

Valtion
taloudellinen
tutkimuskeskus

Tutkimukset 146

Ääriarvomenetelmät sähkömarkkinoiden
riskinarvioinnissa

VATT TUTKIMUKSET

146

Ääriarvomenetelmät sähkö- markkinoiden riskinarvioinnissa

Tarja Tuovinen

ISBN 978-951-561-852-8 (nid.)
ISBN 978-951-561-853-5 (PDF)

ISSN 0788-5008 (nid.)
ISSN 1795-3340 (PDF)

Valtion taloudellinen tutkimuskeskus

Government Institute for Economic Research

Arkadiankatu 7, 00100 Helsinki, Finland

Email: etunimi.sukunimi@vatt.fi

Oy Nord Print Ab

Helsinki, huhtikuu 2009

Ääriarvomenetelmät sähkömarkkinoiden riskinarvioinnissa

Valtion taloudellinen tutkimuskeskus
VATT Tutkimukset 146/2009

Tarja Tuovinen

Tiivistelmä

Tutkimuksessa verrataan ääriarvoteoriaan perustuvia menetelmiä muihin yleisesti käytettyihin riskinarviointimenetelmiin. Menetelmiä sovelletaan sähköpörssi Nord Poolin Suomen aluehinnan tunnittaisiin sähkön hinnan vaihteluihin vuosina 2000–2007. Tulosten mukaan ääriarvomenetelmillä pystytään arvioimaan ja enustamaan tulevaisuuden riskit tarkemmin kuin muilla menetelmillä. Yleisesti käytetyt normaali- ja t-jakaumaan perustuvat menetelmät aliarvioivat riskin suuruutta, kun taas pelkästään jakauman häntään sovitettava ääriarvojakauma kuvaa sähkön hinnan suuretkin vaihtelut hyvin.

Asiasanat: sähkö – hinnat, riskinarviointi, aikasarja-analyysi, jakaumat

Abstract

This study compares the methods based on the extreme value theory to other commonly used risk assessment methods. These methods are applied to the prices changes of the electricity exchange Nord Pool's Finnish hourly prices in 2000–2007. The results reveal that extreme value methods can assess and forecast the risks more precisely than other methods used. Commonly used methods which are based on normal and t-distributions underestimate the magnitude of the risks, while the extreme value distribution fitted to the tail of the distribution describes well also the large price changes.

Key words: electricity – prices, risk assessment, time series analysis, probability distributions

Yhteenveto

Tämä tutkimus on Helsingin yliopistossa toukokuussa 2008 hyväksytty tilastotieteen pro gradu -tutkielma. Tutkielman ohjaajana on toiminut Pentti Saikkonen.

Tutkielmassa verrataan ääriarvoteoriaan perustuvia menetelmiä normaali- ja t-jakaumamenetelmiin riskinarvioinnissa. Menetelmiä sovelletaan sähkön hinnan nousujen aiheuttamien riskien arviointiin ja ennustamiseen pohjoismaisilla sähkömarkkinoilla. Tarkoituksena on selvittää, voidaanko ääriarvomenetelmillä ennustaa muita menetelmiä tarkemmin tulevaisuuden suuret sähkön tukkuhinnan nousut. Tutkimusaineistona käytetään sähköpörssi Nord Poolin tunnitaita Suomen aluehintaa vuosilta 2000–2007.

Tutkielmassa esitetään ääriarvoteorian yleiset tulokset ja ääriarvomenetelmissä käytettävät jakaumat. Ääriarvoteoria on kehitetty kuvaamaan harvoin tapahtuvia poikkeavia havaintoja, joiden tarkka mallintaminen on riskien arvioinnin kannalta erityisen tärkeää. Riskin mittaamiseen käytetään yleisesti käytettyä value at risk -suuretta (VaR), joka vastaa riskikohteen jakauman häntäkvantiilia annetulla todennäköisyydellä. Tutkielmassa sovelletaan McNeilin ja Freyn (2000) esittämää kaksivaiheista estimointimenetelmää, jossa sähkön hinnan aikasarja ensin suodatetaan aikasarjamallilla ja toisessa vaiheessa mallin estimoituihin jäännöksiin sovitetaan ääriarvojakauma. Sovitetun jakauman kvantiilien ja aikasarjamallin ehdollisen odotusarvon ja varianssin avulla voidaan laskea ehdolliset riskisuureet. Tässä tutkielmassa ehdollisissa menetelmissä sähkön hinnan logaritimuutoksia mallinnetaan Byströmin (2005) tapaan ensin AR-GARCH-mallilla. Ehdollisessa ääriarvomenetelmässä AR-GARCH-mallin jäännöksiin sovelletaan kynnyksarvon ylitys -menetelmää, jossa suuriin hinnan nousuihin sovitetaan yleistetty Pareto-jakauma. Vertailussa käytettävissä ehdollisissa menetelmissä AR-GARCH-mallin jäännösten oletetaan noudattavan normaali- ja t-jakaumaa. Neljäs vertailtava menetelmä on ehdollistamaton ääriarvomenetelmä, jossa yleistetty Pareto-jakauma sovitetaan suoraan sähkön hinnan muutoksiin. Menetelmien hyvyttä mitataan eri menetelmien antamien riskiennusteiden tarkkuutta testaamalla.

Menetelmien vertailut suoritettiin kahdella eri ajanjaksolla, joista ensimmäisellä (v. 2000–2004) estimoinnit suoritettiin koko ajanjaksolle ja toisella (v. 2005–2007) menetelmiä käytettiin ennustamaan seuraavan päivän riskisuureet. Ehdolliset menetelmät pystyvät tuottamaan markkinatilanteeseen reagoivat riskiestimaatit. Normaali- ja t-jakaumamenetelmät aliarvioivat riskin suuruutta, joten näillä menetelmillä ei pystytä varautumaan suuriin hinnan nousuihin. Ohuthäntäisinä jakaumina ne eivät sovellu kuvaamaan sähkön hinnan paksuhäntäistä ja vinoa jakaumaa. Ääriarvoteoriaan perustuvilla menetelmillä voidaan suurtenkin hinnan nousujen aiheuttamat riskit arvioida ja ennustaa tarkasti. Pelkästään jakauman häntään sovitettava ääriarvojakauma kuvaa hyvin sähkön hinnan suuria muutoksia.

Sisältö

1	Johdanto	1
2	Ääriarvoteoria	3
2.1	Ääriarvojakaumat	4
2.1.1	Yleistetty ääriarvojakauma ja jaksomaksimimenetelmä	4
2.1.2	Yleistetty Pareto-jakauma ja kynnsarvon ylitys -menetelmä	9
3	Yleistetyn Pareto-jakauman ominaisuudet ja parametrien estimointi	13
3.1	Kvantiilikuviot	13
3.2	Keskiylitysfunktio	13
3.3	Suurimman uskottavuuden estimointi	15
4	Value at risk	17
4.1	Historiallinen simulaatio	17
4.2	Value at risk normaali- ja t-jakauman avulla	18
4.3	Value at risk kynnsarvon ylitys -menetelmällä	19
4.4	Menetelmien testaus	20
5	GARCH-mallit	23
5.1	AR-GARCH-mallit	26
6	Nord Poolin sähkön hinta	28
6.1	Spot- ja aluehintojen määräytyminen	28
6.2	Aineiston kuvailu	29
7	Sähkön hinnan mallinnus ja VaR	33
7.1	Aikaisempia tutkimuksia	34
7.2	Havaintojakson mallinnus	36
7.2.1	Jäännösten mallinnus	39
7.2.2	Havaintojakson tulosten vertailu ja testaus	42
7.3	Ennustemenetelmä	47
7.3.1	Ennustejakson tulosten vertailu ja testaus	49
8	Johtopäätöksiä	53

Lähteet

56

Liitteet

59

1 Johdanto

Sähkömarkkinat ovat viime aikoina alkaneet vapautua. Sääntelyä on purettu monissa maissa, minkä vuoksi riskien arvioinnin ja hallinnan tärkeys on korostunut myös energiamarkkinoilla. Myös Suomen sähköstä merkittävä osa ostetaan ja myydään pohjoismaisessa sähköpörssissä Nord Poolissa, joka on maailman ensimmäinen monikansallinen sähköpörssi. Sähkö on nykyisin muiden hyödykkeiden tapaan sijoituskohde, joten sähkömarkkinoiden riskit kiinnostavat muitakin kuin sähköön tuottajia ja ostajia. Sähkömarkkinoilla hintojen vaihtelut ovat useita muita hyödyke- ja rahoitusmarkkinoiden vaihteluita suurempia, koska sähköä ei voi juurikaan varastoida ja sähköön kysynnän ja tarjonnan tulee olla aina yhtä suurta. Tämän vuoksi poikkeustilanteissa, kuten sähköön tuonnin äkillisesti vähentyessä, sähköön hinta voi nousta nopeasti paljon. Suurten poikkeavien hintojen ja hinnan muutosten ennustaminen on riskien arvioinnin ja hallinnan kannalta erittäin tärkeää. Perinteisillä tilastollisilla menetelmillä poikkeavien havaintojen mallintaminen on vaikeaa, koska menetelmät perustuvat keskimmäisten havaintojen mallinnukseen ja käytetyt jakaumat ovat usein ohuthäntäisiä. Harvoin esiintyvien tapahtumien mallinnukseen on kehitetty 1900-luvulta alkaen tilastollisia menetelmiä, jotka keskittyvät jakauman häntien mallintamiseen. Nämä menetelmät perustuvat ääriarvoteoriaan.

Tässä tutkielmassa esitetään ääriarvoteoriaan perustuva riskinarviointimenetelmä, jota sovelletaan Nordpoolin sähkönn tunnittaiseen spot-hintaan. Sähkönn hinnan muutoksen mallinnukseen käytetään aluksi aikasarjamallinnusta ja mallin jäännöksiin sovitetaan yleistetty Pareto-jakauma kynnysarvon ylitys -menetelmällä. Tällä menetelmällä saadaan estimoitua ehdolliset riskisuureet ja sen esittivät ensimmäisinä McNeil ja Frey (2000). Ehdollisen ääriarvomenetelmän riskisuureita verrataan perinteisiin ehdollisiin normaali- ja t-jakaumamenetelmiin sekä ehdollistamattomaan ääriarvomenetelmään, jossa kynnysarvon ylitys -menetelmää sovelletaan suoraan sähkönn hinnan muutoksiin. Menetelmien paremmuutta mitataan testeillä, jotka testaavat riskiennusteiden tarkkuutta ja hyvyttä.

Tutkielma jakaantuu teoriaosaan (luvut 2–5) ja empiiriseen osaan (luvut 6–7). Luvussa 2 esitetään ääriarvoteorian keskeiset tulokset sekä ääriarvojakaumat. Yleistetyn Pareto-jakauman sovittamisen menetelmä, eli kynnysarvon ylitys -menetelmä, ja parametrien estimointi kuvataan luvussa 3. Tutkielmassa käytettävä riskimittari

ja sen estimointitavat esitetään luvussa 4. Luvussa 5 esitetään yleiset tulokset AR- ja GARCH-malleista sekä kuvataan tutkielmassa käytettävä aikasarjamalli. Tutkielmassa käytettävä aineisto kuvataan luvussa 6 ja saadut tulokset esitetään ja tulkitaan luvussa 7. Tutkielman johtopäätökset esitetään luvussa 8.

2 Ääriarvoteoria

Useimmat tilastolliset menetelmät keskittyvät tutkimaan jakaumien keskimmäisiä havaintoja ja keskilukuja, mutta ne eivät huomioi juurikaan jakauman häntien eli ääriarvojen (minimi tai maksimi) käyttäytymistä (Smith, 2001). Usein ääriarvoja kutsutaan poikkeaviksi havainnoiksi (outlier) ja niiden vaikuttavuutta pyritään pienentämään esimerkiksi erilaisilla robusteilla menetelmillä (Smith, 2001; Embrechts ym., 1997, 2). Moniin tutkittaviin ilmiöihin liittyy kuitenkin poikkeavia havaintoja ja usein ne ovat juuri mielenkiinnon kohteena olevia asioita. Esimerkiksi hydrologit haluavat mallintaa harvinaisten suurtulvien esiintymistodennäköisyyksiä ja vakuutusyhtiöiden täytyy varautua vakuuttamaan myös suuret vahingot.

Perinteiset menetelmät toimivat hyvin alueilla, missä on paljon havaintoja, mutta ne sopivat usein huonosti kuvaamaan empiirisen jakauman häntiä ja antavat siten epätarkkoja tuloksia ääri-ilmiöistä. Ääriarvoteoria (extreme value theory, EVT) on kehitetty mallintamaan ääriarvoja eli harvinaisia ja poikkeavia havaintoja käyttämällä vain aineiston suuria tai pieniä havaintoja (Byström, 2005). Ääriarvoteorian avulla pystytään myös ekstrapoloimaan havaitun aineiston ulkopuolelle paremmin kuin useimmilla muilla menetelmillä (Fernandez, 2003).

Ääriarvoteorian historia alkaa 1920-luvulta, jolloin useat tutkijat alkoivat kehittää ääriarvojen tilastollista teoriaa. Ensimmäisenä ääriarvojen rajajakauman ja asymptotiset tulokset ratkaisivat Fisher ja Tippett (1928). He johtivat maksimin rajarvolauseet (limit laws for maxima), joiden mukaan riippumattomien ja samoin jakautuneiden (independent and identically distributed, IID) havaintojen maksimi noudattaa asymptoottisesti joko Fréchetin, Weibullin tai Gumbelin jakaumaa. Tuloksen todistuksen sekä riittävät ja välttämättömät ehdot johti Gnedenko vuonna 1943. Jenkinson (1955) ja von Mises (1936) osoittivat, että nämä kolme ääriarvojakaumaa ovat yleistetyn ääriarvojakauman (generalised extreme value distribution, GEV) erikoistapauksia. Ensimmäisen ja tunnetuimman kirjan ääriarvoteoriasta ja sen sovelluksista kirjoitti Emil Gumbel vuonna 1958. Ensimmäiset ääriarvoteorian sovellusalat olivat tekniikka, ekologia ja hydrologia.

Pelkkien ääriarvojen mallintamisen sijasta Balkema ja de Haan (1974) sekä Pickands (1975) ovat kehittäneet menetelmän tietyn kynnsarvon ylittävien havaintojen mal-

lintamiseen (peaks over threshold, POT). He osoittivat ylityksien noudattavan yleistettyä Pareto-jakaumaa (generalised pareto distribution, GPD), kun kynnsarvo on tarpeeksi suuri. Huolimatta ääriarvoteorian pitkästä historiasta, sitä on käytetty riskienhallintaan ja taloudellisten ilmiöiden kuvaamiseen vasta viime aikoina, vaikka usein taloudellisia ilmiöitä kuvaavat parhaiten paksuhäntäiset jakaumat ja äärimmäiset tapahtumat (Byström, 2005; Rachev, 2003, 38). Embrechts ym. (1997) tarjoavat kattavan esityksen ääriarvoteorian käytöstä rahoitus- ja vakuutusallalla.

2.1 Ääriarvojakaumat

Ääriarvojen mallintamiseen käytetään kahta erilaista menetelmää. Ääriarvoteorian alkuvaiheessa kehitettiin jakaumatulokset satunnaismuuttujajoukon maksimiarvolle. Tätä lähestymistapaa kutsutaan jaksomaksimimenetelmäksi (block maxima) ja siihen liittyvää jakaumaa yleistetyksi ääriarvojakaumaksi. Toinen tapa on mallintaa satunnaismuuttujajoukon suurimpia arvoja, jotka ylittävät määrätyn kynnsarvon. Tätä menetelmää kutsutaan tässä tutkielmassa kynnsarvon ylitys -menetelmäksi ja siinä satunnaismuuttujat noudattavat yleistettyä Pareto-jakaumaa. Tässä tutkielmassa keskitytään käsittelemään vain maksimia ja oikeanpuoleista häntää, mutta tulokset voidaan yleistää koskemaan minimiä ja vasenta häntää käyttämällä yhtälöä

$$\min(X_1, \dots, X_n) = -\max(-X_1, \dots, -X_n). \quad (1)$$

2.1.1 Yleistetty ääriarvojakauma ja jaksomaksimimenetelmä

Olkoot X_1, X_2, \dots, X_n riippumattomia ja samoin jakautuneita (IID) satunnaismuuttujia, joiden kertymäfunktio on $F(x) = P(X \leq x)$ ja olkoon $M_n = \max(X_1, X_2, \dots, X_n)$. Klassisessa ääriarvoteoriassa tutkitaan M_n :n käyttäytymistä, kun n kasvaa rajatta.

Koska X_1, X_2, \dots, X_n ovat IID-havaintoja, voidaan maksimin (M_n) kertymäfunktio kirjoittaa muodossa

$$\begin{aligned} P(M_n \leq x) &= P(X_1 \leq x, \dots, X_n \leq x) \\ &= \prod_{i=1}^n F(x) = F^n(x), \quad x \in \mathbb{R}, \quad n \in \mathbb{N}. \end{aligned}$$

Äärimmäiset tapahtumat sijaitsevat jakauman ylärajan läheisyydessä, joten M_n :n asymptoottinen käyttäytyminen riippuu funktion F oikeanpuoleisesta hännästä. Merkitään jakauman F oikeanpuoleista päätepistettä x_F :llä:

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}.$$

Kaikille $x < x_F$ pätee siten

$$P(M_n \leq x) = F^n(x) \rightarrow 0, \quad n \rightarrow \infty$$

ja jos funktiolla F on äärellinen päätepiste x_F , pätee kaikille $x \geq x_F$

$$P(M_n \leq x) = F^n(x) = 1.$$

Siten maksimi $M_n \xrightarrow{P} x_F$ kun $n \rightarrow \infty$ ja $x_F \leq \infty$.

Nämä tulokset eivät kerro mitään maksimin M_n jakaumasta suurilla n . Lauseessa 2.1 määritellään satunnaismuuttujajonon standardoidun maksimin asymptoottinen jakauma eli maksimin raja-arvolause. Lause tunnetaan Fisher-Tippett-lauseena. Se kuvaa maksimin käyttäytymisen otoskoon kasvaessa samoin kuin keskeinen raja-arvolause kuvaa keskiarvon käyttäytymisen.

Lause 2.1 (Fisher ja Tippett 1928) *Olkoon (X_n) jono IID-satunnaismuuttujia. Jos on olemassa vakiot $\sigma_n > 0$ ja $\mu_n \in \mathbb{R}$ sekä jokin ei-degeneroituva kertymäfunktio H , siten että*

$$\frac{M_n - \mu_n}{\sigma_n} \xrightarrow{d} H, \quad (2)$$

niin H on tyypiltään jokin seuraavista kertymäfunktioista:

$$\text{Fréchet :} \quad \Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ \exp\{-x^{-\alpha}\}, & x > 0 \end{cases} \quad \alpha > 0. \quad (3)$$

$$\text{Weibull :} \quad \Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0. \quad (4)$$

$$\text{Gumbel :} \quad \Lambda(x) = \exp\{-e^{-x}\}, \quad x \in \mathbb{R}. \quad (5)$$

Jakaumia Φ_α , Ψ_α ja Λ kutsutaan standardeiksi ääriarvojakaumiksi. Kun $X > 0$, näillä jakaumilla on riippuvuussuhde

$$X \sim \Phi_\alpha \iff \log X^\alpha \sim \Lambda \iff -\frac{1}{X} \sim \Psi_\alpha.$$

Jakaumat (3), (4) ja (5) voidaan yhdistää hyödylliseksi yksiparametriseksi jakaumaksi, jota kutsutaan yleistetyksi ääriarvojakaumaksi (generalised extreme value distribution, GEV). Yleistetyn ääriarvojakauman yleisin esitysmuoto on seuraava.

Määritelmä 2.1 (Yleistetty ääriarvojakauma) *Määritellään kertymäfunktio*

$$H_\xi(x) = \begin{cases} \exp\{- (1 + \xi x)^{-1/\xi}\} & \text{jos } \xi \neq 0, \\ \exp\{-e^{-x}\} & \text{jos } \xi = 0, \end{cases} \quad (6)$$

missä $1 + \xi x > 0$.

Tällöin jakauman H_ξ määrittelyjoukko on

$$\begin{aligned} x &> -\xi^{-1}, & \text{kun } \xi > 0, \\ x &< -\xi^{-1}, & \text{kun } \xi < 0, \\ x &\in \mathbb{R}, & \text{kun } \xi = 0. \end{aligned}$$

Jakaumaan (6) voidaan lisätä sijainti- ja skaalaparametrit $\mu \in \mathbb{R}$ ja $\sigma > 0$ sijoittamalla argumentin x paikalle $(x - \mu)/\sigma$ eli

$$H_{\xi,\mu,\sigma}(x) = H_\xi\left(\frac{x - \mu}{\sigma}\right).$$

Tällöin jakauman määrittelyjoukko on $1 + \xi \frac{x - \mu}{\sigma} > 0$. Gumbel-jakauma ($\xi = 0$) voidaan tulkita yleisen ääriarvojakauman raja-arvoksi, kun $\xi \rightarrow 0$, joten kertymäfunktio voidaan kirjoittaa lyhyemmin muodossa (Embrechts ym., 1997, 294)

$$H_{\xi,\mu,\sigma}(x) = \exp\left\{- \left(1 + \xi \frac{x - \mu}{\sigma}\right)^{-1/\xi}\right\}, \quad 1 + \xi \frac{x - \mu}{\sigma} > 0. \quad (7)$$

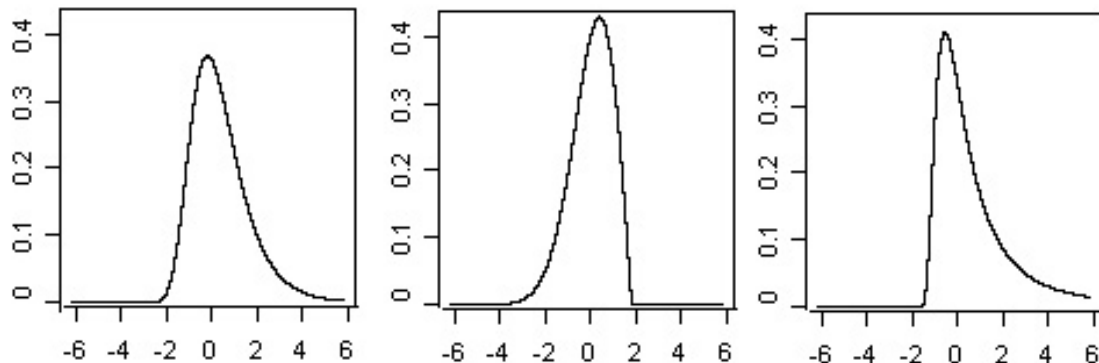
Kun jakauman määrittelee vain muotoparametri ξ , yleistetyn ääriarvojakauman kertymäfunktio on

$$H_\xi(x) = e^{-(1+\xi x)^{-1/\xi}}, \quad (8)$$

kaikille $\xi \in \mathbb{R}$, kun $1 + \xi x > 0$. Yleistetyn ääriarvojakauman (8) ja jakauminen (3)–(5) välillä on seuraavanlainen yhteys (Embrechts ym., 1997, 293–294):

$$H_\xi(x) = \begin{cases} \Phi_\alpha(1 + x/\alpha), & \text{kun } x > -\alpha, \xi = 1/\alpha > 0, \\ \Psi_\alpha(-(1 - x/\alpha)), & \text{kun } x < \alpha, \xi = -1/\alpha < 0, \\ \Lambda(x), & \text{kun } x \in \mathbb{R}, \xi = 0. \end{cases}$$

Kuva 1: Yleistetyn ääriarvojakauman H_ξ tiheysfunktiot eri muotoparametrin ξ arvoilla. Vasemmalla Gumbel- ($\xi = 0$), keskellä Weibull- ($\xi = -0.5$) ja oikealla Fréchet-jakauma ($\xi = 0.5$).



