

Kva påverkar straumprisen?

Ei empirisk analyse av Nord Pool

Henning Stokke

Rettleiar: Professor Gunnar Stensland

Masteroppgåve i finansiell økonomi (FIE)

NORGES HANDELSHØYSKOLE

Dette sjølvstendige arbeidet er gjennomført som eit ledd i siviløkonomutdanninga ved Norges Handelshøyskole (NHH) og er godkjent som såleis. Godkjenninga inneber ikkje at NHH står inne for dei metodar som er brukt, dei resultat som er framkome eller dei konklusjonar som er trekte i arbeidet.

Samandrag

Denne empiriske analysen ser på kva for faktorar som har påverka spotprisdanninga på den nordiske kraftbørsen Nord Pool frå 1995 til 2011. Ved hjelp av økonomiske teknikkar blir det gjennom sju regresjonsmodellar inngåande studert på kva måte faktorar som fyllingsgrad, turnover (omsetting), temperatur og oljepris påverkar spotprisen på Nord Pool.

Hovudfunnet er at fundamentale hydrologiske faktorar påverkar spotprisen. Mellom anna har det vist seg i denne tidsperioden at om temperaturen aukar, så vil spotprisen bli redusert. Fyllingsgrad i vassmagasina viser seg også å påverke på same vis. Når fyllingsgrada aukar, så vil spotprisen bli redusert. Effekten er sterkast når det er lav kapasitet i vassmagasina. Meir konkret viser denne utredninga at om fyllingsgrad aukar med 10 % ved låg kapasitet i vassmagasina, så blir spotprisen i gjennomsnitt redusert med cirka 7 øre/kWh. Likeeins finn eg at om fyllingsgrad aukar med 10 % ved høg kapasitet i vassmagasina, så blir spotprisen i gjennomsnitt redusert med cirka 6 øre/kWh.

I tillegg kan det sjå ut som at nokre andre faktorar er med på å påverke spotprisdanninga. Oljeprisen viser seg i mindre grad å påverke spotprisen. Her går påverknaden derimot i motsatt retning i forhold til dei andre variablane. Det vil seie at når oljeprisen aukar med 10 US dollar pr fat, så vil spotprisen i gjennomsnitt auke med cirka 1 øre/kWh.

Ein skal likevel vere forsiktig med å tolke desse koeffisientane veldig pedantisk då det heftar stor usikkerheit ved dei, og i og med at dei er svært enkle å justere ved å legge til eller fjerne enkelte variablar i modellen.

Forord

Dette sjølvstendige arbeidet utgjer 30 studiepoeng og er siste del av masterstudiet i finansiell økonomi ved Norges Handelshøyskole. For meg personleg er det også ein viktig milestolpe i den forstand at det markerar slutten på fem år med studiar.

Grunnen til at eg har valt å skrive om Nord Pool er fordi eg alltid har vore svært interessert i energibransjen og spesielt fornybar energi. Slike studiar er ofte av ein samfunnsøkonomisk karakter, så for å finne ein meir finansretta innfallsvinkel vart ei analyse av Nord Pool valt. Og etter å ha studert denne kraftbørsen opplever eg at den i stor grad fungerer på same måte som ein kva som helst annan finansiell marknadsplass, sjølv om det er straum og ikkje aksjar som blir omsett.

Ein veikskap ved oppgåva er omfanget av data. Det hadde styrka oppgåva å ha inkludert fleire ulike variablane, og då spesielt temperaturdata frå dei andre nordiske landa. I tillegg hadde det vore ein fordel å ha inkludert variablar med enno lengre tidsseriar.

Først og fremst vil eg takke rettleiaren min, Gunnar Stensland, for god rettleiing, nyttige faglege diskusjonar og innspel. Eg vil også rette ei stor takk til Nord Pool Spot for tilgang til data gjennom deira FTP-serverar, og til Meteorologisk institutt for tilgang til klimadata gjennom databasen eKlima.

Eg vil også takke vennar og familie for god hjelp og støtte i samband med korrekturlesing.

Bergen, 14. juni 2012



Henning Stokke

Innhold

| | |
|--|-----------|
| SAMANDRAG | 2 |
| FORORD | 3 |
| 1. INNLEIING | 7 |
| 1.1 NORD POOL..... | 7 |
| 1.1.1 <i>Historie</i> | 8 |
| 1.1.2 <i>Den finansielle marknaden</i> | 9 |
| 1.1.3 <i>Den fysiske marknaden</i> | 9 |
| 1.2 NORD POOL-OMRÅDET..... | 10 |
| 1.2.1 <i>Elektrisitetsproduksjon</i> | 10 |
| 1.2.2 <i>Overføringsnett</i> | 11 |
| 1.3 TIDLEGARE FORSKING | 13 |
| 2. METODE | 17 |
| 2.1 KORRELASJON..... | 17 |
| 2.2 REGRESJON | 17 |
| 2.3 FORKLARINGSGRAD | 17 |
| 2.4 ENDOGENITET | 18 |
| 2.5 CHOW TEST | 18 |
| 2.6 RESET TEST | 19 |
| 2.7 HETEROSKEDASTISITET..... | 19 |
| 2.8 AUTOKORRELASJON..... | 20 |
| 2.9 AUTOKORRELASJONSFUNKSJON (ACF)..... | 21 |
| 2.10 PARTIELL AUTOKORRELASJONSFUNKSJON (PACF)..... | 21 |
| 2.11 INFORMASJONSKRITERIUM..... | 21 |
| 2.12 STASJONÆRITET | 22 |
| 2.13 BOX PLOT | 22 |
| 2.14 HP-FILTER..... | 23 |
| 3. SKILDRING AV DATA | 25 |
| 3.1 SYSTEMPRISEN | 25 |

| | | |
|-----------|-----------------------------------|-----------|
| 3.1.1 | <i>Skewness og kurtosis</i> | 28 |
| 3.1.2 | <i>Volatilitet</i> | 29 |
| 3.1.3 | <i>Autokorrelasjon</i> | 31 |
| 3.1.4 | <i>Stasjonærhet</i> | 32 |
| 3.2 | TURNOVER..... | 33 |
| 3.3 | TEMPERATUR..... | 34 |
| 3.4 | FYLLINGSGRAD..... | 35 |
| 3.5 | OLJEPRIS..... | 38 |
| 4. | EMPIRISK ANALYSE | 39 |
| 4.1 | MODELLAR..... | 39 |
| 4.1.1 | <i>Korrelasjon</i> | 40 |
| 4.1.2 | <i>Regresjonsmodell 1</i> | 41 |
| 4.1.3 | <i>Regresjonsmodell 2</i> | 42 |
| 4.1.4 | <i>Regresjonsmodell 3</i> | 43 |
| 4.1.5 | <i>Regresjonsmodell 4</i> | 45 |
| 4.1.6 | <i>Regresjonsmodell 5</i> | 47 |
| 4.1.7 | <i>Regresjonsmodell 6</i> | 49 |
| 4.1.8 | <i>Regresjonsmodell 7</i> | 52 |
| 4.1.9 | <i>Oppsummering</i> | 53 |
| 4.2 | REGRESJONSVALIDITET..... | 54 |
| 4.2.1 | <i>Heteroskedastisitet</i> | 54 |
| 4.2.2 | <i>Chow test</i> | 56 |
| 4.2.3 | <i>RESET test</i> | 58 |
| 5. | KONKLUSJON | 59 |
| 6. | REFERANSAR | 61 |

Tabelliste:

| | |
|---|----|
| <i>Tabell 1: Elektrisitetsproduksjon i Norden etter energikjelder</i> | 10 |
| <i>Tabell 2: Elektrisitetsproduksjon i Norden i TWh</i> | 11 |
| <i>Tabell 3: Deskriptiv statistikk</i> | 25 |
| <i>Tabell 4: Forklaring</i> | 40 |
| <i>Tabell 5: Korrelasjonsmatrise</i> | 40 |
| <i>Tabell 6: Resultat frå regresjon av modell 1</i> | 41 |
| <i>Tabell 7: Resultat frå regresjon av modell 2</i> | 42 |
| <i>Tabell 8: Resultat frå regresjon av modell 3</i> | 43 |
| <i>Tabell 9: Resultat frå regresjon av modell 4</i> | 46 |
| <i>Tabell 10: Resultat frå regresjon av modell 5</i> | 48 |
| <i>Tabell 11: Resultat frå regresjon av modell 6</i> | 49 |
| <i>Tabell 12: Ulike informasjonskriterium</i> | 51 |
| <i>Tabell 13: Resultat frå regresjon av modell 7</i> | 52 |
| <i>Tabell 14: Resultat frå Chow-testen</i> | 56 |

Figurliste:

| | |
|---|----|
| <i>Figur 1: Overføringsnett</i> | 12 |
| <i>Figur 2: Døgnpris</i> | 26 |
| <i>Figur 3: Vekepris</i> | 26 |
| <i>Figur 4: Fordeling av observasjonar (%)</i> | 28 |
| <i>Figur 5: Volatilitet</i> | 30 |
| <i>Figur 6: ACF</i> | 31 |
| <i>Figur 7: Omsetting</i> | 33 |
| <i>Figur 8: Temperatur i grader celsius</i> | 34 |
| <i>Figur 9: Fyllingsgrad</i> | 35 |
| <i>Figur 10: Fyllingsgrad over og under gjennomsnitt (gjennomsnitt = 0)</i> | 36 |
| <i>Figur 11: Faktisk fyllingsgrad og gjennomsnittleg fyllingsgrad</i> | 37 |
| <i>Figur 12: Oljepris (\$) pr fat</i> | 38 |
| <i>Figur 13: Box plot av systemprisen</i> | 44 |
| <i>Figur 14: Modelltilpassing</i> | 47 |
| <i>Figur 15: Modelltilpassing</i> | 50 |
| <i>Figur 16: Partiell autokorrelasjonsfunksjon</i> | 51 |
| <i>Figur 17: Plott av feilledd</i> | 55 |
| <i>Figur 18: Feilledd frå regresjon av dei 444 første observasjonane</i> | 57 |
| <i>Figur 19: Feilledd frå regresjon av dei 444 siste observasjonane</i> | 57 |

1. Innleiing

Samanlikna med ei rekkje andre land har elektrisitetsprisane i Noreg vore relativt låge. Dette er truleg ein av grunnane til at vi forbrukar svært mykje straum, i tillegg til at det er relativt kaldt i dette landet. Men kva avgjer egentleg denne prisen?

Denne masterutredninga tek for seg den nordiske kraftbørsen, Nord Pool, og søkjer å finne ut kva for faktorar som påverkar spotprisen og korleis denne blir danna.¹ Spesielt vil eg sjå på korleis hydrologiske forhold spelar inn i denne prisdanninga. Utredninga vil ta form som ei empirisk analyse og vil i stor grad vere deskriptivt orientert.

Eg vil inngåande studere utviklinga sidan Nord Pool blei etablert, og deretter forsøke å trekke dei store linjene og å sjå samanhenger ved hjelp av økonomiske teknikkar. Dette vil kunne gi ei auka evne til å kunne svare på problemstillinga om kva som påverkar straumprisen. Formålet med å vinne denne innsikta er at ein då vil vere betre i stand til å skode framover, og i så måte få ei betre forståing for den framtidige pris- og marknadsutviklinga på Nord Pool.

Utredninga er bygt opp på følgjande vis. I første del vil eg sjå nærmare på Nord Pool og korleis denne marknadsplassen operer. I andre del vil eg presentere relevant teori og metodar. I tredje del vil eg skildre datamaterialet. I fjerde del vil den empiriske analysen bli presentert og i femte og siste del vil oppsummering og konklusjon bli lagt fram.

1.1 Nord Pool

Den nordiske energibørsen har gradvis vakse fram og er i dag ein svært viktig marknadsplass for aktørar i energibransjen. For betre å forstå Nord Pool og korleis denne marknadsplassen har utvikla seg er det fruktbart først å sjå på historia. Deretter vil eg gå inn på korleis marknaden fungerer sett ut i frå ein praktisk og geografisk ståstad, før eg ser på eit utval av kva som er gjort av tidlegare forskning på området.

¹ Spotprisen blir ofte også omtala som systemprisen og desse to nemningane vil vere synonyme i denne oppgåva.

1.1.1 Historie

I 1992 vart Statnett Marked offisielt opna som eit resultat av tanken om gradvis deregulering av elektrisitetsmarknaden som opna opp for i den nye energilova frå 1991. I starten var dette ein medlemsbasert marknadsplass for kjøp og sal av straum for dei største aktørane i marknaden. 1. januar 1993 vart det endra til at det ikkje lenger var nødvendig å vere medlem for å delta, noko som opna vegen for dei mindre aktørane.

1996 blei Statnett Marked utvida til også å omfatte Sverige. Eigarskapen blei då endra til at Statnett eigde 50 % og Svenska Kraftnät eigde 50 %. Namnet Nord Pool ASA, den nordiske kraftbørsen, tok også over. I 1998 blei Finland inkludert, og i 2000 blei Danmark med. Vidare vart spotmarknaden, eller den fysiske marknaden, skilt ut som eit eige selskap i 2002 og det fekk namnet Nord Pool Spot AS. I 2005 vart auksjonsområdet Kontek i Tyskland opna, og eit år seinare opna Nord Pool Spot kontor i Berlin.

Segmenteringa frå 2002 utvikla seg vidare og i 2008 vart den finansielle delen av Nord Pool ASA seld til NASDAQ OMX Commodities. I 2010 kjøpte NASDAQ OMX Commodities resten av Nord Pool ASA og endra namnet til NASDAQ OMX Commodities Europe. Nord Pool Spot, som tek for seg den fysiske handelen i straum og som er hovudtemaet i denne oppgåva, vart ikkje råka av denne endringa og driv si verksemd uavhengig av NASDAQ OMX Commodities Europe.

I 2010 blei det opna eit eige auksjonsområde i Estland, og Nord Pool Spot inngjekk ein avtale om å levere dei tekniske løysingane for den litauiske marknaden. I dag er Nord Pool Spot eigd av Statnett SF (30 %), Svenska Kraftnät (30 %), Fingrid Oyj (20 %) og Energinet.dk (20 %). 350 selskap frå 18 land handlar på marknaden og i 2011 vart det omsatt 316 TWh straum (Nord Pool Spot, 2011).²

I og med at eigarstrukturen er noko kompleks og har vore stadig i endring dei siste åra blir ofte Nord Pool litt upresist brukt som ein felles term for alle operasjonelle aktivitetar på den nordiske energimarknaden.³ Vidare plar mange aktørar gjerne å skilje mellom den finansielle

² Wattime (Wh) er ei vanleg måleining for energi og straumforbruk. I denne oppgåva bruker eg denne i ulike størrelsar. Ein TWh (terawattime) er det same som 1000 GWh (gigawattime), ein GWh er det same som 1000 MWh (megawattime), og ein MWh er det same som 1000 kWh (kilowattime).

³ Når eg vidare i denne oppgåva omtalar Nord Pool er det meir presist Nord Pool Spot som blir vist til.

marknaden (NASDAQ OMX Commodities Europe) og den fysiske marknaden (Nord Pool Spot). Fordi denne inndelinga er både intuitiv og mykje brukt vil eg halde meg til den i dei påfølgande avsnitta.

1.1.2 Den finansielle marknaden

Den finansielle marknaden er plassen for derivathandelen. Denne marknadsplassen er strukturert slik at ein har ein del som går under namnet Eltermin, der det blir det handla futures- og forwardkontraktar, og ein del som går under namnet Elopsjon, der det blir handla opsjonar og andre finansielle instrument. Den finansielle marknaden tilbyr også oppgjerstenestar (clearing) for dei ulike aktørane. Oppgjer vil seie at NASDAQ OMX Commodities Europe tek på seg risikoen knytt til handelen og garanterar at aktørane får sitt oppgjer sjølv om ein av partane ikkje klarer å møte sin skyldnad.

Eit anna relativt nytt produkt som har kome på den finansielle marknaden, er Contracts of difference. Det har vakse fram i samband med utviklinga av fleire ulike prisområde i Nord Pool-området. Dette finansielle instrumentet er også eit hjelpemiddel for større aktørar knytt til risikostyring. Ved å inngå ei slik kontrakt kan ein sikre seg mot eventuelle prisdifferansar mellom systemprisen og den aktuelle områdeprisen.

1.1.3 Den fysiske marknaden

Den fysiske marknaden er, som den finansielle marknaden, også delt i to med Elspot og Elbas. Elspot er namnet på marknaden til spotprisen pr time. Her blir 74 % av straumkonsumet i dei nordiske landa handla. På Elspot blir straumprisen for kvar einaste time det påfølgande døgnet notert kvar dag cirka klokka 13. Før klokka 12 er alle aktørane nøydt til å melde inn kor mykje dei planlegg å kjøpe og selje av straum kvar time det neste døgnet. Basert på aggregert tilbod og aggregert etterspørsel reknar Nord Pool ut likevektsprisen, også kalla systemprisen. Det spesielle med den er at systemprisen ikkje tek omsyn til flaskehalsar og kapasiteten i overføringsnett. Dette problemet har Nord Pool løyst ved å dele inn i områdeprisar, og å la differansen i områdeprisane og systemprisen reflektere overskot eller manglande kapasitet i linjenettet.

Det er vanskeleg for straumselskapa å vite nøyaktig kor mykje straum ein skal produsere i det neste døgnet for å møte konsumbehovet. Straum er ei råvare (commodity) som ikkje kan

lagrast, og følgjeleg må all straum som er produsert også forbrukast, anten her eller i utlandet. Det eksisterar difor ein marknad i tillegg til Elspot. Denne blir kalla Elbas. Dette er den heilt kortsiktige produksjons- og forbruksplanlegginga, og her blir spotprisen for kvar klokke-time handla opp til ein time før levering. Deretter blir siste steg i pris- og kapasitetsvurderingane før levering gjort av dei respektive nasjonale linjeeigarane (TSO). Her blir tilbod og etterspørsel balansert og vidare blir prisen satt etter levering frå TSO (Nord Pool Spot, 2011).

1.2 Nord Pool-området

Nord Pool-området omfattar Norden.⁴ I denne delen vil det bli sett nærmare på kva typar energikjelder som er dominerande i Nord Pool-området og vidare korleis dette blir overført mellom dei ulike landa.

1.2.1 Elektrisitetsproduksjon

| Elektrisitetsproduksjon i Norden etter energikjelder | | | | |
|--|----------|--------------------|--------------|-----------------|
| | Termiske | Fossile brennstoff | Hydrologiske | Andre fornybare |
| Noreg | 0 % | 4 % | 95 % | 1 % |
| Sverige | 38 % | 5 % | 46 % | 11 % |
| Danmark | 0 % | 72 % | 0 % | 28 % |
| Finland | 28 % | 40 % | 17 % | 14 % |

Tabell 1: Elektrisitetsproduksjon i Norden etter energikjelder

Kjelde: (European Network of Transmission System Operators for Electricity, 2010)⁵

Tabell 1 gir ei betre forståing av kvifor hydrologiske energikjelder spelar ei så sentral rolle i energiproduksjonen i Norden. Noreg er på topp med 95 % av sin elektrisitetsproduksjon frå vasskraft eller hydrologiske kjelder. Vidare kan ein sjå at Sverige har noko under halvparten av sin elektrisitetsproduksjonen frå vasskraft, og resten frå fossile brennstoff og termiske energikjelder (atomkraft). Finland har cirka 17 % av elektrisitetsproduksjonen frå vasskraft, 14 % frå andre fornybare kjelder som mellom anna vindmøller og solenergi, og resten frå

⁴ Norden omfattar landa Noreg, Sverige, Danmark, Finland og Island. I denne analysen vil eg ikkje inkludere Island når eg omtalar Norden fordi Island ikkje er ein del av Nord Pool.

⁵ Tabellen, oversetjinga og utrekningane er mine eigne, men talmaterialet er henta ENTSOE. Grunnen til at summen av elektrisitetsproduksjon i Finland summerar seg til 99 % og ikkje 100 %, er fordi det i utgangspunktet var med ei eiga kolonne som representerte ikkje-identifiserbare energikjelder. Av plassomsyn vart denne fjerna.

fossile brennstoff og termiske energikjelder. Danmark har ikkje noko elektrisitetsproduksjon frå vasskraft, men har cirka 28 % frå andre fornybare energikjelder der vindmøller utgjer brorparten, og resten frå fossile brennstoff.

| Elektrisitetsproduksjon i Norden i TWh | | | |
|---|-------------|------------|------------------------|
| | Hydrologisk | Totalt | Hydrologisk ift totalt |
| Noreg | 117 | 123 | 95 % |
| Sverige | 66 | 145 | 46 % |
| Danmark | 0 | 37 | 0 % |
| Finland | 13 | 77 | 17 % |
| SUM | 196 | 383 | 51 % |

Tabell 2: Elektrisitetsproduksjon i Norden i TWh

Kjelde: (European Network of Transmission System Operators for Electricity, 2010)

Frå tabell 2 ser ein at 51 % av all elektrisitetsproduksjon i Norden kjem frå hydrologiske energikjelder, og at fyllingsgrad og andre hydrologiske variablar difor naturleg nok spelar ei sentral rolle i dette systemet. Det er også interessant å sjå dette ut i frå absolutte størrelsar. Tabellen gir i så måte eit overblikk over asymmetrien i produksjon av elektrisk straum og befolkning. Ein kan sjå at Noreg produserar meir straum enn både Finland og Danmark til saman. Meir spesifikt kan ein seie at Noreg produserar meir enn tre gongar så mykje elektrisitet som i Danmark, og er heller ikkje langt unna å produsere nesten dobbelt så mykje elektrisitet som i Finland. I tillegg ser ein at Noreg ikkje er langt unna Sverige sin straumproduksjon. Dette talmaterialet bekreftar at Noreg er på topp i Norden når det gjeld produksjon av elektrisitet per innbyggjar.

