|
| Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| a_0 | 0.01 | 4.02 | 0.00** | α_0 | 0.0002 | 1.16 | 0.25 |
| a_1 | 0.08 | 1.19 | 0.23 | α_1 | 0.15 | 1.99 | 0.048* |
| | | | | β_1 | 0.8 | 8.19 | 0.00** |

Tabell 4: Restleddene til GARCH(1,1)-modellen følger en student t-fordeling. (**) indikerer $p\text{-verdi} \leq 0.01$, (*) indikerer $p\text{-verdi} \leq 0.05$.

Det viser seg at kun konstantleddet, representert ved a_0 , har signifikant effekt på avkastningen (se tabell 1). Den laggede avkastningen, a_1 , har ikke signifikant effekt. I den betingede variansligningen har vektet gjennomsnittlig varians i det lange løp, representert ved α_0 , ingen signifikant effekt på variansen. α_1 og β_1 har derimot signifikant effekt på variansen ved henholdsvis 5% og 1% signifikansnivå. For at tradisjonelle GARCH-modeller skal kunne brukes, må koeffisientene i variansligningen være positive, et kriterie som viser seg å være oppfylt, se tabell 4. Det viser seg at kriteriet for at prosessene som testes er stasjonære også er oppfylt, representert ved at $\alpha_1 + \beta_1 < 1$ er oppfylt. En GARCH(1,1)-modell ser derfor ut til å passe bra til dataene, da det er viktigere at α_1 og β_1 gir signifikante verdier, enn α_0 . Av figur 4 i delkapittel 3.2.1 så vi tendenser til asymmetri i modellen. Dette er det ikke sikkert en tradisjonell GARCH(1,1)-modell klarer å fange opp. Det beregnes derfor også TGARCH(1,1)- og EGARCH(1,1)-modeller, da begge disse tar høyde for asymmetri, og kan fange opp effekter som en enkel GARCH-modell kan ha problemer med å fange opp. Tabell 5 viser resultater fra TGARCH(1,1)-modellen:

Tabell 5: Resultater fra TGARCH(1,1)-modell med student t-fordelte restledd

| Hovedligning | | | | Variansligning | | | |
|--------------|--------|---------|---------|----------------|--------|---------|---------|
| Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| a_0 | 0.01 | 3.76 | 0.00** | α_0 | 0.0004 | 1.56 | 0.12 |
| a_1 | 0.08 | 1.27 | 0.21 | α_1 | -0.01 | -0.11 | 0.91 |
| | | | | β_1 | 0.75 | 8.96 | 0.00** |
| | | | | λ_1 | 0.25 | 1.08 | 0.28 |

Tabell 5: Restleddene til TGARCH(1,1)-modellen følger en student t-fordeling. (**) indikerer p-verdi ≤ 0.01 , (*) indikerer p-verdi ≤ 0.05 .

Av delkapittel 2.4 fremgår det at en av antagelsene for å kunne bruke TGARCH-modeller er at α_1 må være positiv. Ved estimeringene av TGARCH(1,1) i tabell 5 ses det tydelig at denne antagelsen ikke oppfylles. Dette tyder på at TGARCH(1,1) ikke er en passende modell til datasettet. En negativ α_1 kan være en indikasjon på at EGARCH(1,1) er mer passende, da den tillater for negativ effekt av restleddene i foregående periode, representert ved α_1 . Derfor ble det valgt å estimere en EGARCH(1,1)-modell:

Tabell 6: Resultater fra EGARCH(1,1)-modell med normalfordelte restledd

| Hovedligning | | | | Variansligning | | | |
|--------------|--------|---------|---------|----------------|--------|---------|---------|
| Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| α_0 | 0.01 | 3.94 | 0.00** | α_0 | -0.83 | -2.88 | 0.004** |
| α_1 | 0.02 | 0.22 | 0.83 | α_1 | -0.23 | -2.55 | 0.011* |
| | | | | β_1 | 0.85 | 16.70 | 0.00** |
| | | | | λ_1 | 0.25 | 3.13 | 0.002** |

Tabell 6: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer p-verdi ≤ 0.01 , (*) indikerer p-verdi ≤ 0.05 .

I tabell 6 for EGARCH(1,1) ser vi at alle koeffisientene unntatt en har signifikant effekt på både avkastning og volatilitet. Den eneste variabelen som ikke har signifikant effekt er lagget avkastning ved OSEBX, representert ved α_1 . For å finne ut hvilken GARCH-modell som passer best til dataene må koeffisientenes signifikansnivå ses i sammenheng med log-likelihood og informasjonskriteriene, som vises i tabell 7.

Tabell 7: Valg av beste modell basert på informasjonskriterier

| Modell | Log-likelihood | AIC | HQC | SBC |
|-------------|----------------|--------|--------|--------|
| GARCH(1,1) | 393.68 | -2.94 | -2.90< | -2.86< |
| TGARCH(1,1) | 394.96 | -2.94< | -2.90 | -2.84 |
| EGARCH(1,1) | 385.74 | -2.88 | -2.84 | -2.80 |

Tabell 7: (<) indikerer hvilken modell som foretrekkes basert på de ulike informasjonskriteriene.

Log-likelihood er estimeringsmetoden til modellen og forteller oss hvor bra modellen er. Denne verdien skal helst være høyest mulig, mens informasjonskriteriene helst skal være så negative som mulig. Av tabell 7 ser man at informasjonskriteriene HQC og SBC favoriserer den tradisjonelle GARCH(1,1)-modellen. Dette er ikke uventet, da spesielt SBC ofte favoriserer modellene med færrest variabler. AIC derimot favoriserer TGARCH(1,1)-modellen, som vi allerede har utelukket. Grunnen til at AIC velger TGARCH(1,1) kan skyldes at denne modellen gir høyest log-likelihood verdi, etterfulgt av GARCH(1,1)-modellen. Valg av modell til videre arbeid faller da på EGARCH(1,1)-modellen, selv om ingen av informasjonskriteriene favoriserer modellen. Begrunnelsen for valget skyldes variablenes signifikansnivå.

Log-likelihood og informasjonskriterier er oppgitt i tabell 7, og disse virker å favorisere GARCH(1,1)- og TGARCH(1,1)-modellene. Koeffisientenes signifikansnivå får det derimot til å se ut som EGARCH(1,1)-modellen passer best til datasettet vårt. Ved bruk av EGARCH(1,1)-modellen har alle koeffisientene utenom en, signifikant effekt. EGARCH-modellene har de mest signifikante koeffisientene og ser derfor ut til å estimere den beste modellen. Totalt sett tyder dette på at en modell av typen EGARCH vil være best til å predikere avkastningen og volatiliteten til OSEBX. Det tas derfor utgangspunkt i ulike versjoner av EGARCH når vi videre ser på om forklaringskraften til modellen økes ved inkludering av korrelasjon mellom bransjeindekser og OSEBX, som ekstra forklaringsvariabel i både hoved-og variansligning.

Denne oppgaven fokuserer på korrelasjonskoeffisientens effekt til å predikere fremtidig avkastningen ved OSEBX. Det estimeres følgelig en EGARCH(1,1)-modell uten et autoregressivt ledd i hovedligningen, som vil si at man fjerner lagget avkastning som variabel fra modellen. Avkastningen avhenger altså da bare av et konstantledd og et restledd.

Tabell 8: Resultater fra EGARCH(1,1)-modell med normalfordelte restledd

| Hovedligning | | | | Variansligning | | | |
|--------------|--------|---------|---------|----------------|--------|---------|---------|
| Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| α_0 | 0.01 | 8.15 | 0.00** | α_0 | -0.84 | -2.92 | 0.004** |
| | | | | α_1 | -0.23 | -2.61 | 0.01** |
| | | | | β_1 | 0.85 | 17.00 | 0.00** |
| | | | | λ_1 | 0.26 | 3.87 | 0.00** |

Tabell 8: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer p-verdi ≤ 0.01 , (*) indikerer p-verdi ≤ 0.05

Ved fjerning av lagget avkastning som variabel ser man tydelig at alle parameterne ble signifikante ved 1% signifikansnivå, som var forventet. Spørsmålet er hvordan EGARCH-modellene reagerer når modellen utvides ved inkludering av korrelasjonsledd, som er illustrert i tabellene 9-16. Estimeringene i disse tabellene tar utgangspunkt i EGARCH(1,1)-modeller med og uten lagget avkastning og med og uten lagget korrelasjon i variansligning. Tabell 9 tar først for seg utvidete EGARCH(1,1)-modeller med lagget avkastning, mens tabell 11 ikke tar høyde for lagget avkastning, og fokuserer kun på korrelasjon som forklaringsvariabel.

De påfølgende sidene tar for seg fire ulike utvidete EGARCH(1,1) modeller med korrelasjon som ekstra forklaringsvariabel.

Hver tabell nedenfor inneholder 8 modeller, der modellnummeret indikerer hvilken korrelasjonsvariabel som er inkludert:

1. Korrelasjon mellom avkastningen til bransjeindeksen energi og OSEBX sin avkastning.
2. Korrelasjon mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX sin avkastning.
3. Korrelasjon mellom avkastningen til bransjeindeksen finans og OSEBX sin avkastning.
4. Korrelasjon mellom avkastningen til bransjeindeksen helsevern og OSEBX sin avkastning.
5. Korrelasjon mellom avkastningen til bransjeindeksen industri og OSEBX sin avkastning.

6. Korrelasjon mellom avkastningen til bransjeindeksen informasjonsteknologi og OSEBX sin avkastning.
7. Korrelasjon mellom avkastningen til bransjeindeksen konsumvarer og OSEBX sin avkastning.
8. Korrelasjon mellom avkastningen til bransjeindeksen materialer og OSEBX sin avkastning.

Tabellene under viser resultatene av de ulike utvidelsene av EGARCH(1,1)-modellen, som ble introdusert i delkapittel 4.6. Utvidelsen av modellene utgjøres av korrelasjonen mellom hver bransjeindeks og hovedindeksen i forrige periode. Tabell 9 tar for seg resultatene fra estimeringer av den utvidede EGARCH(1,1)-modellen med korrelasjon i både hoved-og variansligning.