Yleistetyn ääriarvojakauman oikeanpuoleisen hännän käyttäytymisen määrää muotoparametri ξ . Jakauma on ohuthäntäinen kun $\xi = 0$ (Gumbel), lyhythäntäinen eli ylhäältä rajoitettu kun $\xi < 0$ (Weibull) ja paksuhäntäinen kun $\xi > 0$ (Fréchet). Yleistetyn ääriarvojakauman tiheysfunktioiden kuvaajat eri muotoparametrin arvoilla on esitetty kuvassa 1. Ääriarvoteorian sovelluksissa kiinnostuksen kohteena on juuri paksuhäntäinen Fréchet-jakauma.

Standardoidun maksimin asymptoottinen jakauma on yleistetty ääriarvojakauma, jos raja-arvossa H (kaava 2) on ei-degeneroituneen jakauman kertymäfunktio. Muotoparametri ξ määrää havaintojen X_i jakauman oikean hännän käyttäytyminen. Jos havaintojen jakauman F oikea häntä vähenee eksponentiaalisesti, jakauma H_ξ on Gumbel-jakauma ja $\xi = 0$. Tällaisia ohuthäntäisiä jakaumia ovat normaali-, lognormaali-, eksponentti- ja gammajakaumat. Mikäli F :n oikeanpuoleinen häntä vähenee potenssifunktion vauhtia, on H_ξ Fréchet-jakauma ja $\xi > 0$. Tällöin F on paksuhäntäinen jakauma, kuten Pareto-, Cauchy- ja t-jakauma. Jos jakauman F oikea häntä on rajoitettu, H_ξ on Weibull-jakauma ja $\xi < 0$. Tällöin jakauman F kantaja on rajoitettu, kuten tasa- ja betajakaumalla.

Yleistetyn ääriarvojakauman odotusarvo on äärellinen, kun $\xi < 1$ ja varianssi, kun $\xi < \frac{1}{2}$. Yleisesti jakauman k :s momentti on olemassa, kun $\xi < \frac{1}{k}$. Siten Gumbel- ja Weibull-jakauman kaikki momentit ovat äärellisiä. Yleistetyn ääriarvojakauman

$H_{\xi,\mu,\sigma}$ odotusarvo ja varianssi ovat

$$\begin{aligned} E(Y) &= \mu + \frac{\sigma}{\xi} [\Gamma(1 - \xi) - 1], \quad \xi < 1 \\ \text{Var}(Y) &= \left(\frac{\sigma}{\xi}\right)^2 [\Gamma(1 - 2\xi) - \Gamma^2(1 - \xi)], \quad \xi < \frac{1}{2}, \end{aligned}$$

missä $\Gamma(\cdot)$ on gammafunktio ja μ , σ ja ξ ovat jakauman sijainti-, skaala- ja muoto-parametrit. (Smith, 2001).

Asymptoottiset tulokset on johdettu käyttäen oletusta IID-havainnoista. Yleistetty ääriarvojakauma on maksimin rajajakauma myös joidenkin riippuvuusehtojen vallitessa, kuten autoregressiivisen prosessin tapauksessa. Tällöin vain jakauman sijainti- ja skaalaparametrien arvot muuttuvat (ks. Leadbetter ym., 1983). Yleistetyn ääriarvojakauman parametrien voidaan sallia riippuvan myös ajasta. (Katz ym., 2005).

Yleistetty ääriarvojakauma sovitetaan havaintoihin jaksomaksimimenetelmällä. Jaksomaksimimenetelmää käytetään seuraavasti. Olkoon X_1, X_2, \dots, X_{mn} IID-havainnoja, jotka on jaettu m :ään n :n pituiseen jaksoon. Jotta raja-arvolauseet jaksojen maksimeille $\hat{M}_n^{(j)} = \max(X_{(j-1)n+1}, \dots, X_{(j-1)n+n}), j = 1, \dots, m$ pätevät, täytyy jakson pituuden n olla tarpeeksi suuri. Toisaalta n :n kasvattaminen vähentää parametrien estimoinnin harhaa, mutta m :n kasvattaminen (n :n pienentäminen) pienentää parametriestimaattien varianssia. Kun m ja n on valittu, yleistetty ääriarvojakauma $H_{\xi,\mu,\sigma}$ sovitetaan jaksojen maksimiarvoihin $\hat{M}_n^{(1)}, \dots, \hat{M}_n^{(m)}$ esimerkiksi suurimman uskottavuuden menetelmällä (SU-menetelmä). (Rachev, 2003, 80.)

SU-menetelmässä yleistetyn ääriarvojakauman (7) parametrit estimoidaan maksimoimalla log-uskottavuusfunktio numeerisesti parametrien suhteen. Esimerkiksi Gumbel-jakauman tapauksessa ($\xi = 0$), log-uskottavuusfunktio on

$$l(0, \mu, \sigma; \mathbf{X}) = -n \log \sigma - \sum_{i=1}^n \exp \left\{ -\frac{X_i - \mu}{\sigma} \right\} - \sum_{i=1}^n \frac{X_i - \mu}{\sigma}.$$

Derivoimalla funktio parametrien μ ja σ suhteen saadaan uskottavuusyhtälöt Gumbelin tapauksessa:

$$\begin{aligned} n - \sum_{i=1}^n \exp \left\{ -\frac{X_i - \mu}{\sigma} \right\} &= 0, \\ n + \sum_{i=1}^n \frac{X_i - \mu}{\sigma} \left(\exp \left\{ -\frac{X_i - \mu}{\sigma} \right\} - 1 \right) &= 0. \end{aligned}$$

Yhtälöt täytyy ratkaista numeerisesti. Kun $\xi \neq 0$, yhtälöt ovat ylläolevia monimutkaisemmat. (Embrechts ym., 1997, 318).

Jaksomaksimimenetelmä on ääriarvoteorian klassinen lähestymistapa, jossa oletetaan, että maksimin jakauma on $H_{\xi,\mu,\sigma}$. Menetelmän käyttö vaatii suuren määrän havaintoja ja sen avulla saadut kvantiiliestimaatit ovat usein epäluotettavia. Usein jaksomaksimimenetelmän sijasta käytetään kynnsarvon ylitys -menetelmää, jonka käyttömahdollisuudet ovat monipuolisemmat ja joka ei keskity pelkästään suurimpaan havaintoon tietyllä aikavälillä (de Rozario, 2002). Kynnsarvon ylitys -menetelmä esitetään seuraavassa kappaleessa.

2.1.2 Yleistetty Pareto-jakauma ja kynnsarvon ylitys -menetelmä

Uudempi lähestymistapa ääriarvojen mallintamiseen ei keskity pelkästään suurimpiin arvoihin, vaan ottaa huomioon kaikki arvot, jotka ylittävät määrätyn kynnsarvon u . Tätä menetelmää kutsutaan kynnsarvon ylitys -menetelmäksi. Ylityksien jakauma ehdolla, että havainto on suurempi kuin kynnsarvo u on

$$F_u(x) = P(X - u \leq x \mid X > u) = \frac{F(x + u) - F(u)}{1 - F(u)}, \quad (9)$$

kun $0 \leq x < x_F - u$. Balkema ja de Haan (1974) sekä Pickands (1975) osoittivat, että ylitysten jakaumaa F_u voidaan approksimoida yleistetyllä Pareto-jakaumalla, kun jakaumakonvergenssi (2) pätee ja u lähestyy oikeaa päätepistettä x_F . Toisin sanoen yleistetty Pareto-jakauma on ylitysten rajajakauma, kun u lähestyy x_F :ää eli

$$\lim_{u \rightarrow x_F} \sup_{0 \leq x < x_F - u} |F_u(x) - G_{\xi,\sigma(u)}(x)| = 0, \quad (10)$$

missä yleistetyn Pareto-jakauman kertymäfunktio $G_{\xi,\sigma(u)}$ määritellään seuraavasti.

Määritelmä 2.2 (Yleistetty Pareto-jakauma) *Määritellään kertymäfunktio*

$$G_{\xi,\sigma(u)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\sigma(u)}\right)^{-1/\xi}, & \text{kun } \xi \neq 0 \\ 1 - e^{-x/\sigma(u)}, & \text{kun } \xi = 0, \end{cases} \quad (11)$$

missä

$$\begin{aligned} x &\geq 0, & \text{jos} & \quad \xi \geq 0, \\ 0 \leq x &\leq -\frac{\sigma(u)}{\xi}, & \text{jos} & \quad \xi < 0 \end{aligned}$$

ja skaalaparametri $\sigma(u) > 0$. Jakaumaa $G_{\xi, \sigma(u)}$ kutsutaan yleistetyksi Pareto-jakaumaksi.

Tulos (10) tarkoittaa, että suurilla kynnysarvon u arvoilla on olemassa jokin $\sigma(u)$ ja ξ (ei riipu u :sta), joilla yleistetty Pareto-jakauma on hyvä approksimaatio ylitysten jakaumalle F_u . Skaalaparametri σ on kynnysarvon u funktio, koska se estimoidaan ylityksien y_i avulla, joten skaalaparametriin vaikuttavat ylityksien määrä ja sijainti. Muotoparametri ξ kuvaa jakauman hännän vähenemisnopeutta, joten se ei riipu valitusta kynnysarvosta, kunhan havaintoja on tarpeeksi. Tästä eteenpäin merkitään $\sigma(u) = \sigma$, eli parametrin riippuvuutta kynnysarvosta ei esitetä eksplisiittisesti. Pickands (1975) osoitti, että yleistetyn Pareto-jakauman ja yleistetyn ääriarvojakauman parametri ξ on identtinen molemmissa jakaumissa ja skaala- ja sijaintiparametrien σ ja μ välille löydetään täsmällinen vastaavuus (ks. Davison ja Smith, 1990).

Kuten yleistetyn ääriarvojakauman tapauksessa, voidaan jakaumaan (11) lisätä sijaintiparametri μ , eli $G_{\xi, \mu, \sigma}(x) = G_{\xi, \sigma}(x - \mu)$. Yleistetty Pareto-jakauma (11) on yleistetyn ääriarvojakauman (6) tavoin yleistetty esitysmuoto, joten se sisältää useita jakaumia erikoistapauksina. Esimerkkejä yleistetyn Pareto-jakauman erikoistapauksista ovat tasa- ($\xi = -1$), eksponentti- ($\xi = 0$) ja Pareto-jakauma ($\xi > 0$). Jakaumien tiheysfunktioiden kuvaajat eri ξ :n arvolla on esitetty kuvassa 2. Jakauma on paksuhäntäinen, kun $\xi > 0$.

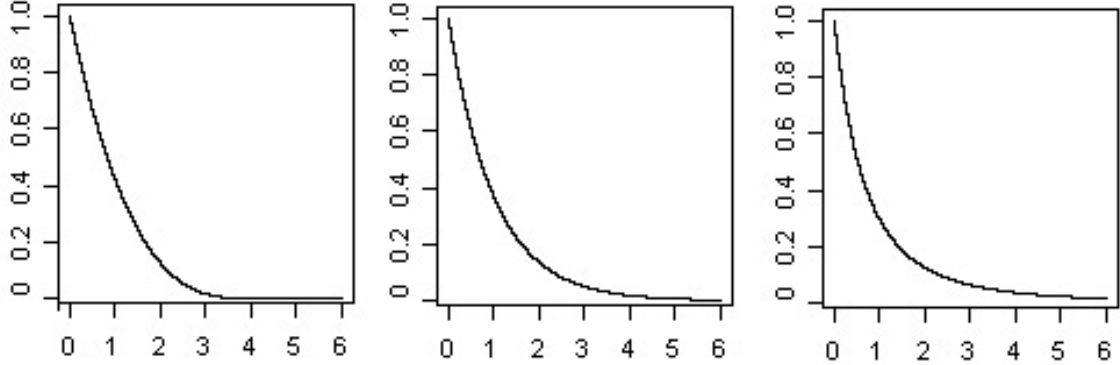
Yleistetyn Pareto-jakauman kertymäfunktio parametrin arvolla $\xi = 0$ ($G_{0, \sigma}$) voidaan tulkita $G_{\xi, \sigma}$:n raja-arvoksi, kun $\xi \rightarrow 0$ (Embrechts ym., 1997, 162). Siten voimme merkitä

$$G_{\xi, \sigma}(x) = 1 - \left(1 + \xi \frac{x}{\sigma}\right)^{-1/\xi}, \quad x \in D(\xi, \sigma), \quad (12)$$

missä

$$x \in D(\xi, \sigma) = \begin{cases} [0, \infty), & \text{jos } \xi \geq 0, \\ [0, -\sigma/\xi], & \text{jos } \xi < 0. \end{cases}$$

Kuva 2: Yleistetyn Pareto-jakauman $G_{\xi,\sigma}$ tiheysfunktioit eri muotparametrin ξ arvoilla, kun $\sigma = 1$. Vasemmalla ylhäältä rajoitettu ($\xi = -0,25$), keskellä ohuthäntäinen ($\xi = 0$) ja oikealla paksuhäntäinen jakauma ($\xi = 0.5$).



Tässä tutkielmassa tarkoitetaan kertymäfunktion (12) määrittelemää jakaumaa, kun puhutaan yleistetystä Pareto-jakaumasta parametreilla ξ ja σ .

Kuten yleistetyllä ääriarvojakaumalla, yleistetyllä Pareto-jakaumalla on äärellinen odotusarvo, kun $\xi < 1$ ja variassi, kun $\xi < \frac{1}{2}$:

$$E(Y) = \frac{\sigma}{1-\xi}, \quad \xi < 1,$$

$$Var(Y) = \frac{\sigma^2}{(1-\xi)^2(1-2\xi)}, \quad \xi < \frac{1}{2}.$$

Kynnysarvon ylitys -menetelmää käytetään tässä tutkielmassa seuraavasti. Olkoon X_1, X_2, \dots, X_n IID-havaintoja, joiden kertymäfunktion on F ja joille on olemassa vakiot $c_n > 0$ ja $d_n \in \mathbb{R}$ siten, että tulos (2) pätee. Valitaan korkea kynnysarvo u ja merkitään ylitysten lukumäärää N_u :lla ja ylityksiä $Y_i = X_i - u$, kun $X_i > u$. Kynnysarvon u valinta on tärkeää, jotta asymptoottiset tulokset pätsivät. Menetelmä u :n valitsemiseksi esitetään seuraavassa kappaleessa. Tästä eteenpäin estimaatteja merkitään kirjoittamalla merkinnän yläpuolelle hattu, esimerkiksi $\hat{\xi}$. Tuloksen (10) mukaan ylitysten Y_1, Y_2, \dots, Y_{N_u} jakaumaa F_u voidaan approksimoida yleistetyllä Pareto-jakaumalla ($F_u(y) \approx G_{\xi,\sigma}(y)$), kun u on tarpeeksi suuri. Kun $y \geq 0$, kaavan (9) mukaan

$$G_{\hat{\xi},\hat{\sigma}}(y) \approx \hat{F}_u(y) = \frac{\hat{F}(y+u) - \hat{F}(u)}{1 - \hat{F}(u)}, \quad (13)$$

missä \hat{F} on havaintojen empiirinen kertymäfunktion. Parametrien ξ ja σ estimaattien $\hat{\xi}$ ja $\hat{\sigma}$ laskemiseen käytetään suurimman uskottavuuden menetelmää, joka kuvataan

seuraavassa kappaleessa.

Kertymäfunktiota $F(u)$ approksimoidaan havaintojen (x_i) empiirisellä kertymäfunktiolla pisteessä u , joka on kynnyksarvon u alittavien havaintojen lukumäärä jaettuna havaintojen lukumäärällä n , joten $\hat{F}(u) = 1 - \frac{N_u}{n}$ ja $1 - \hat{F}(u) = \frac{N_u}{n}$. Nämä ovat kohtuullisen hyviä estimaatteja, kun havaintojen lukumäärä n on tarpeeksi suuri. Kun yhdistetään nämä estimaatit ja $G_{\hat{\xi}, \hat{\sigma}}$ (kaava 12) saadaan yhtälöstä (13) ratkaistua

$$\begin{aligned}\hat{F}(y+u) &= \left[1 - \hat{F}(u)\right] G_{\hat{\xi}, \hat{\sigma}}(y) + \hat{F}(u) = \frac{N_u}{n} G_{\hat{\xi}, \hat{\sigma}}(y) + 1 - \frac{N_u}{n} \\ &= 1 - \frac{N_u}{n} \left(1 - G_{\hat{\xi}, \hat{\sigma}}(y)\right) \\ &= 1 - \frac{N_u}{n} \left(1 + \hat{\xi} \frac{y}{\hat{\sigma}}\right)^{-1/\hat{\xi}},\end{aligned}\tag{14}$$

kun $y \geq 0$. Alkuperäisten havaintojen $X_i = Y_i + u$ kertymäfunktion $F(x)$:n oikeanpuoleista häntää voidaan siten estimoida kaavalla

$$\hat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \hat{\xi} \frac{x-u}{\hat{\sigma}}\right)^{-1/\hat{\xi}} \quad \text{kaikilla } x > u.$$

Kääntämällä kertymäfunktio $\hat{F}(x)$ saadaan esitys q :nille kvantiiliestimaatille \hat{x}_q :

$$\hat{x}_q = \hat{F}^{-1}(q) = u + \frac{\hat{\sigma}}{\hat{\xi}} \left[\left(\frac{n}{N_u} (1-q) \right)^{-\hat{\xi}} - 1 \right],\tag{15}$$

kun $\hat{\xi} \geq 0$.

3 Yleistetyn Pareto-jakauman ominaisuudet ja parametrien estimointi

Tässä luvussa käsitellään yleistetyn Pareto-jakauman ominaisuuksia, joiden avulla oikean mallin valinta helpottuu. Jotta yleistetty Pareto-jakauma sopisi hyvin mallinnettavaan aineistoon, on kynnysarvon u valinta tärkeä. Valinnan apuna käytetään usein keskiylitysfunktiota. Estimoidun mallin häntäparametrin ξ hyvyttä voidaan arvioida kvantiilikuvion avulla tai vertaamalla eri kynnysarvoilla estimoitujen parametrien arvoja ja luottamusvälejä. Luvun lopussa esitetään suurimman uskottavuuden menetelmä yleistetyn Pareto-jakauman parametrien estimointiin.

3.1 Kvantiilikuviot

Kvantiilikuvion (QQ-plot) avulla voidaan graafisesti tarkastella yleistetyn Pareto-jakauman sopivuutta estimoitavaan aineistoon. Mikäli malli on oikein spesifioitu, kvantiilikuvion pisteiden tulisi sijaita suoralla. Kvantiilikuviossa esitetään mallinukseen käytettävän jakauman teoreettisten kvantiilien ja aineiston empiirisen jakauman kvantiilien riippuvuus seuraavasti. Järjestetään ylitykset Y_1, \dots, Y_n siten, että $Y_{1,n} = \max(Y_1, \dots, Y_n) \geq Y_{2,n} \geq \dots \geq Y_{n,n} = \min(Y_1, \dots, Y_n)$. Kun ylityksiin Y_1, \dots, Y_n sovitettava jakauma on yleistetty Pareto-jakauma $G_{\xi, \sigma}$ kvantiilikuvion pisteet ovat

$$\left\{ (Y_{k,n}, G_{\hat{\xi}, \hat{\sigma}}^{-1}(p_{k,n})) : k = 1, \dots, n \right\},$$

missä $p_{k,n} = (n - k + 1)/(n + 1)$.

Jos kvantiilikuvio ei ole suora vaan konkaavi, sovitettava malli aliarvioi jakauman häntää ja paksumpi häntäinen (suurempi ξ) jakauma voisi sopia paremmin. Vastaa- vasti jos kuvaaja on konvekksi, pienempi ξ :n arvo parantaisi mallin sopivuutta.

3.2 Keskiylitysfunktio

Toinen graafinen tapa tutkia mallin sopivuutta aineistoon on keskiylitysfunktio (mean excess function). Se kuvaa odotettua ylityksen suuruutta, kun kynnysarvo

on ylitetty. Keskiylitysfunktion kuvaajan avulla nähdään sopiiko yleistetty Pareto-jakauma kuvaamaan aineistoa ja mikä olisi sopiva kynnsarvon u valinta.

Määritelmä 3.1 (Keskiylitysfunktio) *Keskiylitysfunktio määritellään yhtälöllä*

$$e(u) = E(X - u \mid X > u)$$

ja se voidaan esittää kertymäfunktion avulla seuraavasti (Embrechts ym., 1997)

$$e(u) = \frac{1}{1 - F(u)} \int_u^{x_F} (1 - F(x)) dx, \quad 0 < u < x_F, \quad (16)$$

kun jakaumalla F on äärellinen odotusarvo.

Yleistetylle Pareto-jakaumalle (12) keskiylitysfunktio voidaan johtaa kaavasta (16), kun $\xi < 1$:

$$\begin{aligned} e(u) &= \frac{1}{(1 + \xi \frac{u}{\sigma})^{-1/\xi}} \int_u^{x_F} \left(1 + \xi \frac{x}{\sigma}\right)^{-1/\xi} dx \\ &= \frac{\sigma}{\xi (1 + \xi \frac{u}{\sigma})^{-1/\xi}} \int_u^{x_F} \frac{\xi}{\sigma} \left(1 + \xi \frac{x}{\sigma}\right)^{-1/\xi} dx \\ &= \frac{\sigma}{\xi (1 + \xi \frac{u}{\sigma})^{-1/\xi}} \left[\frac{x}{\sigma} - \frac{\xi}{\xi - 1} \left(1 + \xi \frac{x}{\sigma}\right)^{-\frac{1}{\xi} + 1} \right] \\ &= \frac{\sigma}{\xi (1 + \xi \frac{u}{\sigma})^{-1/\xi}} \left[0 + \frac{\xi}{\xi - 1} \left(1 + \xi \frac{u}{\sigma}\right)^{-\frac{1}{\xi} + 1} \right] \\ &= \frac{\sigma}{\xi - 1} \left(1 + \xi \frac{u}{\sigma}\right) \\ &= \frac{\sigma + \xi u}{1 - \xi}, \quad \text{kun } \sigma + \xi u > 0, \end{aligned} \quad (17)$$

koska kertymäfunktion arvo päätepisteessä x_F on yksi ja siten $1 - G_{\xi, \sigma}(x_F) = (1 + \xi \frac{x_F}{\sigma})^{-\frac{1}{\xi}} = 1 - 1 = 0$. Rajoitus $\xi < 1$ takaa, että paksuhäntäisellä jakaumalla on äärellinen odotusarvo. Kaavasta (17) huomataan, että keskiylitysfunktio on lineaarinen kynnsarvon u suhteen. Siten pisteet $(u, e(u))$ sijaitsevat suoralla, kun yleistetty Pareto-jakauma on hyvä approksimaatio ylitysten jakaumalle. Oikean kynnsarvon valitsemiseksi täytyy etsiä u :n arvo, josta eteenpäin keskiylitysfunktio on (lähes) suora. Kynnsarvon määrittämiseen liittyy samanlainen valinta harhan ja varianssin välillä kuin jaksomaksimimenetelmässä. Jos kynnsarvoa kasvatetaan, havaintojen lukumäärä vähenee ja estimaattien varianssi kasvaa. Toisaalta, jos kynnsarvo on

liian pieni, yleistetty Pareto-jakauma ei ole hyvä approksimaatio ja estimaatit ovat siten harhaisia.