1.2.2 Overføringsnett

Det er også interessant å sjå nærmare på korleis elektrisiteten flyttar seg i dette området. I og med at straum ikkje kan lagrast er det avgjerande å ha god flyt og kapasitet i overføringsnettet for å kunne påstå at dette området er ein reell og felles integrert marknad.



Figur 1: Overføringsnett

Kjelde: Statnett (2008)

Av figur 1 kan ein sjå at det er mellom Noreg og Sverige at det ligg føre mest overføringskapasitet mellom dei ulike nasjonane i Norden. Figuren er noko stilisert, så brorparten av denne overføringskapasiteten er i dei sørlege deler av desse to landa, og er ikkje uniformt fordelt over heile grensa som figuren kan gi inntrykk av. Vidare er det fleire sjøkablar mellom Noreg og Danmark over Skagerrak og mellom Noreg og Nederland (den såkalla NorNed-kabelen). Desse transporterar likevel ikkje like mykje straum som mellom Noreg og Sverige. I tillegg har Noreg to mindre kablar i nord til høvesvis Finland og Russland.

Det kan sjå ut som at det også er stor grad av integrasjon mellom dei andre landa i Nord Pool-området. Likevel er det ingen land som har i nærleiken same omfang av overføringskapasitet som Noreg og Sverige.

Det kan vere forvirrende at overføringskapasitet er oppgitt i Megawatt (MW), medan straumforbruk ofte blir målt i kWh. Ein kilowattime (kWh) er energimengda som tilsvarar eit effektforbruk på ein kilowatt over ei periode på ein time. Det betyr at NorNed-kabelen mellom Noreg og Nederland teoretisk sett kan overføre 700 MW pr time, eller 6,1 TWh i året, om den blir maksimalt utnytta.⁶ Dette tilsvarar 5 % av årleg produksjon i Noreg om ein tek utgangspunkt i talmaterialet frå tabell 2. Så om utanlandsk kraft ikkje kan fullstendig substituerast med innanlandsk kraftproduksjon, er det i alle fall grunn til å tru at det kan vere eit verdifullt stabiliseringsbidrag i periodar med press i anten produksjons- eller konsumbehovet. Det er eit argument som svekker påstanden om at nasjonale hydrologiske forhold påverkar systemprisen, og eit argument som styrkjer påstanden om at Nord Pool er ein felles integrert marknad.

1.3 Tidlegare forskning

Det er mange faktorar som spelar inn ved fastsetjinga av systemprisen, og det er mange måtar å modellere desse. I tillegg er det vanskeleg å vite kva faktorar som er viktigast, og det er mest uråd å ta omsyn til alt som verkar inn på prisdanninga. Det er likevel gjort mange forsøk på å forklare systemprisen, og her blir eit knippe av dette presentert.

Ein av dei mest kjende og brukte modellane er Samkjøringsmodellen som er utvikla av SINTEF. Dette er ein svært avansert modell som krev mykje data og fleire ulike variablar, og som blir brukt av dei største aktørane i marknaden til produksjonsplanlegging, optimalisering og simulering av kraftsituasjonen (Solland, 2007). Samstundes er det ikkje nødvendigvis slik at alle faktorane er like viktige, eller at det er omfanget og kompleksiteten av data som nødvendigvis gir best resultat.

Det er heller ikkje lett å vite kva for faktorar ein bør bruke og graden av kompleksitet som ein bør legge i modellane. Vallevik og Øyan (2004) finn at fundamentalinformasjon som ver

⁶ $700 \text{ MWh} * 24 * 365 = 6\,132\,000 \text{ MWh} \approx 6,1 \text{ TWh}$

og hydrologisk situasjon har påverknad på prisdanninga, og at dette har størst effekt når forholdet mellom kapasitet og etterspørsel er pressa.

Johnsen (2001) skisserar ein vekebasert modell for etterspørsel og pris, og finn at faktorar som uventa snø eller tilsig samt temperatur verkar inn på prisdanninga. I tillegg er desse effektane størst ved slutten av vintersesongen. Escribano, Peña og Villaplana (2002) brukar ei noko meir avansert tilnærming ved at dei modellerer inn sprangvise volatilitetsmønster, samt sesong og mean-reversion. Dette blir basert på funn som viser at spotprisen er svært volatil, sjølv etter korrigering for sesong. Dei finn og at Nord Pool har den lågaste grada av mean-reversion samanlikna med ei rekkje andre energibørsar.

Det er også nye og meir perifere faktorar som kan spele inn på systemprisen. Bye og Rosendahl (2005) finn mellom anna at kvoteprisen på utslepp av klimagassar kan påverke noko på kort sikt, men mindre i gjennomsnitt over året. I tillegg hevdar dei at utviklinga av kull- og gassprisar, ulike teknologiar og kapasitet i kraftproduksjon og avgrensingar i overføringsnett mellom regionar er vel så viktig påverknadskanalar.

Andre meir hydrologiske, men samstundes perifere, faktorar som kan påverke kraftprisen er vindkraft. Mauritzen (2012) finn mellom anna i si doktorgradsavhandling at vindkraft i Danmark pressar den norske områdeprisen ned. Vidare hevdar han at vindkraft kan redusere prisvariasjonen på kort sikt i den nordiske kraftmarknaden, men at vindkraft samstundes er sterkt avhengig av vasskrafta i systemet. Han hevdar også at det kan ligge potensielle gevinstar i å samordne vindkraft- og vasskraftproduksjon i og med at førstnemnde ikkje kan lagrast i eit vassmagasin, men at det må gjerast vidare undersøkingar på området for å avklare i kor stort omfang dette er mogleg i og med at det er avgrensa overføringskapasitet mellom dei ulike landa i Norden.

Eit anna perspektiv som er verdt å vurdere er i kor stor grad det ligg føre mønster i systemprisen. Li og Flynn (2004) studerar 14 ulike deregulerte energibørsar over heile verda og finn at Spania og Storbritannia (delvis også Skandinavia) i større grad har føreseielege og konsistente elektrisitetsprisar, medan dette ikkje er tilfelle i mellom anna sørlege delar av Australia. Dei hevdar vidare at dette kan forklarast med tanke på faktorar som konsumentåtfærd, konsumplanlegging og risikostyring.

Koopman, Ooms og Carnero (2007) har ein noko anna metodisk innfallsvinkel, men også desse forfattarane finn bevis som trekk i den same retninga. Dei samanliknar fire europeiske

kraftbørsar og finn at den logaritmiske spotprisen på Nord Pool er den som er mest konsistent. For å påvise dette brukar dei avanserte statistiske og autoregressive teknikkar der long memory (vedvarande autokorrelasjon) og sesong blir modellert inn.

Andre innfallsvinklar for å studere elektrisitetsprisar handlar om kapasitetsdynamikk. Det at kraft ikkje kan lagrast vil i utgangspunktet gjere den til ei veldig volatil råvare. På den andre sida kan ein argumentere med at i eit hydrologisk system kan ein likevel lagre kraft i vassmagasin, og dette burde gagne produsentar og/eller konsumentar. Bushnell (1998) viser at det kan ligge økonomiske fordelar for store og dominerande aktørar i å bruke sin posisjon i eit vasskraftbasert kraftsystem, og at desse kan ha innflytelse på prisdanninga.

Marknadsmakt med Nord Pool som bakteppe er det også fleire andre som har studert. Hjalmarsson (2000) forkastar ikkje hypotesen om fullkommen konkurranse, og kan difor hevde at det ikkje ligg føre vesentleg marknadsmakt på Nord Pool. Han viser til lav eigarkonsentrasjon hos produsentane som ei mogleg forklaring på dette. Amundsen, Bergman og Bo (1998) finn også at Nord Pool i stor grad oppfører seg som ein fri marknad. Dei får likevel motsatt konklusjon ved ekskludering av handel mellom dei ulike landa på kraftbørsen i modellane sine, noko som blir påpeikt som noko overraskande då kraftoverføringane mellom landa er av ein relativ beskjeden art.

Steen (2004) tek denne tankegangen eit steg vidare, og ser på marknads- og konkurransesituasjonen knytt til flaskehalsar i linjenettet. Han finn at når det oppstår flaskehalsar, så vil det skape marknadsmakt i det heilt korte tidsrommet for produsentane, som igjen kan påverke spotprisdanninga. Dette meiner han er ein indikasjon om at det ikkje bør tillatast meir eigarskapskonsentrering før kapasiteten er blitt betre.

Det er mange forskarar som er opptatt av marknadsmakt og eigarskapskonsentrering. Flatabø et al. (2003) viser i sin studie at eigarskapen har blitt konsentrert på færre hender på Nord Pool og at denne gradvise utviklinga er drive fram av deregulering i energibransjen og påfølgande restrukturering i form av auke i fusjonar og oppkjøp. Desse forfattarane er bekymra for denne utviklinga.

Eit anna moment å ta stilling til når ein skal modellere elektrisitetsprisen er om marknaden er effektiv og effisient, samt korleis den har utvikla seg dei siste åra. Lundgren, Hellstrøm og Rudholm (2008) ser på dei historiske linjene frå Nord Pool vart etablert og viser at gjennomsnittsprisen har auka i takt med at Finland og Danmark har blitt innlemma.

Volatiliteten har gått i begge retningar etter kvart som fleire ulike aktørar har blitt med, men i sum hevdar desse forfattarane at eksterne sjokk likevel blir betre handtert i ein felles nordisk marknad enn kva som var tilfelle tidlegare med nasjonale marknader.

Johnsen og Gjølborg (2001) påviser at Nord Pool truleg ikkje er effisient. Dei kjem til denne konklusjonen på bakgrunn av studiar av pariteten mellom spot- og futuresprisen samt regresjonsanalysar. Vidare betraktar dei at vassmagasina kan brukast som lager for framtidig elektrisitetsbehov, og observerer at fyllingsgrada på denne måten har innflytelse på derivatprisane. I tillegg finn dei at spotprisen er uvanleg volatil i forhold til andre råvarer, og at det er vanskeleg å forklare dette ut i frå fundamentale marknadsforhold.

2. Metode

I denne delen vil eg kort presentere relevante teoriar og metodar som blir nytta i data- og analysedelen.

2.1 Korrelasjon

Korrelasjon er eit relativt mål på styrka og retninga på den lineære samvariasjonen mellom to variablar. Observert samvariasjon er ein nødvendig, men samtidig ikkje ein tilstrekkeleg, føresetnad for å påvise at det ligg føre kausalitet. Det kan mellom anna vere andre underliggande drivarar som gjer at to variablar samvarierer, og ikkje at ein variabel påverkar den andre (Keller, 2008).

2.2 Regresjon

Regresjon basert på minste kvadratars metode (OLS) er ei metode for å modellere ein samanheng mellom ein avhengig variabel og ein eller fleire uavhengige variablar. Resultatet blir ei regresjonslinje og nokre feilledd som modellen ikkje kan forklare. Metoden går ut på å minimere den kvadrerte variansen eller det kvadrerte vertikale avviket mellom dei observerte datapunkta og dei estimerte datapunkta frå ein lineær approksimasjon.

For at regresjon basert på minste kvadratars metode skal vere BLUE (Best Linear Unbiased Estimators) ved tidsseriar må dei følgjande fem vilkåra vere oppfylt: Lineære parameter, ingen perfekt kollinearitet, homoskedastisitet, ingen seriekorrelasjon og forventingsverdien til feilledda må vere null. Om ein i tillegg ønskjer å gjere statistisk inferens, så må også feilledda vere normalfordelt (Wooldridge, 2009).

2.3 Forklaringsgrad

Forklaringsgrad (R^2) eller determinasjonskoeffisienten fortel kor mykje av utvalsvariansen i den avhengige variabelen som blir forklart av dei uavhengige variablane. Med andre ord er

R^2 eit mål på størrelsen på den forklarte variansen i forhold til den totale variansen og viser kor godt regresjonslinja passar til datamaterialet.

Ved modellbygging er det mest riktig å ta stilling til justert R^2 (adjusted R^2) når ein skal vurdere ulike modellar og å finne ut kva for faktorar ein bør inkludere i ein modell. Ein av grunnane til dette er fordi R^2 aldri vil bli redusert om ein legg til ein ekstra variabel i samband med modellbygging. Justert R^2 tek derimot omsyn til talet på uavhengige variablar. Det er likevel ikkje alltid tilfelle at justert R^2 er meir egna for å vurdere ulike modellar. Om ein til dømes ser på heile populasjonar og ikkje utval, så vil ikkje justert R^2 gi noko meir nyttig informasjon enn R^2 (Wooldridge, 2009).

2.4 Endogenitet

Endogenitet oppstår når det er korrelasjon mellom ein uavhengig variabel og feilledet i ein regresjon. Simultanitet er ein type endogenitet som går ut på at ein eller fleire av dei uavhengige variablane blir felles bestemt med den avhengige variabelen, ofte gjennom ei likevektsmekanisme. Ein mykje brukt metode for å løyse dette problemet er ved hjelp av instrumentregresjon (Wooldridge, 2009).

2.5 Chow test

Ein Chow test er ein test for å sjekke om det ligg føre strukturelle endringar i regresjonsmodellen over tid, til dømes ved at parametrane er ulike i første del av tidsperioden samanlikna med siste del av tidsperioden. I tillegg til å vurdere dette ut i frå dei ulike parameterestimata frå dei to ulike periodane kan Chow-testen i seg sjølv vere eit nyttig numerisk hjelpemiddel for å avklare dette. Reint teknisk handlar det om å samanlikne summen av dei kvadrerte feilleda i dei to ulike tidsperiodane med dei kvadrerte feilleda til heile tidsperioden ved hjelp av ein F-test. Nullhypotesen er at parameterane er like (Chow, 1960).

2.6 RESET test

Ramsey sin RESET test (Regression Specification Error Test) er ein formell test for å sjekke om regresjonsmodellen er riktig spesifisert. Testen baserar seg på å undersøkje om ikkje-lineære kombinasjonar av regresjonslinja er med på å forklare den avhengige variabelen. Om dette er tilfelle, så er modellen misspesifisert (Ramsey, 1969).

Ramsey sin RESET test er eit hjelpemiddel for å vite om modellen faktisk har med reelle faktorar som påverkar. Eit minus ved testen er at om det kjem fram at dette ikkje stemmer, så gir testen likevel ikkje nokre forslag om kva som kan gjerast for å løyse dette problemet. Vidare kan RESET-testen ikkje brukast for å teste om det ligg føre uobserverte utelatne variablar, ei heller om det ligg føre heteroskedastisitet i modellen. Difor er RESET-testen utelukkande ein funksjonell form test (Wooldridge, 2009).

2.7 Heteroskedastisitet

Heteroskedastisitet er at variansen til feilledda i ein regresjonsmodell ikkje er konstante (det motsette av homoskedastisitet). Dette er mest vanleg ved tverrsnittsdata, og det typiske eksempelet for å illustrere heteroskedastisitet er at feilleddet knytt til prediksjonen av produksjonen til ei stor bedrift ofte vil vere større enn feilleddet knytt til prediksjonen av produksjonen til ei mindre bedrift.

Heteroskedastisitet er mindre vanlig i tidsseriar fordi endringar i den avhengige variabelen og endringar i dei uavhengige variablane ofte er av same størrelsesorden over tid. Ein vil ofte sjå i ein modell som til dømes predikerar konsum ut i frå disponibel inntekt at desse to nemnde variablane ofte veks meir eller mindre i same omfang over tid.

Ved heteroskedastisitet er parameterane frå regresjon basert på minste kvadratars metode framleis forventningsrette og konsistente, men dei er ikkje lenger effektive. Det vil seie at variansen til parametrane ikkje er minimert. I tillegg kan standardavvika bli for høge eller for låge. Det vil ha betydning for estimering, då konfidensintervalla ikkje vil bli riktige. Ein måte for å løyse problemet med heteroskedastisitet kan vere ved å bruke robuste standardavvik. Eventuelt kan ein rekonstruere modellen. To mykje brukte formelle testar for

heteroskedastisitet er mellom anna Breusch-Pagan-testen og White-testen (Pindyck & Rubinfeld, 1991).

Ved Breusch-Pagan-testen brukar ein kvadratet av feilledda frå den opphavlege regresjonen som avhengig variabel i ein ny regresjon. Dei uavhengige variablane er dei same i begge regresjonane. Ein ser då på F-verdien frå den siste regresjonen og avgjer då om ein kan forkaste nullhypotesen om homoskedastisitet. Med andre ord kan ein seie at ein vurderar i kor stor grad kvadratet av feilledda frå den opphavlege regresjonen blir forklart ut i frå dei uavhengige variablane (Breusch & Pagan, 1979).

White-testen liknar noko på Breusch-Pagan-testen i den forstand at ein også her brukar kvadratet av feilledda frå den opphavlege regresjonen som avhengig variabel i ein ny regresjon. I tillegg brukar ein regresjonslinja (fitted values) frå den opphavlege regresjonen som ein eigen variabel og kvadrerar denne. Deretter ser ein då i kor stor grad kvadratet av feilleddet frå den opphavlege regresjonen blir forklart ut i frå regresjonslinja og kvadratet av denne. Til slutt må ein kjøre ein F-test av regresjonslinja og kvadratet av denne for vidare å vurdere om ein kan forkaste nullhypotesen om homoskedastisitet. Det kan ein gjere om F-verdien er høgare enn den kritiske grensa (White, 1980).

2.8 Autokorrelasjon

Autokorrelasjon (seriekorrelasjon) oppstår ofte i tidsseriar, og er tilfelle når det er korrelasjon mellom feilledda til dei ulike observasjonane i ein regresjonsmodell. Ein ser det av og til intuitivt om dei tidlegare observerte verdiane har innflytelse på dagens og framtidige observasjonar. Typiske tidsseriar der ein må vere var på dette er om ein følgjer prisen på ein aksje eller BNP for eit land over fleire år.

Autokorrelasjonen kan både vere positiv eller negativ. Positiv autokorrelasjon oppstår om feilledda i ein tidsperiode er positivt korrelert med feilledda i neste tidsperiode og på same vis oppstår negativ autokorrelasjon om feilledda i ein tidsperiode er negativt korrelert med feilledda i neste tidsperiode

Ved autokorrelasjon er parameterane frå regresjon basert på minste kvadratars metode framleis forventningsrette og konsistente, men dei er ikkje lenger effektive. Det vil seie at variansen til parametrane ikkje er minimert. I tillegg vil standardavvika ikkje bli korrekte.

Ved positiv autokorrelasjon vil standardavvika bli for låge og t-verdiane for høge, og ved negativ autokorrelasjon vil standardavvika bli for høge og t-verdiane for låge. Det vil ha betydning for estimering, då konfidensintervalla ikkje vil bli riktige.

Ein kan også skilje mellom første og fleire ordens autokorrelasjon. Første ordens autokorrelasjon er tilfelle om feilledda i ei tidsperiode er direkte korrelert med feilledda i påfølgande tidsperiode. Fleire ordens autokorrelasjon er tilfelle om feilledda i ei tidsperiode er direkte korrelert med opptil fleire påfølgande tidsperiodar. Durbin Watson-testen er eit nyttig hjelpemiddel for å ta stilling til første eller andre ordens autokorrelasjon, medan andre testar blir brukt for å ta vurdere fleire ordens autokorrelasjon (Pindyck & Rubinfeld, 1991).

2.9 Autokorrelasjonsfunksjon (ACF)

Dette verktøyet plottar autokorrelasjonen grafisk og er til hjelp for å identifisere eller estimere tidsseriemodellar. ACF viser autokorrelasjon over ulike lag (ofte tid) og vil då vise korleis tidlegare observasjonar (lag) påverkar dagens observasjon. Det kan mellom anna brukast i vurderinga av stasjonæritet. Om ein tidsserie er stasjonær, så er autokorrelasjonsfunksjonen geometrisk avtakande (Enders, 2010).