Tabell 9: Resultater fra EGARCH(1,1)-modell med utvidet hoved-og variansligning, og normalfordelte restledd

| Modell | Hovedligning | | | | Variansligning | | | |
|--------|--------------|---------|---------|---------|----------------|---------|---------|---------|
| | Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| 1 | a_0 | 0.01 | 0.10 | 0.92 | α_0 | -0.83 | -2.29 | 0.023* |
| | a_1 | 0.02 | 0.20 | 0.85 | α_1 | -0.23 | -2.47 | 0.014* |
| | a_2 | -0.0002 | -0.002 | 1.00 | β_1 | 0.85 | 14.10 | 0.00** |
| | | | | | λ_1 | 0.25 | 2.77 | 0.006** |
| | | | | | γ_1 | -0.0005 | -0.002 | 1.00 |
| 2 | a_0 | -0.02 | -1.07 | 0.29 | α_0 | -7.61 | -3.60 | 0.00** |
| | a_1 | 0.10 | 0.79 | 0.43 | α_1 | -0.13 | -0.59 | 0.56 |
| | a_2 | 0.05 | 1.63 | 0.10 | β_1 | -0.42 | -1.25 | 0.21 |
| | | | | | λ_1 | 0.26 | 1.27 | 0.20 |
| | | | | | γ_1 | -0.69 | -0.98 | 0.33 |
| 3 | a_0 | 0.001 | 0.08 | 0.94 | α_0 | -0.25 | -1.27 | 0.21 |
| | a_1 | 0.03 | 0.35 | 0.73 | α_1 | -0.24 | -3.28 | 0.001** |
| | a_2 | 0.01 | 0.75 | 0.45 | β_1 | 0.87 | 23.50 | 0.00** |
| | | | | | λ_1 | 0.20 | 2.96 | 0.003** |
| | | | | | γ_1 | -0.74 | -3.27 | 0.001** |
| 4 | a_0 | 0.003 | 0.66 | 0.51 | α_0 | -0.77 | -2.50 | 0.013* |
| | a_1 | 0.04 | 13.50 | 0.00** | α_1 | -0.27 | -2.60 | 0.01** |
| | a_2 | 0.2 | 3.07 | 0.002** | β_1 | 0.85 | 16.40 | 0.00** |
| | | | | | λ_1 | 0.27 | 3.20 | 0.002** |
| | | | | | γ_1 | -0.28 | -0.98 | 0.33 |
| 5 | a_0 | -0.001 | -0.08 | 0.94 | α_0 | -0.91 | -2.99 | 0.003** |
| | a_1 | 0.04 | 1.02 | 0.31 | α_1 | -0.19 | -2.26 | 0.024* |
| | a_2 | 0.02 | 0.56 | 0.58 | β_1 | 0.23 | 0.91 | 0.36 |
| | | | | | λ_1 | 0.26 | 3.05 | 0.003** |
| | | | | | γ_1 | 0.23 | 0.91 | 0.36 |
| 6 | a_0 | 0.001 | 0.12 | 0.90 | α_0 | -0.93 | -2.03 | 0.044* |
| | a_1 | 0.05 | 0.58 | 0.57 | α_1 | -0.22 | -2.41 | 0.017* |
| | a_2 | 0.01 | 0.80 | 0.43 | β_1 | 0.86 | 17.70 | 0.00** |
| | | | | | λ_1 | 0.20 | 1.50 | 0.14 |
| | | | | | γ_1 | 0.16 | 0.38 | 0.70 |
| 7 | a_0 | -0.001 | -0.07 | 0.95 | α_0 | -9.69 | -11.00 | 0.00** |
| | a_1 | 0.11 | 1.25 | 0.21 | α_1 | -0.03 | -0.33 | 0.74 |
| | a_2 | 0.01 | 0.59 | 0.55 | β_1 | -0.79 | -6.79 | 0.00** |
| | | | | | λ_1 | 0.20 | 1.57 | 0.12 |
| | | | | | γ_1 | -0.72 | 2.29 | 0.023* |
| 8 | a_0 | 0.009 | 0.83 | 0.41 | α_0 | -0.48 | -1.55 | 0.12 |
| | a_1 | 0.02 | 0.21 | 0.84 | α_1 | -0.24 | -2.87 | 0.004** |
| | a_2 | 0.001 | 0.08 | 0.94 | β_1 | 0.88 | 18.50 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.36 | 0.001** |
| | | | | | γ_1 | -0.33 | -1.70 | 0.09 |

Tabell 9: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer p-verdi ≤ 0.01 , (*) indikerer p-verdi ≤ 0.05 . Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med.

Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom

bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX.

Modellene 1, 3 og 4, utvidede EGARCH-modeller med henholdsvis korrelasjon mellom bransjeindeksen energi og OSEBX, korrelasjonen mellom bransjeindeksen finans og OSEBX, og til slutt korrelasjonen mellom bransjeindeksen helsevern og OSEBX, finner nærmest bare signifikante verdier. For å velge hvilken modell som forklarer datasettet best tas det i bruk informasjonskriterier. Av disse tre modellene ser modell 4 mest lovende ut på bakgrunn av p-verdier, tett etterfulgt av modell 3. Vi legger merke til at modell 4 som ser på korrelasjonen mellom bransjeindeks helsevern og OSEBX er den eneste modellen der korrelasjonskoeffisienten har signifikant effekt på avkastningen til OSEBX. Dette må likevel testes ved bruk av informasjonskriter, som presenteres i tabell 10.

Tabell 10: Valg av EGARCH(1,1)-modell basert på informasjonskriterier

| Modell | Log-likelihood | AIC | HQ | SBC |
|--------|----------------|---------|---------|---------|
| 1 | 385.744 | -2.862 | -2.818 | -2.753 |
| 2 | 371.732 | -2.756 | -2.712 | -2.647 |
| 3 | 395.822 | -2.938< | -2.895< | -2.830< |
| 4 | 387.296 | -2.874 | -2.830 | -2.765 |
| 5 | 387.120 | -2.872 | -2.829 | -2.764 |
| 6 | 386.441 | -2.867 | -2.823 | -2.759 |
| 7 | 368.842 | -2.734 | -2.690 | -2.625 |
| 8 | 387.265 | -2.873 | -2.830 | -2.765 |

Tabell 10: Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX. (<) indikerer hvilken modell som foretrekkes basert på log-likelihoodverdi og de ulike informasjonskriteriene

Modell 3, som inkluderer korrelasjonen mellom bransjeindeksen finans og OSEBX, favoriseres av alle informasjonskriteriene og er følgelig den beste modellen å bruke til

estimering og prediksjon av OSEBX sin avkastning og volatilitet. Vi legger dog merke til at modell 4, som inkluderer korrelasjonen mellom helsevern og OSEBX, har flere signifikante koeffisienter enn modell 3 som informasjonskriteriene favoriserer. Grunnen til at informasjonskriteriene favoriserer modell 3 skyldes nok i stor grad at korrelasjonskoeffisienten i variansligningen, representert ved γ_1 , har signifikant effekt på variansen. Som nevnt i delkapittel 1.1 fokuserer GARCH-modeller på prediksjon av volatilitet, og informasjonskriteriene vil derfor vektlegge signifikante verdier i variansligningen høyere enn signifikante verdier i hovedligningen. Det er også verdt å nevne at finanssektoren utgjør en vesentlig større del av OSEBX, enn helsesektoren (se vedlegg B for sammensetningen av sektorer i OSEBX). Bransjeindeksen for finans utgjør hele 18% av OSEBX, i motsetning til bransjeindeksen for helsevern som på sin side bare utgjør 0.16% av hovedindeksen. På bakgrunn av sektorsammensetningen til OSEBX velger vi å stole på informasjonskriteriene, og den beste modellen vil derfor inneholde korrelasjonen mellom OSEBX og bransjeindeksen finans. Denne modellen omtales som modell a) i resten av oppgaven.

Vi tar utgangspunkt i modellene som ble estimert i tabell 9, og estimerer nye utvidede EGARCH(1,1)-modeller. Forskjellen er at vi i denne omgangen ekskluderer lagget avkastning i hovedligningen, for å se om dette endrer koeffisientens signifikansnivå og valg av beste modell. Resultatene av disse estimeringene er vist i tabell 11.

Tabell 11: Resultater fra EGARCH(1,1)-modell, uten lagget avkastning i hovedligning, men med lagget korrelasjon i variansligning

| Modell | Hovedligning | | | | Variansligning | | | |
|--------|--------------|---------|---------|---------|----------------|--------|---------|---------|
| | Variabel | Koeff. | t-verdi | p-verdi | Variabel | Koeff. | t-verdi | p-verdi |
| 1 | a_0 | 0.01 | 1.70 | 0.09 | α_0 | -0.83 | -2.28 | 0.023* |
| | a_2 | -0.0004 | -0.05 | 0.96 | α_1 | -0.23 | -2.59 | 0.01** |
| | | | | | β_1 | 0.85 | 16.60 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.87 | 0.00** |
| | | | | | γ_1 | -0.006 | -0.03 | 0.98 |
| 2 | a_0 | 0.003 | 0.41 | 0.68 | α_0 | -0.35 | -0.78 | 0.44 |
| | a_2 | 0.01 | 1.00 | 0.32 | α_1 | -0.22 | -2.52 | 0.012* |
| | | | | | β_1 | 0.89 | 13.8 | 0.00** |
| | | | | | λ_1 | 0.27 | 4.36 | 0.00** |
| | | | | | γ_1 | -0.45 | -2.10 | 0.037* |
| 3 | a_0 | 0.0007 | 0.15 | 0.88 | α_0 | -0.25 | -1.25 | 0.21 |
| | a_2 | 0.01 | 2.46 | 0.014* | α_1 | -0.24 | -3.34 | 0.001** |
| | | | | | β_1 | 0.87 | 23.60 | 0.00** |
| | | | | | λ_1 | 0.20 | 3.54 | 0.00** |
| | | | | | γ_1 | -0.75 | -3.50 | 0.001** |
| 4 | a_0 | 0.005 | 0.82 | 0.41 | α_0 | -0.79 | -2.53 | 0.012* |
| | a_2 | 0.02 | 1.29 | 0.20 | α_1 | -0.27 | -2.52 | 0.012* |
| | | | | | β_1 | 0.84 | 15.90 | 0.00** |
| | | | | | λ_1 | 0.28 | 3.35 | 0.001** |
| | | | | | γ_1 | -0.29 | -0.94 | 0.35 |
| 5 | a_0 | 0.002 | 0.30 | 0.76 | α_0 | -0.94 | -3.05 | 0.003** |
| | a_2 | 0.01 | 1.20 | 0.23 | α_1 | -0.19 | -2.26 | 0.025* |
| | | | | | β_1 | 0.87 | 20.00 | 0.00** |
| | | | | | λ_1 | 0.27 | 3.96 | 0.00** |
| | | | | | γ_1 | 0.23 | 0.94 | 0.35 |
| 6 | a_0 | 0.005 | 0.42 | 0.68 | α_0 | -0.93 | -2.03 | 0.044* |
| | a_2 | 0.01 | 0.56 | 0.58 | α_1 | -0.23 | -2.44 | 0.015* |
| | | | | | β_1 | 0.85 | 17.50 | 0.00** |
| | | | | | λ_1 | 0.23 | 2.10 | 0.036* |
| | | | | | γ_1 | 0.13 | 0.32 | 0.75 |
| 7 | a_0 | 0.01 | 2.24 | 0.026* | α_0 | -0.55 | -1.34 | 0.18 |
| | a_2 | -0.006 | -0.53 | 0.60 | α_1 | -0.24 | -2.79 | 0.006** |
| | | | | | β_1 | 0.86 | 15.70 | 0.00** |
| | | | | | λ_1 | 0.22 | 3.80 | 0.00** |
| | | | | | γ_1 | -0.46 | -1.70 | 0.091 |
| 8 | a_0 | 0.01 | 4.17 | 0.00** | α_0 | -0.48 | -1.57 | 0.12 |
| | a_2 | 0.0007 | 0.28 | 0.78 | α_1 | -0.24 | -2.88 | 0.004** |
| | | | | | β_1 | 0.88 | 18.70 | 0.00** |
| | | | | | λ_1 | 0.26 | 4.17 | 0.00** |
| | | | | | γ_1 | -0.33 | -1.79 | 0.074 |

Tabell 11: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer $p\text{-verdi} \leq 0.01$, (*) indikerer $p\text{-verdi} \leq 0.05$. Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom

bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX.