Käytännössä keskiylitysfunktiota (17) estimoidaan otoskeskiylitysfunktiolla

$$\hat{e}(u) = \frac{1}{N_u} \sum_{i=1}^n (X_i - u)^+, \quad (18)$$

missä $(X_i - u)^+ = \max(0, X_i - u)$. $\hat{e}(u)$ on siis ylityksien summa jaettuna ylityksien lukumäärällä. Kynnysarvon u valitsemiseksi piirretään otoskeskiylitysfunktio u :n suhteen ja valitaan u siten, että $\hat{e}(u)$ on likimain lineaarinen kaikilla $x \geq u$. Kaavasta (17) nähdään, että keskiylitysfunktion kuvaajan kulmakerroin on $\xi/(1-\xi)$. Positiivinen kulmakerroin merkitsee paksuhäntäistä jakaumaa ($0 < \xi < 1$), negatiivinen kulmakerroin lyhythäntäistä jakaumaa ($\xi < 0$) ja vaakasuora kuvaaja ohuthäntäistä jakaumaa ($\xi = 0$). Kynnysarvon valinnassa voidaan myös estimoida jakauman muotoparametri ξ useilla eri kynnysarvoilla. Tällöin pyritään valitsemaan kynnysarvo, jolla parametrin arvo on vakiintunut ja jonka luottamusväli on mahdollisimman pieni.

3.3 Suurimman uskottavuuden estimointi

Yleistetyn Pareto-jakauman sovittamiseksi aineistoon täytyy estimoida muoto- ja skaalaparametrit ξ ja σ . Parametrit voidaan estimoida useilla eri menetelmillä, mutta tässä tutkielmassa esitetään vain suurimman uskottavuuden menetelmä. SU-menetelmässä jakauman log-uskottavuusfunktio maksimoidaan parametrien suhteen. Seuraavien tulosten todistukset on esittänyt Smith (1987). Hosking ja Wallis (1987) osoittivat, että SU-menetelmän säännöllisyys ehdot toteutuvat, estimaatit ovat tarkentuvia, asympotoottisesti tehokkaita ja normaalisti jakautuneita yleistetyn Pareto-jakauman tapauksessa, erityisesti kun $\xi > -1/2$.

Olkoot IID-havainnoista $\mathbf{X} = (X_1, \dots, X_n)$ lasketut ylitykset $\mathbf{Y} = (Y_1, \dots, Y_{N_u})$ ($Y_i = X_i - u$, $X_i > u$) ja $G_{\xi, \sigma}$ yleistetyn Pareto-jakauman (12) kertymäfunktio parametreilla ξ ja σ . Vastaava tiheysfunktio on

$$g_{\xi, \sigma}(y) = G'_{\xi, \sigma}(y) = \frac{1}{\sigma} \left(1 + \xi \frac{y}{\sigma} \right)^{-1/\xi-1}, \quad y \in D(\xi, \sigma). \quad (19)$$

Uskottavuusfunktio on

$$L(\xi, \sigma; \mathbf{Y}) = \prod_{i=1}^n g_{\xi, \sigma}(Y_i),$$

ehdolla $1 + \xi \frac{Y_i}{\sigma} > 0$ ja maksimoitava log-uskottavuusfunktio

$$\begin{aligned} l(\xi, \sigma; \mathbf{Y}) &= \sum_{i=1}^n \log \left[\frac{1}{\sigma} \left(1 + \xi \frac{Y_i}{\sigma} \right)^{-1/\xi - 1} \right] \\ &= -n \log \sigma - \left(\frac{1}{\xi} + 1 \right) \sum_{i=1}^n \log \left(1 + \frac{\xi}{\sigma} Y_i \right). \end{aligned} \quad (20)$$

Jakauman parametrit on ratkaistava maksimoimalla log-uskottavuusfunktio numeerisesti. Alkuarvot ξ_0 ja σ_0 voidaan valita otoskeskiylitysfunktion avulla, koska teorettisesti keskiylitysfunktion kulmakerroin on $\xi/(1 - \xi)$ ja vakio $\sigma/(1 - \xi)$. Sovittamalla suora otoskeskiylityskuvioon, voidaan ratkaista ξ_0 ja σ_0 .

Kun $\xi > -1/2$, voidaan SU-estimaateille $\hat{\xi}_n$ ja $\hat{\sigma}_n$ osoittaa asymptoottinen normaalisuus

$$n^{1/2} \left(\hat{\xi}_n - \xi, \frac{\hat{\sigma}_n}{\sigma} - 1 \right) \xrightarrow{d} \mathbf{N}(\mathbf{0}, \mathbf{M}^{-1}), \quad n \rightarrow \infty,$$

missä

$$\mathbf{M}^{-1} = (1 + \xi) \begin{bmatrix} 1 + \xi & -1 \\ -1 & 2 \end{bmatrix}.$$

Estimaateilla $\hat{\xi}_n$ ja $\hat{\sigma}_n$ on siten SU-estimaattien tavanomaiset asymptoottiset ominaisuudet eli ne ovat tarkentuvia ja asymptoottisesti normaalisti jakautuneita.

4 Value at risk

Value at risk (VaR) on riskinhallinnassa yleisesti käytetty riskin mittari. Riski voi liittyä esimerkiksi hinnan tai tuoton poikkeukselliseen muutokseen. VaR kertoo mahdollisen tappion suuruuden, joka ylittyy tietyllä aikavälillä pienellä todennäköisyydellä α ,

$$P(\text{tappio} > VaR) \leq \alpha,$$

ja se ilmaistaan usein rahamääräisenä. Esimerkiksi VaR laskettuna todennäköisyydellä 5 % yhden päivän ajanjaksolla kertoo, että tappio ylittää VaR:n keskimäärin viisi kertaa sadan päivän aikana. Matemaattisesti VaR on riskikohteen jakauman q :s kvantiili

$$VaR_q(X) = F^{-1}(1 - \alpha) = F^{-1}(q), \quad (21)$$

kun $q = 1 - \alpha$.

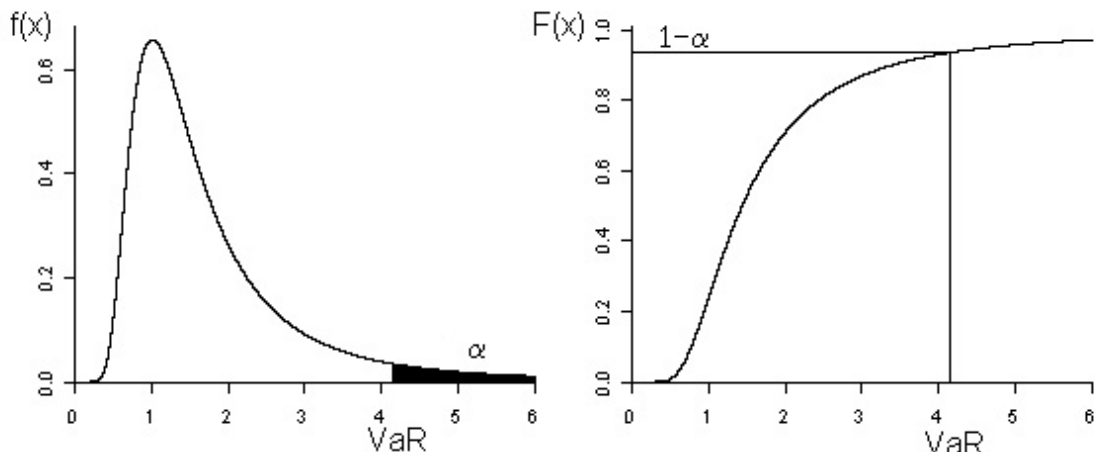
Kuva 3 havainnollistaa VaR:n määritelmää. VaR voidaan estimoida kolmella eri tavalla: historiallisella simulaatiolla, parametrisilla malleilla tai Monte Carlo simuloinnilla. Tässä tutkielmassa esitetään historiallinen simulaatio ja kolme parametrista tapaa, joissa käytetään normaali- ja t-jakaumaa sekä ääriarvoteorian kynnysarvon ylitys -menetelmää. Laajemman esityksen VaR:n estimoinnista ja käyttötavoista tarjoaa mm. Jorion (2007).

4.1 Historiallinen simulaatio

Historiallinen simulaatio (HS) on ei-parametrinen menetelmä, jossa ei tehdä mitään oletusta riskikohteen jakaumasta. Siinä käytetään riskikohteen historiallisia havaintoja VaR:n laskemiseen. Aineisto X_1, \dots, X_n järjestetään suuruusjärjestykseen $X^{(1)} \geq \dots \geq X^{(n)}$, jolloin $VaR_{1-\alpha}(X)$ on $X^{(i)}$, jota suurempia havaintoja $X^{(1)}, \dots, X^{(i-1)}, X^{(i)}$ aineistossa on α %.

Historiallinen simulaatio on helppo tapa laskea VaR, jos historiallista aineistoa on saatavilla. Historiallinen simulaatio saallii myös ei-lineaariset riippuvuudet sekä paksumat hännät, koska jakaumaoletusta ei tarvita. Se on myös hyvin intuitiivinen VaR:n estimointitapa. Ongelmia menetelmän käyttöön voi tuoda riittämätön aineisto. Pie-

Kuva 3: Value at risk



nellä määrällä havaintoja ei voida estimoida kvanttileja tarkasti, koska havaintomäärät hännissä ovat vähäiset ja keskivirheet siten suuret. Suurten kvantiilien estimointi on vaikeaa, koska ekstrapolointi havaintojen ulkopuolelle ei ole mahdollista. Menetelmä ei myöskään erota suuren ja pienen volatilitteen jaksoja toisistaan, joten ylitykset voivat klusteroitua yli ajan. Menetelmässä oletetaan tulevaisuus samanlaiseksi kuin menneisyys, mikä ei välttämättä ole realistista.

4.2 Value at risk normaali- ja t-jakauman avulla

Eräs tapa VaR:n estimointiin on parametrinen menetelmä, jossa riskikohteen oletetaan noudattavan normaalijakaumaa eli $F_X = N(\mu, \sigma^2)$. Jakauman parametrit estimoidaan joko historiallisesta aineistosta tai eteenpäin katsovasta optioaineistosta. Riskikohteen VaR luottamustasolla q on

$$\text{VaR}_q = F_X^{-1}(q) = \mu + \sigma\Phi^{-1}(q), \quad (22)$$

missä $\Phi^{-1}(q)$ on standardinormaalijakauman q :s kvantiili.

Normaalijakauman hännät ovat ohuet, joten jos riskikohteen jakauma onkin paksuhäntäinen, menetelmä aliarvioi riskiä suurilla luottamustasoilla. Jos riskikohteen jakauma olisi esimerkiksi paksumpihäntäinen t-jakauma ja käytetään normaalijakaumaa, pätee suurilla q :n arvoilla $F_N^{-1}(q) < F_t^{-1}(q)$ ja VaR_N on siten pienempi. Normaalijakaumaan perustuva menetelmä on hyvin yleinen tapa estimoida VaR, koska se on helppo ja nopea käyttää. Toinen vaihtoehto on olettaa riskikohde t-

jakautuneeksi ja käyttää t-jakauman kvanttiileja VaR:n estimointiin samaan tapaan kuin normaalijakauman tapauksessa (kaava 22). Ehdollistamattomina menetelminä normaali- ja t-jakaumamenetelmät eivät kuitenkaan huomioi heteroskedastisuutta, joten ilman aineiston suodattamista menetelmät eivät useinkaan toimi hyvin, kun käytetään taloudellisia muuttujia.

4.3 Value at risk kynnsarvon ylitys -menetelmällä

Kolmas parametrinen menetelmä VaR:n estimoimiseksi on olettaa riskikohteen kynnsarvon ylityksien noudattavan yleistettyä Pareto-jakaumaa jollakin suurella kynnsarvolla. Kynnsarvon valinta ja jakauman estimointi on kuvattu luvussa 3. Yhdistämällä VaR:n määritelmä (21) ja yleistetyn Pareto-jakauman kvantiiliestimaattori (15) saadaan

$$VaR_{q,\xi,\hat{\sigma}} = u + \frac{\hat{\sigma}}{\hat{\xi}} \left[\left(\frac{n}{N_u} \alpha \right)^{-\hat{\xi}} - 1 \right], \quad \text{kun } VaR_{q,\xi,\hat{\sigma}} > u, \quad (23)$$

missä $\alpha = 1 - q$ on todennäköisyys, että VaR:a suurempi tappio tapahtuu.

Tämä ehdollistamaton menetelmä sallii paksuhäntäiseen jakauman (kun $\xi > 0$) ja se perustuu häntien estimointiin, joten kvantiiliestimaattorit ovat tarkempia kuin ohuthäntäisillä jakaumilla. Kynnsarvon ylitys menetelmä kattaa useita erilaisia jakaumia, joista valitaan aineistoon parhaiten sopiva. Menetelmä sallii myös häntien epäsymmetrisyyden toisin kuin esimerkiksi normaalijakauma. Usein taloudellisten muuttujien jakaumat ovat hyvin vinoja, jolloin äärimmäisiä havaintoja esiintyy vain jakauman toisessa päässä. Tällöin yleistetty Pareto-jakauma sopii aineistoon paremmin kuin normaalijakauma. Ehdollistamattolla menetelmällä lasketut VaR-estimaatit eivät kuitenkaan reagoi esimerkiksi varianssin kasvuun, jolloin volatiilissa markkinatilanteessa useita VaR:n ylityksiä voi tapahtua lyhyellä aikavälillä.

Taloudellisiin aikasarjoihin liittyy usein autokorrelaatiota sekä heteroskedastisuutta, mikä tulisi ottaa huomioon ääriarvoteoriaa sovellettaessa. Ääriarvomenetelmät olettavat havaintojen olevan riippumattomia ja samoin jakautuneita, joten muuttuja täytyy suodattaa ennen ääriarvojakauman sovittamista. Yksi lähestymistapa on suodattaa muuttuja ensin AR- tai GARCH-mallilla, jolloin mallin residuaalit ovat lähes IID-havaintoja. Tämän jälkeen sovitaan yleistetty Pareto-jakauma standar-

doituihin residuaaleihin ja lasketaan näiden avulla ehdollistettu VaR (McNeil ja Frey, 2000). Tämä menetelmä ottaa huomioon riippuvuudet, varianssin vaihtelun sekä paksut hännät, joten VaR-estimaatit vaihtelevat ajassa ja sopivat siten kuvaamaan taloudellisia ilmiöitä. GARCH-mallinnus kuvataan seuraavassa luvussa.

4.4 Menetelmien testaus

Eri VaR:n estimointimenetelmiä voidaan verrata keskenään testaamalla menetelmän tarkkuutta. Tässä tutkielmassa testataan VaR:n ylittävien havaintojen ehdollistamatonta (unconditional) sekä ehdollista peittävyyttä (conditional coverage). Ehdollistamattossa peittävyudessa testataan ylitysten suhteellisen osuuden oikeellisuutta, kun taas ehdollisessa peittävyudessa ylitysten tulee myös olla riippumattomia toisistaan. Testaamiseen käytettävät testit esitetään tässä kappaleessa.

Ylitysten ehdollistamatonta peittävyyttä, eli ylitysten lukumäärän poikkeavuutta odotetusta, voidaan testata olettamalla ylityksien lukumäärä binomijakautuneeksi. Tällöin testisuure on muotoa (Fernandez, 2003)

$$BT = \frac{\frac{N}{n} - \alpha}{\sqrt{\frac{\alpha(1-\alpha)}{n}}}, \quad (24)$$

missä N on menetelmällä havaittujen ylitysten lukumäärä ja $\alpha = 1 - q$ on odotettu ylitysten todennäköisyys. Esimerkiksi 95 %:n VaR:n ylittäviä havaintoja tulisi olla 5 % kaikista havainnoista. Nollahypoteesin pätiessä testisuure BT noudattaa approksimatiivisesti $N(0,1)$ -jakaumaa. Ylitysten lukumäärän tarkkuutta testataan yksisuuntaisella testillä, jossa nollahypoteesina on $N/n = \alpha$. Vastahypoteeseina on, että menetelmä joko aliarvioi systemaattisesti kvantiileja eli ylityksiä on odotettua enemmän, jos $N/n > \alpha$ tai että menetelmä systemaattisesti yliarvioi kvantiileja eli ylityksiä odotettua vähemmän, jos $N/n < \alpha$. Koska testisuureen laskemiseen ei tarvita jakaumaoletusta, samaa testiä voidaan käyttää jokaiselle menetelmälle. Binomitesti testaa vain ylitysten ehdollistamatonta peittävyyttä eli sitä onko ylitysten suhteellinen osuus oikea koko ajanjaksolla. Kuitenkin tärkeää olisi myös, että menetelmä antaisi oikean ylitysten määrän lyhyemmällä ajanjaksolla, jolloin menetelmä pystyy erottamaan korkean ja matalan hajonnan ajanjaksot (Chan ja Gray, 2006).

Christoffersen (1998) osoitti, että ehdollisen peittävyuden testaus voidaan jakaa eh-

dollistamattoman peittävyiden ja riippumattomuuden testaamiseen. Kun menetelmän tulokset ovat ehdollisesti peittäviä, ylitykset tapahtuvat tasaisesti yli ajan ja ylitysten suhteellinen osuus on oikea. Christoffersenin esittämät uskottavuusosamäärätestit eivät perustu mihinkään havainnoista tehtävään jakaumaoletukseen, joten ne voidaan laskea jokaiselle VaR:n estimointimenetelmälle.

Määritellään VaR:n ylitys seuraavasti:

$$I_t = \begin{cases} 1, & \text{kun } y_t > VaR_q \\ 0, & \text{kun } y_t \leq VaR_q, \end{cases}$$

missä y_t on aikasarjan havaittu arvo ja VaR_q on valitulla menetelmällä laskettu q :s häntäkvantiili. Ehdollistamattoman peittävyiden uskottavuusosamäärätestin nollahypoteesi on $E(I_t) = \alpha$, kun $\alpha = 1 - q$ ja vastahypoteesi $E(I_t) \neq \alpha$, annettuna riippumattomuus. Uskottavuusfunktio on

$$L(\pi; I_t) = (1 - \pi)^{N_0} \pi^{N_1},$$

missä $\pi = E(I_t)$, N_1 on ylityksien lukumäärä, N_0 VaR:n alittavien havaintojen lukumäärä ja $t = 1, 2, \dots, T$. Nollahypoteesi on $\pi = \alpha = 1 - q$. Uskottavuusosamäärätesti ehdollistamattomalle peittävyydelle saadaan kaavalla

$$LR_{uc} = -2 \log \left[\frac{L(\alpha; I_t)}{L(\hat{\pi}; I_t)} \right],$$

missä $\hat{\pi} = N_1 / (N_0 + N_1)$ on π :n SU-estimaatti. Testisuure LR_{uc} noudattaa asymp-totottisesti χ^2 -jakaumaa vapausasteella 1. Tässä testissä kuten binomitestissäkin vain ylitysten suhteellinen osuus määrää testin tuloksen, mutta ylitysten riippuvuus ei.

Riippumattomuutta voidaan testata olettamalla ylitysten noudattavan Markovin ketjua. Markovin ketjussa havainto riippuu vain edellisestä havainnosta, mutta ei mistään muusta havainnosta. Tällöin prosessin I_t todennäköisyysjakauman määrittelevät siirtymätodennäköisyydet $\pi_{ij} = P(I_t = j \mid I_{t-1} = i)$, jotka voidaan esittää matriisina

$$\mathbf{\Pi}_1 = \begin{bmatrix} 1 - \pi_{01} & \pi_{01} \\ 1 - \pi_{11} & \pi_{11} \end{bmatrix}.$$

Tämän prosessin uskottavuusfunktiota voidaan approksimoida lausekkeella

$$L(\mathbf{\Pi}_1; I_t) = (1 - \pi_{01})^{N_{00}} \pi_{01}^{N_{01}} (1 - \pi_{11})^{N_{10}} \pi_{11}^{N_{11}},$$

missä N_{ij} on niiden havaintojen lukumäärä, joissa arvo i seuraa arvo j . Kun ehdollistetaan ensimmäinen havainto, voidaan ratkaista siirtymätodennäköisyyksien matriisiin SU-estimaatti

$$\hat{\mathbf{\Pi}}_1 = \begin{bmatrix} \frac{N_{00}}{N_{00}+N_{01}} & \frac{N_{01}}{N_{00}+N_{01}} \\ \frac{N_{10}}{N_{10}+N_{11}} & \frac{N_{11}}{N_{10}+N_{11}} \end{bmatrix}.$$

Testin nollahypoteesi on, että prosessi I_t on riippumaton, jolloin siirtymätodennäköisyyksien matriisi voidaan esittää muodossa

$$\mathbf{\Pi}_2 = \begin{bmatrix} 1 - \pi_2 & \pi_2 \\ 1 - \pi_2 & \pi_2 \end{bmatrix}.$$

Tällöin uskottavuusfunktio nollahypoteesin pätiessä on

$$L(\mathbf{\Pi}_2; I_t) = (1 - \pi_2)^{(N_{00}+N_{10})} \pi_2^{(N_{01}+N_{11})}$$

ja SU-estimaatti $\hat{\pi}_2 = (N_{01} + N_{11}) / (N_{00} + N_{10} + N_{01} + N_{11})$. Riippumattomuutta testaavan uskottavuusosamäärän testisuure on

$$LR_{ind} = -2 \log \left[\frac{L(\hat{\mathbf{\Pi}}_2; I_t)}{L(\hat{\mathbf{\Pi}}_1; I_t)} \right]$$

ja se noudattaa asymptoottisesti $\chi^2(1)$ -jakaumaa. Testisuure LR_{ind} ei riipu oikeasta ylitysten suhteellisesta osuudesta α , joten se testaa vain ylitysten riippumattomuutta.

VaR:n estimointimenetelmän hyvyttä arvioitaessa mielenkiinnon kohteena on testata oikeaa ylitysten osuutta ja riippumattomuutta yhdessä. Tällöin nollahypoteesina on ehdollinen peittävyys, eli $\pi = \alpha$, ja vastahypoteesina on Markovin ketjua noudattava prosessi. Ehdollista peittävyyttä testaavan uskottavuusosamäärän testisuure on siten

$$LR_{cc} = -2 \log \left[\frac{L(\alpha; I_t)}{L(\hat{\mathbf{\Pi}}_1; I_t)} \right]$$

ja se noudattaa asymptoottisesti $\chi^2(2)$ -jakaumaa. Kun sarjan I_t ensimmäinen havainto poistetaan, uskottavuusosamäärätesteille pätee

$$LR_{cc} = LR_{uc} + LR_{ind}.$$

(Christoffersen, 1998).