2.10 Partiell autokorrelasjonsfunksjon (PACF)

I motsetning til ACF så viser PACF korleis ein enkelt tidlegare observasjon påverkar dagens observasjon (ein korrigerar då for autokorrelasjon frå eventuelle andre lag). Dette verktøyet er nyttig for å avgjere kor mange lag ein skal ha med når ein skal inkludere eit eller fleire autoregressivt ledd i ein regresjonsmodell. Nummeret i rekka på det siste laget som er statistisk signifikant viser talet på autoregressive ledd ein bør inkludere i ein regresjonsmodell (Enders, 2010).

2.11 Informasjonskriterium

Ei anna metode for å vurdere kor mange autoregressive ledd som ein bør ta med i ein modell er å sjå på ulike informasjonskriterium. I denne oppgåva har eg sett på Akaike information criterion (AIC), Schwarz-Bayesian information criterion (SBIC) og Finite prediction error

(FPE). Generelt kan ein seie at desse kriterier vurderer i kor stor grad informasjon går tapt når ein prøver å skildre røynda i ein modell. Eller, med andre ord, ein veg nøyaktigheit mot kompleksitet (Enders, 2010).

2.12 Stasjonæritet

Det er hensiktsmessig å vite om den avhengige variabelen er stasjonær eller ikkje når ein skal lage ein regresjonsmodell. Ein prosess er stasjonær om dei statistiske eigenskapane ikkje blir endra over tid. Kvit støy er eit klassisk døme på ein prosess som er stasjonær. Ved kvit støy er prosessen ikkje korrelert over tid, har ein forventingsverdi som er null og variansen er konstant.

Om ein lager ein modell som ikkje tek omsyn til at den avhengige variabel ikkje er stasjonær, kan ein få spuriøse regresjonar, det vil seie ein regresjon utan meining. Ein metode for å unngå dette problemet ved ikkje-stasjonære data er å differensiere datamaterialet.

For å teste stasjonæritet kan ein (augmented) Dickey-Fuller test vere eit nyttig hjelpemiddel. Denne testen har ei nullhypotese som påstår at feilledda er ikkje-stasjonære (unit root), og ei alternativhypotese som påstår at feilledda er stasjonære (Wooldridge, 2009).

2.13 Box plot

Eit box plot viser ulikskapar mellom ulike populasjonar eller grupper utan å kategorisere dette i ei bestemt statistisk fordeling. Dei ulike gruppene blir delt i fire kvartil med ein median som skil det andre og tredje kvartilet. Observasjonar som ligg utanfor den heiltrokne boksa (50 % av observasjonane er innanfor boksa) er uteliggjarar (outliers). Box plotet viser på denne måten både spreing og eventuelle skeivskapar i datamaterialet. Vidare kan det gi ein peikepinn på om det ligg føre sesongvariasjonar, mellom anna ved å observere om dei ulike gruppene har ulik median (McGill, et al., 1978).

2.14 HP-filter

Hodrick-Prescott-filteret (HP-filteret) er ei mykje brukt univariat metode for å skilje ut trendkomponenten i ein tidsserie. I omgrepet univariat ligg det at ein tek utgangspunkt i berre ein enkelt variabel. Ein kan då påvise underliggande trendar i datamaterialet ved å "glatte ut" ein tidsserie. Teknikken er mellom anna mykje brukt i makroøkonomi i samband med å fastsetje konjunkturar. Likningane under viser korleis filteret er bygt opp.

$$(1) \quad y_t = \tau_t + c_t$$

Ein kan tenke seg at ein har ein tidsserie (y_t) som kan dekomponerast i ein trendkomponent (τ_t) og ein stasjonær komponent (c_t). Ofte er det trenden ein er ute etter og denne linja kan ein finne ved numerisk iterasjon av programmeringsproblemet i (2).

$$(2) \quad \text{Min} \sum_{t=1}^T (y_t - \tau_t)^2 + \lambda \sum_{t=2}^{T-1} [(\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1})]^2$$

Reint teknisk består HP-filteret av ei likning med to ledd der målet er å minimere summen av desse ledda. Det første leddet er kvadratet av differansen mellom den originale tidsserien (y_t) og trenden (τ_t), eller, med andre ord, kvadratet av den stasjonære komponenten (c_t). Det andre leddet er kvadratet av endringa i veksten til trenden eller, med andre ord, kvadratet av trenden differensiert to gongar (Hodrick & Prescott, 1997).

Intuitivt kan ein forklare denne likninga ved å førestille seg at det første leddet vil forsøke i størst mogleg grad å tilpasse trenden til den originale tidsserien, og at det andre leddet vil forsøke i størst mogleg grad å glatte ut trenden. Ein kan seie at det eksisterar eit avvegingsproblem knytt til best mogleg tilpassing eller utglating der λ er med på å bestemme kor mykje vekt som skal tillegast desse to kryssande interessene. Eventuelt kan ein tenke seg at λ reflekterar kostnaden eller straffa av å tillate fluktuasjonar i trendkomponenten. λ er eit eksogent parameter som kan ha ein verdi frå null til uendeleg og avgjer kor stor vekt variasjonar i vekstraten (det andre leddet) som skal tillatast. I grensetilfellet der λ er uendeleg stor vil ein få ei lineær trend med konstant vekst. I det motsette tilfellet vil ein ved $\lambda = 0$ få ei trend som vil være identisk med den originale tidsserien og trendlinja vil difor ikkje vere synleg (Bjørnland, et al., 2004).

Det spesielle med HP-filteret er at ein sjølv kan bestemme kor glatt trendlinja skal vere. Det gjer ein når ein fastsetter den nemnte λ -verdien. Det å avgjere kva verdi λ skal ha er ikkje

trivielt. Det avhenger av kva type data ein har og kor hyppige observasjonar ein har innanfor det aktuelle tidsrommet. Det er og mykje fagleg skjønn i slike vurderingar. Ein metode for å regne ut ein passende λ er å bruke tommelfingerregelen under.

$$(3) \quad \lambda = 100 \cdot (\text{talet på observasjonar i eit år})^2$$

Ved empirisk bruk av HP-filteret bør ein vere særleg oppmerksom på tre forhold. For det første er resultata veldig sensitive for endringar i λ , i og med at det som nemnt ligg føre ein del skjønn og usikkerheit i denne verdien, og difor bør denne vere gjenstand for ytterlige faglege vurderingar om ein får særlege ulogiske resultat. For det andre kan HP-filteret gi misvisande resultat om ein har svært lange tidsseriar då det ikkje nødvendigvis klarer å avdekke eventuelle skift i datamaterialet på ein korrekt måte. For det tredje må ein vere varsam med å tolke endepunkta i trenden. Det er fordi det ikkje ligg føre trendestimat for $t+1$ i slutten av perioden og for $t-1$ i byrjinga av perioden. Då vil i staden dei faktiske observasjonane bli lagt til grunn og det fører til at dei faktiske observasjonane vil bli relativt meir vektlagt i starten og i slutten av perioden. Ein praktisk konsekvens av dette er at evna til eventuell framtidig prognosering vil vere avgrensa.

3. Skildring av data

I denne delen vil eg gå djupare i datamaterialet som skal brukast i den empiriske analysen. Først vil eg ta for meg systemprisen, som er hovudfokuset og den avhengige variabelen i regresjonsmodellane, og som difor vil bli handsama noko meir utførleg enn dei andre variablane. Deretter vil eg kort sjå på dei fire andre uavhengige variablane.

3.1 Systemprisen

Daglige og vekelege data for systemprisen i NOK pr MWh er henta frå Nord Pool sin FTP-serverar. Døgnprisen er eit aritmetisk gjennomsnitt av timeprisen i alle dei 24 timane i eit døgn og den vekebaserte prisen er eit aritmetisk gjennomsnitt av timeprisen i alle dei 168 timane i ei veke.

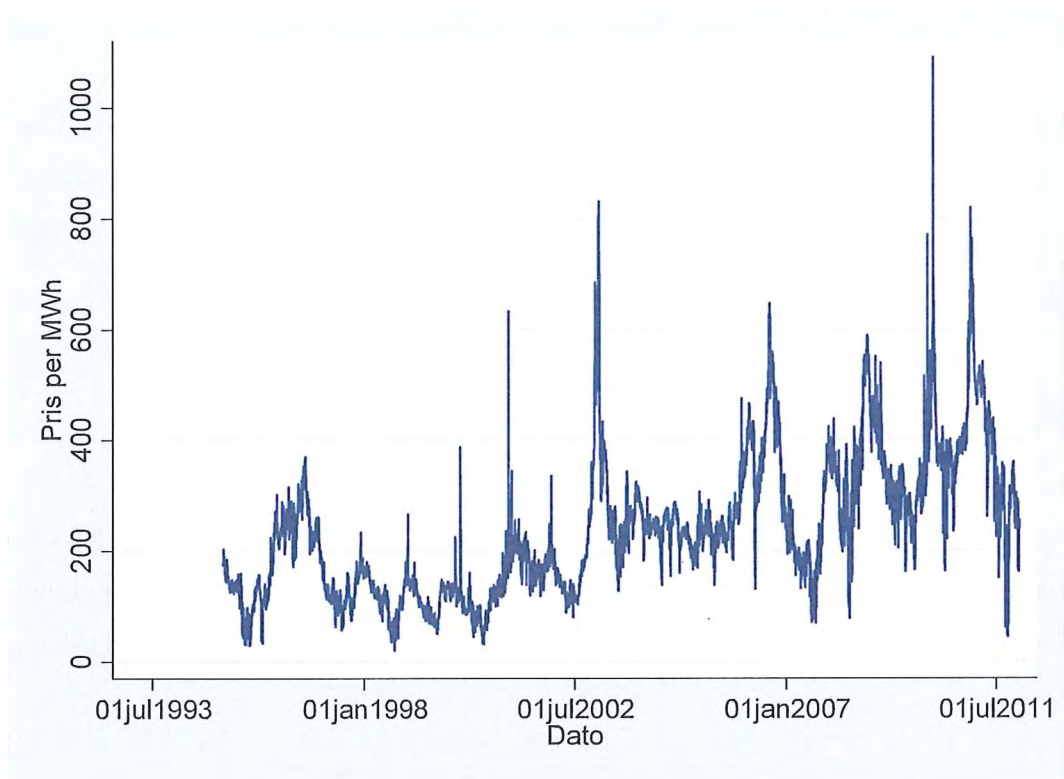
| Variabel | Tal på obs | Gjsnitt | Std avvik | Min | Maks |
|----------|------------|---------|-----------|-------|---------|
| Døgnpris | 6209 | 239.79 | 125.58 | 21.27 | 1090.02 |
| Vekepris | 887 | 239.80 | 123.68 | 37.5 | 751.72 |

Tabell 3: Deskriptiv statistikk

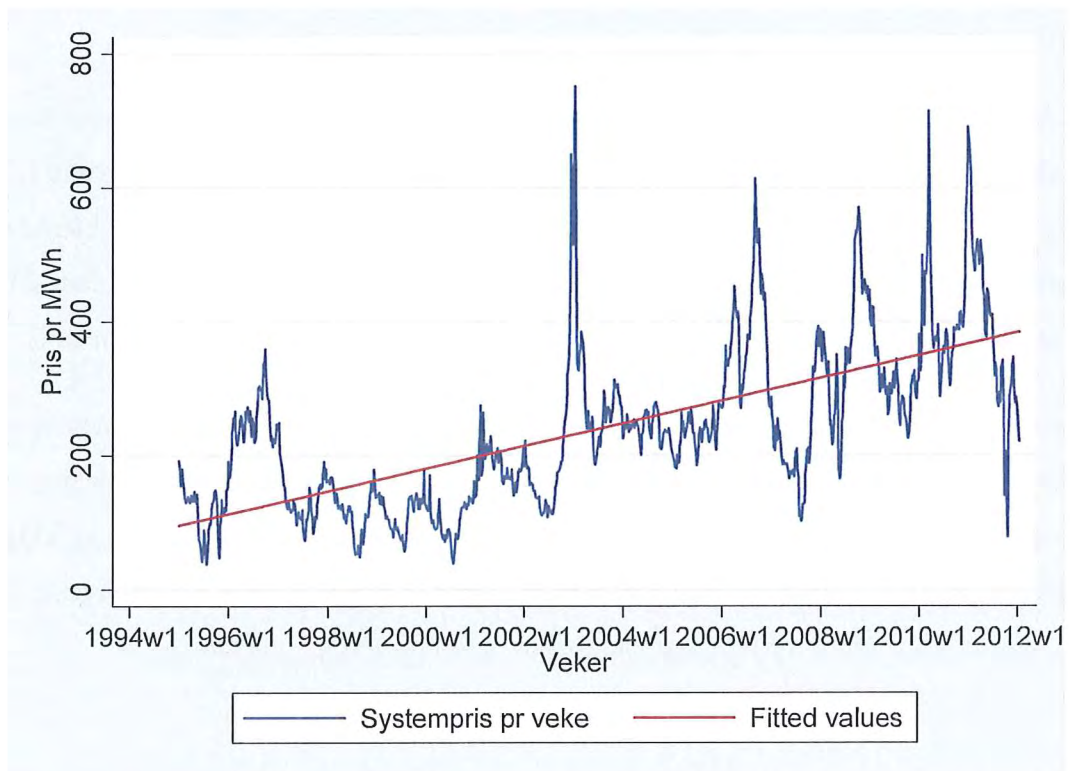
Av tabell 3 ser ein at det ikkje er store forskjellar i døgnprisen og vekeprisen, noko som heng saman med at dei er bygde opp av akkurat same datamaterialet. Gjennomsnitt og standardavvik er ikkje veldig ulike, men maksimale noteringar avviker noko. Dette kjem frå "straumkrise" vinteren 2010 då det var nokre timar der prisen var ekstremt høg, og dette blir i større grad fanga opp i døgnprisen som baserar seg på eit kortare aritmetisk gjennomsnitt.⁷

Døgnprisen er tatt med her for betre å vise volatiliteten i denne variabelen. I resten av oppgåva er det vekepris som blir studert og som er fokuset. Grunnen til at vekeprisen og ikkje døgnprisen blir sett nærmare på er fordi førstnemnte ikkje inneheld sesongeffektar knytt til den enkelte vekedag (straumforbruket er ulikt både i mellom dei fem vekedagane og i helgene), og fordi dei andre uavhengige variablane også ligg føre på vekeform.

⁷ Maksimalverdiane ville nok ha vore enno høgare om eg hadde studert timeprisar i staden, men då ville også talet på observasjonar vore betydeleg meir enn 6209. Samstundes som ein då må ta høgde for sesongeffektar gjennom dagen.



Figur 2: Døgnpris



Figur 3: Vekepris

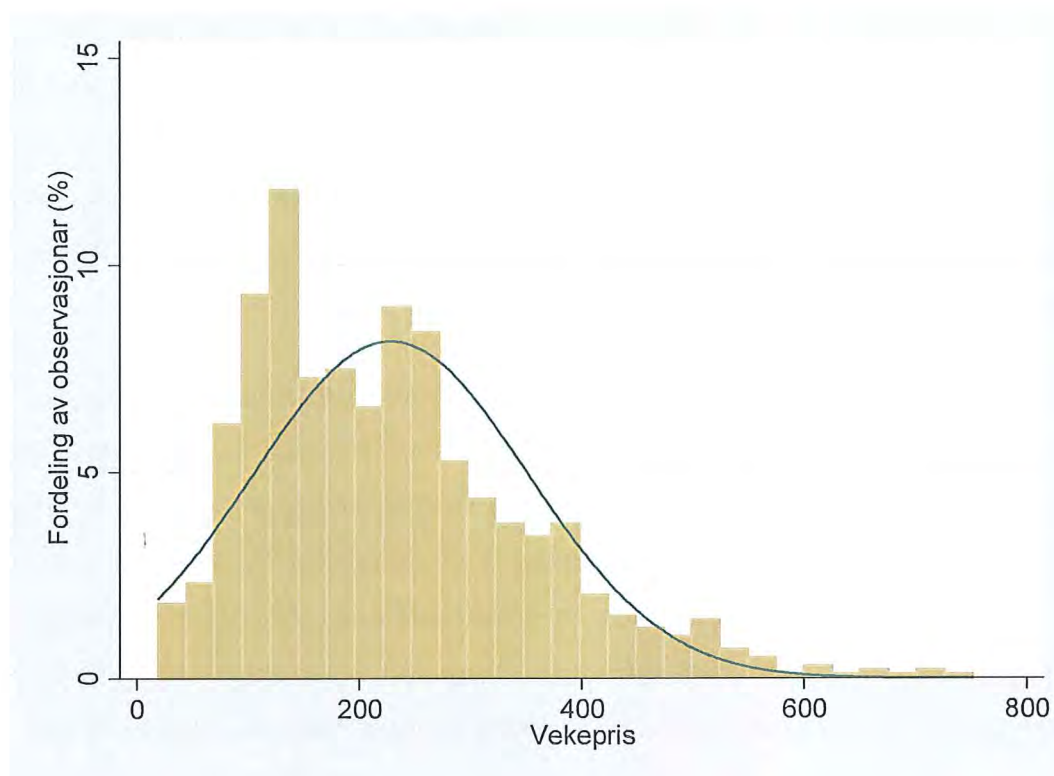
For å betre forstå talmaterialet kan det også vere interessant å sjå dette grafisk. Når ein observerer figur 2, døgnprisen, og figur 3, vekeprisen, ser ein at begge prisane har auka gradvis frå 1995 til 2011. Ved å kjøre ein enkel regresjon med ei tidstrend for den vekebaserte prisutviklinga i figur 3 vert dette bekrefta. Parameteren til denne tidstrenden viser ein verdi på 0,3254 NOK per MWh. Det betyr at systemprisen i gjennomsnitt vil auke med 1,7 øre/kWh per år.⁸ Vidare kan ein ved å studere det grafiske plottet også observere at systemprisen har fleire og meir markerte uteliggjarar, samt er noko meir volatil i den siste delen av perioden.

For å ta omsyn til fenomenet sommar- og vintertid har eg brukt same metode som Lucia og Torrò (2008). Dette er for å unngå hull i datamaterialet. Nord Pool har ikkje notert prisen i det tidsrommet der vi stiller klokka ein time fram ved sommartid, og ved vintertid er prisen det dobbelte i den timen vi skur klokka tilbake. Ved sommartid har eg difor teke eit aritmetisk gjennomsnitt av timeprisen før og etter det tidsrommet der vi stiller klokka, og ved vintertid delar eg den oppgitte timeprisen i det tidsrommet der vi stiller klokka på to. På denne måten oppstår det ei jamnare prisutvikling og ein får ingen hopp eller utelatne punkt.

Generelt har eg brukt aritmetiske gjennomsnitt dei få gongane då det er manglar i datamaterialet. Dette gjeld mellom anna ved turnover- og temperaturvariabelen, der det av ikkje oppgitte grunnar er nokre få plassar som manglar observasjonar.

⁸ $0,0003254 \text{ NOK pr kWh} * 52 = 0,017 \text{ NOK per kWh}$

3.1.1 Skewness og kurtosis



Figur 4: Fordeling av observasjonar (%)

Med tanke på å avklare dei statistiske eigenskapane til systemprisen er det fruktbart å vurdere i kor stor grad vekeprisen er normalfordelt. Av figur 4 ser ein korleis alle observasjonane er fordelt i prosent (summen av alle søylene vil vere 100 %). Ein ser at vekeprisen på mange måtar ser ut til å vere lognormal og vidare å ha ei leptokurtisk form, som er ein karakteristikk som kan brukast om ein meiner at fordelinga av observasjonane er noko spiss og har tilhøyrande tjukke/lange halar. Samstundes er det ikkje heilt utelukka at den har ei tilnærma normalfordelt form eller har noko av det same mønsteret som ei normalfordeling.