Av tabell 11 kan det se ut til at modell 3 passer best til vårt datasett. Dette skyldes i hovedsak at modell 3 er den eneste modellen med 5 signifikante koeffisienter. Det bør også legges merke til at lagget korrelasjon i modell 3 er blitt signifikant, mens det motsatte har skjedd ved lagget korrelasjon i modell 4, sammenlignet med EGARCH-modellene i tabell 9. I modell 4 har altså effekten av å fjerne lagget avkastning i modellen medført at koeffisienten til lagget korrelasjon i hovedligningen har gått fra å være signifikant, til å bli ikke-signifikant. Igjen er det bransjeindeksene finans og helsevern som skiller seg ut som gode forklaringsvariabler i estimeringene. Tabell 12 viser log-likelihood og informasjonskriterier til den utvidede EGARCH(1,1)-modellen uten lagget avkastning i hovedligningen, men med lagget korrelasjon i variansligningen.

Tabell 12: Valg av modell basert på informasjonskriterier

| Modell | Log-likelihood | AIC | HQ | SBC |
|--------|----------------|--------|--------|--------|
| 1 | 385.72 | -2.87 | -2.83 | -2.77 |
| 2 | 389.75 | -2.90 | -2.86 | -2.80 |
| 3 | 395.76 | -2.95< | -2.91< | -2.85< |
| 4 | 387.06 | -2.88 | -2.84 | -2.78 |
| 5 | 386.90 | -2.88 | -2.84 | -2.78 |
| 6 | 386.18 | -2.87 | -2.83 | -2.78 |
| 7 | 388.22 | -2.89 | -2.85 | -2.79 |
| 8 | 387.24 | -2.88 | -2.84 | -2.79 |

Tabell 12: Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX. (<) indikerer hvilken modell som foretrekkes ifølge informasjonskriteriene

Av tabell 12 ser man at alle informasjonskriteriene fortsatt favoriserer modell 3, selv om den laggede avkastningen til OSEBX fjernes som variabel. Det ser altså ut som modell 3 er best igjen til å estimere avkastning og varians. Modell 3 inkluderer den utvidede EGARCH(1,1)-modellen med korrelasjonskoeffisienten mellom bransjeindeksen finans og OSEBX som ekstra forklaringsvariabel. Denne modellen vil omtales som b) i resten av oppgaven.

Tabellene 13-16 tar for seg resultatene av utvidede EGARCH(1,1)-modeller når korrelasjonskoeffisienten fjernes i variansligningen. I utgangspunktet er det avkastningen som ønskes estimert, og derfor fjernes korrelasjonen mellom avkastningene til bransjeindeksene og OSEBX fra variansligningen for å se om resultatene bedres og om dette endrer utfallet for valg av modell. Tabell 13 presenterer en EGARCH(1,1)-modell med bare utvidet hovedligning.

Tabell 13: Resultater fra EGARCH(1,1)-modell med utvidet hovedligning

| Modell | Hovedligning | | | | Variansligning | | | |
|--------|--------------|---------|---------|---------|----------------|--------|---------|---------|
| | Variabel | Koeff. | t-verdi | t-sanns | Variabel | Koeff. | t-verdi | t-sanns |
| 1 | a_0 | 0.01 | 0.12 | 0.91 | α_0 | -0.83 | -2.29 | 0.023* |
| | a_1 | 0.02 | 0.19 | 0.85 | α_1 | -0.23 | -2.52 | 0.012* |
| | a_2 | -0.0002 | -0.002 | 1.00 | β_1 | 0.85 | 13.80 | 0.00** |
| | | | | | λ_1 | 0.25 | 2.55 | 0.011* |
| 2 | a_0 | 0.0008 | 0.05 | 0.96 | α_0 | -0.91 | -2.60 | 0.01** |
| | a_1 | 0.04 | 0.45 | 0.65 | α_1 | -0.23 | -2.54 | 0.012* |
| | a_2 | 0.02 | 0.58 | 0.57 | β_1 | 0.84 | 13.90 | 0.00** |
| | | | | | λ_1 | 0.24 | 2.58 | 0.01** |
| 3 | a_0 | 0.004 | 0.25 | 0.80 | α_0 | -0.86 | -2.68 | 0.008** |
| | a_1 | 0.03 | 0.31 | 0.75 | α_1 | -0.22 | -2.16 | 0.032* |
| | a_2 | 0.009 | 0.38 | 0.71 | β_1 | 0.85 | 15.20 | 0.00** |
| | | | | | λ_1 | 0.25 | 2.98 | 0.003** |
| 4 | a_0 | 0.004 | 0.05 | 0.96 | α_0 | -0.83 | -3.00 | 0.003** |
| | a_1 | 0.05 | 0.04 | 0.97 | α_1 | -0.24 | -1.15 | 0.25 |
| | a_2 | 0.01 | 0.11 | 0.91 | β_1 | 0.86 | 18.40 | 0.00** |
| | | | | | λ_1 | 0.24 | 0.59 | 0.56 |
| 5 | a_0 | -0.003 | -0.31 | 0.76 | α_0 | -0.79 | -3.00 | 0.003** |
| | a_1 | 0.04 | 8.42 | 0.00** | α_1 | -0.22 | -2.62 | 0.009** |
| | a_2 | 0.02 | 1.45 | 0.15 | β_1 | 0.86 | 18.90 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.76 | 0.00** |
| 6 | a_0 | 0.002 | 0.06 | 0.95 | α_0 | -0.83 | -2.91 | 0.004** |
| | a_1 | 0.04 | 0.73 | 0.47 | α_1 | -0.23 | -2.68 | 0.008** |
| | a_2 | 0.01 | 0.29 | 0.78 | β_1 | 0.85 | 17.30 | 0.00** |
| | | | | | λ_1 | 0.23 | 3.37 | 0.001** |
| 7 | a_0 | 0.01 | 1.01 | 0.31 | α_0 | -0.83 | -2.90 | 0.004** |
| | a_1 | 0.01 | 1.34 | 0.18 | α_1 | -0.24 | -2.62 | 0.009** |
| | a_2 | -0.009 | -0.34 | 0.73 | β_1 | 0.85 | 16.90 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.33 | 0.001** |
| 8 | a_0 | 0.01 | 1.10 | 0.27 | α_0 | -0.83 | -2.90 | 0.004** |
| | a_1 | 0.01 | 0.18 | 0.86 | α_1 | -0.23 | -2.48 | 0.014* |
| | a_2 | -0.003 | -0.16 | 0.87 | β_1 | 0.85 | 17.00 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.12 | 0.002** |

Tabell 13: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer $p\text{-verdi} \leq 0.01$, (*) indikerer $p\text{-verdi} \leq 0.05$. Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX.

I tabell 13 ser alle modellene lovende ut utenom modell 4, som er utvidet med korrelasjonen mellom bransjeindeksen helsevern og OSEBX, da denne har færre signifikante verdier enn de andre modellene. Modell 5, som ser på korrelasjonen mellom bransjeindeksen industri og OSEBX, er modellen med flest signifikante verdier. I likhet med estimeringene som er gjort tidligere vil vi også vektlegge informasjonskriteriene ved valg av endelig modell, disse verdiene er presentert i tabell 14.

Tabell 14: Valg av modell basert på informasjonskriterier

| Modell | Log-likelihood | AIC | HQ | SBC |
|--------|----------------|--------|--------|--------|
| 1 | 385.74 | -2.87 | -2.83 | -2.77 |
| 2 | 386.60 | -2.88< | -2.84< | -2.78< |
| 3 | 385.91 | -2.87 | -2.83 | -2.78 |
| 4 | 386.48 | -2.87 | -2.84 | -2.78 |
| 5 | 386.57 | -2.88 | -2.84 | -2.78 |
| 6 | 386.22 | -2.87 | -2.83 | -2.78 |
| 7 | 385.96 | -2.87 | -2.83 | -2.78 |
| 8 | 385.76 | -2.87 | -2.83 | -2.77 |

Tabell 14: Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX. (<) indikerer hvilken modell som foretrekkes basert på log-likelihoodverdi og de ulike informasjonskriteriene

Alle informasjonskriteriene favoriserer modell nummer 2, som bruker korrelasjonen mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX som forklaringsvariabel, i tillegg til lagget avkastning. Vi legger også merke til at det er denne modellen som gir høyest log-likelihood verdi. Vi vil omtale modell 2 i tabell 13 som modell c) i resten av oppgaven.

Den siste utvidede EGARCH(1,1)-modellen inkluderer en utvidet hovedligning uten lagget avkastning. Den utvidede hovedligningen inkluderer da korrelasjonen mellom avkastningene til bransjeindeksene og OSEBX, mens variansligningen er lik den opprinnelige EGARCH(1,1)-modellen. Resultatene fra disse estimeringene er gjengitt i tabell 15.