5 GARCH-mallit

Seuraavat tulokset esittävät laajemmin mm. Hamilton (1994) sekä Franses ja van Dijk (2000). Astetta p oleva autoregressiivinen prosessi (merkitään $AR(p)$) määritellään yhtälöllä

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \epsilon_t,$$

missä virhetermi ϵ_t on valkoista kohinaa:

$$E(\epsilon_t) = 0$$

$$E(\epsilon_t \epsilon_\tau) = \begin{cases} \sigma^2 & \text{kun } t = \tau, \\ 0 & \text{muulloin.} \end{cases}$$

$AR(p)$ -prosessi on heikosti stationaarinen, kun yhtälön

$$1 - \phi_1 z - \phi_2 z^2 - \dots - \phi_p z^p = 0 \tag{25}$$

juuret sijaitsevat yksikköympyrän ulkopuolella ja vahvasti stationaarinen, jos $\epsilon_t \sim \text{IID}(0, \sigma^2)$

AR -malleilla pyritään mallintamaan aikasarjan käyttäytymistä, kun oletetaan, ettei prosessin varianssi muutu ajassa. Sen sijaan $ARCH$ - ja $GARCH$ -malleilla ((generalised) autoregressive conditional heteroscedasticity) mallinnetaan aikasarjan varianssia tai hajontaa. $ARCH$ - ja $GARCH$ -mallit ovat epälineaarisia malleja, jotka sallivat ehdollisen varianssin vaihtelun ja klusteroitumisen. Mallien lähtökohta on, että virhetermin varianssi ajanhetkellä t riippuu aiemmista neliöidyistä virhetermeistä.

Olkoon u_t jonkin mallin virhetermi (esim. $AR(p)$) ja $\Omega_{t-1} = \{u_{t-j}, j \geq 1\}$ ajanhetkellä $t-1$ havaittu informaatiojoukko. Tällöin $ARCH(q)$ -prosessi on

$$u_t = \sigma_t \epsilon_t,$$

$$\sigma_t^2 \equiv E(u_t^2 | \Omega_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2 + \alpha_2 u_{t-2}^2 + \dots + \alpha_q u_{t-q}^2, \tag{26}$$

missä $\epsilon_t \sim \text{IID}(0,1) \perp \{u_{t-j} \geq 1\}$, $\alpha_0 > 0$ ja $\alpha_i \geq 0$ kaikilla $i = 1, 2, \dots, q$, jotta ehdollinen varianssi olisi positiivinen. $ARCH$ -prosessin ehdollisen varianssin vaihtelun määrää prosessi u_t menneisyys. Koska u_t :n ehdollinen varianssi on u_{t-1} :n funktio, u_t ja u_{t-1} eivät ole riippumattomia, mutta ne ovat korreloimattomia. Tämä

nähdään käyttämällä iteroitujen odotusarvojen lakia

$$\begin{aligned} E(u_t u_{t-1}) &= E[E(u_t u_{t-1} \mid \Omega_{t-1})] = E[u_{t-1} \sigma_t E(\epsilon_t \mid \Omega_{t-1})] \\ &= E[u_{t-1} \sigma_t E(\epsilon_t)] = 0, \end{aligned}$$

koska $\epsilon_t \perp \Omega_{t-1}$ ja $E(\epsilon_t) = 0$. Samoin voidaan osoittaa, että $E(u_t u_s) = 0 \forall s < t$.

Jos ARCH(q)-prosessi on heikosti stationaarinen, u_t^2 :n ehdollistamaton odotusarvo σ^2 on riippumaton t :stä. Tämä nähdään laskemalla

$$\sigma^2 = E(u_t^2) = E(\sigma_t^2 \epsilon_t^2) = E(\sigma_t^2) E(\epsilon_t^2) = E(\sigma_t^2),$$

koska $E(\epsilon_t^2) = 1$ ja ottamalla odotusarvo puolittain yhtälöstä (26) saadaan

$$E(\sigma_t^2) = \alpha_0 + \alpha_1 E(u_{t-1}^2) + \dots + \alpha_q E(u_{t-q}^2) = \alpha_0 + \sigma^2 \sum_{i=1}^q \alpha_i,$$

joten

$$\sigma^2 = \frac{\alpha_0}{1 - \sum_{i=1}^q \alpha_i}.$$

Jotta σ^2 olisi olemassa ja positiivinen, täytyy päteä $\sum_{i=1}^q \alpha_i < 1$, jolloin prosessi u_t on sekä vahvasti että heikosti stationaarinen.

Käytännössä ARCH-mallit eivät useinkaan kuvaa hyvin taloudellisia aikasarjoja. Yleistetyt ARCH-mallit eli GARCH-mallit ottavat paremmin huomioon ehdollisen varianssin käyttäytymisen, joka on tyypillistä taloudellisille aikasarjoille. GARCH-malleja on käytetty laajasti useisiin taloudellisiin aikasarjoihin. Niiden suosio perustuu siihen, että ne kuvaavat hajonnan pysyvyyttä ja jakauman paksuja häntiä ja että ne ovat matemaattisesti yksinkertaisia ja siten helposti laskettavissa (Verhoeven ja McAleer, 2000).

GARCH(p,q)-prosessi voidaan kirjoittaa muodossa

$$\begin{aligned} u_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &\equiv E(u_t^2 \mid \Omega_{t-1}) = \alpha_0 + \sum_{i=1}^q \alpha_i u_{t-i}^2 + \sum_{j=1}^p \delta_j \sigma_{t-j}^2, \end{aligned}$$

missä $\alpha_0 > 0$, $\alpha_i \geq 0$ ja $\delta_j \geq 0 \forall i, j$, jotta σ_t^2 olisi positiivinen. Yksinkertaisin ja suosituin malli on GARCH(1,1)-prosessi, jossa ehdollinen varianssi on

$$\sigma_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2. \quad (27)$$

GARCH(1,1)-mallin ehdollinen varianssi on siten viivästetyn ehdollistamattoman ja ehdollisen varianssin lineaarikombinaatio. ARCH-parametri α_1 mittaa edellisen ajanhetken jäännöksen vaikutusta ja tuo siten hajontaan äkillisiä muutoksia. GARCH-parametri δ_1 puolestaan mittaa edellisen ehdollisen varianssin vaikutusta nykyiseen arvoon ja tuo hajontaan tasaisempia muutoksia. Kun parametri α_1 on pieni ja δ_1 suuri, ehdolliseen varianssiin tullut shokki häviää hitaasti.

Kun prosessi on heikosti stationaarinen, u_t :n ehdollistamaton varianssi σ^2 (eli u_t^2 ehdollistamaton odotusarvo) saadaan ottamalla odotusarvo molemmin puolin yhtälöstä (27), jolloin

$$\sigma^2 = \alpha_0 + \alpha_1 \sigma^2 + \delta_1 \sigma^2.$$

Ratkaisemalla yhtälö σ^2 :n suhteen saadaan

$$\sigma^2 = \frac{\alpha_0}{1 - \alpha_1 - \delta_1}.$$

Jotta varianssi olisi olemassa, täytyy päteä $\alpha_1 + \delta_1 < 1$. Tällöin prosessi u_t on sekä vahvasti että heikosti stationaarinen. Termi $\alpha_1 + \delta_1$ mittaa ehdollisen varianssin pysyvyyttä. Jos summa on lähellä ykköstä, varianssiin tullut shokki kuoleutuu hitaasti.

Usein käytettäessä GARCH(1,1)-mallia taloudellisten tiheäjaksoisten aikasarjojen kuvaamiseen, parametrien α_1 ja δ_1 summa on lähellä ykköstä. GARCH-mallia tilanteessa $\alpha_1 + \delta_1 = 1$ kutsutaan integroiduksi GARCH-malliksi (IGARCH). Kun parametrien summa on suurempi tai yhtä suuri kuin yksi, ehdollistamaton varianssi ei ole äärellinen ja siten malli ei ole heikosti stationaarinen. Se voi kuitenkin olla vahvasti stationaarinen, mikäli pätee

$$E[\log(\alpha_1 \epsilon_t^2 + \delta_1)] < 0. \tag{28}$$

Ehdon tarkastaminen käytännössä on ongelmallista, koska se edellyttää jäännöksen ϵ_t jakauman tuntemista. Nelson (1990) on määritellyt GARCH(1,1)-prosessin parametrien avulla tilanteet, joissa prosessi on epästationaarinen, vahvasti tai heikosti stationaarinen. Jäännöksen u_t neljännen momentin olemassaololle riittävä ehto on $(k\alpha_1^2 + 2\alpha_1\delta_1 + \delta_1^2) < 1$, missä k on ϵ_t :n neljäs momentti. Normaalijakaumaoletuksen tapauksessa $k = E(\epsilon_t^4) = 3$ (Verhoeven ja McAleer, 2000).

5.1 AR-GARCH-mallit

Yhdistämällä AR- ja GARCH-prosessit saadaan malli, joka huomioi sekä aikasarjan aikariippuvuuden että volatilitiitin klusteroitumisen. Tässä tutkielmassa AR-GARCH-mallilla tarkoitetaan mallia, joka sisältää aikasarjan omia viivästettyjä arvoja ja GARCH(1,1)-jäännökset:

$$\begin{aligned}y_t &= c + \sum_{i=1}^p \phi_i y_{t-i} + u_t \\u_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2,\end{aligned}\tag{29}$$

missä σ_t^2 on u_t :n ehdollinen varianssi ja $\epsilon_t \sim \text{IID}(0,1)$. Ehdollinen varianssi on positiivinen, kun $\alpha_0 > 0$, $\alpha_1 \geq 0$ ja $\delta_1 \geq 0$. Prosessi on heikosti ja vahvasti stationaarinen, kun

$$1 - \phi_1 z - \dots - \phi_n z^n \neq 0, \quad \text{kun } |z| \leq 1$$

ja $\alpha_1 + \delta_1 < 1$ sekä vahvasti stationaarinen, kun ehto (28) pätee.

Mallin (29) parametrit voidaan estimoida SU-menetelmällä, kun oletetaan, että jäännökset ϵ_t ovat normaali- tai t-jakautuneita. Vaikka normaalisuusoletus ei usein päde empiirisille aikasarjoille, normaaliseen uskottavuuteen perustuvalla menetelmällä saadaan estimoitua parametrit tarkentuvasti ja estimaatit ovat asympotoottisesti normaalijakautuneita, kun malli on vahvasti stationaarinen ja GARCH(1,1)-prosessin neljännen momentin ehto pätee. (Franses ja van Dijk, 2000, 172). Tätä menetelmää kutsutaan suurimman kvasi-uskottavuuden menetelmäksi. Kvasi-uskottavuuden estimaatit ovat tarkentuvia myös ilman 4. momenttia koskevaa ehtoa, eli esimerkiksi IGARCH(1,1)-mallin tapauksessa. (Francq ja Zakoïan, 2004). Kun jäännösten oikea jakauma ei ole normaalijakauma, jota käytetään uskottavuusfunktion laskemiseen, täytyy keskivirheet laskea robusteilla menetelmillä. (Franses ja van Dijk, 2000, 172-173). Usein empiirisissä aikasarjoissa toisen tai neljännen momentin ehdot eivät päde. Tällöin GARCH-parametrien t-arvot eivät kuvaa oikein parametrien merkittävyyksiä (Verhoeven ja McAleer, 2000). ARMA-GARCH-mallien kvasi-uskottavuuden estimaattien tarkentuvuuden ja asympotoottisen normaalisuuden täsmälliset ehdot ja tulokset antavat Francq ja Zakoïan (2004).

Kun aikasarja sisältää poikkeavia havaintoja, ne voivat vaikuttaa haitallisesti AR-

GARCH-mallin kvasi-uskottavuuden parametreihin. Verhoeven ja McAleer (2000) toteavat tutkimuksessaan, että poikkeavat havainnot dominoivat AR(1)-GARCH(1,1)-mallin estimaatteja, jolloin ne pienentävät AR- ja ARCH-parametrien ja suurentavat GARCH-parametrin estimaatteja. Heidän mukaansa poikkeavat havainnot voivat aiheuttaa neljännen momentin ehdon hylkäämisen ja parametrien t-arvojen pienene-
misen. Ne myös kasvattavat hajonnan ennusteita ja voivat aiheuttaa autoregressii-
visyyttä ja heteroskedastisuutta. (Verhoeven ja McAleer, 2000).

Häntäkvantiilit eli VaR:t voidaan estimoida myös käyttämällä yllä olevaa AR-GARCH-mallia. Kun malli on sovitettu aineistoon, oletetaan jäännöksen ϵ_t noudattavan esimerkiksi standardinormaali- tai t-jakaumaa varianssilla yksi. Tällöin ehdollinen kvantiili $x_{t,q}$ saadaan kertomalla oletetun jakauman kvantiili ehdollisella hajonnalla $\hat{\sigma}_t$ ja lisäämällä ehdollinen odotusarvo (vrt. kaava (22)):

$$\hat{x}_{t,q} = \hat{\mu}_t + \hat{\sigma}_t F^{-1}(q), \quad (30)$$

missä $\hat{\mu}_t = \hat{c} + \sum_{i=1}^p \hat{\phi}_i y_{t-i}$ ja $F^{-1}(q)$ on q :s kvantiili normaali- tai t-jakaumasta odotusarvolla nolla ja varianssilla yksi. Tätä laskentatapaa merkitään AR-GARCH-malliksi, kun oletetaan normaalijakautuneet jäännökset ja AR-GARCH-t-malliksi kun oletetaan t-jakauma. Tämä menetelmä ottaa huomioon aikasarjan aikariippu-
vuudet, muttei äärimmäisiä havaintoja yhtä hyvin kuin ääriarvomenetelmät.

AR-GARCH-mallin jäännöksiä voidaan mallintaa ääriarvojakaumalla, joka ottaa huomioon jakauman paksut hännät ja mahdollisen vinouden. Tällöin normaalisuus-
oletuksella estimoidun AR-GARCH-mallin (29) jäännöksiä $\hat{\epsilon}_t$ mallinnetaan aluksi kynnysarvon ylitys -menetelmällä (kappale 2.1.2). Koska jäännökset ovat lähes IID-havainnoja, ääriarvoteorian soveltaminen on perusteltua. Jäännösten jakauman eh-
dollistamattomat kvantiiliestimaatit \hat{x}_q^c voidaan laskea kaavan (15) avulla, kun ja-
kauman parametrit on estimoitu. Sen jälkeen alkuperäisen aikasarjan ehdolliset kvantiiliestimaatit $\hat{x}_{t,q}$ saadaan skaalaamalla jäännösten kvantiilit seuraavasti

$$\hat{x}_{t,q} = \hat{\mu}_t + \hat{\sigma}_t \hat{x}_q^c, \quad (31)$$

missä $\hat{\mu}_t$ ja $\hat{\sigma}_t$ ovat AR-GARCH-mallin estimoitu ehdollinen odotusarvo ja hajonta. Tätä kvantiilien estimointimenetelmää kutsutaan AR-GARCH-EVT-malliksi.

6 Nord Poolin sähkön hinta

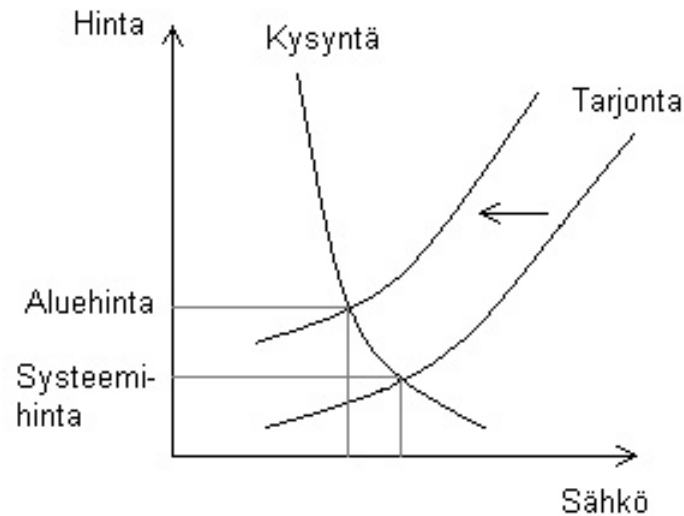
Nord Pool on pohjoismainen sähköpörssi, jossa sähkön tuottajat, vähittäismyyjät ja suuret sähkön kuluttajat ostavat ja myyvät sähköä. Pörssin markkina-alueeseen kuuluvat Suomi, Ruotsi, Norja ja Tanska, mutta kaupankäyntiin osallistuu sähköyhtiöitä myös mm. Hollannista ja Iso-Britanniasta. Nord Pool on perustettu vuonna 1993 ja se on maailman ensimmäinen monikansallinen sähköpörssi. Ensimmäiset suomalaiset jäsenet liittyivät mukaan vuoden 1996 lopulla. Pohjoismaisesta sähköstä n. 60 % ja suomalaisesta n. 40 % myytiin Nord Poolin kautta vuonna 2006. Pörssissä käydään kauppaa kolmenlaisilla markkinoilla. Spot-markkinoilla määrätään seuraavan päivän tukkuhinnat tunneittain kysynnän ja tarjonnan perusteella. Johdannaismarkkinoilla käydään kauppaa sähköfutuureilla, -forwardeilla ja -optioilla. Elbas- eli tasesähkömarkkinat ovat jälkimarkkinat, jossa kauppaa voi käydä saman päivän aikana toimitettavasta sähköstä. Nord Poolin Internet-sivut tarjoavat lisätietoja sähköpörssin toiminnasta ja hinnan määräytymisestä (Nord Pool).

6.1 Spot- ja aluehintojen määräytyminen

Spot-markkinoilla toimijat antavat puolilta päivin tarjouksensa siitä, kuinka paljon ja millä hinnalla he ovat valmiita myymään tai ostamaan sähköä seuraavana päivänä kullakin tunnilla. Tämän jälkeen osto- ja myyntitarjoukset yhdistetään kysyntä- ja tarjontakäyriksi, joiden leikkauspisteessä määräytyy tunnittainen systeemihinta. Hinta ei huomioi sähkön siirtorajoituksia. Systeemihinnasta muodostetaan jokaiselle markkina-alueelle aluehinta, jossa siirtorajoitukset otetaan huomioon. Jos siirtokapasiteettia ei ole riittävästi tai alueella on yli- tai alituotantoa, aluehinnat poikkeavat systeemihinnasta. Aluehinnan määräytymistä siirtokapasiteetin pienentyessä kuvataan kuviossa 4. Tällöin tarjontakäyrää siirretään vasemmalle siirtokapasiteetin rajoituksen verran. Uudessa leikkauspisteessä määräytyy aluehinta. Tässä tutkimuksessa sähkön hinnan mallinnukseen käytetään Suomen tunnittaista aluehintaa ja sitä käsitellään aitona spot-hintana, vaikka normaalitilanteessa se on lyhytaikainen futuurihinta seuraavalle päivälle.

Spot-hinnoissa on usein suurta vaihtelua ja poikkeavia arvoja, koska sähköä ei voi

Kuva 4: Aluehinnan määräytyminen siirtokapasiteetin aletessa



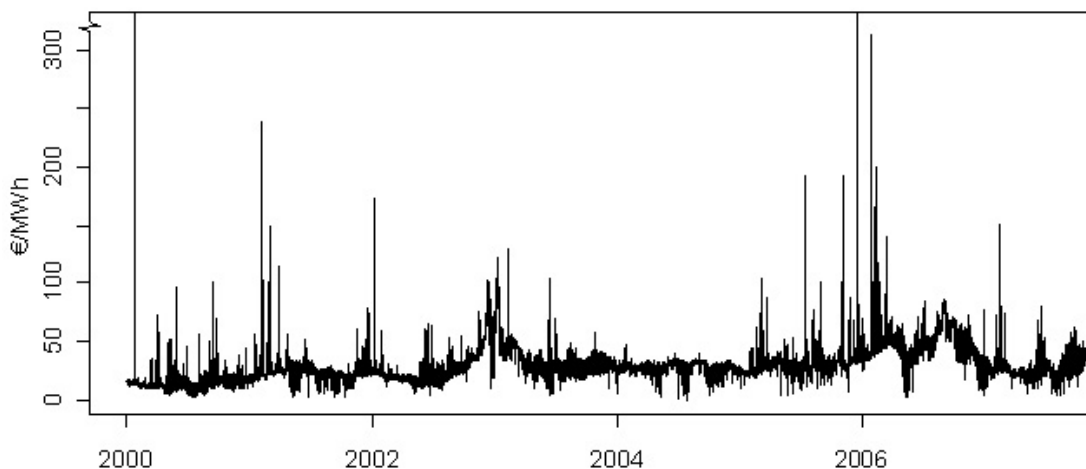
juurikaan varastoida ja sitä tuotetaan ja toimitetaan kysynnän mukaan. Kysynnän ja tarjonnan on oltava koko ajan tasapainossa, jotta sähkönjakelu ei keskeytyisi. Siten kysynnässä tai tarjonnassa tapahtuvat äkilliset muutokset (esim. voimalaitoksen sulkeminen) vaikuttavat nopeasti hintaan. Lisäksi sähkön kysyntä on hyvin joustamatonta (Chan ja Gray, 2006). Ääriarvoteoria sopii siten hyvin kuvaamaan sähkömarkkinoita. Spot-hinnoilla on myös hyvin suuri merkitys sähkön johdannaisten hinnoittelussa.

Markkinahinnoissa esiintyy hyvin paljon kausivaihtelua, joka johtuu kysynnän vaihtelusta sekä varastojen puuttumisesta. Sähkön kysyntä on suurempaa päivällä ja arkin kuin yöllä ja viikonloppuisin. Sähkön kulutus riippuu myös lämpötilasta, joten sähkön hinnat ovat korkeammat talvella kuin kesällä. (Byström, 2005). Seuraavassa kappaleessa kuvataan tutkimuksessa käytettävä sähkömarkkina-aineisto.