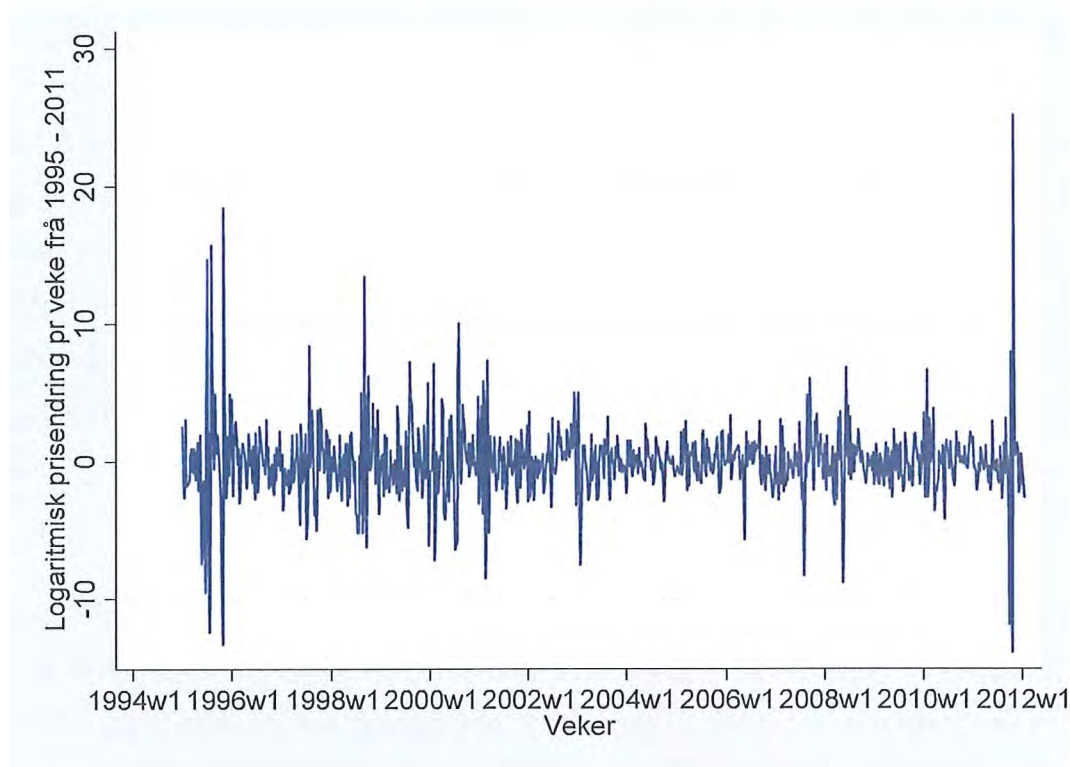
Dette kan også vurderast basert på verdiane ein finn frå å måle skewness og kurtosis. Eg finn at systemprisen har ein kurtosis på 3,94 og skewness på 0,95. I ei perfekt normalfordeling vil kurtosis vere tre og skewness null, så den formelle testen indikerer at systemprisen ikkje er langt unna å vere normalfordelt. Difor konkluderer eg, under noko tvil, at systemprisen er tilnærma normalfordelt.

3.1.2 Volatilitet

For å få ei betre forståing knytt til kor mykje eit aktivum på ein børs svingar kan det vere greitt å studere kor mykje prisen endrar seg frå periode til periode. Dette blir kalla volatilitet og er eit mål på risiko. Elektrisitet er som nemnt ei råvare som ikkje kan lagrast. Det vil kunne tilseie at elektrisitetsprisen vil svinge relativt mykje eller vere relativt volatil samanlikna med andre råvarer. Men, i eit hydrobasert kraftsystem, så kan produsentane likevel lagre noko av denne krafta i reservoar eller magasin i fjellet. Det burde i så tilfelle gi produsentane meir marknadsmakt, og difor gi rom for i større grad både å påverke prisutviklinga samt redusere volatiliteten.

Ein kan vidare argumentere med at sjølv om Nord Pool er ein mindre aktør i forhold til dei store energibørsane i verda, så er den likevel noko unik med tanke på det relative store innslaget av hydrologiske kraftkjelder i forhold til total produksjon. Mellom anna vart det påvist i innleiinga at cirka halvparten av energiproduksjonen i Nord Pool-området kjem frå hydrobaserte energikjelder. Desse forholda burde difor tilseie, samanlikna med andre energibørsar, at volatiliteten isolert sett burde vere lågare på Nord Pool. Motargumentet er at ein kan for så vidt også styre produksjonen frå til dømes termiske energikjelder, sjølv om vassmagasin likevel kanskje er enklare å justere, spesielt i det korte tidsrommet. Ei vidare analyse knytt til det å samanlikne volatilitet på ulike energibørsar i verda samt marknadsmakt på Nord Pool vil likevel ligge utanfor rammene for denne oppgåva.

Ved å betrakte figur 5 under kan det verke som om systemprisen er meir eller mindre uniform fordelt i den forstand at det ikkje ligg føre nokre tydelege mønster i korkje positiv eller negativ retning. Det ser likevel ut som at når det først oppstår prisendringar av ein relativ lik karakter, så klumpar desse seg saman. Det ser altså ut til at store endringar ofte blir følgt av store endringar, og at små endringar ofte blir følgt av små endringar. Dette gjeld uavhengig av forteikn. Dette fenomenet, volatilitetsklynger, er noko som ikkje er uvanleg på ein børs eller marknadsplass. Det er likevel mest viktig å ta stilling til om ein skal trendframskrive eller lage prognosar.



Figur 5: Volatilitet

Av figuren ser ein den logaritmiske prisendringa pr veke. Standardavviket til den logaritmiske prisendringa er 2,68 % pr veke. Dette gir ein årleg annualisert volatilitet på 194 %. Denne verdien finn ein ved å multiplisere standardavviket med kvadratrotta av 52.⁹

Lucía og Schwartz (2002) fann ein årleg annualisert volatilitet på 189 % i sin studie. Denne artikkelen er nokre år gamal samt at det var daglege prisendringar som var utgangspunktet i studien, men dette kan likevel gi ein svak indikasjon på at volatiliteten har auka marginalt. Den kalde vinteren i 2010 med påfølgande høge straumprisar kan kanskje vere med på å forklare noko av dette. Motargumentet finn ein ved å studere figuren ovanfor. Av den kan det sjå ut som at volatilitetsklyngene har blitt færre, om ein ser vekk i frå slutten av 2011.

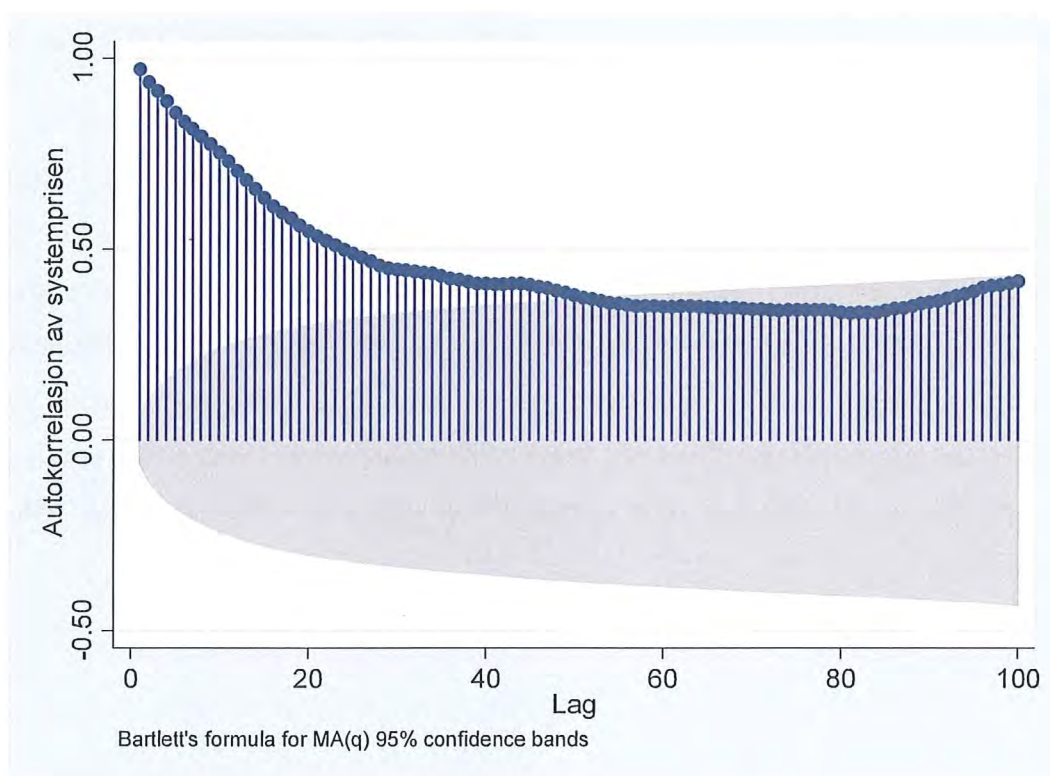
For at desse uttrekingane skal vere gyldig føreset ein at volatiliteten frå veke til veke er uavhengig, og at systemprisen følgjer ein tilnærma Wiener-prosess. Det er kanskje ein relativ

⁹ Grunnen til at volatiliteten veks med kvadratrotta av tida, og ikkje veks lineært med tida, er at ein ut i frå finasteori forventar at ein del fluktusjonar vil kansellere kvarandre ut, slik at avviket etter 52 veker ikkje vil vere 52 gonger avstanden frå utgangspunktet. Denne måten å rekne ut volatiliteten tek omsyn til denne problemstillinga då eit standardavvik seier noko om spreining, men ikkje noko om retning på prisendringa i og med at det er utrekna basert på kvadratet (og er difor positivt) av avviket frå førre periode.

sterk føresetnad og kanskje mindre sannsynleg i og med at det, som drøfta meir inngåande i den empiriske analysen, er både tendensar til sesong og autokorrelasjon i datamaterialet. Likevel, denne volatiliteten er ganske høg tatt i betraktning at volatiliteten til hovudindeksen på Oslo Børs, som jamvel gjenspeilar svært mange og godt diversifiserte aksjar, har vore 21,78 % i tidsrommet 1980 – 2005 (Bjerkseth, 2006).

3.1.3 Autokorrelasjon

I og med at systemprisen kan tolkast som ein tidsserie er det interessant å undersøkje om det ligg føre autokorrelasjon i dette datamaterialet. Eit autokorrelasjonsplott (ACF) kan gi eit overblikk over situasjonen.



Figur 6: ACF

Av figur 6 ser ein at det er stor grad av positiv autokorrelasjon i systemprisen når ein studerar 100 lag og dette er ein indikasjon på at kravet om at feilledda skal vere tilfeldige truleg ikkje er oppfylt. Reint praktisk vil det seie at systemprisen i tidlegare veker forklarar mykje av systemprisen denne veka. Dette er ikkje eit uvanleg funn når det gjeld energiprisar. Deng (2006) finn også autokorrelasjon i systemprisen, og meiner dette blir skapt ved at det blir lagra kraft i vassmagasin samt på grunn av at det er sesongelement i systemprisen.

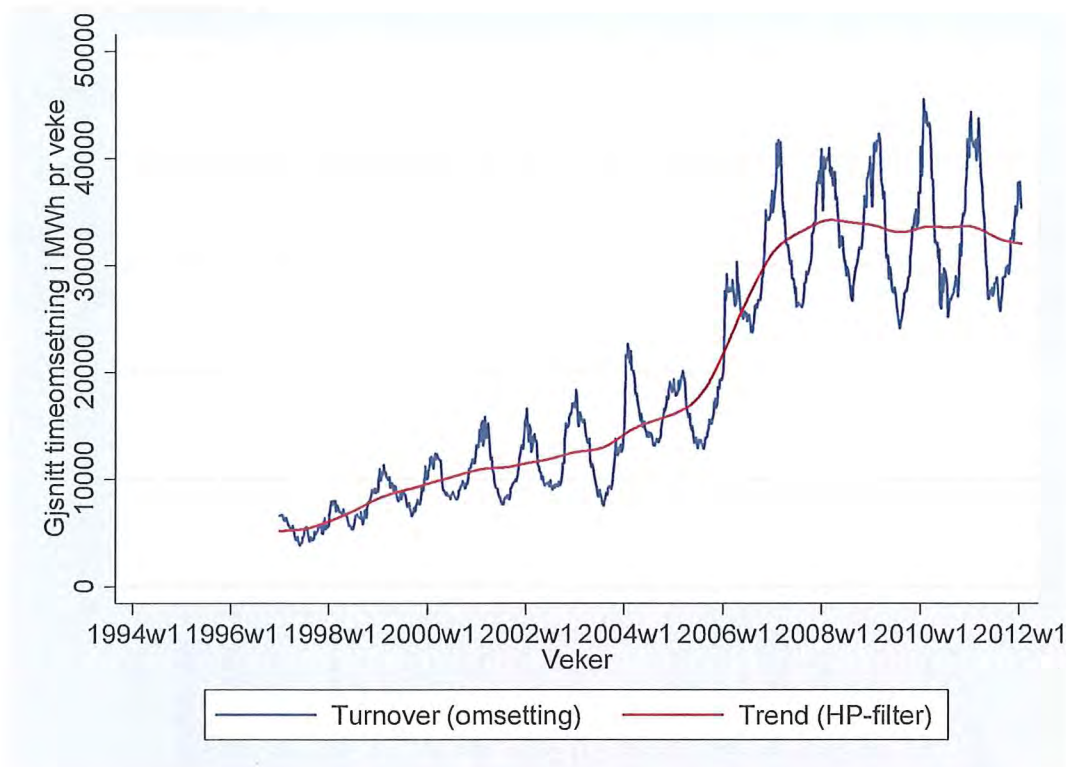
Ein kan og observere at autokorrelasjonen er relativt bratt avtakande til og med lag 20, for deretter å vere meir eller mindre jamn. Det betyr at laga nærmast i tid vil ha mest å seie for prisutviklinga, samt at tidlegare sjokk ikkje forsvinn. Dette er noko ein må ta omsyn til ved modellering for å unngå å ende opp med ein spuriøs regresjon. Ein måte å ta omsyn til dette er å legge inn ein tidsvariabel som noko grovkorna korrigerar for autokorrelasjon i form av å ta omsyn til at observasjonar seinare i tid ofte har høgare verdi enn observasjonar i byrjinga. Ein anna metode for å justere for autokorrelasjon er ved å differensiere. Dette må sjåast i samanheng med om systemprisen er stasjonær og blir drøfta nærmare i neste avsnitt.

3.1.4 Stasjonærheit

Ved å betrakte systemprisen i figur 3 er det nærliggande å spørje seg om den er stasjonær. Dickey-Fuller testar av systemprisen med ulike lag viser at ein må forkaste nullhypotesen om unit root ved 5 % signifikansnivå. Dette indikerar at systemprisen er stasjonær.

Lucía og Schwartz (2002) finn i sin studie av systemprisen at den er stasjonær. Weron (2006) finn derimot at ein ofte kan trekkje motsatt konklusjon med tanke på om elektrisitetsprisar er stasjonære. I tillegg er Dickey-Fuller testar av ein slik karakter at dei er relativt enkle å justere i den forstand at ein då anten kan få bekrefte eller avkrefta det resultatet ein søker etter. Dette er likevel eit meir viktig spørsmål å ta stilling til om ein skal lage prognoser for framtidige spotprisar. Eg brukar difor konklusjonen frå denne testen og påstår at systemprisen er stasjonær, men vedgår at det ligg noko usikkerheit bak denne konklusjonen.

3.2 Turnover



Figur 7: Omsetting

Turnover (omsetting) er også henta frå Nord Pool sin FTP-serverar, sjå figur 7. Den viser gjennomsnittleg omsetting i timen av straum i MWh pr veke. Måten Nord Pool har rekna ut dette er å summere omsetning pr time over ei veke, for deretter å dele dette på 168 (som er talet på timer i ei veke).

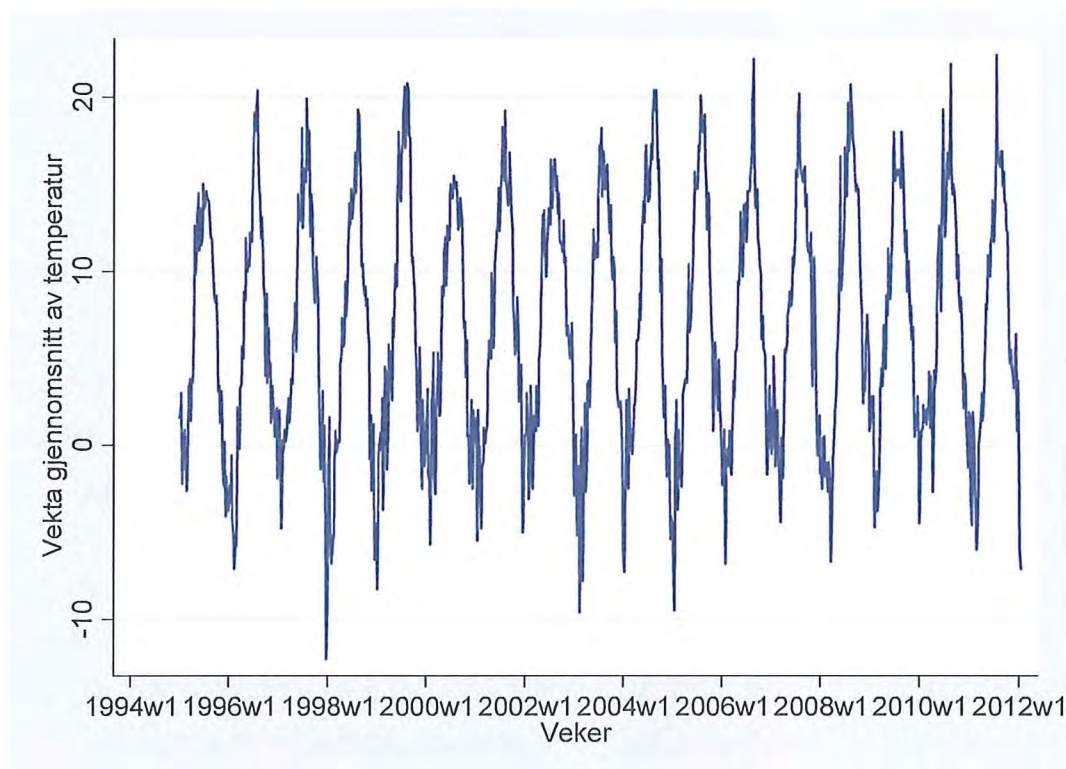
Grunnen til at dei første observasjonane manglar i figuren er fordi Nord Pool starta å registrere omsatt straum 1. januar 1997. Variabelen ser ut til å ha hatt ein relativ stor auke siste åra. Dette kan kanskje være ein indikasjon på at marknadsplassen Nord Pool har vokse, eller at samla straumforbruk har vokse, eller ein kombinasjon av desse to. Ein kan også sjå at denne variabelen har kraftige sesongtendensar gjennom året. Det kan forklarast gjennom høgare sal av straum om vinteren enn om sommaren.

For å jamne ut denne sesongkomponenten har eg brukt eit HP-filter med $\lambda = 270\,400$. Denne λ -verdien finn eg ved å bruke tommelfingerregelen frå likning (3). Eg har brukt 52 som talet på observasjonar i eit år.

Bakgrunnen for at eg brukar eit HP-filter er fordi eg ønskjer å unngå problem knytt til simultanitet. I dette tilfelle kan ein risikere at det oppstår simultanitet knytt til at det er mogleg at systemprisen og turnover gjensidig påverkar kvarandre. Det er ikkje nødvendigvis slik at kausaliteten berre går ein veg og at det dermed berre er omsatt straum som kan påverke systemprisen, det kan også vere andre vegen slik at systemprisen verkar inn på mengda straum som blir omsatt.

Ved å bruke eit HP-filter for å glatte ut trenden i turnover kan ein ta omsyn til dette fenomenet. Og så lenge det er fleire sesongelement i dei andre uavhengige variablane så vil sesongeffekten i systemprisen framleis bli forklart (Kube & Wadhwa, 2007).