Tabell 15: Resultater fra EGARCH(1,1)-modell med utvidet hovedligning uten lagget avkastning

| Modell | Variabel | Hovedligning | | | Variansligning | | | |
|--------|----------|--------------|---------|---------|----------------|--------|---------|---------|
| | | Koeff. | t-verdi | t-sanns | Variabel | Koeff. | t-verdi | t-sanns |
| 1 | a_0 | 0.01 | 1.76 | 0.08 | α_0 | -0.84 | -2.90 | 0.004** |
| | a_2 | -0.0004 | -0.06 | 0.95 | α_1 | -0.23 | -2.60 | 0.01** |
| | | | | | β_1 | 0.85 | 16.80 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.86 | 0.00** |
| 2 | a_0 | 0.001 | 0.24 | 0.81 | α_0 | -0.93 | -2.84 | 0.005** |
| | a_2 | 0.02 | 2.35 | 0.02** | α_1 | -0.23 | -2.58 | 0.01** |
| | | | | | β_1 | 0.84 | 14.70 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.67 | 0.00** |
| 3 | a_0 | 0.006 | 0.95 | 0.35 | α_0 | -0.86 | -2.77 | 0.006** |
| | a_2 | 0.008 | 1.03 | 0.30 | α_1 | -0.23 | -2.43 | 0.016* |
| | | | | | β_1 | 0.85 | 15.60 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.84 | 0.00** |
| 4 | a_0 | 0.006 | 2.35 | 0.02* | α_0 | -0.85 | -3.11 | 0.002** |
| | a_2 | 0.01 | 4.98 | 0.00** | α_1 | -0.24 | -2.69 | 0.008** |
| | | | | | β_1 | 0.85 | 18.00 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.81 | 0.00** |
| 5 | a_0 | -0.0004 | -0.007 | 1.00 | α_0 | -0.81 | -2.91 | 0.004** |
| | a_2 | 0.02 | 0.21 | 0.84 | α_1 | -0.22 | -2.72 | 0.007** |
| | | | | | β_1 | 0.86 | 17.90 | 0.00** |
| | | | | | λ_1 | 0.27 | 3.02 | 0.003** |
| 6 | a_0 | 0.005 | 0.47 | 0.64 | α_0 | -0.85 | -3.00 | 0.003** |
| | a_2 | 0.01 | 0.67 | 0.50 | α_1 | -0.23 | -2.60 | 0.01** |
| | | | | | β_1 | 0.85 | 17.10 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.99 | 0.00** |
| 7 | a_0 | 0.02 | 1.84 | 0.066 | α_0 | -0.84 | -2.91 | 0.004** |
| | a_2 | -0.01 | -0.64 | 0.53 | α_1 | -0.24 | -2.58 | 0.01** |
| | | | | | β_1 | 0.85 | 16.90 | 0.00** |
| | | | | | λ_1 | 0.25 | 3.51 | 0.001** |
| 8 | a_0 | 0.01 | 4.00 | 0.00** | α_0 | -0.83 | -2.94 | 0.004** |
| | a_2 | -0.003 | -0.77 | 0.44 | α_1 | -0.23 | -2.61 | 0.01** |
| | | | | | β_1 | 0.85 | 17.20 | 0.00** |
| | | | | | λ_1 | 0.26 | 3.86 | 0.00** |

Tabell 15: Restleddene til EGARCH(1,1)-modellen følger en normalfordeling. (**) indikerer $p\text{-verdi} \leq 0.01$, (*) indikerer $p\text{-verdi} \leq 0.05$. Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX.

Av resultattabell 15 ser vi at det er flere modeller som virker å passe bra til våre data. Det er verdt å merke seg modell 2, 4 og 8, som tar hensyn til bransjeindeksene forbruksvarer, helsevern og materialer. Modell 2 og 4 er spesielt interessante da de innebærer signifikant effekt på avkastning av korrelasjonskoeffisienten. Helsevern må også trekkes fram da de estimerte koeffisientene er alle sammen signifikante. Alle koeffisientene i variansligningen er signifikante ved 1%, i tillegg til at koeffisientene i hovedligningen er signifikante ved 1% for lagget korrelasjon og 5% ved konstantleddet. Tabell 16 viser log-likelihood og informasjonskriteriene til EGARCH(1,1)-modellen med utvidet hovedligning og uten lagget avkastning.

Tabell 16: Valg av modell basert på informasjonskriterier

| Modell | Log-likelihood | AIC | HQ | SBC |
|--------|----------------|---------|---------|---------|
| 1 | 385.72 | -2.877 | -2.844 | -2.795 |
| 2 | 386.51 | -2.883< | -2.850< | -2.801< |
| 3 | 385.86 | -2.878 | -2.845 | -2.796 |
| 4 | 386.24 | -2.881 | -2.848 | -2.799 |
| 5 | 386.37 | -2.882 | -2.849 | -2.800 |
| 6 | 386.03 | -2.879 | -2.846 | -2.798 |
| 7 | 385.95 | -2.878 | -2.846 | -2.797 |
| 8 | 385.74 | -2.877 | -2.844 | -2.796 |

Tabell 16: Modellnummeret refererer til hvilken korrelasjon modellen er utvidet med. Modellene er 1: korrelasjonen mellom bransjeindeksen energi og OSEBX, 2: korrelasjonen mellom bransjeindeksen forbruksvarer og OSEBX, 3: korrelasjonen mellom bransjeindeksen finans og OSEBX, 4: korrelasjonen mellom bransjeindeksen helsevern og OSEBX, 5: korrelasjonen mellom bransjeindeksen industri og OSEBX, 6: korrelasjonen mellom bransjeindeksen IT og OSEBX, 7: korrelasjonen mellom bransjeindeksen konsumvarer og OSEBX, 8: korrelasjonen mellom bransjeindeksen materialer og OSEBX. (<) indikerer hvilken modell som foretrekkes basert på log-likelihoodverdi og de ulike informasjonskriteriene

Alle informasjonskriteriene favoriserer igjen modell 2, der eneste forklaringsvariabel i hovedligningen er korrelasjonen mellom forbruksvarer og OSEBX. Denne modellen vil bli omtalt som modell d) i restene av oppgaven.

Ut ifra tabell 9-16 ser det dermed ut til at to bransjeindekser skiller seg ut som bedre enn de resterende bransjeindeksene ved estimering av OSEBX sin avkastning. De to aktuelle bransjene er henholdsvis forbruksvarer og finans. I de neste delkapitlene vil det beste

modellene fra hver EGARCH(1,1)-modell, inkludert den opprinnelige EGARCH(1,1)-modellen, testes for autokorrelasjon i restleddene. Disse vil brukes til å predikere avkastningen til OSEBX i perioden januar-desember 2018, og til slutt sammenlignes på bakgrunn av modellens prediksjonsnøyaktighet. I løpet av delkapittel 4.4 har vi gitt de beste modellene i hver tabell en bokstav for å gjøre det enklere å holde styr på hvilke modeller som inneholder de ulike variablene:

- Den originale EGARCH(1,1)-modellen med lagget avkastning
 - a) Utvidet EGARCH(1,1)-modell med korrelasjon mellom avkastningen til bransjeindeksen finans og OSEBX i både hovedlingning og variansligning
 - b) Utvidet EGARCH(1,1)-modell med korrelasjon mellom avkastningen til bransjeindeksen finans og OSEBX i både hovedlingning og variansligning, men uten lagget avkastning
 - c) Utvidet EGARCH(1,1)-modell med korrelasjon mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX i hovedligningen.
 - d) Utvidet EGARCH(1,1)-modell med korrelasjon mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX i hovedligningen, men uten lagget avkastning

4.5 Test for autokorrelasjon i restleddene

Tabell 17-21 viser resultater av tester for om det finnes autokorrelasjon mellom restledd og laggede restledd i de fem EGARCH(1,1)-modellene vi fant at var best, basert på signifikansnivå, informasjonskriterier og log-likelihood-estimering. Ljung-Box-tester viste at det ikke var autokorrelasjon mellom restledd og laggede restledd i noen av de 5 EGARCH(1,1)-modellene vi har estimert (se tabell 17-21).

Tabell 17: Ljung-Box test av standardiserte restledd ved original EGARCH(1,1)-modell

| | | |
|-------|---------|--------|
| Q(5) | 3.3915 | 0.6398 |
| Q(10) | 4.7088 | 0.9098 |
| Q(20) | 13.8836 | 0.8363 |
| Q(50) | 39.9666 | 0.8442 |

Tabell 17: * indikerer at det er autokorrelasjon mellom restledd og laggede restledd i dataene. Det ser dermed ut til at vi ikke har problemer med autokorrelasjon i den opprinnelige EGARCH(1,1)-modellen.

Tabell 18: Ljung-Box test av de standardiserte restleddene ved modell a)

| | | |
|-------|---------|--------|
| Q(5) | 2.9681 | 0.5632 |
| Q(10) | 7.2842 | 0.6076 |
| Q(20) | 14.3330 | 0.7639 |
| Q(50) | 36.7200 | 0.9021 |

Tabell 18: * indikerer at det er autokorrelasjon mellom restledd og laggede restledd i dataene. Det ser dermed ut til at vi ikke har problemer med autokorrelasjon i modell a).

Tabell 19: Ljung-Box test av de standardiserte restleddene ved modell b)

| | | |
|-------|---------|--------|
| Q(5) | 3.6481 | 0.6011 |
| Q(10) | 7.8071 | 0.6477 |
| Q(20) | 14.7360 | 0.7913 |
| Q(50) | 36.9060 | 0.9157 |

Tabell 19: * indikerer at det er autokorrelasjon mellom restledd og laggede restledd i dataene. Det ser dermed ut til at vi ikke har problemer med autokorrelasjon i modell b).

Tabell 20: Ljung-Box test av de standardiserte restleddene ved modell c)

| | | |
|-------|---------|--------|
| Q(5) | 4.1517 | 0.3859 |
| Q(10) | 12.4460 | 0.1893 |
| Q(20) | 19.392 | 0.4320 |
| Q(50) | 42.062 | 0.7482 |

Tabell 20: * indikerer at det er autokorrelasjon mellom restledd og laggede restledd i dataene. Det ser dermed ut til at vi ikke har problemer med autokorrelasjon i modell c).