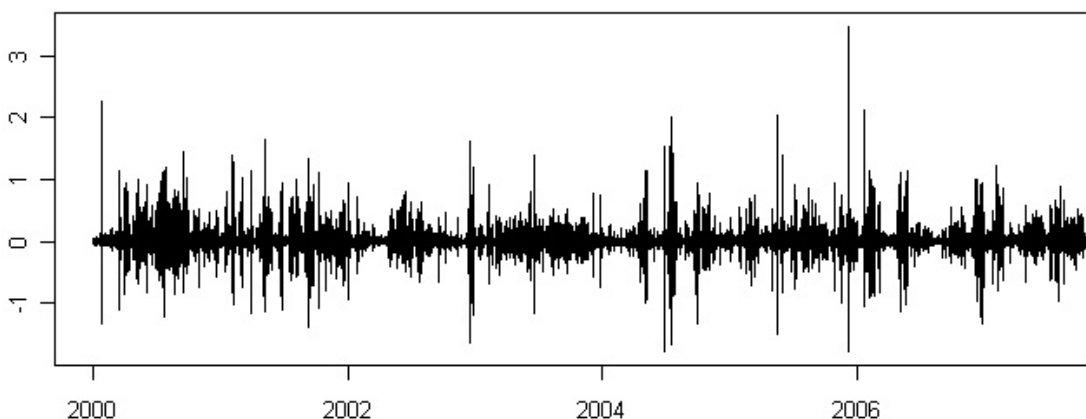
6.2 Aineiston kuvailu

Käytettävä aineisto on Nord Poolin Suomen tunnitaiset aluehinnat megawattitunnilta (€/MWh) tammikuusta 2000 marraskuuhun 2007. Tuntihintoja aineistossa on yhteensä 68880 kappaletta. Kesä- ja talviaikaan siirtymisen takia maaliskuu- ja lokakuun viimeisen sunnuntain kolmas hinta on laskettu keskiarvona edellisestä ja seu-

Kuva 5: Sähkön hinta (€/MWh) 1.1.2000–9.11.2007. Arvoasteikko katkaistu



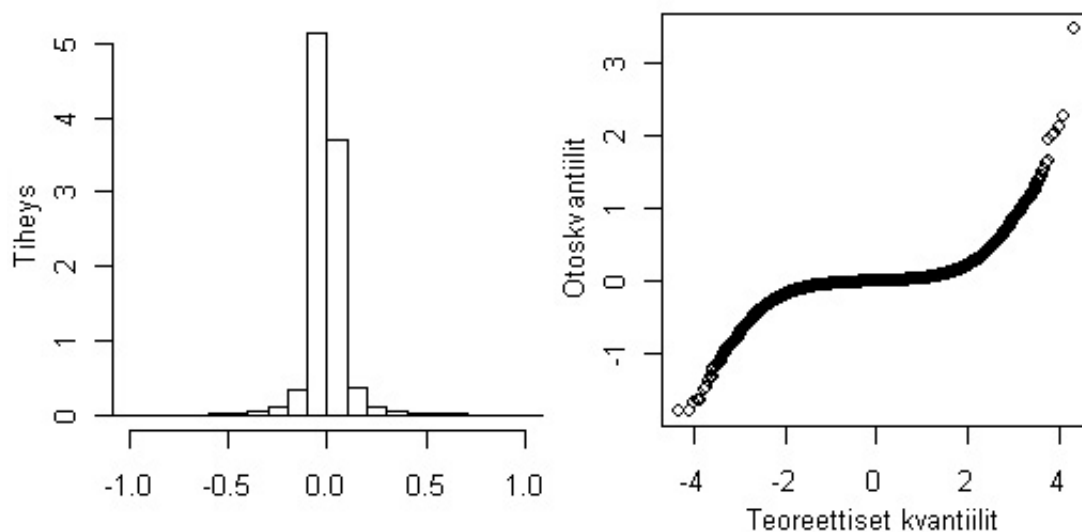
Kuva 6: Sähkön hinnan logaritmuutokset 1.1.2000–9.11.2007.



raavasta tuntihinnasta. Sähkön hinnan aikasarja on piirrettynä kuvassa 5. Kuvion asteikko on katkaistu arvosta 320 €/MWh, jota suurempia havaintoja aineistossa on neljä (776,85, 337,45 ja 337,34 €/MWh tammikuussa 2000 sekä 1147,45 €/MWh joulukuussa 2005). Kuviosta nähdään, että sähkönhinnan vaihtelu voi olla todella suurta jopa tunnista toiseen ja se on hyvin heteroskedastista. Suuret havainnot ovat myös klusteroituneet eli lyhyessä ajassa poikkeavia havaintoja on useita ennen kuin hinta palautuu normaalille tasolle. Aikasarjassa näkyy myös autokorrelaatiota sekä kausivaihtelua.

Koska hinnat vaihtelevat suuresti, käytän mallinnukseen hinnan logaritmuutoksia $r_t = \log(P_t) - \log(P_{t-1})$. Tässä tutkielmassa tarkastellaan vain jakauman oikean puoleista häntää (suuria hinnan nousuja). Suuria hinnan pudotuksia voidaan käsitellä ääriarvoteoriolla samoin kuin nousujakin käyttämällä kaavaa (1). Esimerkik-

Kuva 7: Sähkön hinnan logaritmimuutoksen histogrammi ja kvantiilikuvio



si sähköä paljon kuluttavat teollisuuslaitokset ovat kiinnostuneita hinnan suurista nousuista, kun taas sähkön tuottajia kiinnostaa enemmän hinnan alenemiset. Tun-
tihintojen logaritmimuutosten aikasarja on esitetty kuvassa 6. Hinnan muutosten
jakauma on selvästi paksuhäntäinen, koska poikkeavia havaintoja on paljon. Kuvaa-
jassa näkyy myös heteroskedastisuutta ja volatilitietin klusteroitumista, jonka takia
GARCH-malli voisi kuvata hinnan muutoksia hyvin.

Paksuhäntäisyys ilmenee myös hinnan muutosten histogrammista ja kvantiilikuvios-
ta (kuva 7). Histogrammin asteikko on katkaistu sekä ylä- että alapäässä, joten
pienimmät ja suurimmat havainnot eivät kuviossa näy. Jakauman normaalisuutta
testataan kvantiilikuviolla. Jos kvantiilikuvion pisteet ovat suoralla, havainnot ovat
normaalisia. Kuviossa 7 nähdään, ettei hinnan muutokset ole normaalisia, vaan ja-
kauman hännät ovat normaalijakaumaa pidemmät ja paksummat.

Tunnittaisten hinnan logaritmimuutosten tunnusluvut on esitetty taulukossa 1 koko
aineistolle sekä estimointijaksolle eli vuosille 2000–2004 (in-sample). Vuosia 2005–
2007 käytetään havaintojakson ulkopuoliseen ennustamiseen ja testaukseen (ns. out-
of-sample). Tunnusluvut vahvistavat käsityksen suuresta vaihtelusta, paksuista hän-
nistä ja ei-normaalisuudesta, koska hinnan muutosten hajonta, vinous ja huipuk-
kuus ovat suuria. Normaalisuusoletus hylätään Jarque-Bera-testillä. Korkeat Ljung-
Box-testisuureet (Q) osoittavat voimakasta autokorrelaatiota ja suuret Q^2 -testisuu-
reet puolestaan heteroskedastisuutta ja poikkeavien havaintojen klusteroitumista

Taulukko 1: Hinnan logaritminmuutosten tunnusluvut

Tunnusluku	Koko aineisto		Havaintojakso	
		p-arvo		p-arvo
Alkupvm	1.1.2000		1.1.2000	
Loppupvm	9.11.2007		31.12.2004	
Havaintoja	68879		43848	
Keskiarvo	$1.358 \cdot 10^{-5}$		$6,919 \cdot 10^{-6}$	
Mediaani	-0,0031		-0,0034	
Minimi	-1,787		-1,785	
Maksimi	3,475		2,271	
Hajonta	0,109		0,112	
Vinous	1,544		1,149	
Huipukkuus	60,97		44,62	
Jarque-Bera	$1,07 \cdot 10^7$	0,000	$3,65 \cdot 10^6$	0,000
Q(6)	7674	0,000	5874	0,000
Q(24)	26444	0,000	19489	0,000
$Q^2(6)$	9798	0,000	8838	0,000
$Q^2(24)$	13405	0,000	13318	0,000

(GARCH-vaikutus). Testisuureet lasketaan kaavalla

$$Q(h) = (n(n+2)) \sum_{j=1}^h \frac{\rho_j^2}{n-j},$$

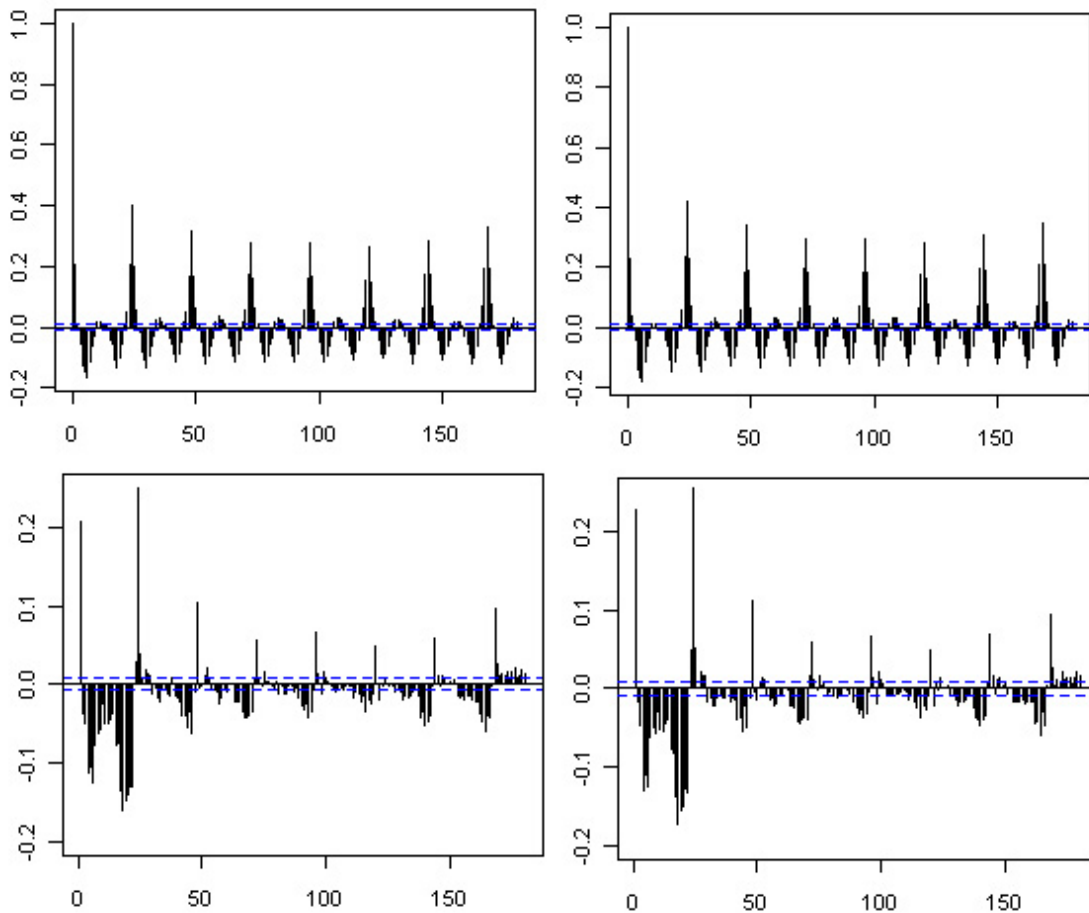
missä n on otoskoko ja ρ_j on havaintojen j :s autokorrelaatio. Q^2 -testissä ρ_j lasketaan neliöidyistä havainnoista. Ljung-Box-testisuureet noudattavat $\chi^2(h)$ -jakaumaa nollahypoteesin IID-havainnoista pätiessä.

Taulukosta 1 nähdään myös, että vuoden 2005 joulukuussa tapahtuneella suurimmalla hinnan nousulla (3,48) on suuri vaikutus muuttujan vinouteen, huipukkuuteen ja Ljung-Box-testisuureisiin. Kuvasta 8 nähdään, että hinnan muutos on autokorreloitunut ja siinä esiintyy vuorokausittaista (viive 24) ja viikottaista (viive 168) kausivaihtelua.

7 Sähkön hinnan mallinnus ja VaR

Tässä kappaleessa sovelletaan aiemmin esitettyä AR-GARCH-EVT-menetelmää sähkön tunnitusten hinnan muutosten kvantiilien estimointiin. Kappaleessa esitetään aikaisempia tutkimuksia VaR:n estimoinnista ääriarvoteorian avulla. Aineistoon sopivan AR-GARCH-EVT-mallin valinta kuvataan kappaleessa 7.2. Valittua menetelmää käytetään kvantiilien estimointiin havaintojaksolla sekä sen ulkopuolella VaR:n ennustamiseen. Ennustemenetelmä kuvataan kappaleen lopussa. Lisäksi ääriarvoteoriaan perustuvan menetelmän tuloksia verrataan tutkielmassa esitettyihin muihin VaR:n estimointimenetelmiin.

Kuva 8: Sähkön hinnan logaritmuutoksen autokorrelaatiot (ylhällä) ja osittaisautokorrelaatiot (alhaalla). Vasemmalla koko aineistolle, oikealla estimoitijaksolle laskettuna



7.1 Aikaisempia tutkimuksia

Ääriarvoteoriaa on käytetty VaR:n laskemiseen 1990-luvulta alkaen. Ensimmäisissä tutkimuksissa ääriarvojakaumat sovitettiin taloudellisiin aikasarjoihin jaksomaksimimenetelmällä ja myöhemmin kynnsarvon ylitys -menetelmää käyttäen. Näissä ehdollistamattomissa menetelmissä unohdetaan aikasarjoissa esiintyvät aikariippuvuudet ja varianssin vaihtelu. Myöhemmin ääriarvoteoriaa on käytetty ehdollisten VaR:ien estimointiin käyttämällä kaksivaiheista mallinnusta, jossa aikasarja ensin suodatetaan esimerkiksi AR- ja GARCH-malleilla, jotta ääriarvoteorian oletukset toteutuisivat mallin jäännöksille. Toisessa vaiheessa ehdollistettu VaR estimoidaan sovittamalla mallin residuaaleihin ääriarvojakauma. Ääriarvoteoriaa on käytetty vasta hyvin vähän sähkömarkkinoiden riskien arviointiin. Seuraavaksi esitellään tutkimuksia, joissa ääriarvoteoriaa on sovellettu sekä sähkön hinnan että muiden taloudellisten muuttujien riskien mittaamiseen.

Longin (2000) kuvaa jaksomaksimimenetelmän käyttöä taloudellisten muuttujien VaR:ien laskemiseen. Hän sovittaa yleistetyn ääriarvojakauman Yhdysvaltojen osakemarkkinaindekseihin ja toteaa, että ääriarvojakauma sopii hyvin kuvaamaan indeksien paksuhäntäisiä jakaumia. Longin vertaa ääriarvomenetelmällä saatujen kvantiilien tarkkuutta normaalijakaumamenetelmällä sekä historiallisella simulaatiolla estimoitujen kvantiilien tarkkuuteen. Hänen mukaansa historiallinen simulaatio ja ääriarvomenetelmä antavat saman suuruisia kvantiileja, kun taas normaalijakauman kvantiilit ovat liian alhaiset, joten ne aliarvioivat riskin suuruutta. Bali (2003) tutkii Yhdysvaltojen joukkovelkakirjojen tuottoja sekä jaksomaksimi- että kynnsarvon ylitys -menetelmällä, jotka hänen mukaansa ovat luotettavampia ja tarkempia VaR:n estimointitapoja kuin normaalijakaumaan perustuva menetelmä.

McNeil ja Frey (2000) esittävät ensimmäisinä kaksivaiheisen ehdollistetun VaR:n estimointimenetelmän, jossa muuttujaa mallinnetaan ensin AR-GARCH-mallilla. Estimoidun mallin jäännösten jakauman hännän kuvaamiseen he käyttävät kynnsarvon ylitys -menetelmää. Menetelmän hyvyttä he testaavat useilla eri tuottosarjoilla ja havaitsevat, että ehdollistettu ääriarvomenetelmä (AR-GARCH-EVT) tuottaa parhaat kvantiiliestimaatit, mutta myös ehdollistettuun t-jakaumaa perustuva estimointi (AR-GARCH-t) toimii hyvin, jos tuottojen jakauma on symmetrinen. McNeil ja Frey vertaavat myös alkuperäisiin tuottosarjoihin sovellettua yleistettyä Pareto-

jakaumaa eli ehdollistamatonta ääriarvomenetelmää ehdollistettuihin menetelmiin ja toteavat, että aineiston suodattaminen AR-GARCH-mallilla parantaa kvantiilies-timaatteja.

Fernandez (2003) käyttää kaksivaiheista VaR-estimointia Yhdysvaltojen ja Chilen osakkeiden, valuuttakurssien ja kuparin hinnan tuottojen mallintamiseen. Hän es-timoi GARCH-mallit sekä havaintojaksolle että ennustamalla sen ulkopuolelle ja vertailee ehdollisia ja ehdollistamattomia kvantiileja. Hän toteaa ehdollisen ääriar-vomenetelmän olevan paljon parempi estimointitapa kuin ehdollinen normaalijakau-mamenetelmä, mutta myös empiirisiin kvantiileihin (historiallinen simulaatio) ja t-jakaumaan perustuvat ehdollisen menetelmät toimivat joissakin tapauksissa hyvin. T-jakaumamenetelmä pystyy kuvaamaan tarkasti erityisesti symmetrisiä jakaumia ja historiallinen simulaatio antaa tarkat VaR-estimaatit matalemmilla kvantiileilla.

Andrews ja Thomas (2002) soveltavat kynnsarvon ylitys -menetelmää sähkön hin-nan muutosten riskien arviointiin. Heidän mukaansa menetelmä sopii kuvaamaan sähkömarkkinoiden voittoja ja tappioita normaalijakaumaa paremmin. Samaa me-netelmää käyttää myös de Rozario (2002) Australian sähköpörssin sähkön hinnan muutoksien mallintamiseen. Hänen mukaansa yleistettyyn Pareto-jakaumaan perus-tuva ehdollistamaton VaR on yhtä hyvä estimointitapa kuin historiallinen simu-laatio, mutta normaalijakaumalla estimoidut VaR:t aliarvioivat riskiä. Kuitenkin de Rozario toteaa, että ääriarvoteorian IID-oletus ei ole voimassa, joten kausivaih-telun, varianssin vaihtelun sekä klusteroitumisen huomioonottava menetelmä voisi parantaa estimoinnin tarkkuutta.

McNeilin ja Freyn ehdottamaa kaksivaiheista ehdollista estimointimenetelmää käyt-tää Byström (2005) tutkiessaan Nord Poolin suuria tunnittaisia hinnan suhteelli-sia muutoksia. Hänen mukaansa aiemmin esitetty AR-GARCH-EVT-malli viipeil-lä 1, 24 ja 168 on tarkempi kvantiilien estimointimenetelmä kuin normaali- tai t-jakaumaan perustuva menetelmä. Byström testaa menetelmän hyvyttä sekä ha-vaintojaksolla että ennustamalla jakson ulkopuolella. Samaan johtopäätökseen pää-tyvät myös Chan ja Gray (2006), jotka käyttävät useiden sähköpörssien päivittäisten hinnan muutosten suodattamiseen eksponentiaalista GARCH-mallia (EGARCH) ja residuaalien kvantiilien estimointiin kynnsarvon ylitys -menetelmää. He vertaavat menetelmää myös historialliseen simulaatioon. Heidän mukaansa historialliseen si-

mulaatioon perustuvat estimaatit ovat myös hyviä, mutta VaR:n ylitykset tapahtuvat klustereina, koska menetelmä ei salli estimaattien vaihtelua ajassa, kuten AR-GARCH-mallit.

7.2 Havaintojakson mallinnus

Tässä kappaleessa estimoinnit tehdään havaintojaksolle 1.1.2000–31.12.2004 sovit- tamalla AR-GARCH-EVT-malli kaikkiin ajanjakson havaintoihin. Mallia ei käyte- tä tulevien havaintojen ennustamiseen, vaan kaikki havainnot ovat mukana para- metrien estimoinnissa. Havaintojakson estimointeja käytetään eri menetelmien hy- vyyden vertailuun. Kuitenkin riskien hallinnassa on kiinnostavampaa tietää, kuinka hyvin menetelmällä pystytään ennustamaan tulevaisuuden riskejä. Tätä tapaa käy- tetään menetelmien vertailuun seuraavassa kappaleessa, jossa mallit estimoidaan uudelleen joka päivä ja seuraavan päivän toteutuneita havaintoja verrataan ennus- tettuihin kvantiileihin.

Ehdolliset häntäkvantiilit estimoidaan kaksivaiheisesti AR-GARCH-EVT-menetel- mällä käyttämällä kynnsarvon ylitys -menetelmää. Vertailun vuoksi estimoidaan myös ehdollistamattomat hinnan muutoksen kvantiilit, sovittamalla yleistetty Pare- to-jakauma suoraan valitun kynnsarvon ylittäviin hinnan muutoksiin. Vartailussa on mukana myös AR-GARCH-n- ja AR-GARCH-t-menetelmät (ks. kappale 4.2).

Autokorrelaatio- ja osittaisautokorrelaatiofunktioiden (kuva 8) perusteella estimoin suurimman kvasi-uskottavuuden menetelmällä hinnan logaritimuutokselle r_t AR(4)- GARCH(1,1)-mallin, jossa on mukana myös kausiviipeet 24 ja 168. Tässä tutkiel- massa käytetty AR-GARCH-malli on siten muotoa

$$\begin{aligned} r_t &= c + \phi_1 r_{t-1} + \phi_2 r_{t-2} + \phi_3 r_{t-3} + \phi_4 r_{t-4} + \phi_5 r_{t-24} + \phi_6 r_{t-168} + u_t \\ u_t &= \sigma_t \epsilon_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 u_{t-1}^2 + \delta_1 \sigma_{t-1}^2, \end{aligned} \tag{32}$$

missä σ_t^2 on u_t :n ehdollinen varianssi ja $\epsilon_t \sim \text{IID}(0,1)$. Mallin avulla pyrin pois- tamaan hinnan logaritimuutoksen sarjasta suurimman osan aikariippuvuuksista, jolloin mallin jäännökset $\hat{\epsilon}_t$ ovat lähes IID-havaintoja, kuten kynnsarvon ylitys - menetelmä vaatii. Käyttämässäni mallissa on enemmän AR-viiveitä kuin Byströmin

(2005) käyttämässä mallissa.

Mallin (32) parametriestimaatit sekä normaali- että t-jakaumaoletuksella esitetään taulukossa 2. Normaalijakaumamallin parametrit on estimoitu suurimman kvasi-uskottavuuden menetelmällä rajoittamalla prosessi heikosti stationaariseksi, eli AR-osa on stationaarinen ja $\alpha_1 + \delta_1 < 1$. Samaa rajoitusta on käytetty estimoitaessa parametrit t-jakaumaoletuksella. Rajoitus on tehty ennustejaksolla tarvittavien ennusteiden parantamiseksi, koska ehdollisen varianssin ennusteet kasvavat nopeasti ennustehorisontin kasvaessa, jos $\alpha_1 + \delta_1 > 1$ (ks. kappale 7.3). Suoritin myös estimoinnit havaintojaksolle ilman stationaarisuusrajoitusta (ei esitetä tässä), jolloin ehdollisten VaR:n tarkkuus oli hieman parempi kuin rajoitetussa tapauksessa. Vertailtavuuden vuoksi rajoitusta käytetään kuitenkin myös havaintojakson estimoinneissa.

T-jakauman vapausaste ν on myös estimoitu aineistosta. Molemmissa tapauksissa GARCH-parametrien estimaattien summa on lähes yksi, joten hinnan logaritmi-muutoksen ehdollistamaton varianssi voi olla ääretön. Tämä nähdään myös hinnan muutosten ja jäännösten paksuhäntäisistä jakaumista (taulukot 1 ja 3, kuva 7). T-jakauman estimoitu vapausaste tuottaa myös paksuhäntäisen jakauman. Ilman stationaarisuusrajoitusta estimointi tuottaa äärettömän varianssin. Sekä normaali- että t-jakaumalla estimoiduissa AR-GARCH-malleissa useimmat AR-kertoimet ja kaikki GARCH-kertoimet ovat merkitseviä. Normaali- ja t-jakaumamallien tulokset ovat muuten samansuuntaiset, mutta α_1 - ja δ_1 -parametriestimaattien suuruudet poikkeavat toisistaan. Normaalijakaumaoletuksella estimoidun mallin ehdollisessa varianssissa esiintyy enemmän pysyvyyttä kuin t-jakaumamallissa, koska parametrin α_1 estimaatti on pieni ja δ_1 suuri. Normaalijakaumaoletuksella estimoidun mallin jäännöksiä tarkastellaan seuraavassa kappaleessa.