3.3 Temperatur



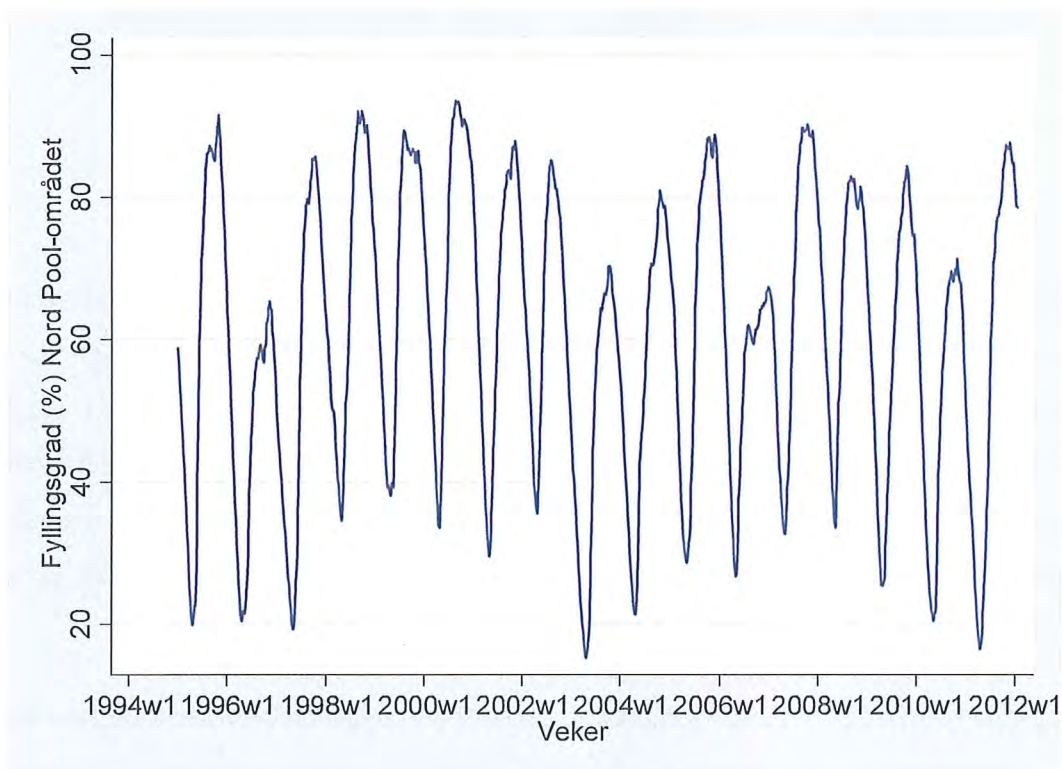
Figur 8: Temperatur i grader celsius

Temperatur er samla inn frå eKlima som er ein ver- og klimabasert database frå Meteorologisk institutt. Eg har samla inn vekelege temperaturnoteringar frå Oslo, Bergen og Tromsø. Dette er så vekta etter folketal, sjå figur 8. Grunnen til at eg har valt desse byane er for å prøve å fange opp noko av det relativt spreidde og varierte klima- og busetjingsmønsteret i Noreg, samstundes som dette er vekta etter folketal. Ved å fokusere

utelukkande på byar sør og aust i landet ville eg mest sannsynlig ikkje ha fanga opp denne dynamikken. Samstundes ville det nok også ha styrka analysen å ta med fleire byar og også å inkludere resten av Norden.

På den andre sida kan det argumenterast med at heile temperaturmaterialet likevel skal vektast ned til ein enkelt variabel, og at det difor er avgrensa kor mykje informasjon om ulike temperaturområde og temperatursvingingar som kan reflekterast i denne eine variabelen. Vidare kan det leggest til at gjennomsnittstemperaturen i Norden på mange måtar ikkje er så veldig heterogen i den forstand at heile området meir eller mindre ligg på same breiddegrad (bortsett frå Danmark), og difor deler eit noko liknande temperert klima i sør og eit mildare arktisk klima i nord.

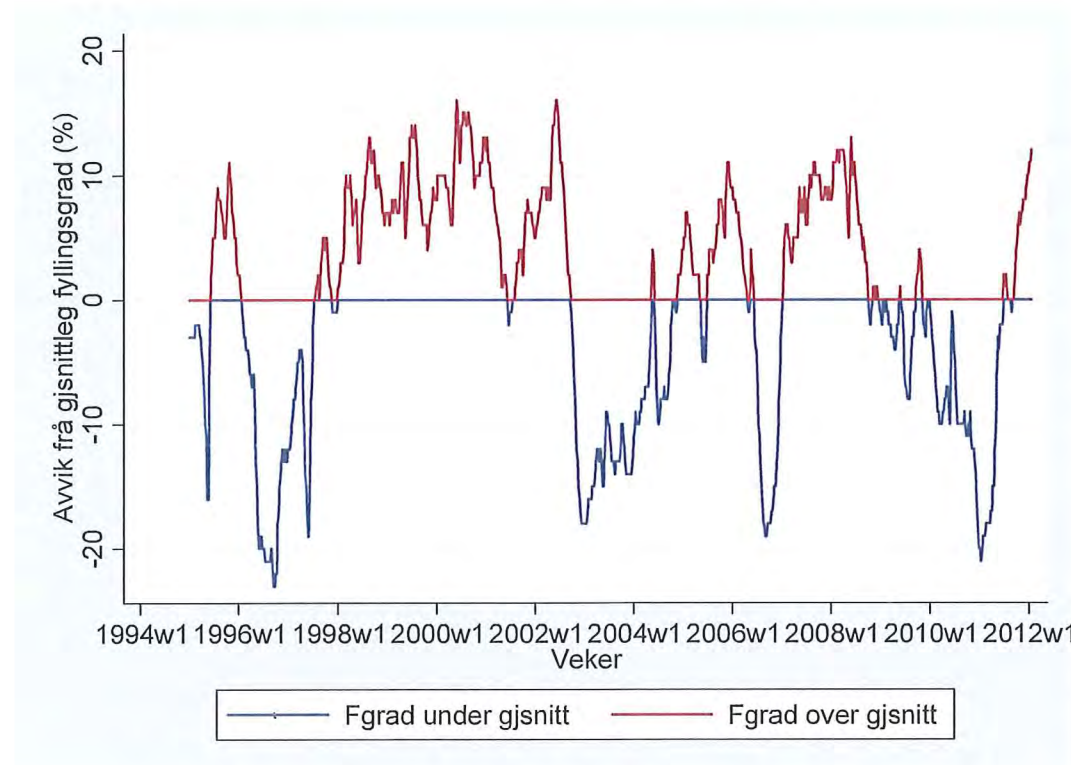
3.4 Fyllingsgrad



Figur 9: Fyllingsgrad

Fyllingsgrad i prosent for Noreg, Sverige og Finland er henta frå Nord Pool sin FTP-server og vist i figur 9 ovanfor. Dette er rekna ut i frå summen av fyllingsgrad i GWh pr veke i dei ulike nordiske landa delt på summen av total kapasitet i GWh. Denne måten å rekne ut total

yllingsgrad i Norden gir då Noreg, som har 95 % av elektrisitetsproduksjonen frå vasskraft eller hydrologiske kjelder i følgje tabell 1, ei relativ stor vektning.

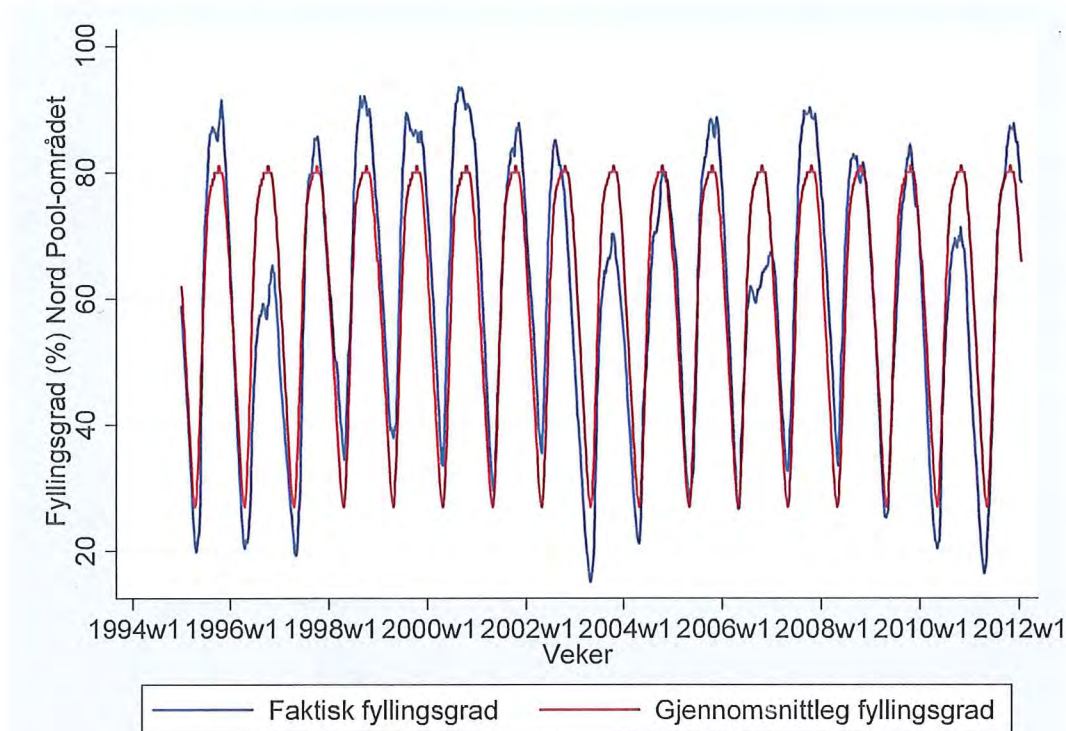


Figur 10: Fyllingsgrad over og under gjennomsnitt (gjennomsnitt = 0)

Også fyllingsgrad viser seg å ha store sesongbaserte svingingar. For betre å forstå denne dynamikken og å sjå på kva måte låg og høg fyllingsgrad verkar inn på systemprisen har eg forsøkt å splitte denne i to komponentar, på same måte som Lucia og Torró (2008). Eg har då generert to nye variablar. Ein variabel som inneheld negative avvik frå gjennomsnittleg fyllingsgrad pr veke, og ein variabel som inneheld positive avvik frå gjennomsnittleg fyllingsgrad pr veke, sjå figur 10. Merk at gjennomsnittleg fyllingsgrad pr veke er ein størrelse som endrar seg frå veke til veke gjennom året.

Meir inngåande forklart er desse to variablane rekna ut på bakgrunn av differansen mellom observert fyllingsgrad i kvar einaste veke frå 1995 til 2011 og gjennomsnittleg fyllingsgrad pr veke. Då har eg først rekna ut gjennomsnittleg fyllingsgrad i veke 1, gjennomsnittleg fyllingsgrad i veke 2, gjennomsnittleg fyllingsgrad i veke 3 osv. Ein kan då seie at ein på denne måten finn ut kva fyllingsgrada villa ha vore i eit ”gjennomsnittså”. Dei to nye variablane, merka som høvesvis over og under i den empiriske analysen, vil difor då

innehalde høvesvis den positive og negative differansen mellom faktisk observert fyllingsgrad og fyllingsgraden i dette konstruerte ”gjennomsnittsåret”.



Figur 11: Faktisk fyllingsgrad og gjennomsnittleg fyllingsgrad

Ein kan så samanstille den faktiske fyllingsgradane med den konstruerte fyllingsgradane i dette ”gjennomsnittsåret”, for å få ei betre grafisk forståing knytt til kor store avvika reelt sett er, sjå figur 11.

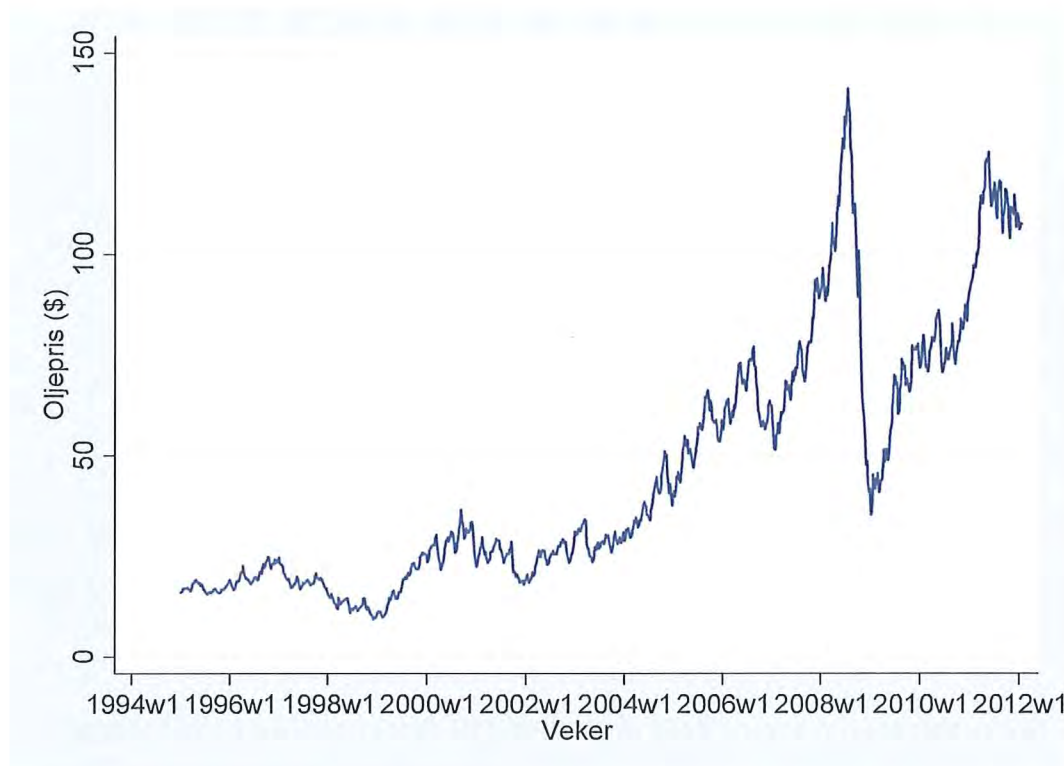
Ein kan sjå av figuren at i dei fleste åra avviker faktisk observert fyllingsgrad med ”gjennomsnittsåret” når ein er i ytterpunkta av fyllingsgradintervallet. Det vil seie at når faktisk observert fyllingsgrad er anten veldig høg eller veldig låg, noko som ein av figuren kan få inntrykk av at oppstår meir eller mindre i dei same vekene i løpet av året¹⁰, så er det sannsynleg at avviket frå ”gjennomsnittsåret” då er relativt stort. Men ”gjennomsnittsåret” er også rekna ut av det dette datamaterialet, noko som er ein indikasjon på at det ligg føre stor heterogenitet knytt til kva som er veldig høg og veldig låg fyllingsgrad i dei ulike åra. Reint

¹⁰ Det kan også bekreftast frå det empiriske datamaterialet, og av å observere figur 9, at fyllingsgrad er ein tilnærma form for ein periodisk funksjon gjennom året.

praktisk betyr det at eit lågt nivå på fyllingsgrada i prosent eit år ikkje nødvendigvis samsvarar med eit lågt nivå på fyllingsgrad i eit anna år.

3.5 Oljepris

Oljeprisen i dollar pr fat er henta frå Thomson Reuters, sjå figur 12, og den er tatt med for å modellere inn ein alternativ energipris. Energiprisar har ein tendes til å trekkje i same retning, og det kan vere nyttig å ta med ein faktor som tek omsyn til den generelle energiprisutviklinga for nærliggande substituttar.



Figur 12: Oljepris (\$) pr fat

Dei siste åra har det vore konsensus om at oljeprisen er høg og at den kanskje har skifta til eit vedvarande høgt nivå. Av figuren kan ein få inntrykk av at dette skiftet kan ha kome rundt 2005 eller 2006. Etter dette har oljeprisen vore jamt høgare en tidlegare, bortsett frå ved det midlertidige fallet i månedane etter Lehman Brothers konkursen 15. september 2008. Om ein skal tolke dette skiftet må det bety at olje og der i gjennom energi i større grad enn tidlegare er dyrare og ein meir pressa faktor.

4. Empirisk analyse

Wooldridge (2009) skisserer dei to mest vanlege framgangsmåtane når ein stegvis skal bygge ein empirisk økonometrisk modell. Ein kan anten byrje med mange variablar og stegvis fjerne dei variablane som ikkje er signifikante eller ein kan byrje med nokre få variablar og stegvis legge til fleire etter kvart. Vidare presiserar han at ulempa med desse stegvise metodane er at den endelege modellen ofte avhenger av rekkefølga når dei ulike variablane vart lagt til eller droppa. I verste fall kan det føre til at ikkje-signifikante variablar framstår som signifikante

Eg har brukt den sistnemnde metoden og har dermed byrja med dei aller enklaste spesifikasjonane, og vidare utvida desse med fleire faktorar og andre innfallsvinklar etter kvart. Eg har også valt ikkje berre å presentere dei mest komplekse og endelege versjonane av modellane i oppgåva. Det er fordi eg opplever at både dei enklaste modellane og nokre av dei alternative modellane kvar for seg gir ulik innsikt i fundamentale trekk ved systemprisen. Likevel, av omsyn til oppgåva sitt omfang, så er ikkje alle modellane og spesifikasjonane inkludert.

I siste del av den empiriske analysen vil eg sjå nærmare på regresjonsvaliditet og modelltilpassing.

Alt av analysearbeid er gjennomført ved hjelp av statistikkprogrammet Stata.

4.1 Modellar

I denne delen vil eg ta for meg sju regresjonsmodellar. Desse modellane er med fordi dei alle på sin måte seier noko fundamentalt om systemprisen. Før eg presenterar modellane vil eg kort sjå på korrelasjon. Dette er nødvendig for å få ein peikepinn på korleis dei ulike modellane bør spesifiserast.

4.1.1 Korrelasjon

Når ein skal bygge ein regresjonsmodell er det mange faktorar som kan trekkjast inn. For å få ein indikasjon på kva faktorar som er dei viktigaste med tanke på å forklare systemprisen kan det vere ein ide å studere ei korrelasjonsmatrise.

| Forklaring på variabelforkortingar | | | |
|------------------------------------|----------------------------|--------|--------------------------------------|
| sysp | Systempris pr veke | hpturn | Gjnsnitt oms glatta ut med HP-filter |
| temp | Vekta gjnsnitt temperatur | olje | Oljepris (\$) per fat |
| turn | Gjnsnitt omsetting pr veke | fgrad | Fyllingsgrad Nord Pool-området |

Tabell 4: Forklaring

| | sysp | temp | turn | hpturn | olje | fgrad |
|--------|-------|-------|-------|--------|-------|-------|
| sysp | 1 | | | | | |
| temp | -0.10 | 1 | | | | |
| turn | 0.73 | -0.18 | 1 | | | |
| hpturn | 0.69 | 0.06 | 0.95 | 1 | | |
| olje | 0.63 | 0.11 | 0.81 | 0.89 | 1 | |
| fgrad | -0.24 | 0.29 | -0.13 | -0.05 | -0.02 | 1 |

Tabell 5: Korrelasjonsmatrise

Korrelasjonsmatrisa, sjå tabell 5, seier noko om den lineære samvariasjonen mellom dei ulike faktorane. Ein ser at systemprisen er negativt korrelert med fyllingsgrad og temperatur, noko som er forventa. At absoluttverdien til fyllingsgrad er høgare enn temperatur er heller ikkje uventa, og fortel oss at fyllingsgrad og systempris i større grad samvarierer enn temperatur og systempris. Oljeprisen viser sterk positiv korrelasjon med systemprisen. Dette er i tråd med intuisjonen om at energiprisar ofte trekk i same retning, og at den aktuelle konjunktursituasjonen ofte vil indikere denne retninga.

Turnover viser også ein sterk positiv korrelasjon med systemprisen. Dette er noko uventa då ein skulle tru at auka omsetning eller likviditet på ein børs ville føre til ein nedgang i prisen på det omsette aktivumet. På ei anna side er det grunn til å tru at denne variabelen fangar opp ein underliggende trend knytt til ei generell prisauke på straum over tid. Den relativ høge korrelasjonskoeffisienten kan kanskje også forklarast i at det er tydelege sesongelement knytt til årstidene i både systemprisen og turnover-variabelen.

Ein ser også at oljepris og turnover er sterkt korrelert med kvarandre. Dette indikerer at ein får eit multikollinearitetsproblem om ein inkluderar begge variablane i ein regresjonsmodell,

og ein må difor nøye seg med å bruke berre ein av desse. Tommelfingerregelen seier at om to variablar har ein korrelasjonskoeffisient på 0,6 eller høgare, då vil det oppstå multikollinearitet om ein inkluderar begge variablane i ein regresjonsmodell.

4.1.2 Regresjonsmodell 1

Korrelasjonsmatrisa viser kva for faktorar som samvarierer mest med systemprisen, og det er eit godt utgangspunkt med tanke på å bygge ein deskriptiv modell.

$$(4) \quad \text{sysp} = \beta_0 + \beta_1 \cdot \text{fgrad} + \beta_2 \cdot \text{sesfakt}$$

Modellen ovanfor er enkel, og gir eit første overblikk over korleis dei viktigaste faktorane spelar inn. Sysp er systemprisen på vekebasis frå 1995 til og med 2011 og er den avhengige variabelen i denne og alle påfølgande regresjonar. Fgrad viser til fyllingsgrad i prosent i same tidsrom. Sesfakt er ein dummy som har verdien 1 om det er sommar og verdien 0 om det er vinter. Sommar er definert å gjelde frå veke 18 – 40 og vinter frå veke 41 - 17 (Johnsen, et al., 1999).

| sysp | Coef. | Std. Err. | T | P> t |
|---------------------|--------|-----------|------------------------|------|
| fgrad | -1.20 | 0.1985 | -6.07 | 0.0 |
| sesfakt | -27.04 | 8.3323 | -3.25 | 0.0 |
| _cons | 324.01 | 12.4627 | 26 | 0.0 |
| Number of obs = 887 | | | Adj R-squared = 0.0593 | |

Tabell 6: Resultat frå regresjon av modell 1

Denne modellen har svært låg forklaringskraft, men likevel er alle koeffisientane signifikante. På grunn av den låge forklaringsgrada bør ein vere varsam med å tolke resultata veldig strengt. Frå tabell 6 ser ein at forteikna til begge parameterane er negative og det kan difor gi ein peikepinn på korleis desse faktorane verkar inn på systemprisen. Dei negative forteikna indikerer, ceteris paribus, at systemprisen er lågare om sommaren enn om vinteren, og at om fyllingsgrada aukar så vil systemprisen reduserast.