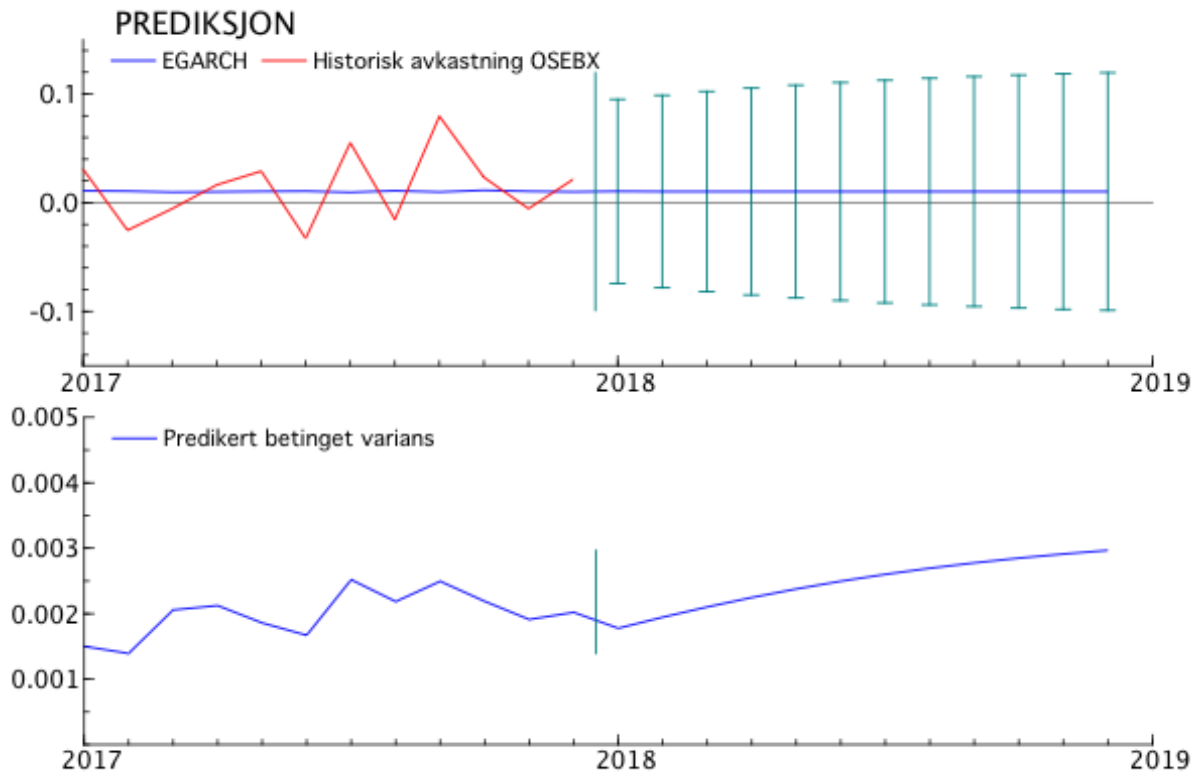
Tabell 21: Ljung-Box test av de standardiserte restleddene ved modell d)

| | | |
|-------|---------|--------|
| Q(5) | 5.6821 | 0.3384 |
| Q(10) | 14.3570 | 0.1573 |
| Q(20) | 21.5580 | 0.3650 |
| Q(50) | 44.5690 | 0.6903 |

Tabell 17: * indikerer at det er autokorrelasjon mellom restledd og laggede restledd i dataene. Det ser dermed ut til at vi ikke har problemer med autokorrelasjon i modell d).

4.6 Predikert avkastning OSEBX

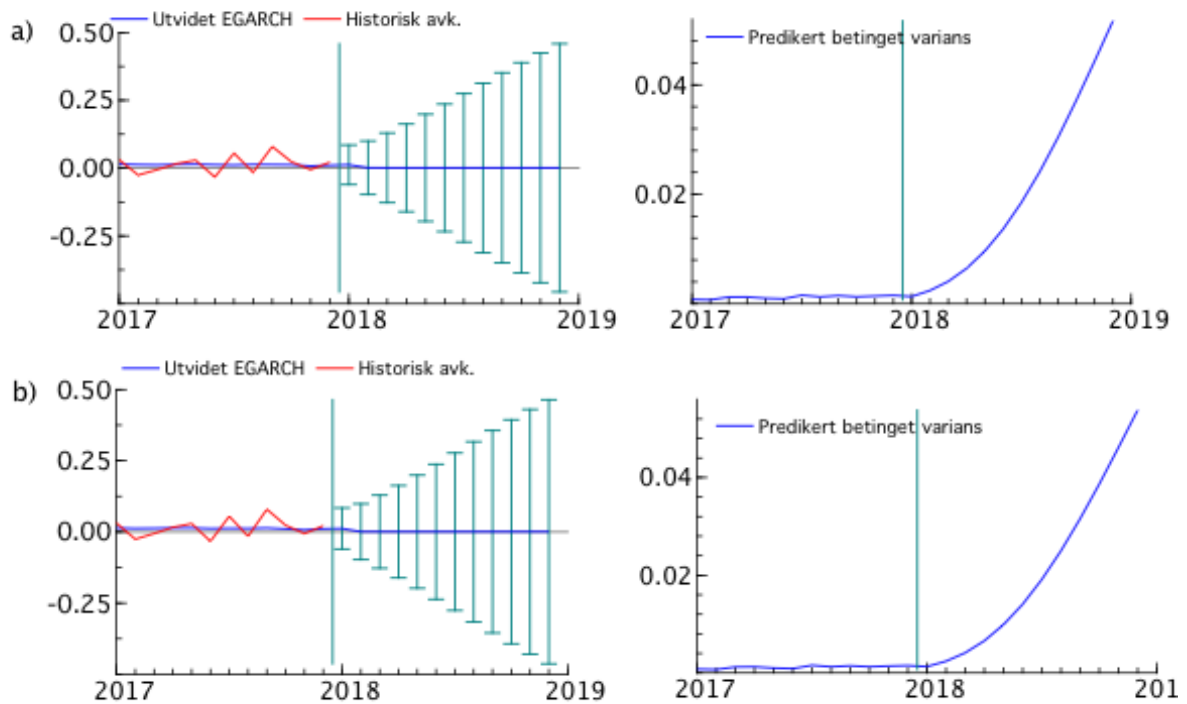
Figurene 7-9 viser predikert avkastning og varians i perioden januar-februar 2018 ved bruk av de fem beste EGARCH(1,1)-modellene fra delkapittel 4.3. Hvis mer nøyaktige verdier er ønskelig, finnes de predikerte verdiene for alle fem modellene i vedlegg C.



Figur 7: Figuren viser historisk og predikert månedlig avkastning og varians til OSEBX for perioden januar 2017 – desember 2018 for en opprinnelig EGARCH(1,1)-modell. Figuren øverst viser historisk avkastning i rødt og predikert avkastning i blått. Den nederste figuren viser predikert varians.

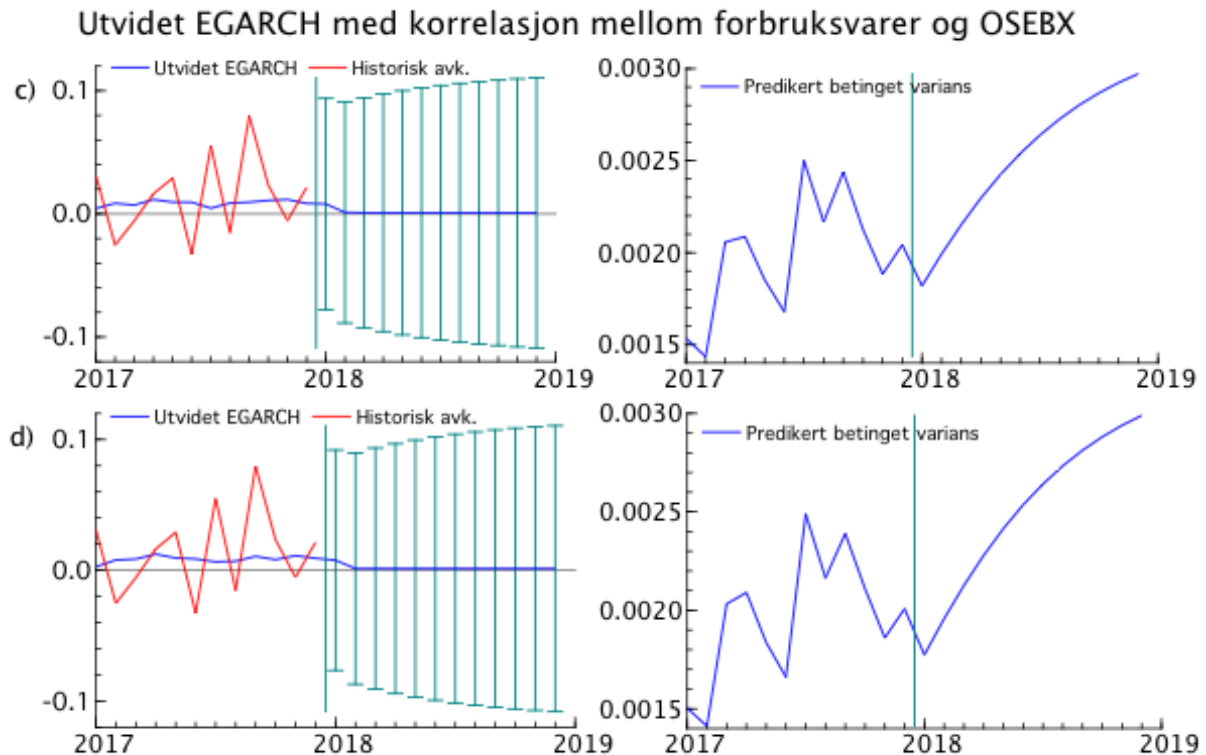
Figur 7 viser at det er predikert positiv månedlig avkastning til OSEBX i hele 2018. Den predikerte avkastningen er konstant i hele perioden og utgjør 1.04%. Det er verdt å legge merke til at usikkerheten til estimatene øker jo lengre frem i tid modellen brukes til å predikere, som illustreres ved at barene blir større og større.

Utvidet EGARCH med korrelasjon mellom finans og OSEBX



Figur 8: Figuren viser historisk og predikert månedlig avkastning og varians til OSEBX for perioden januar 2017 – desember 2018 for to av de beste utvidede EGARCH(1,1)-modellene ifølge informasjonskriteriene fra tabell 10 og 12. Figurene har til felles at de begge er utvidet med korrelasjonen mellom avkastningen til bransjeindeksen finans og OSEBX, både i hoved- og variansligning. Forskjellen mellom figurene er at figur b) er estimert uten lagget avkastning. Figurene til venstre viser historisk avkastning i rødt og predikert avkastning i blått. Figurene til høyre viser predikert varians.

a) i figur 8 viser den utvidede EGARCH(1,1)-modellen med korrelasjon mellom avkastningen til bransjeindeksen finans og OSEBX. Her predikeres OSEBX sin avkastning til 1.24% i januar 2018, med avtakende avkastning til den stabiliserer seg på 0.12% i mars 2018. Det er litt vanskelig å se dette av figuren, men avkastningen er positiv i hele predikeringsperioden. Figur b) viser også en utvidet EGARCH-modell, men her er den laggede avkastningen til OSEBX fjernet. b) viser at det ved fjerning av lagget avkastning fra modellen forandrer ikke påvirker prediksjonene i særlig grad. Ved modellen i figur b) er predikert avkastning estimert til 1.20% i januar 2018, før den synker raskt til 0,07 %. Etter dette forblir predikert avkastning uforandret og vil være 0.07% i perioden februar-desember 2018. Det er også verdt å legge merke til hvordan usikkerheten til estimatene i begge modellene øker jo lengre fram i tid avkastningen predikeres, illustrert ved barene.