Normaalijakaumaoletuksella estimoituja parametreja käytetään myös ääriarvomenetelmän ehdollisen odotusarvon ja varianssin laskemiseen. Vertailussa käytettävässä AR-GARCH-n-menetelmässä kvantiilit estimoidaan olettamalla jäännökset $\epsilon_t \sim N(0,1)$ -jakautuneiksi ja AR-GARCH-t-menetelmässä $t(\nu)$ -jakautuneiksi odotusarvolla 0 ja varianssilla 1. Kun malli (32) on estimoitu, voidaan normaali- ja t-jakauman tilanteessa ehdolliset häntäkvantiilit $\hat{x}_{t,q}$ estimoida kertomalla ennustetut ehdolliset hajonnat $\hat{\sigma}_t$ oletetun jakauman kvantiililla z_q ja lisäämällä ehdollinen odo-

Taulukko 2: AR-GARCH-mallin parametriestimaatit

Parametri	Normaalijakauma		t-jakauma	
	Kerroin	Keskivirhe	Kerroin	Keskivirhe
c	$-1,48 \cdot 10^{-4}$	$2,67 \cdot 10^{-4}$	$-5,14 \cdot 10^{-4}$	$8,58 \cdot 10^{-5}$
ϕ_1	0,2155	0,012	0,2373	0,0067
ϕ_2	-0,0035	0,015	-0,0173	0,0041
ϕ_3	-0,0079	0,011	$-1,46 \cdot 10^{-4}$	0,0030
ϕ_4	-0,0814	0,0083	-0,0288	0,0023
ϕ_5	0,3878	0,019	0,3414	0,0134
ϕ_6	0,1148	0,014	0,1399	0,0093
α_0	$4,41 \cdot 10^{-5}$	$1,26 \cdot 10^{-5}$	$2,11 \cdot 10^{-4}$	$7,66 \cdot 10^{-6}$
α_1	0,1295	0,018	0,6412	0,0016
δ_1	0,8700	0,018	0,3588	0,0016
ν			2,924	

tusarvo (kaava 30). Esimerkiksi normaalijakauman 95 %:n kvantiili $z_q = 1,645$ ja ehdollinen VaR (AR-GARCH-n) on

$$VaR_{t,0.95}^N = \hat{\mu}_t + 1,645\hat{\sigma}_t, \quad (33)$$

missä $\hat{\mu}_t = \hat{c} + \sum_i^4 \hat{\phi}_i r_{t-i} + \hat{\phi}_5 r_{t-24} + \hat{\phi}_6 r_{t-168}$ ja $\hat{\sigma}_t$ ovat estimoidun AR-GARCH-mallin ehdollinen odotusarvo ja hajonta. Siten jäännösten kvantiiliestimaatit skaalataan hinnan muutoksen odotusarvolla ja hajonnalla, jolloin saadaan hinnan muutoksen ehdolliset kvantiiliestimaatit. T-jakauman kvantiilit voidaan estimoida samalla tavalla skaalaamalla enisn jakauman varianssi ykköseksi, eli t-jakauman kvantiili saadaan kaavalla

$$z_q = \sqrt{\frac{\nu - 2}{\nu}} F_\nu^{-1}(q),$$

missä F_ν on t-jakauman kertymäfunktio vapausasteella $\nu > 2$.

Ehdollistamattomalla ääriarvomenetelmällä (lyhennetään eEVT) hinnan muutoksen VaR saadaan sovittamalla yleistetty Pareto-jakauma suoraan hinnan logaritmi-muutoksiin ja käyttämällä kaavaa (15). Esimerkiksi 95 % ehdollistamaton VaR on siten

$$VaR_{0.95}^{eEVT} = \hat{x}_{0.95} = u + \frac{\hat{\sigma}}{\hat{\xi}} \left[\left(\frac{n}{N_u} (0.05) \right)^{-\hat{\xi}} - 1 \right].$$

Ehdollistamaton kvantiili riippuu vain jakauman oikeanpuoleisesta hännästä, eikä se vaihtelee ajassa kuten ehdollistetut kvantiilit. Jos yleistetty Pareto-jakauma sopii hyvin kuvaamaan hinnan muutosten suuria arvoja, ehdollistamaton menetelmä voi antaa tarkkoja tuloksia, mutta VaR:n ylitykset voivat tapahtua klustereina, koska estimaatti ei reagoi volatilitiiteetin kasvuun.

Ehdollistettujen AR-GARCH-EVT-kvantiilien saamiseksi täytyy AR-GARCH-mallin (32) jäännösten $\hat{\epsilon}_t$ häntäkvantiilit estimoida sovittamalla niihin yleistetty Pareto-jakauma kynnyksisarvon ylitys-menetelmällä ja laskemalla jäännösten jakauman avulla hinnan muutoksen ehdolliset häntäkvantiilit samaan tapaan kuin normaali- ja t-jakauman tapauksessa. Tällöin ehdollinen hajonta kerrotaan jäännösten estimoidulla häntäkvantiililla (15) ja lisätään ehdollinen odotusarvo $\hat{\mu}_t$ (kaava 31). Esimerkiksi hinnan muutoksen 95 %:n ehdollinen VaR on

$$VaR_{t,0.95}^{EVT} = \hat{\mu}_t + VaR_{0.95}^{\hat{\epsilon}} \hat{\sigma}_t,$$

missä $VaR_{0.95}^{\hat{\epsilon}}$ on jäännösten häntäkvantiili, joka on saatu soveltamalla kynnyksisarvon ylitys -menetelmää AR-GARCH-mallin jäännöksiin (kaava 15).

Aluksi täytyy varmistaa, että AR-GARCH-mallin (32) jäännökset $\hat{\epsilon}_t$ sopivat ääriarvomallinnukseen, eli jäännökset eivät ole kovin autokorreloituneita ja heteroskedastisia. Tämän jälkeen sovitaan yleistetyn Pareto-jakauman valitsemani kynnyksisarvon ylittäviin jäännöksiin. Kynnyksisarvon valinta ja yleistetyn Pareto-jakauman parametrien estimointi kuvataan seuraavassa kappaleessa.

7.2.1 Jäännösten mallinnus

Kynnyksisarvon ylitys -menetelmää sovelletaan mallin (32) estimoituihin standardoituihin jäännöksiin

$$\hat{\epsilon}_t = \frac{r_t - \hat{\mu}_t}{\hat{\sigma}_t},$$

missä mallin estimoitu ehdollinen odotusarvo on

$$\hat{\mu}_t = \hat{c} + \sum_i^4 \hat{\phi}_i r_{t-i} + \hat{\phi}_5 r_{t-24} + \hat{\phi}_6 r_{t-168}$$

ja hajonta

$$\hat{\sigma}_t = \sqrt{\hat{\alpha}_0 + \hat{\alpha}_1 \hat{u}_{t-1}^2 + \hat{\delta}_1 \hat{\sigma}_{t-1}^2}.$$

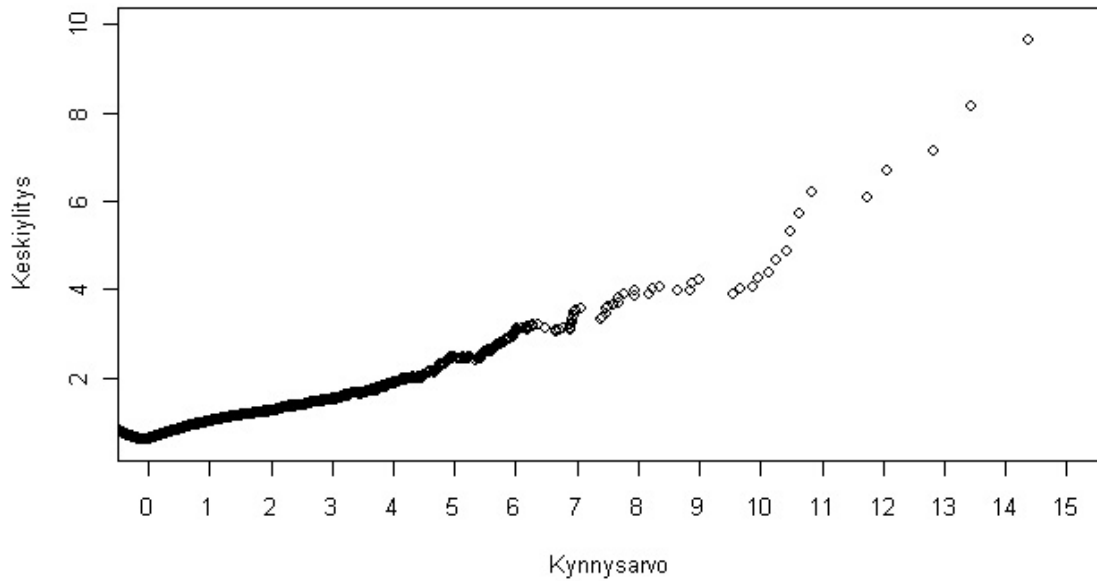
Normaalijakaumaan perustuvalla suurimman kvasi-uskottavuuden menetelmällä esitimoitujen jäännösten tunnusluvut on esitetty taulukossa 3. Tunnusluvuista nähdään, että jäännösten jakauma on paksuhäntäinen, oikealle vino ja ei-normaalinen. Ljung-Box-testisuureet hylkäävät nollahypoteesin autokorreloimattomista jäännöksistä, mutta kausivaihtelu on alkuperäistä aikasarjaa selvästi heikompaa (vrt. kuva 8 ja liite 1). Myös GARCH-vaikutus, eli neliöityjen jäännösten autokorrelaatio (Q^2), on pienempi kuin alkuperäisessä aikasarjassa (ks. liite 1). AR-GARCH-malli on siten pienentänyt aikariippuvuuksia ja varianssin vaihtelua verrattuna alkuperäisen hinnan muutoksen aikariippuvuuksiin. Koska jäännöksessä esiintyy vähemmän aikariippuvuuksia kuin hinnan muutoksessa ja ääriarvoteoria sallii myös autoregressiivisen prosessin, voidaan yleistetty Pareto-jakauma sovittaa jäännösten ylityksiin.

Taulukko 3: Standardoitujen jäännösten tunnusluvut

Tunnusluku	Kerroin	p-arvo
Keskiarvo	-0,0055	
Hajonta	1,099	
Mediaani	-0,0024	
Minimi	-21,993	
Maksimi	34,434	
Vinous	0,1291	
Huipukkuus	53,548	
Jarque-Bera	$5,22 \cdot 10^6$	0,000
Q(6)	231,1	0,000
Q(24)	805,2	0,000
$Q^2(6)$	84,2	0,000
$Q^2(24)$	183,4	0,000

Ensin täytyy valita sopiva kynnsarvo u , jonka yli meneviin havaintoihin yleistetty Pareto-jakauma sovitetaan. Kynnsarvon valintaan käytän keskiylitysfunktiota ja muotoparametrin estimaatteja eri kynnsarvoilla sekä valinnan hyvyyden kuvaamiseen kvantiilikuviota. Otoskeskiylitysfunktion kuvaaja on lähes suora kynnsarvosta nolla lähtien, mutta pieni kulmakertoimen muutos tapahtuu kynnsarvon 1,5 jälkeen (kuva 9). Kun kynnsarvo kasvaa, ylityksien lukumäärä vähenee ja keskiylitysfunktion ja siten jakauman parametrien epävarmuus kasvaa. Koska otoskeskiylitysfunktion ja siten jakauman parametrien epävarmuus kasvaa. Koska otoskeskiylitysfunktion ja siten jakauman parametrien epävarmuus kasvaa.

Kuva 9: Otoskeskiylitysfunktio

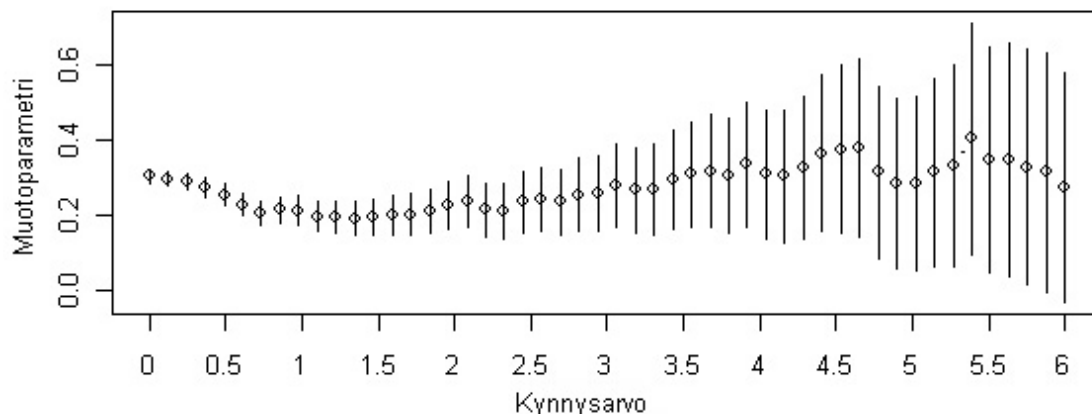


litysfunktion kulmakerroin on positiivinen, jäännösten jakauma on paksuhäntäinen ($0 < \xi < 1$). Kynnysarvon määrittämiseksi estimoin SU-menetelmällä muotoparametrin ξ arvon ja 95 % luottamusvälin useilla kynnysarvon arvoilla. Kuviossa 10 nähdään, että muotoparametrin arvo vakiintuu kynnysarvon 1,5 jälkeen, mutta tämän jälkeen estimaatin luottamusväli kasvaa.

Vertailemalla eri kynnysarvoilla estimoitujen jakaumien kvantiilikuvioita valitsin kynnysarvoksi 1,7. Tällöin muotoparametrin arvo on lähellä kuviossa 10 nähtävää vakiotasoa, mutta estimaatin hajonta on pieni. SU-menetelmällä estimoidut yleistetyn Pareto-jakauman parametriestimaatit on esitetty taulukossa 4. Estimointiin käytettyjä eli kynnysarvon ylittäviä havaintoja on aineistossa 1726 eli 3,95 % kaikista havainnoista. Kvantiilikuviosta (kuva 11) nähdään, että estimoitu jakauma sopii melko hyvin kuvaamaan jäännösten oikeanpuoleista häntää, koska vain kaksi suurinta havaintoa ei sijaitse suoralla.

Suoritin samanlaisen analyysin myös sähkön hinnan logaritmuutoksen kynnysarvon valitsemiseksi ja sovitin yleistetyn Pareto-jakauman suoraan kynnysarvoa suurempiin hinnan logaritmuutoksiin estimoidakseni ehdollistamattomat häntäkvanttiilit. Estimoitujen parametrien arvot on esitetty taulukossa 4. Logaritmuutoksiin sovitettu yleistetty Pareto-jakauma on hieman jäännöksen jakaumaa ohuthäntä-

Kuva 10: Muotoparametrin estimaatit ja 95 % luottamusvälit eri kynnyksarvoilla



Taulukko 4: Jäännösiin ja hinnan logaritminmuutoksiin sovitetun yleistetyn Pareto-jakauman parametrit

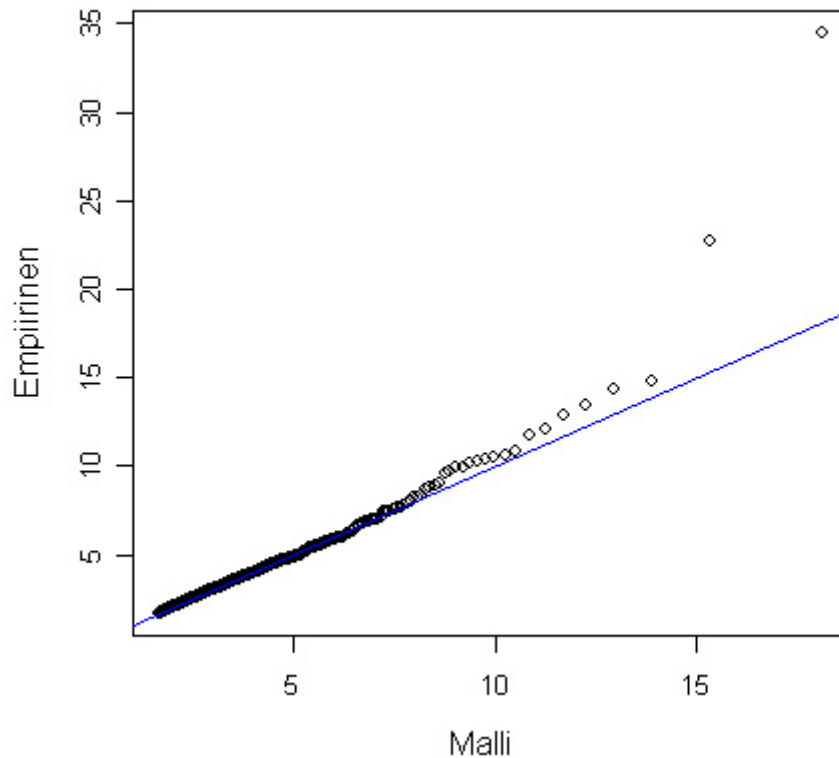
Parametri	Jäännökset		Hinnan logaritminmuutos	
	Kerroin	Keskihajonta	Kerroin	Keskihajonta
u	1,7		0,28	
ξ	0,203	0,0276	0,163	0,0461
σ	0,942	0,0342	0,183	0,0108

sempi, koska muotoparametri on pienempi. Seuraavassa kappaleessa vertaan ääriarvomenetelmällä laskettuja ehdollistettuja häntäkvantileja ehdollistamattomiin sekä normaali- ja t-jakaumalla estimoituihin ehdollisiin kvantileihin. Hyvyyden mittarina käytän toteutuneita kvantiilestimaattien ylityksiä eri riskitasoilla.

7.2.2 Havaintojakson tulosten vertailu ja testaus

Estimoin havaintojakson jokaiselle tunnille VaR:t edellisissä kappaleissa esitetyillä neljällä menetelmällä käyttäen 95, 99, 99,9 ja 99,95 prosentin kvantiilipisteitä. VaR:n ylitys tapahtuu, kun kyseisen tunnin hinnan logaritminmuutos on suurempi kuin estimoitu VaR. Mikäli VaR:n estimointimenetelmä antaa tarkkoja estimaatteja, havaittujen ylitysten lukumäärä on lähellä odotettua ylitysten määrää eli ehdollistamaton peittävyys on oikea. Esimerkiksi 95 % häntäkvantiilin ylittävien havaintojen

Kuva 11: Sovitetun jakauman kvantiilikuvio



odotettu lukumäärä on $0,05 \cdot 43679 = 2184$. Lisäksi menetelmän tulisi reagoida volatiliiteetin vaihteluihin, jottei ylitykset tapahtuisi klustereina. Siten ylityksien tulisi olla riippumattomia. Hyvän menetelmän tulosten tulee olla sekä riippumattomia että antaa oikea ylitysten osuus, jolloin ehdollinen peittävyys on hyvä. Menetelmien hyvyyttä testataan kappaleessa 4.4 esitetyillä testeillä.

Odotettujen ja havaittujen ylitysten lukumäärät sekä ehdollistamattoman peittävyyden binomitestisuuret (kaava (24)) ja p-arvot on esitetty taulukossa 5. Taulukosta huomataan, että normaalijakaumaan perustuva ehdollinen menetelmä (AR-GARCH-n) yliarvioi VaR:a 95 % luottamustasolla, mutta ohuthäntäisenä jakaumana aliarvioi suurempia kvantiileja eli ylityksiä tapahtuu paljon enemmän kuin on odotettu. Tämä on hyvin odotettua, koska jäännösten jakauma on paksuhäntäisempi kuin oletettu normaalijakauma. Estimoitu t-jakauma (AR-GARCH-t) on myöskin ohuthäntäinen, koska menetelmä aliarvioi kaikki kvantiilit, jolloin ylityksiä tapahtuu liikaa. Menetelmien tarkkuutta mittaava testi hylkää nollahypoteesin oikeasta ylitysten määrästä kaikilla häntäkvantiileilla sekä normaali- että t-jakauman

Taulukko 5: Menetelmien arviointi ja testaus havaintojaksolla. Ylityksien lukumäärät ja ehdollistamaton peittävyys

	Odotettu	AR-GARCH-n	AR-GARCH-t	eEVT	AR-GARCH-EVT
95 %					
Ylitykset	2184	1814	3403	2980	2165
BT		-8,12	26,76	17,48	-0,416
p-arvo		0,000	0,000	0,000	0,339
99 %					
Ylitykset	437	917	1104	426	430
BT		23,09	32,09	-0,519	-0,327
p-arvo		0,000	0,000	0,302	0,372
99,9 %					
Ylitykset	47	475	195	47	45
BT		65,30	22,91	0,503	0,200
p-arvo		0,000	0,000	0,307	0,421
99,95 %					
Ylitykset	22	397	106	23	23
BT		80,30	18,01	0,248	0,248
p-arvo		0,000	0,000	0,402	0,402

Taulukossa AR-GARCH-n tarkoittaa kvantiilien ehdollista estimointimenetelmää normaalijakaumaoletuksella ja AR-GARCH-t t-jakaumaoletuksella. Ääriarvoteoriaan perustuvista menetelmistä eEVT on ehdollistamaton ja AR-GARCH-EVT ehdollistettu estimointimenetelmä.

tapauksessa. Näitä menetelmiä tarkemmin toimii ehdollistamaton ääriarvomenetelmä (eEVT). Nollahypoteesi oikeasta ylitysten määrästä hylätään vain 95 prosentin kvantiilin tapauksessa. Suuriin hinnan logaritminuutoksiin sovitettu yleistetty Pareto-jakauma kuvaa hyvin jakauman oikeaa häntää. Ehdollistamaton menetelmä ei kuitenkaan reagoi hinnan logaritminuutoksen volatiliteetin kasvuun, joten ylitykset tapahtuvat klustereina. Tarkimmat tulokset antaa ehdollinen ääriarvomenetelmä (AR-GARCH-EVT), jossa AR-GARCH-parametrit estimoitii normaalijakaumaoletuksella ja mallin jäännöksiin sovitettiin yleistetty Pareto-jakauma. Ylitysten lukumäärä ei poikkea merkittävästi odotetusta millään kvantiililla. Menetelmällä

pystytään estimoimaan hinnan logaritmuutoksen jakauman häntää hyvin tarkasti kaikilla suurillakin kvantiileilla, koska havaittujen ylitysten lukumäärä on hyvin lähellä odotettua määrää.

Vaikka menetelmä toimisi hyvin ehdollistamattoman peittävyden osalta, erityisesti tulevaisuuden riskejä ennustettaessa on tärkeää, että menetelmä pystyy reagoimaan markkinatilanteeseen. Tällöin menetelmän ennustamien ylitysten täytyy olla myös riippumattomia ja siten ehdollinen peittävyys tulee olla oikea. Testasin tätä kappaleessa 4.4 esitetyillä uskottavuusosamäärän testeillä. Ehdollistamattoman peittävyyden (LR_{uc}), riippumattomuuden (LR_{ind}) ja ehdollisen peittävyyden (LR_{cc}) testisuureet on esitetty taulukossa 6. Testisuureista nähdään, että millään menetelmällä ehdollisen peittävyyden nollassa oletus ei jää voimaan 5 % merkitsevyystasolla. Suurin syy tähän on se, että riippumattomuusoletus ei toteudu. Aineistokoon ollessa yli 43000 ovat pienetkin poikkeamat hyvin merkitseviä. Menetelmien testisuureista löytyy kuitenkin eroja, joiden avulla voidaan menetelmien paremmuutta vertailla.