I samband med variabelen fyllingsgrad har eg utelate tilsig og snøsmelting. Denne variabelen kunne eg også ha lasta ned frå Nord Pool sin FTP-server. Grunnen til at eg har utelate denne er fordi eg ville fått eit endogenitetsproblem om eg hadde hatt med både fyllingsgrad og tilsig og snøsmelting same modell i og med at tilsig og snøsmelting vil påverke fyllingsgrad i vassmagasina. Dette kunne eventuelt blitt løyst ved å lage

instrumentvariablar i regresjonen, og dermed unngått eventuelle skeive parameter knytt til utelatne variablar. Eg har derimot valt å legge meg på same linje som Steen (2004) som argumenterar med at fyllingsgrad er meir presist å bruke i en modell enn tilsig og snøsmelting.

4.1.3 Regresjonsmodell 2

I neste modell blir to variablar, temperatur og turnover, som korrelerar med systemprisen inkludert.

$$(5) \quad \text{sysp} = \beta_0 + \beta_1 \cdot \text{fgrad} + \beta_2 \cdot \text{temp} + \beta_3 \cdot \text{turn}$$

I og med at sesongfaktor og temperatur har høg innbyrdes korrelasjon inkluderar eg berre sistnemnde når eg no utvidar modellen. Dette for å unngå multikollinearitet.

| sysp | Coef. | Std. Err. | t | P> t |
|---------------------|--------|------------------------|-------|------|
| fgrad | -1.00 | 0.1553 | -6.45 | 0.0 |
| temp | 1.29 | 0.4588 | 2.80 | 0.0 |
| turn | 0.01 | 0.0003 | 29.39 | 0.0 |
| _cons | 143.27 | 11.6321 | 12.32 | 0.0 |
| Number of obs = 783 | | Adj R-squared = 0.5513 | | |

Tabell 7: Resultat frå regresjon av modell 2

Denne modellen har ei forklaringsgrad på 55 %, sjå tabell 7. Dette er vesentleg høgare enn den førre modellen. Mengda av observasjonar er og blitt redusert. Det er fordi det ikkje var mogleg å oppdrive observasjonar knytt til turnover før 1997 som nemnt under skildring av datamaterialet. Ein ser også at alle variablane er signifikante.

Det som er påfallande i denne modellen er at turnover blir veldig signifikant (t-verdi nesten 30), og det er ikkje nødvendigvis bra. Om ein variabel forklarar veldig mykje kan det vere uheldig for dei andre variablane som då kan bli forvrongde samt at dette ofte kan vere ein indikasjon på at det ligg føre autokorrelasjon.

I tillegg kan det vere at denne høge signifikansen ikkje er korrekt. Det kan nemleg vere at denne variabelen eigentleg plukkar opp ein underliggande trend mellom den avhengige og den uavhengige variabelen i staden for å påvise at den avhengige variabelen blir bestemt av den uavhengige variabelen. Dette blir kalla ein spuriøs regresjon (Wooldridge, 2009).

Den underliggende trenden som driv desse variablane kan vere knytt til inflasjon når det gjeld systemprisen. For turnover kan den vere knytt til auka handel og aktivitet som igjen kan vere eit resultat av generell modning, og det at fleire aktørar med tida etablerar seg på Nord Pool. Eit anna argument som styrkjer denne hypotesen er at det nemnte forteiknet til turnover er positivt. Ut i frå økonomisk teori skulle ein forvente at auka sal av straum og auka likviditet i ein marknad pressa ned prisen på eit aktivum, og at forteiknet difor burde ha vore negativt.

Ei enkel løysing på problemet knytt til trend kan vere å legge inn ein tidsvariabel. Det kan til dømes vere ein variabel som byrjar på ein og vidare tel talet på veker frå 1995 – 2011. Denne variabelen vil då representere generell prisauke på straum frå veke til veke i dette tidsrommet. Det vil føre til at noko av denne trendeffekten vil bli isolert vekk frå dei andre variablane, og difor kanskje også redusere noko av signifikansnivået til dei variablane som har trendkomponentar i seg. Dette vil bli testa ut nærmare i den neste modellen.

4.1.4 Regresjonsmodell 3

I denne modellen blir tidsvariabelen (t) inkludert. I tillegg blir fyllingsgrad delt i to variablar som skissert under skildring av data.

$$(6) \quad \text{syp} = \beta_0 + \beta_1 \cdot \text{over} + \beta_2 \cdot \text{under} + \beta_3 \cdot \text{temp} + \beta_4 \cdot \text{turn} + \beta_5 \cdot t$$

Grunnen til at fyllingsgrad blir delt i to er både for å justere for sesong samt for å kunne tolke koeffisientane meir presist. Ein får då ei betre forståing for korleis høg og låg fyllingsgrad og dermed kapasitetsdynamikken i vassmagasina verkar inn på systemprisen.

| syp | Coef. | Std. Err. | T | P> t |
|---------------------|-------|------------------------|--------|------|
| over | -5.71 | 0.6118 | -9.33 | 0.0 |
| under | -7.54 | 0.4815 | -15.67 | 0.0 |
| temp | -0.48 | 0.3736 | -1.30 | 0.2 |
| turn | 0.01 | 0.0005 | 9.89 | 0.0 |
| t | 0.10 | 0.0266 | 3.90 | 0.0 |
| _cons | 96.59 | 7.4803 | 12.91 | 0.0 |
| Number of obs = 783 | | Adj R-squared = 0.7672 | | |

Tabell 8: Resultat frå regresjon av modell 3

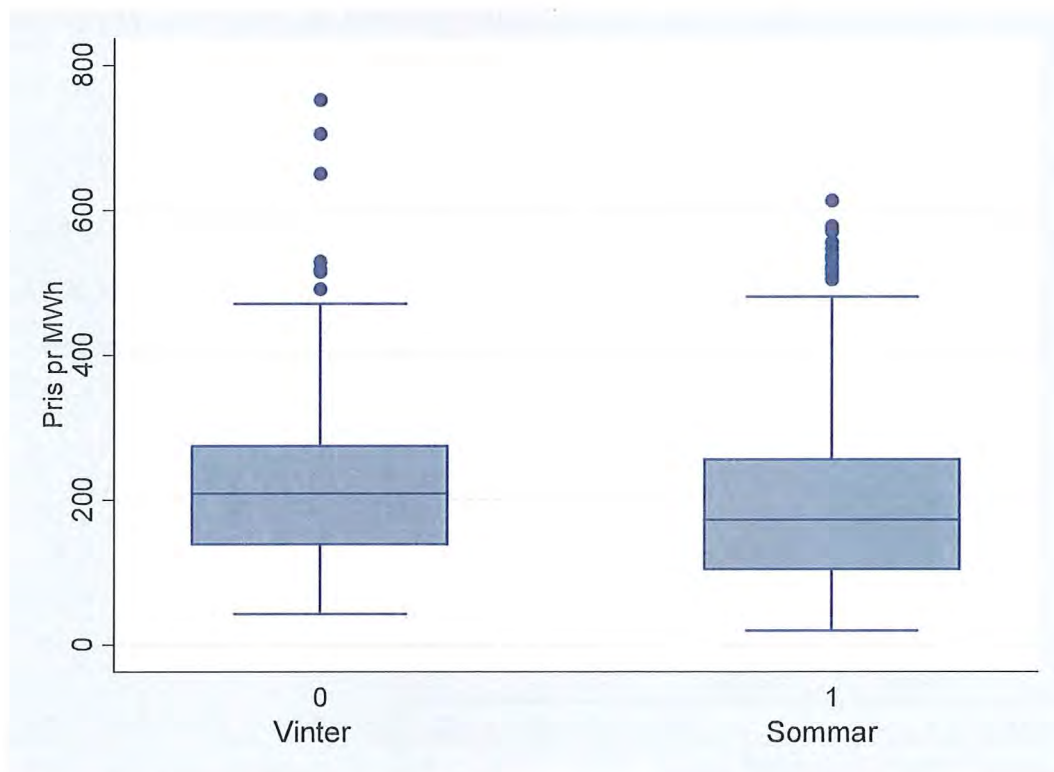
Denne modellen har ei forklaringsgrad på 77 %, noko som er enno høgare enn i dei tidlegare modellane. Ein ser og at alle variablane utanom temperatur er signifikante. Turnover er

framleis signifikant, men t-verdien er mykje mindre enn i førre modell. Det kjem av at det no er blitt lagt til ein tidsvariabel (t) som fangar opp den underliggande trenden.

Ut i frå tabellen ser ein at fyllingsgrad under gjennomsnitt (under) påverkar systemprisen meir enn fyllingsgrad over gjennomsnitt (over). Det ser ein av at under har ein høgare negativ verdi enn over.

Av å observere systemprisen over fleire år, og fordi det vart påvist i første modell at sommar gir signifikant lågare pris, er det naturleg å sjå nærmare på i kor stor grad det ligg føre sesongelement i systemprisen. Det er også naturleg å sette dette i samanheng med eventuelle sesongkomponentar i dei andre variablane.

Det er eit faktum at det er relativt lik temperatur i dei ulike årstidene gjennom fleire år, og difor er det også ikkje ubetydelege sesongelement i temperaturvariabelen. Ved å observere fyllingsgrad og turnover ser ein også at desse variablane går i tilnærma bøljeliknande former, samt inneheld sistnemnde ein mogleg trend. Det er med andre ord ikkje berre den avhengige variabelen som har sesong i seg.



Figur 13: Box plot av systemprisen

Ved å studere eit box plot, sjå figur 13, av systemprisen der ein deler datasettet inn etter sesongfaktoren som introdusert i den første modellen, får ein også ein indikasjon på at det ligg føre sesongelement. Når dummyen har verdien 1 (sommar) kan ein observere at medianen til systemprisen er lågare enn om dummyen har verdien 0 (vinter). Dette bekreftar også funnet frå modell 1 som viser at systemprisen er lågare om sommaren. Box plotet seier også noko om uteliggjarar. Dette er noko som oppstår både om sommaren og om vinteren, men som ein kan observere av figuren så er dei mest avvikande uteliggjarane om vinteren. Dette betyr at dei høgaste og mest ekstreme straumprisane har vore om vinteren.

Neste steg er så å ta stilling til om ein skal korrigere for dette sesongfenomenet eller ikkje. Ein metode er å bruke ein sesongdummy som i den første regresjonsmodellen. Ein anna måte metode er å bruke trigonometriske teknikkar (Torvestad & Johannessen, 2010) for deretter å modellere inn eit periodisk parameter som tek omsyn til sesong. Eg har også forsøkt nokre andre innfallsvinklar for å korrigere for dette fenomenet. Mellom anna har eg forsøkt å bruke sesongfaktorar som er konstruert ut i frå årstider basert på glidande gjennomsnitt. Eg har forsøkt å glatte ut trenden i systemprisen ved hjelp av å sette inn ulike λ -verdier i eit HP-filter, og vidare bruke dei periodiske svingingane som då blir isolert ut og hamnar i restleddet frå estimeringa av trendlinja. Og eg har forsøkt å bruke fleire ulike kategoriske variablar. Ingen av desse forsøka har etter mine vurderingar gitt tilfredsstillande resultat.

Hjalmarsson (2000) tek derimot ikkje omsyn til sesong i det heile og argumenterar med at så lenge dei uavhengige variablane gjenspeiler sesongelementet i den avhengige variabelen, og at sesongen i begge variablane er tilsvarande, så er det ikkje nødvendig å korrigere for dette. Eg konkluderer difor med at det ikkje er nødvendig å skilje ut sesongkomponenten i ein eigen variabel i dei påfølgande modellane.

4.1.5 Regresjonsmodell 4

Som diskutert under skildring av data kan det ligge føre eit endogenitetsproblem (simultanitet). Dette kan føre til skeive, men framleis konsistente, koeffisientar. For å korrigere for dette er eit alternativ å bruke eit HP-filter.

$$(7) \quad \text{sysp} = \beta_0 + \beta_1 \cdot \text{over} + \beta_2 \cdot \text{under} + \beta_3 \cdot \text{temp} + \beta_4 \cdot \text{hpturn} + \beta_5 \cdot t$$

I denne modellen har eg difor bytta ut turnover med ein ny HP-justert versjon av denne variabelen.

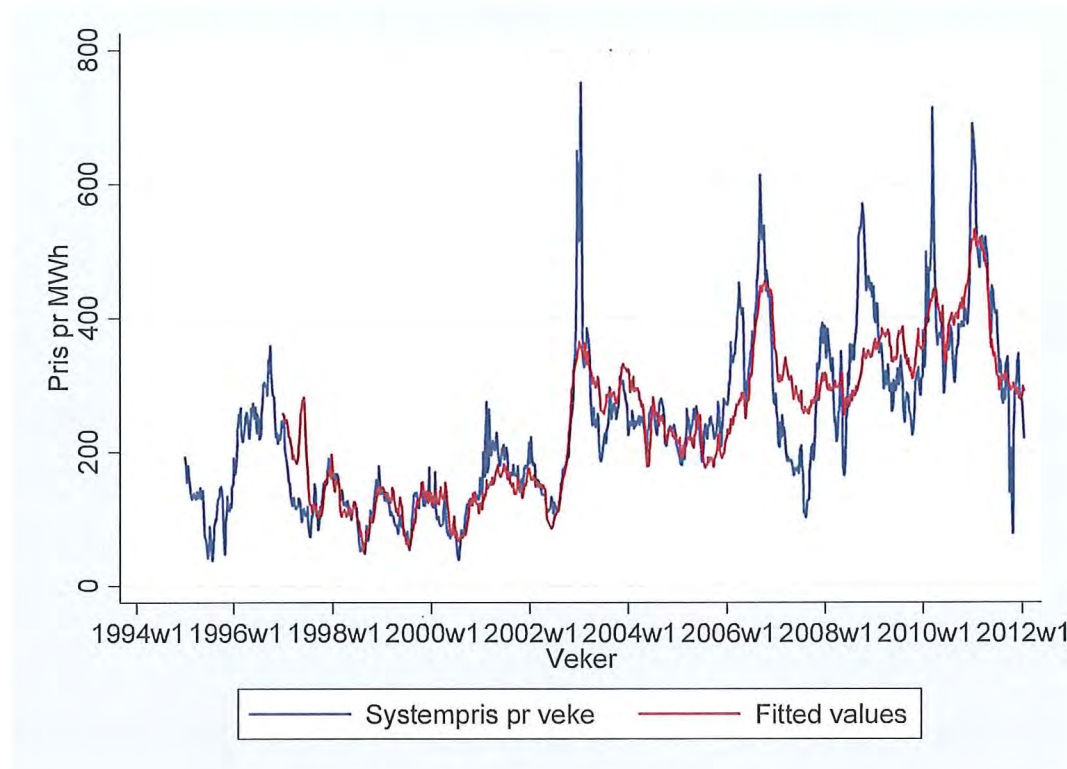
| syp | Coef. | Std. Err. | T | P> t |
|---------------------|----------|-----------|------------------------|------|
| over | -5.42 | 0.6338 | -8.55 | 0.0 |
| under | -7.89 | 0.5022 | -15.71 | 0.0 |
| temp | -2.44 | 0.3257 | -7.48 | 0.0 |
| hpturn | 0.00 | 0.0007 | 6.17 | 0.0 |
| t | 0.14 | 0.0347 | 4.10 | 0.0 |
| _cons | 103.7463 | 8.1198 | 12.78 | 0.0 |
| Number of obs = 783 | | | Adj R-squared = 0.7501 | |

Tabell 9: Resultat frå regresjon av modell 4

Frå tabellen ser vi at denne modellen har ei forklaringskraft på 75 %, noko som er marginalt lågare enn førre modell. Alle parametrane er også signifikante. Både parameteren og signifikansnivået til turnover er no lågare og dei andre parametrane har også endra seg noko. Dette kan tolkast som at dei andre parametrane i større grad forklarar sesongfaktoren i systemprisen i og med at turnover ikkje gjer det lenger. I tillegg er ein no kvitt endogenitetsproblemet som oppstår i og med at volum og pris mest sannsynleg blir bestemt simultant.

Basert på størrelsesorden på over og under ser ein også i denne modellen at fyllingsgrad under gjennomsnittet påverka systemprisen meir enn fyllingsgrad over gjennomsnittet. Det vil seie at ein har ein asymmetri når det er press i kapasiteten i eine retninga, i form av at låg kapasitet i forhold til gjennomsnittleg fyllingsgrad påverkar meir enn høg kapasitet i forhold til gjennomsnittleg fyllingsgrad.

Dette verkar også logisk ut i frå økonomisk intuisjon. I tillegg ser ein at begge desse variablane, som i førre modell, har negative forteikn, noko som indikerar at om fyllingsgrada aukar (uavhengig om den er over eller under gjennomsnitt), så vil systemprisen falle.



Figur 14: Modelltilpassing

Av figuren ovanfor kan ein få eit grafisk innblikk av i kor stor grad regresjonslinja tilpassar seg det observerte datamaterialet. Ein kan sjå at sjølv om forklaringsgrada til modellen er relativ høg, så er regresjonen framleis ikkje perfekt. Det er spesielt dei siste åra, og då dei ekstreme prisane i begge retningar, som modellen har vanskeleg for å plukke opp.

4.1.6 Regresjonsmodell 5

Av korrelasjonsmatrisa kom det fram at turnover og oljepris har ein korrelasjonskoeffisient på over 0,8. Det betyr at dei samvarierer sterkt og at ein får eit problem knytt til multikollinearitet om begge desse variablane blir inkludert i ein modell. Ein bør difor velje å ha med berre ein av desse i ein regresjonsmodell.

$$(8) \quad \text{sysp} = \beta_0 + \beta_1 \cdot \text{over} + \beta_2 \cdot \text{under} + \beta_3 \cdot \text{temp} + \beta_4 \cdot \text{olje} + \beta_5 \cdot t$$

For å forstå korleis oljeprisen, som er ei alternativ energikjelde og som ofte beveger seg i same retning som ei rekke andre fossile energikjelder, påverkar systemprisen blir denne inkludert.

| syp | Coef. | Std. Err. | T | P> t |
|---------------------|--------|------------------------|--------|------|
| over | -6.28 | 0.5686 | -11.04 | 0.0 |
| under | -7.38 | 0.4123 | -17.91 | 0.0 |
| temp | -2.91 | 0.2962 | -9.84 | 0.0 |
| olje | 1.18 | 0.1342 | 8.77 | 0.0 |
| t | 0.19 | 0.0164 | 11.39 | 0.0 |
| _cons | 119.43 | 6.0053 | 19.89 | 0.0 |
| Number of obs = 887 | | Adj R-squared = 0.7582 | | |

Tabell 10: Resultat frå regresjon av modell 5

Av tabellen ovanfor kan ein sjå at denne modellen får ei forklaringsgrad på 76 %. I tillegg er alle variablane signifikante. Ein kan også sjå at talet på observasjonar er høgare her. Det er fordi oljeprisen, i motsetning til turnover, er registrert tilbake til 1995. Også her, på same måte som ved inkludering av turnover i dei tidlegare modellane, får oljeprisen ein svært høg t-verdi om ein ikkje har med ein tidsvariabel (t).

Det er også mogleg å tolke parametrane i tabellen direkte. Ein kan då hevde at om ein aukar temperaturen med 10 grader celsius, ceteris paribus, så vil systemprisen reduserast med cirka 0,03 NOK/kWh.¹¹ Og om ein aukar fyllingsgrada i vassmagasina, så vil systemprisen reduserast med høvesvis cirka 0,06 NOK/kWh og cirka 0,07 NOK/kWh, avhengig av om fyllingsgrad er høvesvis over eller under gjennomsnittleg fyllingsgrad.

Koeffisienten til oljepris blir positiv. Det betyr, ceteris paribus, at om oljeprisen aukar med 10 US dollar pr fat, så vil systemprisen auke med cirka 0,01 NOK/kWh. Dette verkar logisk i og med at energiprisar ofte trekk i same retning og påverkar kvarandre.

Det er viktig å understreke at det heftar stor usikkerheit ved desse tolkingane. For det første så er det relativt enkelt å endre desse resultata ved å legge til eller trekke frå ein eller fleire variablar, noko ein kan få eit inntrykk av ved å sjå på korleis koeffisientane har endra seg i forhold til dei andre modellane. For det andre så er det også eit samspel mellom desse variablane, så å endre berre ein variabel vil også vere noko kunstig. Til dømes så vil ei temperaturauke sannsynligvis også føre til at fyllingsgrad aukar gjennom meir tilsig frå snøsmelting. Koeffisientane seier likevel noko om retning, og delvis noko om størrelsesordenen på endringane isolert sett.