Figur 9: Figuren viser historisk og predikert månedlig avkastning og varians for perioden januar 2017 - desember 2018 for de beste utvidede EGARCH-modellene ifølge informasjonskriteriene fra tabell 14 og 16. Figurene har til felles at de begge er utvidet med korrelasjonen mellom forbruksvarer og OSEBX i hovedligningen, men ikke i variansligningen. Forskjellen mellom figurene er at figur d) er estimert uten lagget avkastning. Figurene til venstre viser historisk avkastning i rødt og predikert avkastning i blått. Figurene til høyre viser predikert varians.

c) og d) i figur 9 viser predikert avkastning og predikert varians til OSEBX ved bruk av utvidet EGARCH(1,1) med korrelasjon mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX. For modellen i figur c) er den predikerte avkastningen 0.76% i januar og 0.13% for de resterende månedene. Modellen illustrert ved figur d) predikerer avkastningen til OSEBX til 0.8% i januar, der denne gradvis avtar og stabiliserer seg på 0.08% i juni 2018. I likhet med de forrige prediksjonsfigurene ser vi også av c) og d) i figur 9 at usikkerheten til estimatene øker i takt med økende predikeringshorisont.

4.7 Evaluering av prediksjonsmodellene

Ved evaluering av prediksjonsmodellene brukes RMSE, MAE og informasjonskriteriene. Disse verdiene er vist for den opprinnelige EGARCH(1,1)-modellen, og for de fire beste utvidede EGARCH(1,1)-modellene.

Tabell 22: Mål på prediksjonsnøyaktighet

| | RMSE | MAE | AIC | HQC | SBC |
|-------------|-----------|-------|----------|----------|----------|
| EGARCH(1,1) | 0.010260 | +Inf. | -2.8768 | -2.8442 | -2.7956 |
| a) | 0.0037505 | +Inf. | -2.9380 | -2.8945 | -2.8297 |
| b) | 0.0051802 | +Inf. | -2.9451< | -2.9070< | -2.8503< |
| c) | 0.0024611 | +Inf. | -2.8758 | -2.8377 | -2.7810 |
| d) | 0.0025202 | +Inf. | -2.8826 | -2.8500 | -2.8014 |

Tabell 22: (<) indikerer hvilken modell som favoriseres på bakgrunn av informasjonskriterier.

RMSE favoriserer modell c), da denne gir RMSE-verdi nærmest null. MAE gir samme output uansett hvilken prediksjonsmodell som brukes. MAE kan oppnå verdien uendelig, «+Inf.», hvis $y_t = 0$ i noen perioder. Ved gjennomgang av data ser det ikke ut til at avkastningen til OSEBX er null, verken i historiske data eller predikerte verdier. At MAE likevel kommer med +Inf. kan skyldes at det er perioder med veldig lav avkastning, der denne er tilnærmet lik null. Dette gjør uansett at MAE ikke kan brukes til å velge hvilke modell som gir beste prediksjoner. Ut i fra informasjonskriteriene er modell b) en klar favoritt. Ser dermed at RMSE og informasjonskriteriene favoriserer ulike modeller. Et interessant funn er at ingen av målene i tabell 22 favoriserer den opprinnelige EGARCH(1,1)-modellen. Det ser dermed ut til at bransjeindeksene til finans og forbruksvarer kan være bedre til å predikere avkastning til OSEBX enn en opprinnelig EGARCH(1,1)-modell.

5 Diskusjon

Hovedfunnene i denne oppgaven er at to utvidede EGARCH(1,1)-modeller har bedre prediksjonskraft for OSEBX sin avkastning enn den opprinnelige EGARCH(1,1)-modellen. Utvidelsen av modellen i denne oppgaven utgjøres av korrelasjon mellom den logaritmiske avkastningen til henholdsvis OSEBX og hver enkelt bransjeindeks sin avkastning. Ved utvidelse av modellen estimeres fire varianter av utvidet EGARCH(1,1). Det som skiller de fire utvidede EGARCH(1,1)-modellene er om lagget avkastningsvariabel inkluderes eller ikke i hovedligningen, og om korrelasjonsleddet inngår i variansligningen eller ikke. Vi fant at to bransjer, representert ved to bransjeindekser, har god prediksjonskraft på OSEBX sin avkastning. Disse bransjeindeksene er OSE40GI - finans og OSE25GI – forbruksvarer. Valg av beste modeller ble gjort på bakgrunn av koeffisientenes signifikansnivå, informasjonskriterier og root mean square error (RMSE). I utgangspunktet var mean absolute error (MAE) også tiltenkt å være en del av beslutningsgrunnlaget for hvilken modell som var best. Dessverre gikk prediksjonsverdiene for alle modellene mot $+\infty$, noe som gjorde MAE unyttig i beslutningsgrunnlaget.

Informasjonskriteriene, uttrykt ved AIC, HQC og SBC, ble funnet til å favorisere en utvidet EGARCH(1,1)-modell med korrelasjonsledd i både hovedligning og variansligning, men uten lagget avkastning i hovedligningen. For avkastningen til OSEBX innebærer dette at avkastningen forklares kun på bakgrunn av lagget korrelasjon, der korrelasjonsvariabelen viser til det lineære forholdet mellom avkastningen til OSEBX og avkastningen til bransjeindeksen finans. Variansen derimot beregnes på bakgrunn av lagget restledd, lagget variansledd og lagget korrelasjon. Dette er et interessant og nyttig funn fordi volatilitet, her representert ved varians, er lettere å predikere enn avkastning.

Root mean square error (RMSE) er også et mål på prediksjonsnøyaktigheten til modeller og favoriserer en annen modell enn informasjonskriteriene. Ved bruk av RMSE er den beste modellen en utvidet EGARCH(1,1)-modell med korrelasjon mellom avkastningen til bransjeindeksen forbruksvarer og OSEBX i hovedligningen, men med variansligning lik den opprinnelige EGARCH(1,1)-modellen. Det er altså bare ligningen som estimerer avkastning som blir utvidet. Dette innebærer at korrelasjonen bare har betydning for predikert avkastning, noe som er en svakhet da GARCH-modeller bygger på at volatilitet og avkastning estimeres simultant.

Som nevnt i introduksjonskapittelet står EMH fortsatt sterkt i økonomisk teori, og det er enda ikke blitt introdusert en modell som er tilstrekkelig for å motbevise EMH. Heller ikke i denne oppgaven klarer vi å motbevise EMH, da modellene våre viser for stor grad av usikkerhet. Vi klarer likevel å komme fram til to EGARCH(1,1)-modeller som kan fungere til å predikere avkastning på bakgrunn av korrelasjon, til og med bedre enn en opprinnelig EGARCH(1,1)-modell.

Det viktigste funnet vårt i denne oppgaven er den utvidede EGARCH(1,1)-modellen, hvor avkastningen til OSEBX kun avhenger av lagget korrelasjon, og ikke lagget avkastning. Den tilhørende variansen avhenger av lagget kvadrert restledd, lagget variansledd og lagget korrelasjon. Ved bruk av denne modellen fant vi en signifikant effekt av korrelasjonsleddet på avkastning. Som nevnt inneholder denne modellen korrelasjonen mellom avkastningen til OSEBX og avkastningen til bransjeindeksen finans. Dette tyder på at bransjeindeksen finans kan være den beste bransjen til å forklare utviklingen til OSEBX. Selv om de estimerte modellene våre er usikre, vil denne usikkerheten reduseres ved bruk av kortere prediksjonshorisonter. Ved videre forskning innenfor bransjeindeksene sin effekt på avkastningen til OSEBX, anbefales derfor å prøve kortere prediksjonshorisonter, da dette kan gi mindre usikre estimater og predikerte verdier. Utgangspunktet i denne oppgaven var en predikeringshorisont på 12 måneder. Det kan derfor være interessant å se på kortere prediksjonshorisonter da jo lengre fram i tid man predikerer, jo mer usikkert vil prediksjonsestimatene bli.

Månedlige korrelasjoner er regnet ut ved bruk av daglige avkastninger mellom hver bransjeindeks sin avkastning og OSEBX sin avkastning. Siden EGARCH-modellene i denne oppgaven bruker månedlige korrelasjoner som en forklaringsvariabel, er det derfor den månedlige avkastning til OSEBX som blir predikert. Månedlig avkastning over 22 år utgjør 264 observasjoner, som utgjør få observasjoner ifølge Timmermann og Granger (2004). Videre hevder Timmermann og Granger at avkastningsdata inneholder mye støy og at forklaringskraften til modeller som ønsker å predikere avkastning vil derfor være lave. Som følge av dette kreves det data fra lange tidsperioder for å kunne identifisere og estimere prediksjonsmodeller (Timmermann og Granger, 2004). Det er vanlig å bruke perioder på 5 år med daglige finansielle data ved slike modeller. I denne oppgaven brukes månedlige data for 22 år, totalt 264 observasjoner. Dette er færre enn det som er vanlig, og modellene i denne oppgaven kan derfor fort bli upresise. Dette er en svakhet ved denne oppgaven. Ideelt sett

hadde bransjeindeksene tilhørende på Oslo Børs vært tilbakeregnet til 1980-tallet. Siden bransjeindeksene bare er tilbakeregnet til 29.12.1995, kan det derfor være av interesse å sjekke om man oppnår bedre GARCH-modeller med korrelasjon som forklaringsvariabel ved bruk av ukentlige data. Jo høyere frekvensen på data er, jo mer volatile vil dataene være. Det er derfor å forvente at modellen vil bli mer volatil ved estimering av eksempelvis ukentlige avkastning og mindre volatil ved estimering av årlig avkastning. Bruk av månedlig data kan derfor også være en styrke i denne oppgaven, da volatiliteten i modellen vår vil være lavere enn tilsvarende volatilitet ved bruk av hyppigere frekvens i data, for eksempel ukentlig.