Talukosta 6 nähdään, että kuten binomitestin tapauksessa (taulukko 5) parhaimman ehdollistamattoman peittävyyden tuottaa ehdollistettu ääriarvomenetelmä (AR-GARCH-EVT) jokaisella kvantiilipisteellä. Riippumattomimmat ylitykset antaa kuitenkin t-jakaumaan perustuva ehdollinen menetelmä (AR-GARCH-t), koska testisuureen LR_{ind} arvot on kaikissa tapauksissa pienimmät. Mutta koska t-jakaumamenetelmän ehdollistamaton peittävyys on hyvin kaukana odotetusta ylitysten määrästä, ehdollisessa peittävydessä AR-GARCH-EVT-menetelmä on kuitenkin tarkin. Vain 99,9 %:n kvantiilin estimoinnissa ehdollistamaton ääriarvomenetelmä (eEVT) on aavistuksen verran tarkempi. Ehdollistamaton ääriarvomenetelmä on ehdollisessa peittävydessä sitä parempi, mitä suurempaa kvantiilia estimoidaan. Normaali-jakaumaan perustuva ehdollinen menetelmä (AR-GARCH-N) toimii erityisen huonosti suurilla kvantiileilla, koska normaalijakuma on liian ohuthäntäinen kuvaamaan AR-GARCH-mallin jäännöksiä.

Havaintojaksolle saadut tulokset ovat samansuuntaisia McNeilin ja Freyn (2000), Fernandezin (2003), Byströmin (2005) ja Chanin ja Grayn (2006) tulosten kanssa. Ehdollinen ääriarvomenetelmä sopii käytetyistä menetelmistä parhaiten kuvaamaan sähkömarkkinoiden suurten hinnan nousujen aiheuttamaa riskiä. Vaikka ehdollistamaton ääriarvomenetelmä antaa melko hyvät kvantiiliestimaatit, havaitut ylitykset

Taulukko 6: Riippumattomuuden ja ehdollisen peittävyden testaus.

	AR-GARCH-n	AR-GARCH-t	eEVT	AR-GARCH-EVT
95 %				
LR_{uc}	69,8***	616,7***	275,6***	0,174
LR_{ind}	63,9***	25,0***	2892***	78,2***
LR_{cc}	134***	642***	3167***	78,5***
99 %				
LR_{uc}	405,1***	723,3***	0,271	0,107
LR_{ind}	49,2***	3,5*	286***	61,0***
LR_{cc}	454***	727***	286***	61,1***
99,9 %				
LR_{uc}	1409***	281,4***	0,247	0,040
LR_{ind}	54,9***	1,75	10,98***	11,3***
LR_{cc}	1464***	283***	11,2***	11,4***
99,95 %				
LR_{uc}	1556***	166,7***	0,061	0,061
LR_{ind}	56,4***	0,429	6,94***	6,94***
LR_{cc}	1612***	167***	7,0**	7,0**

Taulukossa esitetään testisuureiden arvot. Luvun perässä * tarkoittaa merkitsevyyttä 10 %, ** 5 % ja *** 1 %:n luottamustasolla. LR_{uc} testaa ehdollistamatonta peittävyttä, LR_{ind} riippumattomuutta ja LR_{cc} ehdollista peittävyttä.

tapahtuvat usein ryppäinä, kun markkinoiden volatilitteetti kasvaa. Ehdollisten menetelmien antamat kvantiiliestimaatit sitä vastoin reagoivat markkinoiden heilahteluihin ja ylitykset tapahtuvat tasaisemmin väliajoin, varsinkin matalammilla kvanttileilla. Liitteen 2 kuvioissa 13 ja 14 ja on esitetty esimerkkinä vuoden 2004 marras- ja joulukuun tunnittaiset hinnan logaritimuutokset, ehdollistamattomalla ja ehdollisilla menetelmillä estimoidut 99 % VaR:t sekä niiden ylitykset. Kuvioista nähdään, että ehdollistettujen menetelmien antamat VaR:n ylitykset tapahtuvat ajanjaksolla tasaisemmin kuin ehdollistamattoman menetelmän ylitykset. Seuraavassa kappaleessa vertaillaan eri menetelmien riskien ennustamiskykyä ja -tarkkuutta.

7.3 Ennustemenetelmä

Havaintojakson estimoinneissa sekä menneet että tulevat hinnan muutokset ovat tiedossa. Riskeihin varautumisessa on kuitenkin tärkeää pystyä ennustamaan tulevaisuudessa mahdollisesti tapahtuvat suuret hinnan nousut tai laskut. Tässä kappaleessa kuvataan käyttämäni ennustemenetelmä, jolla häntäkvantiilit on estimoitu ennustejaksolle. Seuraavassa kappaleessa eri menetelmillä saatuja tuloksia vertaillaan samaan tapaan kuin havaintojaksolla.

Ennustejakson aikaväli on 2.1.2005–9.11.2007 eli yhteensä 1042 vuorokautta. AR-GARCH-mallin (32) estimointiin käytetään jokaisessa vaiheessa kahden edellisen vuoden tuntihintoja, joten havaintoja ja niistä laskettuja standardoituja jäännöksiä on yhteensä 17544 kappaletta. AR-GARCH-malli estimoidaan kerran päivässä kello 12 ja sovitetulla mallilla ennustetaan seuraavan 24 tunnin (kello 13–12) ehdolliset odotusarvot ja hajonnat, joita tarvitaan VaR-ennusteiden laskemiseen (kaava (31)). Ehdollisen odotusarvon optimaalinen ennuste on (Franses ja van Dijk, 2000, 188)

$$\hat{\mu}_{t+j|t} = \hat{c} + \sum_i^4 \hat{\phi}_i r_{t+j-i} + \hat{\phi}_5 r_{t+j-24} + \hat{\phi}_6 r_{t+j-168}, \quad j = 1, \dots, 24.$$

Ehdollisen varianssin optimaalinen ennuste ajanhetkeen $t + j$ voidaan laskea rekursiivisesti kaavasta

$$\hat{\sigma}_{t+j|t}^2 = \hat{\alpha}_0 + \hat{\alpha}_1 \hat{u}_{t+j-1|t}^2 + \hat{\delta}_1 \hat{\sigma}_{t+j-1|t}^2,$$

missä määritelmän mukaan $\hat{u}_{t+s|t}^2 = \hat{\sigma}_{t+s|t}^2$, kun $s > 0$ ja $\hat{u}_{t+s|t}^2 = u_{t+s}^2$ sekä $\hat{\sigma}_{t+s|t}^2 = \sigma_{t+s}^2$, kun $s \leq 0$. Rekursiokaavaa käyttämällä saadaan ehdollinen varianssi muotoon

$$\hat{\sigma}_{t+j|t}^2 = \hat{\alpha}_0 \sum_{i=0}^{j-1} (\hat{\alpha}_1 + \hat{\delta}_1)^i + (\hat{\alpha}_1 + \hat{\delta}_1)^{j-1} \sigma_{t+1}^2,$$

jolla voidaan laskea ennusteet suoraan σ_{t+1}^2 :n avulla. σ_{t+1}^2 kuuluu informaatiojoukkoon Ω_t , koska se voidaan laskea havainnoista y_t, y_{t-1}, \dots , kun parametrit tunnetaan. (Franses ja van Dijk, 2000, 190–191). Tästä huomataan, että jos prosessi ei ole heikosti stationaarinen, eli $\alpha_1 + \delta_1 > 1$, ehdollisen varianssin ennusteet kasvavat eksponentiaalisesti ennustehorisontin j kasvaessa ja kun $\alpha_0 > 0$. VaR-ennusteet lasketaan kertomalla jäännösten kvantiili ennustetulla ehdollisella hajonnalla $\hat{\sigma}_{t+j}$ ja lisäämällä ehdollinen odotusarvo $\hat{\mu}_{t+j}$. Tämän vuoksi myös VaR-ennusteet kasvavat ennustehorisontin kasvaessa, sillä hajonnan kasvu alkaa dominoida VaR:a suurilla j .

Taulukko 7: AR-GARCH-mallin GARCH-parametrien estimaattien vaihteluvälit ennustejaksolla

Parametri	Normaalijakauma	t-jakauma
	Vaihteluväli	Vaihteluväli
α_0	$1,05 \cdot 10^{-8} - 5,38 \cdot 10^{-4}$	$1,05 \cdot 10^{-8} - 2,89 \cdot 10^{-4}$
α_1	0,048 – 0,616	0,315 – 0,807
δ_1	0,384 – 0,952	0,193 – 0,685
ν		2,70 – 3,21

VaR-ennusteiden kasvaessa niiden ylittävien havaintojen esiintyminen vähenee. Sen vuoksi tässä tutkielmassa GARCH-prosessi on rajoitettu sekä heikosti että vahvasti stationaariseksi.

Ehdollisen normaalijakauma- (AR-GARCH-N) ja ääriarvomenetelmän (AR-GARCH-EVT) AR- ja GARCH-parametrit estimoidaan olettamalla jäännös normaalijakautuneeksi, kun taas ehdollisessa t-jakaumamenetelmässä oletetaan t-jakautuneet jäännökset. Estimointiin käytettävää datakehikkoa siirretään jokaisen estimoinnin jälkeen 24 tuntia eteenpäin ja estimoidaan malli uudelleen. AR-GARCH-malli on estimoitu rajoittamalla se stationaariseksi, eli AR-osan tulee täyttää stationaarisuusehto (25) ja GARCH-parametrien estimaattien summan tulee toteuttaa ehto $\alpha_1 + \delta_1 < 1$. Ennustejaksolla tehtyjen 1042 estimoinnin GARCH-parametrien suurimmat ja pienimmät arvot on esitetty taulukossa 7. Estimaateissa on suurta vaihtelua, joten vain kahden vuoden havaintojen käytöllä saadaan malli kuvaamaan hinnan vaihteluita paremmin kuin sovittamalla malli pidemmän ajanjakson havaintoihin. Hinnan logaritmuutoksen käyttäytyminen on siis ollut erilaista ennustejakson eri vaiheissa.

Ehdollisessa ääriarvomenetelmässä estimoinneista saatuihin standardoituihin jäännöksiin sovitetaan yleistetty Pareto-jakauma käyttämällä kynnsarvona jäännösten empiiristä 95 prosentin kvantiilia. Tämä kynnsarvo on valittu estimoinnin yksinkertaistamiseksi. Kynnsarvo voi siten vaihdella eri estimoinneissa, mutta kynnsarvon ylittäviä havaintoja on jokaisessa 877 kappaletta. Kun jakauman parametrit on estimoitu SU-menetelmällä, jäännösten häntäkvantiilit (eli VaR^e) voidaan laskea kaavalla (15). Näiden avulla hinnan logaritmuutoksen q :s häntäkvantiili jaksolle

$t + 1, \dots, t + 24$ saadaan skaalamalla jäännösten häntäkvantiilit ennustetulla ehdollisilla odotusarvolla ja hajonnalla samaan tapaan kuin aiemmin (ks. kappale 7.2), eli

$$VaR_{t+j,q}^{EVT} = \hat{\mu}_{t+j} + VaR^{\hat{\sigma}}_{t+j}, \quad j = 1, 2, \dots, 24.$$

AR-GARCH-n- ja AR-GARCH-t-kvantiiliestimaatit saadaan kertomalla oletetun jakauman teoreettinen kvantiili ehdollisen hajonnan ennusteella ja lisäämällä ennustettu ehdollinen odotusarvo. Ehdollistamattomat ääriarvokvantiilit estimoidaan käyttämällä kahden edellisen vuoden hinnan logaritimuutoksiin kynnysarvon ylitysmenetelmää ja kynnysarvona 95 %:n empiiristä kvantiilia. Siten yleistetty Pareto-jakauma sovitetaan 877 suurimpaan hinnan nousuun. Eri menetelmillä ennustettuja häntäkvantiileja verrataan jakson toteutuneisiin hinnan muutoksiin ja ylitys tapahtuu, jos $r_t > VaR_t$. Seuraavaksi vertailen menetelmien ylitysten lukumääriä ja niiden tarkkuutta samalla tavalla kuin havaintojakson vertailussa (kappale 7.2.2).

7.3.1 Ennustejakson tulosten vertailu ja testaus

Ennustejaksolla estimoitujen neljän eri menetelmän ylitysten määrät ja binomitesisuureet löytyvät talukosta 8 neljälle eri VaR-tasolle. Testisuureella testataan ehdollistamatonta peittävyyttä, eli onko menetelmän avulla estimoidun VaR:n ylityksiä odotettu määrä havainnoista. Yksisuuntaisessa testauksessa vain muutamassa tapauksessa menetelmän antama ylityksien määrä ei poikkea odotetusta määrästä merkitsevästi. Normaalijakaumaan perustuva AR-GARCH-menetelmä aliarvioi VaR:n suuruutta suurilla kvantiileilla, joten ylityksiä tapahtuu liian paljon. Normaalijakauma ei kuvaa hyvin hinnan logaritimuutosten paksua oikeaa häntää. Paksumpihäntäinen t-jakauma toimii normaalijakaumaa paremmin, mutta aliarvioi silti riskin suuruutta suurimmilla kvantiileilla.

Ehdollistamattoman ääriarvomenetelmän (eEVT) ylitysten lukumäärä koko ennustejaksolla on lähimpänä odotettuja määriä. AR-GARCH-mallin ja ääriarvomenetelmän yhdistelmä pystyy kuvaamaan hinnan muutoksen oikeaa häntää melko hyvin, mutta suurimmilla kvantiileilla menetelmä hieman aliarvioi riskiä ja ylitysten määrä on odotettua korkeampi. Koska ennustemenetelmässä jäännöksiin sovitettavan yleisen Pareto-jakauman valinnassa käytettiin yksinkertaistettua kynnysarvoa, estimoitu jakauma ei välttämättä kuvaa parhaalla mahdollisella tavalla jakauman oi-

Taulukko 8: Menetelmien arviointi ja testaus ennustejaksolla. Ylityksien lukumäärät ja ehdollistamaton peittävyys

	Odotettu	AR-GARCH-n	AR-GARCH-t	eEVT	AR-GARCH-EVT
95 %					
Ylitykset	1249	613	1199	1383	1010
BT		-18,5	-1,46	3,88	-6,94
p-arvo		0,000	0,072	0,000	0,000
99 %					
Ylitykset	250	344	474	267	258
BT		5,99	14,3	1,09	0,519
p-arvo		0,000	0,000	0,138	0,302
99,9 %					
Ylitykset	25	212	105	9	44
BT		37,4	16,0	-3,20	3,81
p-arvo		0,000	0,000	0,001	0,000
99,95 %					
Ylitykset	12	193	63	4	27
BT		51,1	14,3	-2,40	4,11
p-arvo		0,000	0,000	0,008	0,000

Taulukossa AR-GARCH-n tarkoittaa kvantiilien ehdollista estimointimenetelmää normaalijakaumaoletuksella ja AR-GARCH-t t-jakaumaoletuksella. Ääriarvoteoriiaan perustuvista menetelmistä eEVT on ehdollistamaton ja AR-GARCH-EVT ehdollistettu estimointimenetelmä.

keaa häntää. Havaintojaksolla käytetyllä kynnsarvon valitsemisella olisi jakauman sopivuutta voitu parantaa, mutta sen toteuttaminen yli tuhannessa estimoinnissa olisi ollut hankalaa. Ääriarvomenetelmät pystyvät kuitenkin ennustamaan perinteisiä normaali- ja t-jakaumamenetelmiä tarkemmin tulevaisuuden suurten hinnan nousujen todennäköisyyksiä ja riskejä.

Riippuvuuden ja ehdollisen peittävyuden testauksen tulokset on esitetty taulukossa 9. Tulokset ovat hyvin sopusoinnussa havaintojaksolle saatujen tulosten kanssa (taulukko 6). Mikään menetelmä ei pysty ennustamaan häntäkvantileja niin, että hypo-

Taulukko 9: Riippumattomuuden ja ehdollisen peittävyuden testaus ennustejaksolla.

	AR-GARCH-n	AR-GARCH-t	eEVT	AR-GARCH-EVT
95 %				
LR_{uc}	416,5***	2,2	14,6***	51,4***
LR_{ind}	144,8***	218,9***	1181***	210,6***
LR_{cc}	561,3***	221,1***	1196***	262,1***
99 %				
LR_{uc}	32,1***	160,8***	1,16	0,266
LR_{ind}	94,0***	104,8***	113,9***	71,3***
LR_{cc}	126,1***	265,6***	115,1***	71,6***
99,9 %				
LR_{uc}	534,0***	141,7***	13,6***	11,8***
LR_{ind}	63,2***	20,8***	0,007	24,4***
LR_{cc}	597,2***	162,5***	13,6***	36,2***
99,95 %				
LR_{uc}	697,0***	103,0***	7,88***	12,6***
LR_{ind}	54,3***	18,6***	0,001	5,2**
LR_{cc}	751,3***	121,6***	7,9**	17,8***

Taulukossa esitetään testisuureiden arvot. Luvun perässä * tarkoittaa merkitsevyyttä 10 %, ** 5 % ja *** 1 %:n luottamustasolla. LR_{uc} testaa ehdollistamatonta peittävyttä, LR_{ind} riippumattomuutta ja LR_{cc} ehdollista peittävyttä.

teesit riippumattomuudesta ja oikeasta ehdollisesta peittävydestä jäisivät voimaan kaikilla kvantiilitasoilla. Yllättäen ehdollistamaton ääriarvomenetelmä (eEVT) antaa suurimmilla kvantiileilla riippumattomat ylitykset, vaikkei menetelmä pysty reagoimaan äkillisiin hajonnan vaihteluihin. Riippumattomuus johtuu tässä tapauksessa siitä, että ehdollistamaton VaR on hyvin korkea koko ennustejaksolle, joten testisuureessa tarvittavia peräkkäisiä ylityksiä ei ole lainkaan, mikä vaikuttaa riippumattomuuden testaamiseen. Sähkön hinta ei kahtena peräkkäisenä tuntina nouse niin suuresti, että ne molemmat ylittäisivät hyvin suuren 99,9 % ehdollistamattoman VaR:n.

Eri menetelmien keskinäinen paremmuus voidaan määrätä vertailemalla testisuureiden arvoja. Parhaimman ehdollisen peittävyuden antaa ehdollinen ääriarvome-

netelmä (AR-GARCH-EVT), kun otetaan huomioon kaikki estimoidut kvantiilit. Menetelmä toimii siis melko hyvin sekä matalilla että korkeilla häntäkvantiileilla. Mutta kuten taulukosta 8 nähtiin, menetelmä yliarvioi VaR:a matalilla ja aliarvioi VaR:a korkeilla kvantiileilla ja lisäksi ylitykset eivät ole riippumattomia, joten ylitysten ehdollinen peittävyys ei ole täysin oikea. Riippuvuuden ja siten myös ehdollisen peittävyyden suhteen parempi menetelmä suurilla kvantiileilla on ehdollistamaton ääriarvomenetelmä. Tämä menetelmä kuitenkin yliarvioi hinnan nousun riskin suuruutta, joten ylityksiä tapahtuu liian harvoin. Ehdollisella normaalijakauma- ja t-jakaumamenetelmällä on ongelmia sekä ehdollistamattoman peittävyyden että riippumattomuuden kanssa, joten ehdollinen peittävyyskään ei ole oikea. Ääriarvoteoriaan perustuvat menetelmät toimivat siten sekä havaintojaksolla että käyttämälläni ennustemenetelmällä saaduissa riskiennusteissa paremmin kuin perinteisiin normaali- ja t-jakaumaan perustuvat menetelmät. Testaukset kuitenkin osoittavat, ettei ääriarvomenetelmien estimaatitkaan ole täysin tarkkoja.

Ennustejaksolle saadut tulokset ovat saman suuntaisia Byströmin (2005) ja Chanin ja Grayn (2006) tulosten kanssa. Ehdollinen ääriarvomenetelmä pystyy ennustamaan normaali- ja t-jakaumaan perustuvia ehdollisia menetelmiä tarkemmin sähkön hinnan nousun riskejä. Chanin ja Grayn (2006) päivittäisiin VaR-ennusteisiin verrattuna tunnittaisen ehdollisen VaR:n ylitykset eivät tapahdu yhtä riippumattomasti. Tähän voi vaikuttaa myös Chanin ja Grayn käyttämä AR-EGARCH-malli, jossa ehdollisen varianssin riippuvuus jäännöksestä on erilaista jäännöksen ollessa positiivinen ja negatiivinen. AR-EGARCH-malli ottaa siten huomioon jäännösten vaikutusvaltaisuuden ehdolliseen varianssiin autokorrelaation ja heteroskedastisuuden lisäksi, mikä perustuu siihen, että kysynnän lisäyksellä on suurempi vaikutus sähkön hintaan kuin kysynnän alenemisella (Chan ja Gray, 2006). Liitteessä 3 on esitetty eri menetelmillä ennustetut 95 %:n VaR:t ja niiden toteutuneet ylitykset ajanjaksolle 15.10.–9.11.2007. Liitteen kuvioista 15 ja 16 nähdään, että valitulla ajanjaksolla ehdollisen ääriarvomenetelmän toteutuneet ylitykset tapahtuvat tasaisemmin väliajoin kuin muiden menetelmien ylitykset. Kuvioista huomataan myös, että ehdollistamattomasta menetelmästä paremmin riskien arviointiin sekä muuttuvassa että tasaisessa markkinatilanteessa.

8 Johtopäätöksiä

Ääriarvoteoria on kehitetty kuvaamaan harvinaisia ja poikkeavia tapahtumia ja sillä pystytään estimoimaan hyvin suuriakin kvantiileja myös havaintojen ulkopuolelle. Ääriarvojakaumat sopivat usein kuvaamaan taloudellisia ilmiöitä, joiden jakaumat ovat hyvin paksuhäntäisiä ja vinoja. Usein talouden toimijoiden on tärkeää tietää, kuinka suuria (tai pieniä) havaintoja voi esiintyä ja mikä on niiden esiintymistodennäköisyys. Ääriarvoteoria tarjoaa menetelmiä riskien arviointiin sekä riskien hallintaan. Vaikka ääriarvomenetelmät ovat jo pitkään olleet käytössä luonnontieteessä, on niitä alettu soveltamaan vasta viime aikoina rahoitus- ja vakuutusalailla.

Tässä tutkielmassa esitettiin ääriarvoteoriaan perustuva menetelmä sähkömarkkinoiden riskien arviointiin ja ennustamiseen. Sovelluskohteena käytin sähkön tunnitusta tukkuhintaa, joka määräytyy Pohjoismaiden sähköpörssissä Nordpoolissa. Tarkastelun kohteena oli sähkön hinnassa tapahtuvat suuret nousut, jotka kiinnostavat erityisesti sähkön ostajia. Koska sähköä ei juuri pystytä varastoimaan, sähkön hinnassa esiintyy hyvin suurta vaihtelua ja poikkeavia havaintoja, joten ääriarvojakaumat sopivat hyvin kuvaamaan sähkön hinnan muutoksia. Riskin mittarina käytin Value at riskia (VaR), joka kertoo riskin suuruuden määrättyllä todennäköisyydellä. Esimerkiksi 95 %:n VaR kertoo, että on 5 %:n todennäköisyys, että seuraava havainto ylittää VaR:n tai toisin sanoen VaR:n ylityksiä tapahtuu keskimäärin 5 tapauksessa 100:sta. Riskien (VaR:ien) määräämiseen käytin McNeilin ja Freyn (2000) kehittämää kaksivaiheista estimointitapaa, jossa riskikohteen aikasarja suodatetaan ennen ääriarvomenetelmän käyttöä. VaR:t saadaan tässä menetelmässä aikasarjamallin ja sen jäännöksiin sovitettavan ääriarvojakauman avulla. Tätä estimointitapaa kutsutaan ehdolliseksi menetelmäksi, koska VaR:n estimaatteihin vaikuttavat aikasarjamallin ehdollinen odotusarvo ja hajonta. Ääriarvomenetelmistä valitsin sovellettavaksi kynnsarvon ylitys -menetelmän, jossa sovitetaan yleistetty Pareto-jakauma valitun kynnsarvon ylittäviin havaintoihin. Hinnan muutoksen aikasarjan suodattamiseen käytin AR-GARCH-mallia, jolla saadaan puhdistettua sähkön hinnassa esiintyvää autokorrelaatiota ja heteroskedastisuutta.