¹¹ Systemprisen er oppgitt i MWh så ein må dele på tusen for å finne kWh

Dei andre koeffisientane har ikkje endra seg vesentleg frå førre modell. Ein kan likevel hevde at i og med at dei faktisk har endra seg noko, så har tolkingar av typen *ceteris paribus* lite verdi fordi det skal så lite til, det er nok med relativt enkle spesifikasjonsendringar, for å påverke dei resterande parameterane i ein regresjonsmodell.

4.1.7 Regresjonsmodell 6

Som tidlegare påvist er det autokorrelasjon i systemprisen, og det kan vere fruktbart å sjå korleis modellen endrar seg ved å inkludere eit autoregressivt ledd av første orden (AR 1).

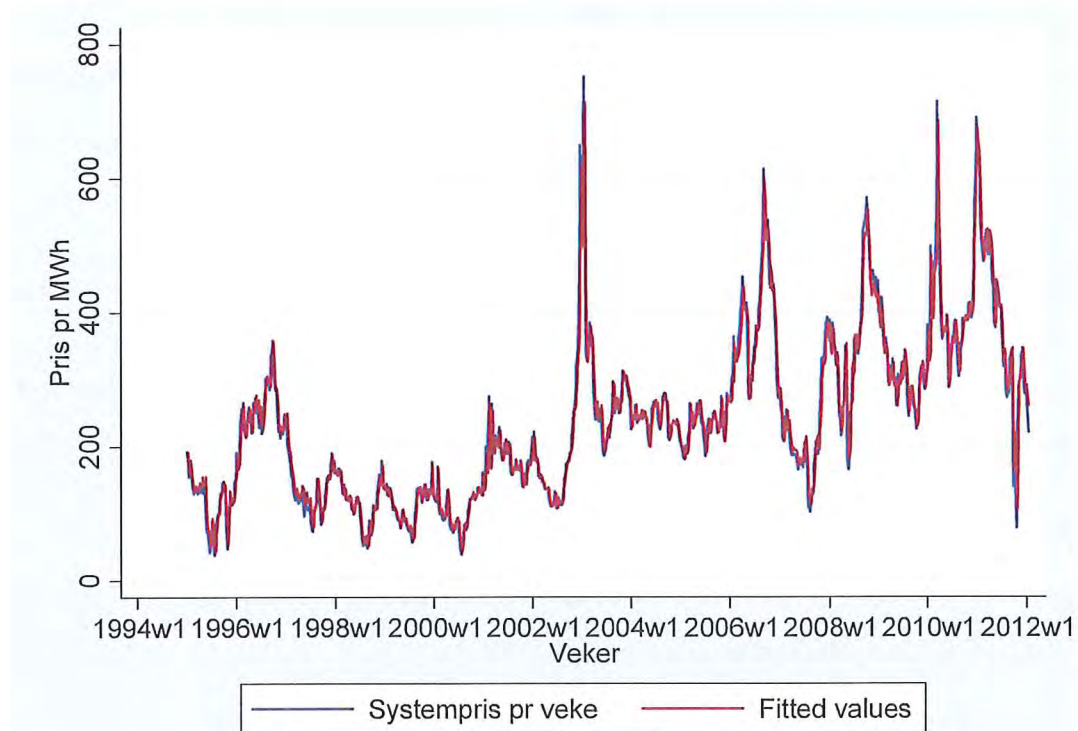
$$(9) \quad \text{syp} = \beta_0 + \beta_1 \cdot \text{over} + \beta_2 \cdot \text{under} + \beta_3 \cdot \text{temp} + \beta_4 \cdot \text{olje} + \beta_5 \cdot \text{syp}_{t-1}$$

Dette betyr i praksis at modellen no er av ein slik karakter at systemprisen førre veke forklarar ein god del av systemprisen denne veka.

| syp | Coef. | Std. Err. | T | P> t |
|---------------------|-------|-----------|------------------------|------|
| over | -0.48 | 0.2935 | -1.64 | 0.1 |
| under | -0.68 | 0.2241 | -3.05 | 0.0 |
| temp | -0.05 | 0.1511 | -0.31 | 0.8 |
| olje | 0.24 | 0.0499 | 4.71 | 0.0 |
| L1.syp | 0.91 | 0.0153 | 59.18 | 0.0 |
| _cons | 11.03 | 3.4949 | 3.15 | 0.0 |
| Number of obs = 886 | | | Adj R-squared = 0.9443 | |

Tabell 11: Resultat frå regresjon av modell 6

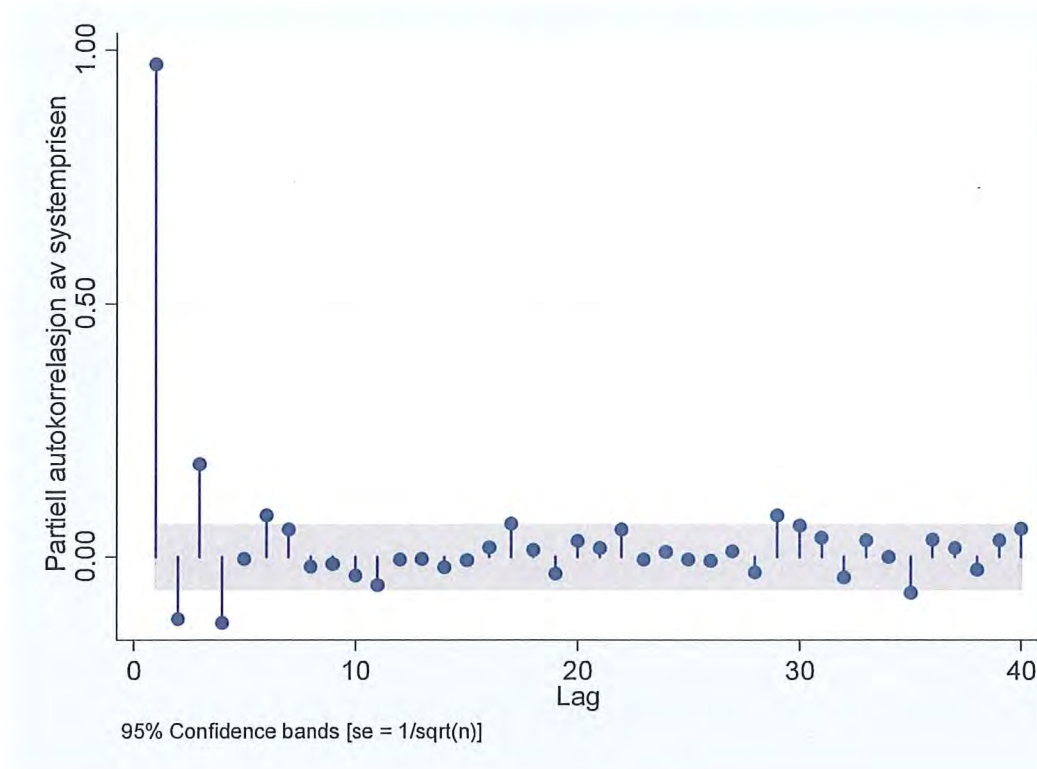
Av tabellen ovanfor kan ein sjå at denne modellen får ei forklaringsgrad på heile 94 %, noko som betyr at denne modellen klarer å fange opp nesten all variasjon i systemprisen. Vidare kan sjå at det autoregressive leddet, oljeprisen, under og konstantleddet er signifikant. I tillegg ser ein at det autoregressive leddet har ein t-verdi på nesten 60, noko som er ekstremt høgt. Det betyr at systemprisen førre veke forklarar svært mykje og vil også såleis vere ein god indikator på systemprisen påfølgande veke.



Figur 15: Modelltilpassing

Av figuren over kan ein også sjå grafisk kor godt modellen er tilpassa. Dei andre parametrane i denne modellen, bortsett frå det autoregressive leddet, gir no ikkje noko økonomisk mening. Så sjølv om nokon av dei framleis er signifikante, så forsøkte eg også (i ein ny modell) å fjerne dei. Eg fekk då ein modell med berre ein uavhengig variabel, nemlig det autoregressive leddet. Også denne modellen fekk ei svært høg forklaringsgrad, og vil difor også tilpasse seg meir eller mindre i same grad som i figuren ovanfor. Dette beviset styrker påstanden om at systemprisen er ein tidsserie og at det ligg føre autokorrelasjon.

Det vart og forsøkt å legge til enno fleire autoregressive ledd når dei opphavlege parametrane frå (9) var fjerna (bortsett frå det autoregressive leddet av første orden). Totalt blei heile 9 ulike modellar forsøkt, det vil seie modellar som var autoregressive av andre og heilt til tiande orden. Resultatet var at i alle modellane så vart det første autoregressive leddet svært signifikant, og også det andre, tredje og fjerde autoregressive leddet (men med betydeleg lågare t-verdi). Dei resterande autoregressive ledda var stort sett ikkje signifikante i dei ulike modellane. Dette er ein indikasjon på at ein AR (4) modell vil vere mest riktig å bruke.



Figur 16: Partiell autokorrelasjonsfunksjon

Det å vite kor mange autoregressive ledd ein skal ta med kan vere utfordrande. Ein enklare og raskare måte å vurdere dette, i staden for å prøve å feile med mange ulike modellar, er ved å studere ein PACF som i figuren over. Denne grafen bekreftar konklusjonane frå førre avsnitt om at ein bør bruke ein modell med fire autoregressive ledd (dette ser ein ved å sjå på kor mange partielle lag som er signifikante).

| Lag | Df | p | AIC | SBIC | FPE |
|---------------------|----|------|---------|---------|--------|
| 0 | | | 12.4777 | 12.4831 | 15364 |
| 1 | 1 | 0.00 | 9.6166 | 9.6274 | 878.92 |
| 2 | 1 | 0.00 | 9.6036 | 9.6198 | 867.59 |
| 3 | 1 | 0.00 | 9.5673 | 9.5890 | 836.67 |
| 4 | 1 | 0.00 | 9.5498 | 9.5769 | 822.18 |
| Number of obs = 883 | | | | | |

Tabell 12: Ulike informasjonskriterium

Det er også andre måtar å vurdere kor mange lag ein bør ha med. Ein kan til dømes bruke ulike informasjonskriterium, sjå tabell 12. Eg finn at ved å vurdere både Akaike information criterion (AIC), Schwarz-Bayesian information criterion (SBIC) og Finite prediction error (FPE) at ein bør bruke ein autoregressiv modell med fire lag. Dette ser ein fordi verdiane er lågast ved fire lag (av plassomsyn er vidare lag fjerna).

Grunnen til at eg likevel har valt ein modell som berre har eit autoregressivt ledd er fordi dei regresjonsmodellane eg prøvde med fleire autoregressive ledd fekk nesten identisk forklaringsgrad og nesten identisk modelltilpassing. Difor meiner eg at omsynet til å spesifisere ein enklast mogleg modell veg tyngst, og difor beheld eg modellen med det autoregressive leddet av første orden.

Konklusjonen er at (9) neppe vil vere brukande til å gi ei god historisk og deskriptiv framstilling, men for prediksjonsformål (spesielt på kort sikt) så vil den vere meir eigna.

4.1.8 Regresjonsmodell 7

Den siste modellen brukar den logaritmiske verdien til systemprisen som avhengig variabel og dei same uavhengige variablane som i modell 5.

$$(10) \quad lsyp = \beta_0 + \beta_1 \cdot over + \beta_2 \cdot under + \beta_3 \cdot temp + \beta_4 \cdot olje + \beta_5 \cdot t$$

Likninga ovanfor blir populært kalla ein log-level modell og ein eigenskap ved denne modellen er at ein kan tolke parameterane relativt. I tillegg kan denne vere meir eigna om ein har heteroskedastisitet i dei tidlegare modellane. Vidare er det grunn til å tru at ein slik modell vil vere meir presis om systemprisen er lognormal, noko som det er noko tvil om som diskutert tidlegare under skildring av datamaterialet.

| lsyp | Coef. | Std. Err. | t | P> t |
|---------------------|--------|------------------------|--------|------|
| over | -0.042 | 0.0023 | -18.04 | 0.0 |
| under | -0.023 | 0.0017 | -13.84 | 0.0 |
| temp | -0.015 | 0.0012 | -12.86 | 0.0 |
| olje | 0.005 | 0.0005 | 8.37 | 0.0 |
| t | 0.001 | 0.0001 | 13.24 | 0.0 |
| _cons | 4.925 | 0.0244 | 202.07 | 0.0 |
| Number of obs = 887 | | Adj R-squared = 0.7923 | | |

Tabell 13: Resultat frå regresjon av modell 7

Denne modellen har ei forklaringsgrad på 79 %. Alle variablane er også signifikante. Det som er påfallande i denne regresjonen er at konstantleddet får ein enorm høg t-verdi, sjå tabell 13. T-verdien er også over tre gongar så høg som den ekstreme t-verdien som det autoregressive leddet fekk i modell 6. Dette er eit signal om at dette kanskje er ein spuriøs regresjon, og at ein difor bør vere varsam med å bruke denne logaritmiske modellen. Om ein skal tolke denne modellen, så kan ein seie at om oljeprisen aukar med ein dollar pr fat, så vil systemprisen, ceteris paribus, auke med 0,5 %.

4.1.9 Oppsummering

I denne delen har eg først sett på korrelasjon for å få ein peikepinn på kva og i kor stor grad ulike faktorar påverkar systemprisen. Då fann eg ut at temperatur og fyllingsgrad har negativ korrelasjon med systemprisen, og at turnover og oljepris har positiv korrelasjon med systemprisen.

Den første modellen inneheldt variablane fyllingsgrad og sesongfaktor. Her vart det påvist systemprisen er lågare om sommaren enn om vinteren, og at om fyllingsgrad aukar så vil systemprisen bli redusert. Dette er i tråd med økonomisk intuisjon og korrelasjonsmatrisa.

I den andre modellen vart temperatur og turnover bytta ut med sesongfaktor. Denne modellen opplevde eg ikkje direkte brukbar og den understreka behovet for å inkludere eit trendledd.

I den tredje modellen vart fyllingsgrad splitta i to komponentar og eit trendledd vart lagt til. Modellen endra seg til å bli vesentleg meir presis i den forstand at ein no kunne få eit betre overblikk over korleis press og kapasitet i fyllingsgrad påverka.

I den fjerde modellen vart turnover bytta ut med ein HP-justert turnovervariabel. Modellen vart no meir realistisk i og med at ein no korrigerar for endogenitet i form av simultanitet i denne variabelen.

I den femte modellen vart den justerte turnovervariabelen bytta ut med oljepris. Begge desse variablane kunne ikkje vere med i den same modellen på grunn av at det då ville ha oppstått multikollinearitet. Ein kan i denne modellen tolke parametrane varsamt.

I den sjette modellen vart det inkludert eit autoregressivt ledd av første orden. Modellen viser at systemprisen førre veke er ein god indikator på systemprisen denne veka. Den modellen opplever eg som best til å prognostisere framover med.

I den sjuande modellen blir systemprisen transformert til logaritmiske verdiar. Dei uavhengige variablane er som i den femte modellen. Her kan parametrane tolkast relativt, og dette er ein noko meir presis modell om systemprisen likevel er lognormal.

Oppsummert er eg av ei oppfatning av at det er modell 4 og modell 5 som er mest reelle i den forstand at dei er mest brukande til å skildre korleis systemprisen blir fastsett. Dei tre

første modellane er ufullstendige og eg opplever dei i større grad som arbeidsmodellar, og som prespesifikasjonar av dei fire siste modellane, sjølv om dei også kartlegg nokre interessante samanhengar. Dei to siste modellane er eg skeptisk til på grunn av svært høge t-verdiar. I modell 6 får det autoregressive leddet svært høg t-verdi og i modell 7 får konstantleddet svært høg t-verdi. Dette gjer desse modellane mindre eigna for å vurdere korleis fundamentale faktorar påverkar systemprisen. Likevel, om fokuset i oppgåva hadde vore å sjå framover og å forsøke å lage prognosar, hadde modell 6 ikkje vore ueigna. Det er mellom anna fordi den har ei svært høg forklaringsgrad.

4.2 Regresjonsvaliditet

Det er viktig å studere validiteten til modellane. Eg vil i denne delen sjå nærmare på feilledd, strukturelle ulikskapar og spesifikasjon. Av omsyn til oppgåva sitt omfang ser eg grundig på ein modell i staden for å vurdere fleire.

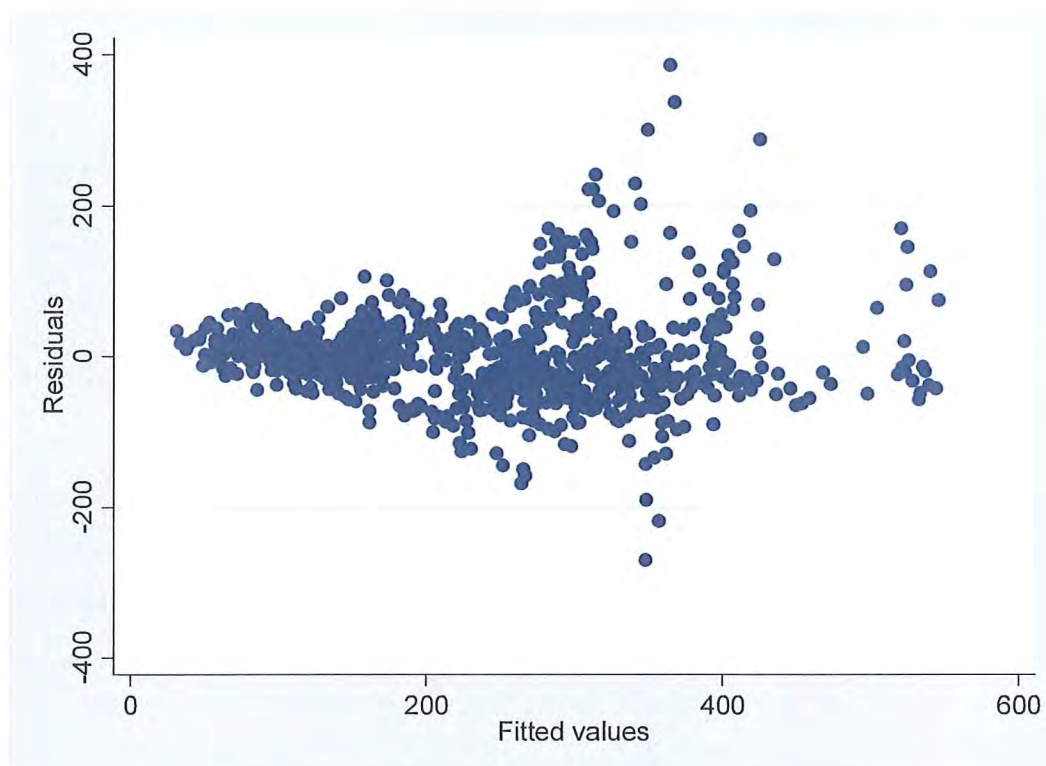
Eg har valt å gå i djupna på modell 5 fordi eg vurderar denne til å vere den som er mest eigna til å skildre systemprisen. Det er basert på at eg opplever at den gir størst innsikt i korleis fundamentale faktorar påverkar. I tillegg har denne modellen marginalt høgare forklaringsgrad enn modell 4. Koeffisienten til turnover i modell 4, sjølv om den vart signifikant, gav heller ikkje noko godt inntrykk av korleis denne variabelen påverka systemprisen i og med at den var svært låg. Det kan vere fordi turnover ikkje påverkar systemprisen, eller at modellen er misspesifisert. Ein siste grunn til at eg ikkje ser på modell 4 er at eg då ikkje hadde fått eit bilde av korleis oljeprisen påverkar, noko som eg opplever som meir interessant enn turnover.

Om fokuset derimot hadde vore på å sjå framover og å forsøke å lage prognoser, ville eg, som tidlegare nemnt, kanskje heller ha sett på modell 6.

4.2.1 Heteroskedastisitet

Ved å sjå på feilledda frå regresjonen plotta mot verdiane som er bestemt ut i frå regresjonsmodellen kan ein vurdere om det ligg føre heteroskedastisitet. Ein må då ta stilling til om feilledda frå regresjonsmodellen aukar eller minkar når ein samordnar desse med dei

predikerte verdiane for systemprisen. Av figur 17 nedanfor kan det sjå ut som om dette er tilfelle då feilledda ser ut til å auke når modellen predikerar store verdjar.



Figur 17: Plott av feilledd

Det er også nødvendig å gjere formelle testar for å avklare om det faktisk er heteroskedastisitet i modellen eller ikkje. Ein metode er Breusch-Pagan-testen som avklarar om variansen frå feilleddet er avhengig av verdiane frå dei uavhengige variablane. Ved å kjøre denne testen får eg ein F-verdi på 11,71. Det er høgare enn den kritiske grensa og det tilseier at eg må forkaste nullhypotesen om homoskedastisitet.