En annen mulig svakhet ved vår modell er at avkastningen til OSEBX inneholder noen ekstreme observasjoner, som modellen vi har utviklet kanskje ikke er i stand til å fange opp. Disse ekstreme observasjonene forekommer i perioden rundt dotcom-krisa på 2000-tallet og finanskrisen i 2008. Det kan derfor være av interesse å innføre dummyvariabler for disse ekstreme variablene slik at man er sikre på at modellen klarer å ta høyde for dem. I denne oppgaven var målet å lage en modell som kunne predikere avkastningen til OSEBX i perioden januar-desember 2018. Målet var altså ikke å lage en generell prediksjonsmodell. Det ble derfor valgt å ikke innføre dummyvariabler for disse ekstremverdiene. I denne oppgaven observeres disse ekstremverdiene til å være negative. En svakhet ved de tradisjonelle GARCH modellene er at de ikke tar høyde for asymmetri i modellen, som oppstår når negative verdier vektlegges i større grad enn positive. Det ble derfor estimert asymmetriske varianter av GARCH, for å fange opp denne effekten, som er en klar styrke i denne oppgaven. De estimerte modellene kunne blitt ytterligere styrket ved inkludering av flere tester på prediksjon. Eksempelvis ved Diebold og Mariano (2002) testen som brukes til å sammenligne prediksjonsverdiene til modeller.

6 Konklusjon

Denne oppgaven viser at korrelasjonen mellom avkastningen til bransjeindeksen forbruksvarer og avkastningen til OSEBX, og korrelasjonen mellom avkastningen til bransjeindeksen for finans og avkastningen til OSEBX kan brukes til å forklare avkastningen til Oslo Børs sin hovedindeks. I introduksjonskapittelet ble det satt et mål om å lage prediksjoner på OSEBX sin avkastning de nærmeste 12 månedene. Dette har de utvidede EGARCH-modellene våre klart å estimere. Derimot var disse avkastningene veldig lave, og i flere måneder er avkastningen predikert til nært null.

Vi kommer fram til at to sektorer, representert ved sine respektive bransjeindekser, forklarer avkastningen til OSEBX bedre enn de resterende åtte bransjeindeksene som inkluderes i analysen. Disse to bransjeindeksene er OSE40GI – finans, og OSE25GI – forbruksvarer. Der det mellom disse to ser ut til at bransjeindeksen finans fanger opp effekten av lagget korrelasjon på avkastning, bedre enn bransjeindeksen for forbruksvarer.

Denne oppgaven har altså funnet to utvidede EGARCH-modeller som kan bidra til å forklare utviklingen til OSEBX. Informasjonskriteriene, AIC, HQC og SBC favoriserer den utvidede EGARCH-modellen som inkluderer bransjeindeksen finans, og ikke lagget avkastning. Denne utvidede EGARCH-modellen predikerer avkastning på bakgrunn av lagget korrelasjon mellom avkastningen til bransjeindeksen finans og avkastningen til OSEBX. Root mean square error (RMSE) favoriserer på sin side en utvidet EGARCH-modell, der forklaringsvariabelen er lagget korrelasjon mellom avkastning til bransjeindeksen forbruksvarer og avkastningen til OSEBX, i tillegg til lagget avkastning.

Referanseliste

- Akaike, H. (1974) A new look at the statistical model identification, *IEEE transactions on automatic control*, 19(6), s. 716-723.
- Akay, A. (2012) Finite-sample comparison of alternative methods for estimating dynamic panel data models, *Journal of Applied Econometrics*, 27(7), s. 1189-1204.
- Andersen, T. G. og Bollerslev, T. (1997) Intraday periodicity and volatility persistence in financial markets, *Journal of empirical finance*, 4(2-3), s. 115-158.
- Andersen, T. G. *et al.* (2003) Modeling and forecasting realized volatility, *Econometrica*, 71(2), s. 579-625.
- Arowolo, W. (2013) Predicting Stock Prices Returns Using Garch Model, *The International Journal of Engineering and Science*, 2(5), s. 32-37.
- Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity, *Journal of econometrics*, 31(3), s. 307-327.
- Bollerslev, T., Chou, R. Y. og Kroner, K. F. (1992) ARCH modeling in finance: A review of the theory and empirical evidence, *Journal of econometrics*, 52(1-2), s. 5-59.
- Bouchaud, J.-P., Matacz, A. og Potters, M. (2001) Leverage effect in financial markets: The retarded volatility model, *Physical review letters*, 87(22), s. 228701.
- Box, G. E. og Pierce, D. A. (1970) Distribution of residual autocorrelations in autoregressive-integrated moving average time series models, *Journal of the American statistical association*, 65(332), s. 1509-1526.
- Brooks, C. (2008) *Introductory Econometrics for Finance*. 2. utg. Cambridge: Cambridge University Press.
- Dickey, D. A. og Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American statistical association*, 74(366a), s. 427-431.
- Diebold, F. X. og Mariano, R. S. (2002) Comparing predictive accuracy, *Journal of Business & Economic Statistics*, 20(1), s. 134-144.
- Eiendom (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE60GI.OSE/overview> (Hentet: 18.02.2018).
- Enders, W. (2015) *Applied econometric time series*. 4. utg. Ed. Chichester, West Sussex: John Wiley & Sons.

Energi (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSE10GL.OSE/overview> (Hentet: 18.02.2018).

Engle, R. (2001) GARCH 101: The use of ARCH/GARCH models in applied econometrics, *Journal of economic perspectives*, 15(4), s. 157-168.

Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica: Journal of the Econometric Society*, s. 987-1007.

Finans (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSE40GL.OSE/overview> (Hentet: 18.02.2018).

Forbruksvarer (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSE25GL.OSE/overview> (Hentet: 18.02.2018).

Forsyning (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSE55GL.OSE/overview> (Hentet: 18.02.2018).

Glosten, L. R., Jagannathan, R. og Runkle, D. E. (1993) On the relation between the expected value and the volatility of the nominal excess return on stocks, *The journal of finance*, 48(5), s. 1779-1801.

Granger, C. W. og Jeon, Y. (2004) Thick modeling, *Economic Modelling*, 21(2), s. 323-343.

Hannan, E. J. og Quinn, B. G. (1979) The determination of the order of an autoregression, *Journal of the Royal Statistical Society. Series B (Methodological)*, s. 190-195.

Hansen, P. R. og Huang, Z. (2016) Exponential GARCH modeling with realized measures of volatility, *Journal of Business & Economic Statistics*, 34(2), s. 269-287.

Hauke, J. og Kossowski, T. (2011) Comparison of values of Pearson's and Spearman's correlation coefficients on the same sets of data, *Quaestiones geographicae*, 30(2), s. 87-93.

Helsevern (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSE35GL.OSE/overview> (Hentet: 18.02.2018).

Hovedindeksen (2017). Tilgjengelig fra:

<https://www.oslobors.no/markedsaktivitet/#/details/OSEBX.OSE/overview> (Hentet: 18.02.2018).

- Hull, J. C. og Basu, S. (2016) *Options, futures, and other derivatives*. Pearson Education India.
- Industri* (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE20GL.OSE/overview> (Hentet: 18.02.2018).
- IT* (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE45GL.OSE/overview> (Hentet: 18.02.2018).
- Konsumvarer* (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE30GL.OSE/overview> (Hentet: 18.02.2018).
- Ljung, G. M. og Box, G. E. (1978) On a measure of lack of fit in time series models, *Biometrika*, 65(2), s. 297-303.
- Mandelbrot, B. B. (1963) The variation of certain speculative prices, *Journal of Business*, (XXXVI), s. 225-236.
- Materialer* (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE15GL.OSE/overview> (Hentet: 18.02.2018).
- Nelson, D. B. (1991) Conditional heteroskedasticity in asset returns: A new approach, *Econometrica: Journal of the Econometric Society*, s. 347-370.
- Niu, H. og Wang, J. (2013) Volatility clustering and long memory of financial time series and financial price model, *Digital Signal Processing*, 23(2), s. 489-498.
- Ratner, B. (2013) The correlation coefficient: Definition, *DM Stat-1 Articles*.
- Ross, G. J. (2013) Modelling financial volatility in the presence of abrupt changes, *Physica A: Statistical Mechanics and its Applications*, 392(2), s. 350-360.
- Schwarz, G. (1978) Estimating the dimension of a model, *The annals of statistics*, 6(2), s. 461-464.
- Telekom* (2017). Tilgjengelig fra:
<https://www.oslobors.no/markedsaktivitet/#/details/OSE50GL.OSE/overview> (Hentet: 18.02.2018).
- Timmermann, A. og Granger, C. W. (2004) Efficient market hypothesis and forecasting, *International Journal of forecasting*, 20(1).
- Țițan, A. G. (2015) The Efficient Market Hypothesis: review of specialized literature and empirical research, *Procedia Economics and Finance*, 32, s. 442-449.

Vedlegg/Appendiks

Vedlegg A: Sammensetningen av OSEBX

Vedlegg A: Viser sammensetningen av de 67 aksjene som inngår i hovedindeksen per 28.05.2018 med vektet proSENTSATS.

Kilde: Oslo Børs

| Selskap | Vektet proSENTSATS av OSEBX |
|---------------------------------|------------------------------------|
| Equinor | 28.63% |
| DNB | 10.30% |
| Telenor | 10.28% |
| Aker BP | 4.46% |
| Norsk Hydro | 4.27% |
| Yara International | 3.93% |
| Marine Harvest | 3.39% |
| Orkla | 3.17% |
| Gjensidige Forsikring | 2.63% |
| Aker | 1.82% |
| SalMar | 1.67% |
| Subsea 7 | 1.66% |
| Lerøy Seafood Group | 1.35% |
| Storebrand | 1.34% |
| Schibsted ser. B | 1.17% |
| TGS – NOPEC Geophysical Company | 1.14% |
| Tomra Systems | 1.09% |
| Schibsted ser. A | 1.04% |
| Entra | 0.88% |
| Bakkafrost | 0.87% |
| Borr Drilling | 0.87% |
| Sparebank 1 SR – Bank | 0.86% |
| Kongsberg Gruppen | 0.86% |
| Wallenius Wilhelmsen Logistics | 0.74% |
| Norwegian Finans Holding | 0.68% |
| Olav Thon Eiendomsselskap | 0.66% |
| DNO | 0.65% |
| Aker Solutions | 0.60% |
| AF Gruppen | 0.56% |
| Petroleum Geo-Services | 0.55% |
| Atea | 0.54% |
| Veidekke | 0.52% |
| Norwegian Air Shuttle | 0.48% |
| EVRY | 0.44% |
| XXL | 0.43% |
| Grieg Seafood | 0.42% |
| Golden Ocean Group | 0.41% |
| Nordic Semiconductor | 0.41% |
| B2Holding | 0.31% |
| Wilh. Wilhelmsen Holding ser. A | 0.30% |