Ei anna metode er å kjøre White-testen. Denne er av ein meir generell karakter enn den førre testen og ved å kjøre denne testen får eg ein F-verdi på 29,63. Nullhypotesen om homoskedastisitet må difor forkastast. Basert på desse to testane har eg derfor grunn til å tru at modellen er heteroskedastisk.

Den praktiske konsekvensen av dette er at koeffisientane i modellen framleis er forventingsrette og konsistente, men standardavvikla kan vere for høge eller for låge. Ein må difor vere forsiktig med å gjere statistisk inferens, og ein bør ta høgde for at det er mogleg at nokre signifikante variablar ikkje er signifikante likevel.

4.2.2 Chow test

For å teste om parametrane endrar seg over tid (strukturelle ulikskapar) brukar eg ein Chow test.

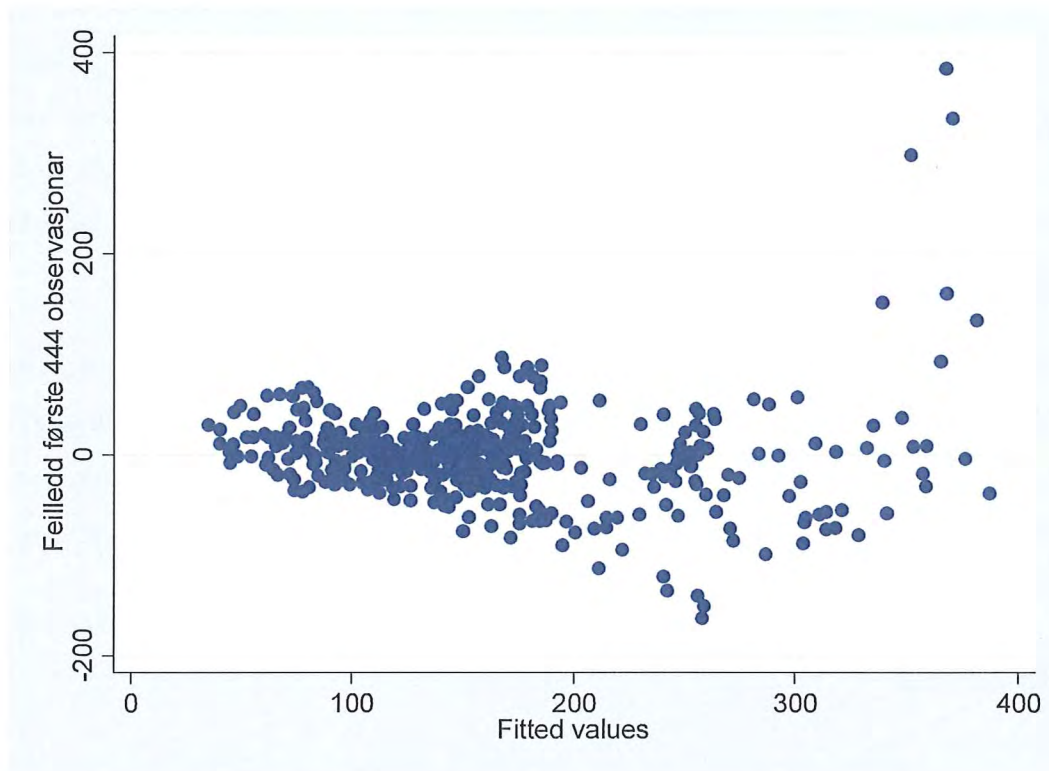
| Parameter | Estimat heile perioda | Estimat første 444 obs | Estimat siste 444 obs |
|-----------|-----------------------|------------------------|-----------------------|
| over | -6.28 | -5.00 | -10.93 |
| under | -7.38 | -7.56 | -6.85 |
| temp | -2.91 | -3.34 | -3.33 |
| olje | 1.18 | 0.92 | 1.93 |
| t | 0.19 | 0.22 | 0.06 |
| _cons | 119.43 | 113.75 | 167.02 |

Tabell 14: Resultat frå Chow-testen

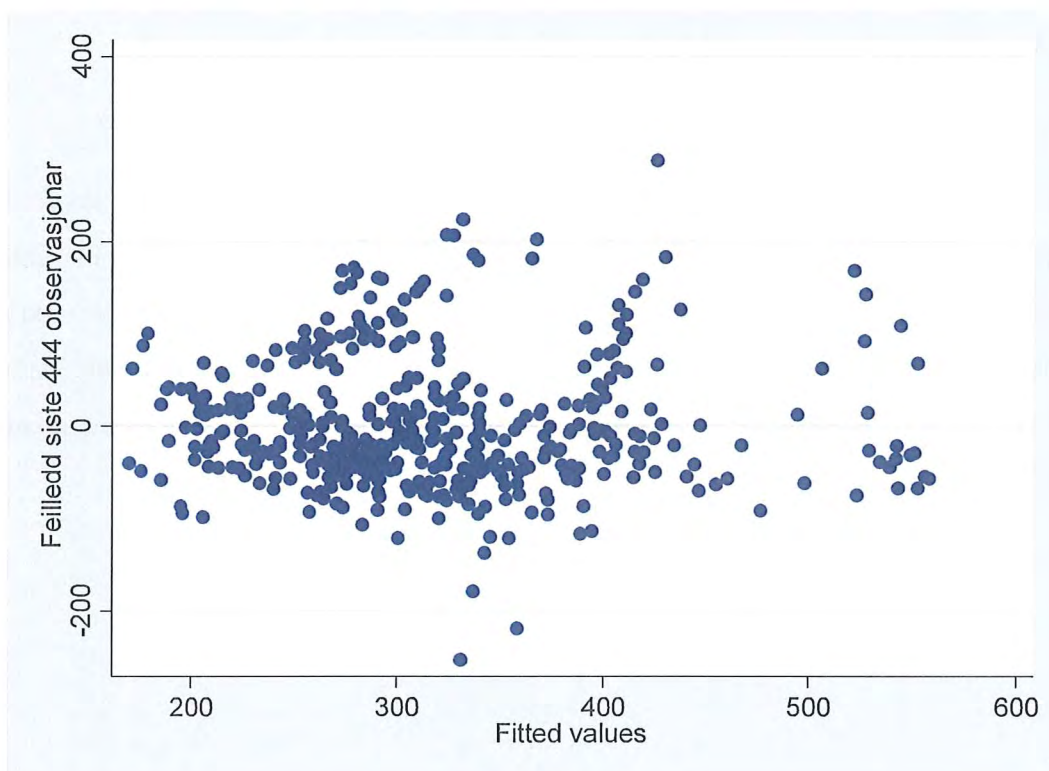
Av å observere tabell 14 kan ein sjå at estimata for dei ulike parameterane er ulike når ein deler datasettet i to tidsperiodar. Det kan sjå ut som at summen av absoluttverdien til estimata i siste tidsperiode er noko høgare enn summen av absoluttverdien til estimata i første tidsperiode, og at verdiane for første tidsperiode ligg nærmare verdiane for heile perioden.

Ved å kjøre ein formell Chow test får eg ein svært høg F-verdi. Den er langt over den kritiske grensa på 2,21 som ein kan lese av frå ein F-tabell og difor forkastar eg nullhypotesen om at parameterane er like. Parameterestimata er difor signifikant ulike over desse to tidsromma. I praksis betyr dette at modellen må reestimerast ofte.

Chow-testen føreset fråvær av heteroskedastisitet. Ghilagaber (2004) hevdar i sin artikkel at ein likevel kan bruke Chow-testen ved heteroskedastisitet om nokre kriterium er oppfylt. Kriteria er at størrelsen på utvala til dei to periodane må vere like og at det må vere relativt moderat og same form for heteroskedastisitet i begge periodane. Om dette er oppfylt vil det då vere betre å bruke den opphavlege spesifikasjonen i staden for å transformere datamaterialet slik at det blir homoskedastisk.



Figur 18: Feilledd frå regresjon av dei 444 første observasjonane



Figur 19: Feilledd frå regresjon av dei 444 siste observasjonane

Det første kriteriet er utan tvil oppfylt i og med at det er 444 observasjonar i begge periodane. Om det andre kriteriet er oppfylt er gjenstand for diskusjon. Ved første augekast kan det sjå ut som at desse to plotta av feilledda i figur 18 og figur 19 er ulike. Likevel kan det sjå ut som at feilledda er relativt jamt fordelt over og under null i begge plotta. Samstundes kan ein hevde at feilledda i den andre perioden i gjennomsnitt har høgare verdi enn feilledda i den første perioden.

Motargumentet er at i og med dette er absolutte verdiar kan ein seie at feilledda relativt sett ikkje er høgare i siste periode. I tillegg kan ein observere at det ligg føre ein del uteliggjarar i begge periodar, og nokre av desse er det kanskje aktuelt å vurdere å fjerne. Samla sett opplever eg at dei to vilkåra er oppfylte, og eg konkluderer difor med at det er mogleg å bruke Chow-testen i dette tilfellet.

4.2.3 RESET test

For å vite om modellen er riktig spesifisert og for å sjekke om det faktiske ligg føre ein lineær relasjon mellom den avhengige og dei uavhengige variablane er det nyttig å vurdere dette i lys av ein RESET test. Ein ser då om ikkje-lineære kombinasjonar av regresjonslinja er med på å forklare systemprisen. Om dette er tilfelle er modellen misspesifisert.

$$(11) \text{ sysp} = \beta_0 + \beta_1 \cdot \text{over} + \beta_2 \cdot \text{under} + \beta_3 \cdot \text{temp} + \beta_4 \cdot \text{olje} + \beta_5 \cdot t + \beta_6 \cdot \hat{S}^2$$

Dette kan gjerast ved å kjøre en regresjon på regresjonslikninga og deretter legge til kvadratet av systemprisen i ein ny variabel som her blir kalla \hat{S}^2 . Så blir det kjørt ein vanlig regresjon av denne modellen, sjå likninga ovanfor. Då finn eg at β_6 blir signifikant, men at parameterverdien blir berre 0,0006. Dette er ein såpass låg verdi i forhold til dei andre parameterane at det er grunn til å tru at denne variabelen ikkje har mykje forklaringskraft. Dette underbygger den valte funksjonsforma og at modellen er riktig spesifisert.

5. Konklusjon

Ved hjelp av økonometriske teknikkar har det gjennom sju ulike regresjonsmodellar inngåande blitt studert på kva måte faktorar som fyllingsgrad, turnover (omsetting), temperatur og oljepris påverkar systemprisen på Nord Pool. Eg opplever at alle desse regresjonsmodellane i denne utredninga fortel noko fundamentalt om systemprisen, sjølv om nokre av modellane kanskje inneheld meir viktig og interessant informasjon. Og trass i at dei ulike modellane har ulike parameterestimater, så er forteikna til dei ulike parametrane i desse modellane i stor grad konsistente.

I denne utredninga er det forsøkt å svare på problemstillinga om kva som påverkar straumprisen. Hovudfunnet er at fundamentale hydrologiske faktorar påverkar denne. Mellom anna så har det vist seg at om temperaturen aukar, så vil systemprisen bli redusert. Fyllingsgrad i vassmagasina viser seg også å påverke på same vis. Når fyllingsgrada aukar, så vil systemprisen bli redusert. Effekten er sterkast når det er låg kapasitet i vassmagasina. Det ser ein av at prisendringa er størst når det er relativt låg fyllingsgrad.

Om ein skal talfeste dette, så viser modell 5 at om temperaturen aukar med 10 grader, så vil systemprisen i gjennomsnitt bli redusert med cirka 3 øre/kWh. I tillegg viser regresjonsmodellen at om fyllingsgrad aukar med 10 % ved låg kapasitet i vassmagasina, så blir systemprisen i gjennomsnitt redusert med cirka 7 øre/kWh. Likeeins finn eg at om fyllingsgrad aukar med 10 % ved høg kapasitet i vassmagasina, så blir systemprisen i gjennomsnitt redusert med cirka 6 øre/kWh.

I tillegg kan det sjå ut som at nokre andre faktorar er med på å påverke spotprisdanninga. Oljeprisen, som er ein eksogen faktor og blir fastsett i ein global kontekst av store internasjonale aktørar og statar, viser seg i mindre grad å påverke systemprisen. Her går påverknaden derimot i motsatt retning i forhold til dei andre variablane. Det vil seie at når oljeprisen aukar, så vil systemprisen også auke. Dette er eit funn som styrker argumentet om at energiprisar trekk i same retning. Om ein ønskjer å talfeste dette funnet, så viser modell 5 at om oljeprisen aukar med 10 US dollar pr fat, så vil systemprisen i gjennomsnitt auke med cirka 1 øre/kWh.

Ein skal likevel vere forsiktig med å tolke desse koeffisientane veldig pedantisk då det ligg føre stor usikkerheit ved dei, og i og med at dei er svært enkle å justere ved å legge til eller

fjerne enkelte variablar i modellen. I tillegg må det presiserast at desse tolkingane er av typen *ceteris paribus*, som betyr at ein endrar ein enkelt variabel samstundes som alle andre variablane er konstante. Ein tek difor ikkje omsyn til den dynamikken som kan oppstå ved at fleire variablar som regel endrar seg simultant.

Det er også på sin plass å nemne at sjølv om Nord Pool si utvikling og prisdanning har vore slik dei første 17 åra av Nord Pool sin historie, så er det ingen garanti for at det er slik prisutviklinga vil vere framover. Det er likevel ikkje usannsynleg at desse mønstera vil halde fram. Mellom anna kan ein påstå at det ville ha vore oppsiktsvekkande om ulike fenomen, som til dømes sesongmønster i form av det at medianprisen er lågare om sommaren enn om vinteren, hadde forsvunne.

Det er generelt svært vanskeleg å vite kva for faktorar som påverkar systemprisen. Ein får ein peikepinn av å studere korrelasjon, men det er ikkje trivielt å vite kva for faktorar som ein bør inkludere i ei korrelasjonsmatrise. Med tanke på vidare forskning hadde det vore interessant å sett på fleire andre påverknadsvariablar til systemprisen. Til dømes ved å konstruere indeksar som reflekterar marknadsmakt og eigarskapskonsentrering i og med at dette er noko som varierar i dei ulike nordiske landa. Andre energivariablar som til dømes kol og gass kunne også ha vore interessant å modellere. Eventuelt kunne ein ha sett nærmare på fleire hydrologiske faktorar som til dømes vind i og med at dette er ei viktig energikjelde for Danmark.

6. Referansar

- Amundsen, E., Bergman, L. & Andersson, B., 1998. *Competition and Prices in the Emerging Nordic Electricity Market*, Stockholm School of Economics: Working paper in Economics and Finance (No. 217).
- Bjerkseth, Ø. S., 2006. *Volatility on Oslo Stock Exchange*, Bergen: Masteroppgåve NHH.
- Bjørnland, H. C., Brubakk, L. & Jore, A. S., 2004. Produksjonsgapet i Norge - en sammenlikning av beregningsmetoder. *Penger og Kreditt*, April 2004, pp. 199-209.
- Breusch, T. & Pagan, A., 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, September 1979, Vol 47, pp. 1287-1294.
- Bushnell, J., 1998. *Water and Power: Hydroelectric Resources in the Era of Competition in the Western U.S.*, Berkley: University of California (PWP-056).
- Bye, T. & Rosendahl, K. E., 2005. Betyr kvoteprisen egentlig noe for kraftprisene?. *Økonomiske analyser*, Årg. 24 (5), pp. 3-13.
- Chow, G. C., 1960. Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, July 1960, Vol 28, pp. 591-605.
- Deng, D., 2006. *Three essays on electricity spot and financial derivative prices at the Nordic Power Exchange*, Gøteborg: Gøteborg University, Department of Economics (doktorgrad).
- Enders, W., 2010. *Applied Econometric Time Series*. Hoboken: Wiley.
- Escribano, A., Peña, J. & Villaplana, P., 2002. *Modeling electricity prices: International evidence*, Universidad Carlos III de Madrid: Departamento de Economía (WP 02-27).
- European Network of Transmission System Operators for Electricity, 2010. *Statistical Yearbook 2010*. [Internett]
Available at: <https://www.entsoe.eu/resources/publications/general-reports/statistical-yearbooks/>
[Funnen 18 04 2012].
- Flatabø, N. et al., 2003. Experience with the Nord Pool design and implementation. *Power Systems*, Mai 2003, Vol 18, pp. 541-547.

- Ghilagaber, G., 2004. Another look at the Chow's test for equality of two heteroscedastic regression models. *Quality and Quantity*, Vol 38 (1), pp. 81-93.
- Gjølberg, O. & Johnsen, T., 2001. *Electricity futures: Inventories and price relationships at Nord Pool*, UMB: Department of Economics and Social Sciences (Working Paper).
- Hjalmarsson, E., 2000. *Nord Pool: A power market without market power*, Goteborg University: Working papers in Economics (No 28).
- Hodrick, R. J. & Prescott, E. C., 1997. Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Money, Credit and Banking*, Vol 29 (1), pp. 1-16.
- Johnsen, T. A., 2001. Demand, generation and price in the Norwegian market for electric power. *Energy Economics*, Vol 23 (3), pp. 227-251.
- Johnsen, T. A., Verma, S. K. & Wolfram, C., 1999. *Zonal pricing and demand-side bidding in the Norwegian electricity market*, Berkeley: UC Berkeley Energy Institute.
- Keller, G., 2008. *Managerial Statistics*. Cengage Learning: South-Western.
- Koopman, S. J., Ooms, M. & Angeles, C. M., 2007. Periodic Seasonal Reg-ARFIMA-GARCH Models for Daily Electricity Spot Prices. *Journal of American Statistical Association*, Vol 102 (477), pp. 16-27.
- Kube, M. & Wadhwa, P., 2007. *Does size matter? The effect of market integration on wholesale prices in the Nordic electricity market*, Stockholm: Master thesis SSE.
- Li, Y. & Flynn, P. C., 2004. Deregulated power prices: comparison of volatility. *Energy Policy*, September 2004, Vol 32 (14), pp. 1591-1601.
- Lucía, J. J. & Schwartz, E., 2002. Electricity prices and power derivatives: Evidence from the Nordic Power Exchange. *Review of Derivatives Research*, Vol 5 (1), pp. 5-50.
- Lucia, J. J. & Torró, H., 2008. *Short-term electricity futures prices: Evidence on the time-varying risk premium*, University of Valencia: Available at SSRN: <http://ssrn.com/abstract=1014035>.
- Lundgren, J., Jørgen, H. & Rudholm, N., 2008. *Multinational electricity market integration and electricity price dynamics*, The Swedish Retail Institute (HUI): HUI Working Papers 16.

Mauritzen, J., 2012. *Windonomics: Empirical essays on the economics of wind power in the Nordic electricity market*, Bergen: Doktorgradsavhandling NHH.

McGill, R., Tukey, J. W. & Larsen, W. A., 1978. Variations of Box Plots. *The American Statistician*, Februar 1978, Vol 32, pp. 12-16.

Nord Pool Spot, 2011. *Company presentation*. [Internett]
Available at: http://nordpoolspot.com/Global/Download%20Center/Annual-report/Nord-Pool-Spot-Company-Presentation_2011.pdf
[Funnen 18 04 2012].

Pindyck, R. S. & Rubinfeld, D. L., 1991. *Econometric Models and Economic Forecast*. New York: McGraw-Hill.

Ramsey, J., 1969. Tests for specification errors in classical linear least-square regression analysis. *Journal of the Royal Statistical Society. Series B (Methodology)*, Vol 31 (2), pp. 350-371.

Solland, K., 2007. *Simulering av kraftsituasjonen i Midt-Norge med Samkjøringsmodellen*, Trondheim: Masteroppgåve NTNU.

Statnett, 2008. *Kraftforbindelser med utlandet*. [Internett]
Available at:
<http://www.statnett.no/Documents/Kraftsystemet/Om%20kraftsystemet/Kraftnettet/Kraftforbindelser%20utland%20080815.pdf>
[Funnen 26 04 2012].

Steen, F., 2004. *Do bottlenecks generate market power? An empirical study of the Norwegian electricity market*, Bergen: SNF project No. 4355.

Torvestad, C. & Johannessen, C. J. B., 2010. *Utfordringer ved innføring av et grønt sertifikatmarked*, Bergen: Masteroppgåve NHH.

Vallevik, I. Ø. & Øyan, S. M., 2004. *Analyse av kraftpriser*, Trondheim: Hovudoppgåve NTNU.

Weron, R., 2006. *Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach*, Hoboken: Wiley.

White, H., 1980. A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica*, May 1980, Vol 48, pp. 817-838.

Wooldridge, J., 2009. *Introductory Econometrics: A Modern Approach*. Cengage Learning: South-Western.