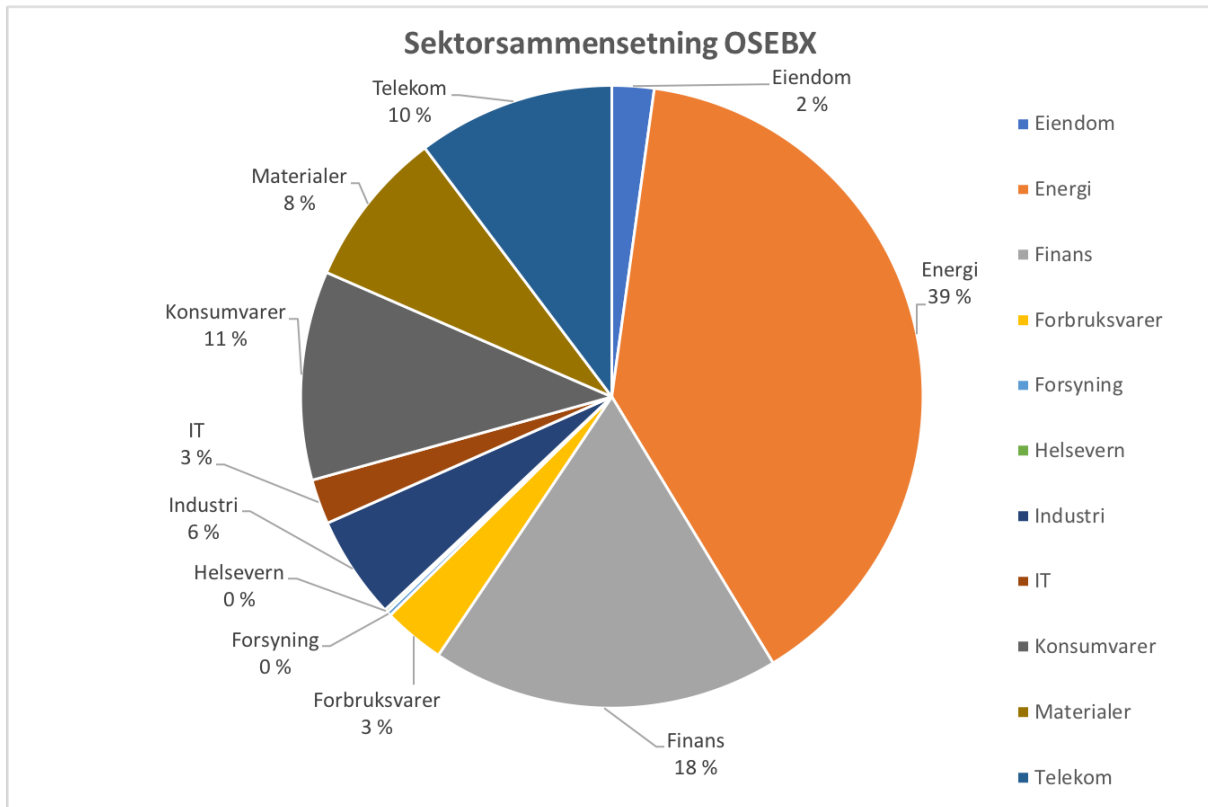
Vedlegg A fortsetter på neste side

Vedlegg A – Fortsettelse fra forrige side

| Selskap | Vektet prosentsats av OSEBX |
|------------------------------|------------------------------------|
| Frontline | 0.29% |
| Stolt Nielsen | 0.25% |
| Norwegian Property | 0.24% |
| Scatec Solar | 0.23% |
| Ekornes | 0.21% |
| Gaming Innovation Group | 0.20% |
| Europris | 0.18% |
| Kongsberg Automotive | 0.18% |
| BW LPG | 0.18% |
| Axactor | 0.18% |
| Hexagon Composites | 0.15% |
| REC Silicon | 0.14% |
| Treasure | 0.14% |
| Otello Corporation | 0.11% |
| Wilh. Wilhelmsen ser. B | 0.10% |
| Link Mobbility Group | 0.10% |
| Nordic Nanovector | 0.10% |
| Questerre Energy Corporation | 0.09% |
| Thin Film Electronics | 0.09% |
| Asetek | 0.08% |
| Seadrill | 0.07% |
| Kitron | 0.07% |
| IDEX | 0.07% |
| Funcom | 0.07% |
| Targovax | 0.04% |
| NEXT Biometrics Group | 0.03% |
| Photocure | 0.02% |
| Totalt | 100% |

Vedlegg B: Sektorsammensetning OSEBX

Figuren viser hvor stor del av hovedindeksen de enkelte sektorene utgjør.



Vedlegg C: Predikert avkastning OSEBX i tallverdi

Vedlegg C.1: Tilhørende prediksjoner for opprinnelig EGARCH(1,1) januar-februar 2018

| Horisont | Prediksjon | Standard Error | Betinget varians |
|----------|------------|----------------|------------------|
| 1 | 0.010425 | 0.042319 | 0.0017773 |
| 2 | 0.010248 | 0.044290 | 0.0019463 |
| 3 | 0.010245 | 0.046041 | 0.0021032 |
| 4 | 0.010245 | 0.047592 | 0.0022472 |
| 5 | 0.010245 | 0.048957 | 0.0023780 |
| 6 | 0.010245 | 0.050154 | 0.0024958 |
| 7 | 0.010245 | 0.051200 | 0.0026009 |
| 8 | 0.010245 | 0.052110 | 0.0026943 |
| 9 | 0.010245 | 0.052901 | 0.0027766 |
| 10 | 0.010245 | 0.053585 | 0.0028489 |
| 11 | 0.010245 | 0.054177 | 0.0029122 |
| 12 | 0.010245 | 0.054688 | 0.0029673 |

Vedlegg C2: Tilhørende prediksjoner for figur a) i perioden januar-februar 2018

| Horisont | Prediksjon | Standard Error | Betinget varians |
|----------|------------|----------------|------------------|
| 1 | 0.012387 | 0.036485 | 0.001316 |
| 2 | 0.0014365 | 0.049353 | 0.0024072 |
| 3 | 0.0011592 | 0.064217 | 0.0040754 |
| 4 | 0.0011522 | 0.080787 | 0.0064498 |
| 5 | 0.0011520 | 0.098690 | 0.0096249 |
| 6 | 0.0011520 | 0.11751 | 0.013645 |
| 7 | 0.0011520 | 0.13682 | 0.018499 |
| 8 | 0.0011520 | 0.15624 | 0.024121 |
| 9 | 0.0011520 | 0.17540 | 0.030401 |
| 10 | 0.0011520 | 0.19402 | 0.037197 |
| 11 | 0.0011520 | 0.21186 | 0.044352 |
| 12 | 0.0011520 | 0.22875 | 0.051705 |

Vedlegg C3: Tilhørende prediksjoner for figur b) i perioden januar-februar 2018

| Horisont | Prediksjon | Standard Error | Betinget varians |
|----------|------------|----------------|------------------|
| 1 | 0.012111 | 0.036049 | 0.0012995 |
| 2 | 0.00070392 | 0.049086 | 0.0024094 |
| 3 | 0.00070392 | 0.064208 | 0.0041227 |
| 4 | 0.00070392 | 0.081109 | 0.0065787 |
| 5 | 0.00070392 | 0.099394 | 0.0098791 |
| 6 | 0.00070392 | 0.11862 | 0.014072 |
| 7 | 0.00070392 | 0.13836 | 0.019143 |
| 8 | 0.00070392 | 0.15818 | 0.025020 |
| 9 | 0.00070392 | 0.17772 | 0.031584 |
| 10 | 0.00070392 | 0.19667 | 0.038680 |
| 11 | 0.00070392 | 0.21480 | 0.046140 |
| 12 | 0.00070392 | 0.23193 | 0.053790 |

Vedlegg C4: Tilhørende prediksjoner for figur c) i perioden januar-februar 2018

| Horisont | Prediksjon | Standard Error | Betinget varians |
|----------|------------|----------------|------------------|
| 1 | 0.0080290 | 0.042889 | 0.0018186 |
| 2 | 0.0011227 | 0.044928 | 0.0019926 |
| 3 | 0.00084465 | 0.046691 | 0.0021520 |
| 4 | 0.00083345 | 0.048229 | 0.0022961 |
| 5 | 0.00083300 | 0.049562 | 0.0024248 |
| 6 | 0.00083298 | 0.050714 | 0.0025387 |
| 7 | 0.00083298 | 0.051704 | 0.0026388 |
| 8 | 0.00083298 | 0.052552 | 0.0027261 |
| 9 | 0.00083298 | 0.053277 | 0.0028018 |
| 10 | 0.00083298 | 0.053895 | 0.0028672 |
| 11 | 0.00083298 | 0.054421 | 0.0029234 |
| 12 | 0.00083298 | 0.054868 | 0.0029715 |

Vedlegg C5: Tilhørende prediksjoner for figur d) i perioden januar-februar 2018

| Horisont | Prediksjon | Standard Error | Betinget varians |
|----------|------------|----------------|------------------|
| 1 | 0.0076482 | 0.042115 | 0.0017737 |
| 2 | 0.0012693 | 0.044245 | 0.0019576 |
| 3 | 0.0012693 | 0.046112 | 0.0021263 |
| 4 | 0.0012693 | 0.047737 | 0.0022788 |
| 5 | 0.0012693 | 0.049144 | 0.0024151 |
| 6 | 0.0012693 | 0.050354 | 0.0025355 |
| 7 | 0.0012693 | 0.051391 | 0.0026411 |
| 8 | 0.0012693 | 0.052277 | 0.0027329 |
| 9 | 0.0012693 | 0.053031 | 0.0028123 |
| 10 | 0.0012693 | 0.053672 | 0.0028806 |
| 11 | 0.0012693 | 0.054214 | 0.0029392 |
| 12 | 0.0012693 | 0.054673 | 0.0029892 |

Vedlegg D: Begreper

- Skjevhet

$$\hat{S}_r = \frac{1}{T-1} \sum_{t=1}^T \frac{(r_t - E(r_t))^3}{\hat{\sigma}^3}$$

Skjevhet måler avviket fra symmetri. Vi sier at en estimator er forventningsrett hvis forventningsverdien, representert ved $E(r_t)$, er lik den virkelige verdien, r_t . Hvis derimot forventningsverdien er ulik den virkelige verdien vil estimatoren være forventningsskjev.

- Kurtose

$$\widehat{K}_r = \frac{1}{T-1} \sum_{t=1}^T \frac{(r_t - E(r_t))^4}{\hat{\sigma}^4}$$

Kurtosen måler hvor fete halene til restleddets fordeling er.