Ehdollistettua ääriarvomenetelmää vertasin normaali- ja t-jakaumiin perustuviin ehdollisiin menetelmiin sekä ehdollistamattomaan ääriarvomenetelmään, jossa kynnsarvon ylitys -menetelmää sovelletaan suoraan sähkön hinnan muutoksiin ilman

aikasarjamallinnusta. Menetelmien vertailut suoritin kahdella eri aikajaksolla, joista ensimmäisellä estimoin kerralla VaR:t koko havaintojaksolle ja toisella käytin menetelmiä ennustamaan seuraavan 24 tunnin VaR:t. Saatuja estimaatteja vertasin toteutuneisiin hinnan nousuihin ja testasin menetelmien antamien riskiestimaattien tarkkuutta. Hyvän riskien arviointimenetelmän tulisi antaa oikea ylityksien osuus koko ajanjaksolle (ehdollistamaton peittävyys) sekä reagoida markkinoilla tapahtuviin muutoksiin, jolloin ylityksien tulee tapahtua myös riippumattomasti (ehdollinen peittävyys).

Menetelmien testausten mukaan ehdollisella ääriarvomenetelmillä voidaan estimoida ja ennustaa sähkön hinnan suuret muutokset tarkemmin kuin ehdollisella normaali- tai t-jakaumamenetelmällä. Myös ehdollistamattomalla ääriarvomenetelmällä pystytään määräämään häntäkvantiilit tarkasti koko tarkasteltavalle ajanjaksolle, mutta lyhyempiä jaksoja tarkastellessa ehdollistamaton menetelmä ei välttämättä toimi yhtä hyvin kuin ehdollistettu menetelmä, koska se ei reagoi hinnan volatiliteetin vaihteluun. Havaintojaksolle tehdyissä estimoinneissa ehdollinen ääriarvomenetelmä oli selvästi muita menetelmiä parempi ehdollistamattomassa peittävyudessa, mutta ennusteissa ei näin selvää eroa syntynyt. Normaali- ja t-jakaumaan perustuvat menetelmät aliarvioivat selvästi riskin suuruutta, jolloin käytännön sovelluksissa voi tapahtua odottamattomia menetyksiä. Täysin oikeaa ehdollistettua peittävyyttä ei saatu millään menetelmällä, koska ylitykset eivät tapahdu täysin riippumattomasti. Kuitenkin ehdollistetun ääriarvomenetelmän ylitykset ovat jakautuneet tasaisesti yli ajan, kun taas ehdollistamattoman menetelmän ylitykset tapahtuvat klustereina.

Tässä tutkielmassa käytin sähkön hinnan muutoksen aikasarjan autokorrelaation ja heteroskedastisuuden suodattamiseen AR-GARCH-mallia. Mallin jäännöksiin jäi kuitenkin jäljelle autokorrelaatiota ja kausivaihtelua. Tämä voi huonontaa kynnyksarvon ylitys -menetelmän sopivuutta jäännöksen mallinnuksessa, koska ääriarvoteoria olettaa riippumattomat ja samoin jakautuneet havainnot. Sähkön hinnassa esiintyy erittäin voimakasta kausivaihtelua, jonka poistaminen AR-GARCH-mallilla ei täysin onnistunut. Jäännösten kausivaihtelun ja sen aiheuttaman autokorrelaation eliminoimiseksi voidaan käyttää flexible Fourier form -menetelmää, jolla mallinnetaan volatiliteetin jaksollisuutta (ks. Andersen ja Bollerslev, 1997, 1998). Estimoidun AR-GARCH-mallin parametrit rajoitin tuottamaan heikosti stationaarisen prosessin. Rajoittaminen huononsi hieman havaintojaksolle estimoitujen VaR:n tarkkuut-

ta, mutta rajoitusta tarvittiin ennustejaksolla ehdollisen varianssin ennustamiseen. Jatkotutkimuksissa tulisi pyrkiä löytämään menetelmiä, joissa GARCH-parametrien heikkoa stationarisuutta ja siten parametrirajoituksia ei tarvita. AR-GARCH-malli ottaa huomioon aikasarjan ehdollisen varianssin aikariippuvuuden, muttei positiivisen ja negatiivisen sokin erilaista vaikuttavuutta ehdolliseen varianssiin. Eksponentiaalinen GARCH-malli (EGARCH) huomioi myös sokin vaikutusvaltaisuuden ja se voisi sopia mallintamaan paremmin sähkön hinnan muutosta, sillä positiivisen ja negatiivisen kysyntäsokin vaikutus hintaan on erilainen (Chan ja Gray, 2006). Sähkön hinnassa esiintyy myös vuodenajasta riippuvaa vaihtelua, sillä lämpötila ja valoisuus vaikuttavat sähkön kysyntään. Sähkön hinnan muutoksen vuodenaikavaihtelun on ottanut huomioon esimerkiksi Byström (2005).

Tässä tutkielmassa tarkasteltiin sähkön hinnan nousujen tunnittaisia riskejä, mutta käytännön sovelluksissa ei välttämättä ole kiinnostavaa tietää mikä on VaR seuraavien tuntien aikana. Sähkön ostajien mielenkiinnon kohteena on seuraavien päivien VaR:t, jolloin suuriin hinnan nousuihin pystytään varautua ja mahdollisesti suojautua riskien aiheuttamilta haitoilta. Tuntiaineistoa käytettäessä VaR:t pystytään aggregoimaan esimerkiksi päivittäisiksi skaalaamalla tuntiestimaatit joko ajan neljänneksellä tai sovitetun yleistetyn Pareto-jakauman muotoparametrin avulla (ks. Byström, 2005; Jorion, 2007, 133).

Saadut tulokset ovat saman suuntaisia kuin aiemmissa tutkimuksissa, joissa on verrattu ääriarvoteoriaan perustuvia riskinarviointimenetelmiä perinteisempiin menetelmiin sähkömarkkinoilla. Ehdollinen ääriarvomenetelmä pystyy mallintamaan tunnittaisen sähkön hinnan muutoksia tarkasti sekä matalilla että korkeilla kvantiileilla. Tutkielmassa esitetyt ääriarvomenetelmät voivat olla hyödyllisiä sähkömarkkinoiden riskien arvioinnissa ja hallinnassa, koska sähkömarkkinat ovat hyvin volatiilit sähkön ominaisuuksien vuoksi.

Lähteet

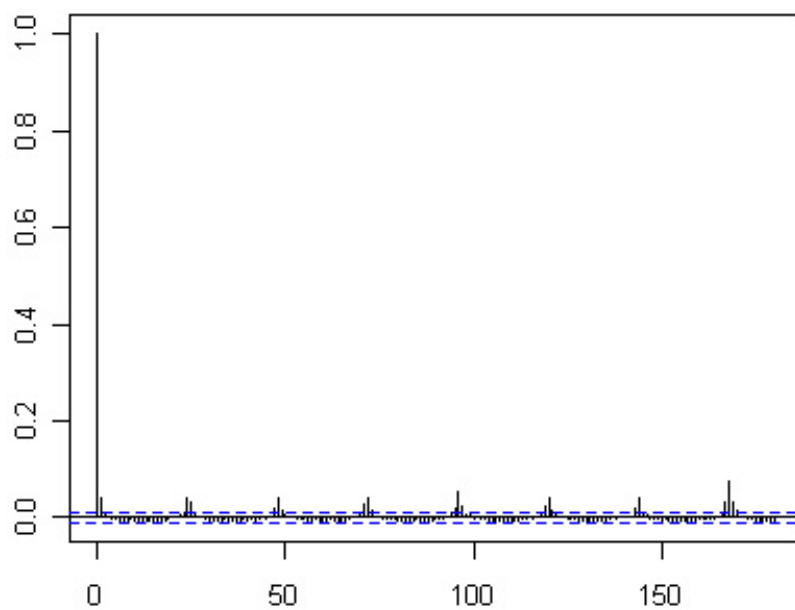
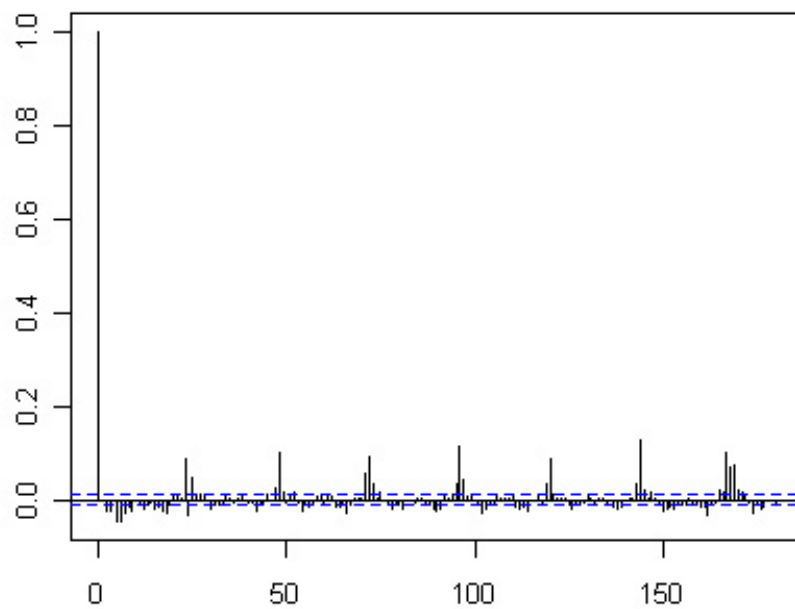
- Andersen, T. G. ja Bollerslev, T. 1997: Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2-3), 115–158.
- Andersen, T. G. ja Bollerslev, T. 1998: Deutsche mark-dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance*, 53(1), 219–265.
- Andrews, N. ja Thomas, M. 2002: At the end of the tail. *Energy & Power Risk Management*, 7(1), 75–77.
- Bali, T. G. 2003: An extreme value approach to estimating volatility and value at risk. *Journal of Business*, 76(1), 83–108.
- Balkema, A. ja de Haan, L. 1974: Residual life time at great age. *Annals of Probability*, 2(5), 792–804.
- Byström, H. N. E. 2005: Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, 14(1), 41–55.
- Chan, K. F. ja Gray, P. 2006: Using extreme value theory to measure value-at-risk for daily electricity spot prices. *International Journal of Forecasting*, 22(2), 283–300.
- Christoffersen, P. F. 1998: Evaluating interval forecasts. *International Economic Review*, 39(4), 841–62.
- Davison, A. C. ja Smith, R. L. 1990: Models for exceedances over high thresholds. *Journal of the Royal Statistical Society. Series B*, 52(3), 393–442.
- Embrechts, P., Mikosch, T. ja Klüppelberg, C. 1997: *Modelling extremal events: for insurance and finance*. Springer-Verlag.
- Fernandez, V. 2003: Extreme value theory and value at risk. *Revista de Analisis Economico*, 18(1), 57–85.
- Fisher, R. A. ja Tippett, L. H. C. 1928: Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Proceedings of the Cambridge Philosophical Society*, 24, 180–190.

- Francq, C. ja Zakoïan, J.-M. 2004: Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4), 605–637.
- Franses, P. H. ja van Dijk, D. 2000: *Non-linear time series models in empirical finance*. Cambridge university press.
- Gnedenko, B. 1943: Sur la distribution limite du terme maximum d’une serie aleatoire. *Annals of Mathematics*, 44(3), 423–453.
- Gumbel, E. J. 1958: *Statistics of extremes*. Columbia University Press.
- Hamilton, J. D. 1994: *Time series analysis*. Princeton.
- Hosking, J. ja Wallis, J. 1987: Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3), 339–349.
- Jenkinson, A. F. 1955: The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Quarterly Journal of the Royal Meteorological Society*, 87, 158–171.
- Jorion, P. 2007: *Value at risk - the new benchmark for managing financial risk*. McGraw Hill, 3. painos.
- Katz, R. W., Brush, G. S. ja Parlange, M. B. 2005: Statistics of extremes: modeling ecological disturbances. *Ecology*, 86(5), 1124–1134.
- Leadbetter, M. R., Lindgren, G. ja Rootzen, H. 1983: *Extremes and related properties of random sequences and processes*. Springer.
- Longin, F. M. 2000: From value at risk to stress testing: the extreme value approach. *Journal of Banking & Finance*, 24(7), 1097–1130.
- McNeil, A. J. ja Frey, R. 2000: Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3-4), 271–300.
- von Mises, R. 1936: La distribution de la plus grande de n valeurs. Teoksessa *Selected Papers of Richard von Mises*, 271–294. American Mathematical Society. Uudelleenjulkaisu.
- Nelson, D. B. 1990: Stationarity and persistence in the GARCH(1,1) model. *Econometric Theory*, 6(3), 318–34.

- Nord Pool 2007: <http://www.nordpool.no/> Haettu 1. huhtikuuta 2009.
- Pickands, J. 1975: Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119–131.
- Rachev, S. T. (toim.) 2003: *Handbook of heavy tailed distributions in finance*. Elsevier Science B. V.
- de Rozario, R. 2002: Estimating value at risk for the electricity market using a technique from extreme value theory. Työpaperi. School of Banking and Finance, University of NSW.
- Smith, R. 1987: Estimating tails of probability distributions. *Annals of Statistics*, 15(3), 1174–1207.
- Smith, R. L. 2001: Extreme value statistics in meteorology and the environment. CBMS course in environmental statistics. Julkaisematon kurssimateriaali.
- Verhoeven, P. ja McAleer, M. 2000: Modelling outliers and extreme observations for ARMA-GARCH processes. Econometric Society World Congress 2000 Contributed Papers 1922, Econometric Society.

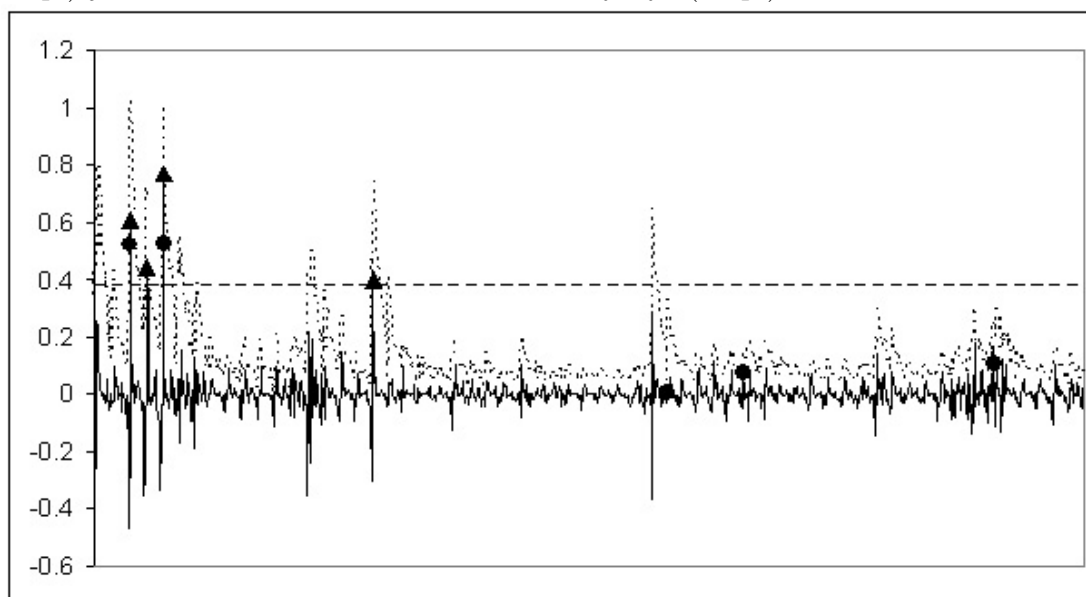
Liite 1: Jäännösten autokorrelaatiofunktiot

Kuva 12: Jäännösten autokorrelaatiofunktiot. Ylhäällä jäännökselle $\hat{\epsilon}_t$, alhaalla neljöllä jäännökselle $\hat{\epsilon}_t^2$

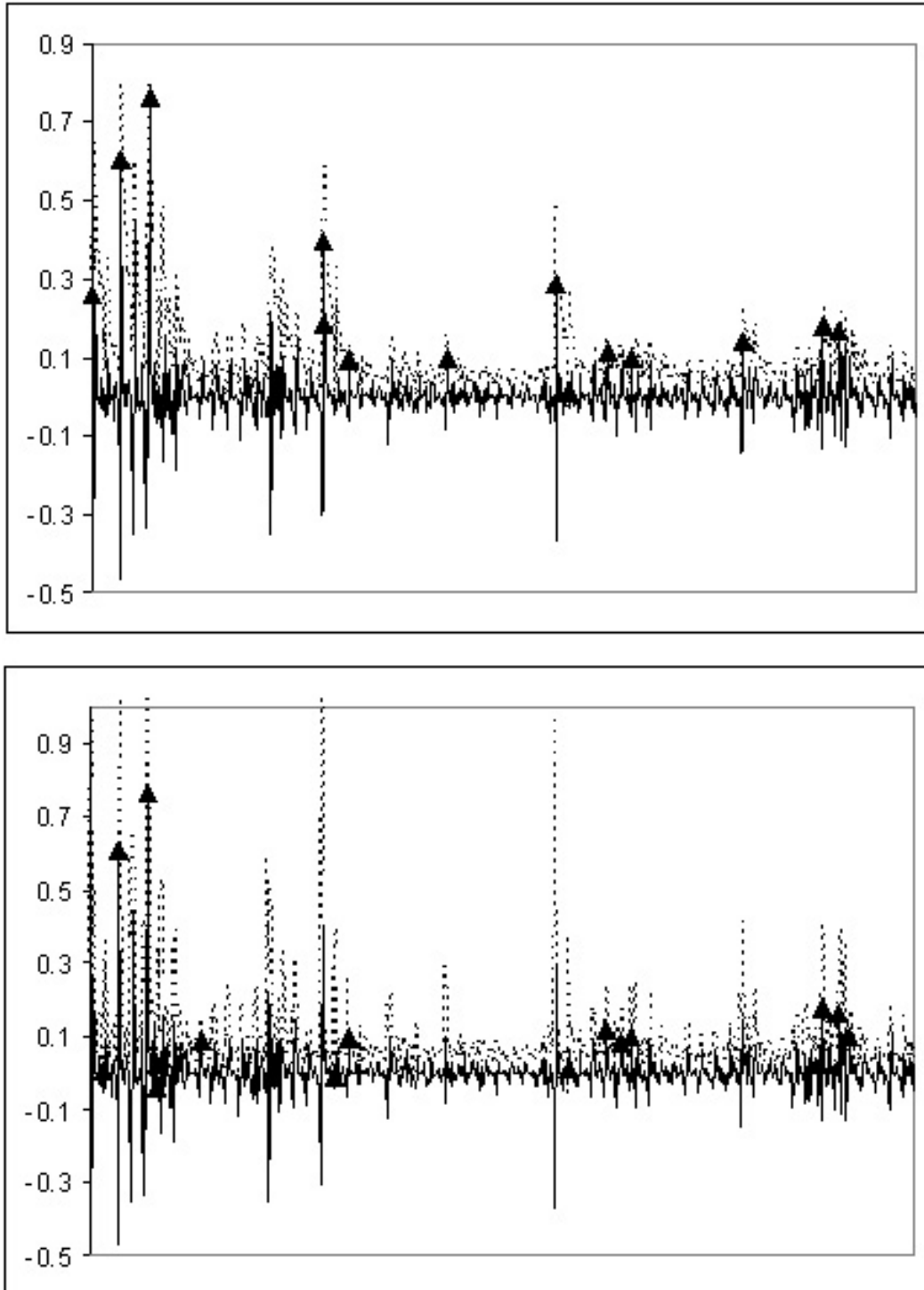


Liite 2: Havaintojakson VaR:t ja ylitykset

Kuva 13: Ehdollisen ja ehdollistamattoman ääriarvomenetelmän 99 % VaR ja ylitykset marras-joulukuussa 2004. Yhtenäinen viiva on hinnan logaritmimuutos, pisteiviiva ehdollinen VaR, katkoviiva ehdollistamaton VaR, ympyrä ehdollisen VaR:n (5 kpl) ja kolmio ehdollistamattoman VaR:n ylitys (4 kpl)

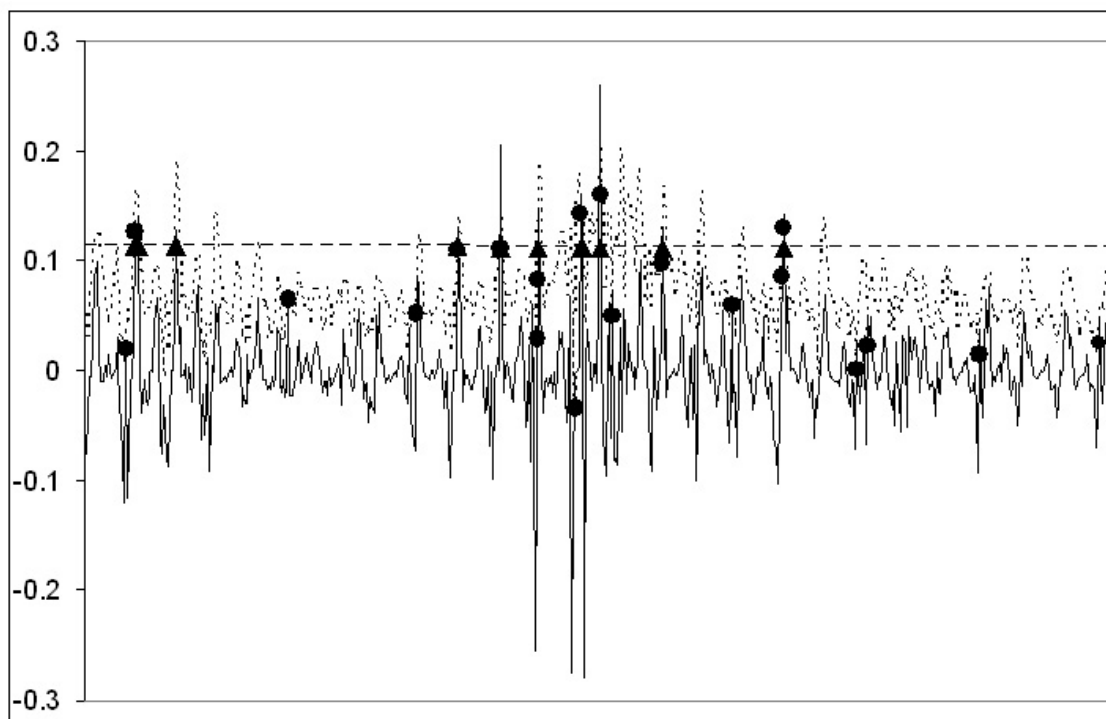


Kuva 14: Normaalijakaumaan (ylhällä) ja t-jakaumaan (alhaalla) perustuvat 99 %:n ehdolliset VaR:t ja ylitykset marras-joulukuussa 2004. Yhtenäinen viiva on hinnan logaritmimuutos, pisteviiva ehdollinen VaR ja kolmio VaR:n ylitys

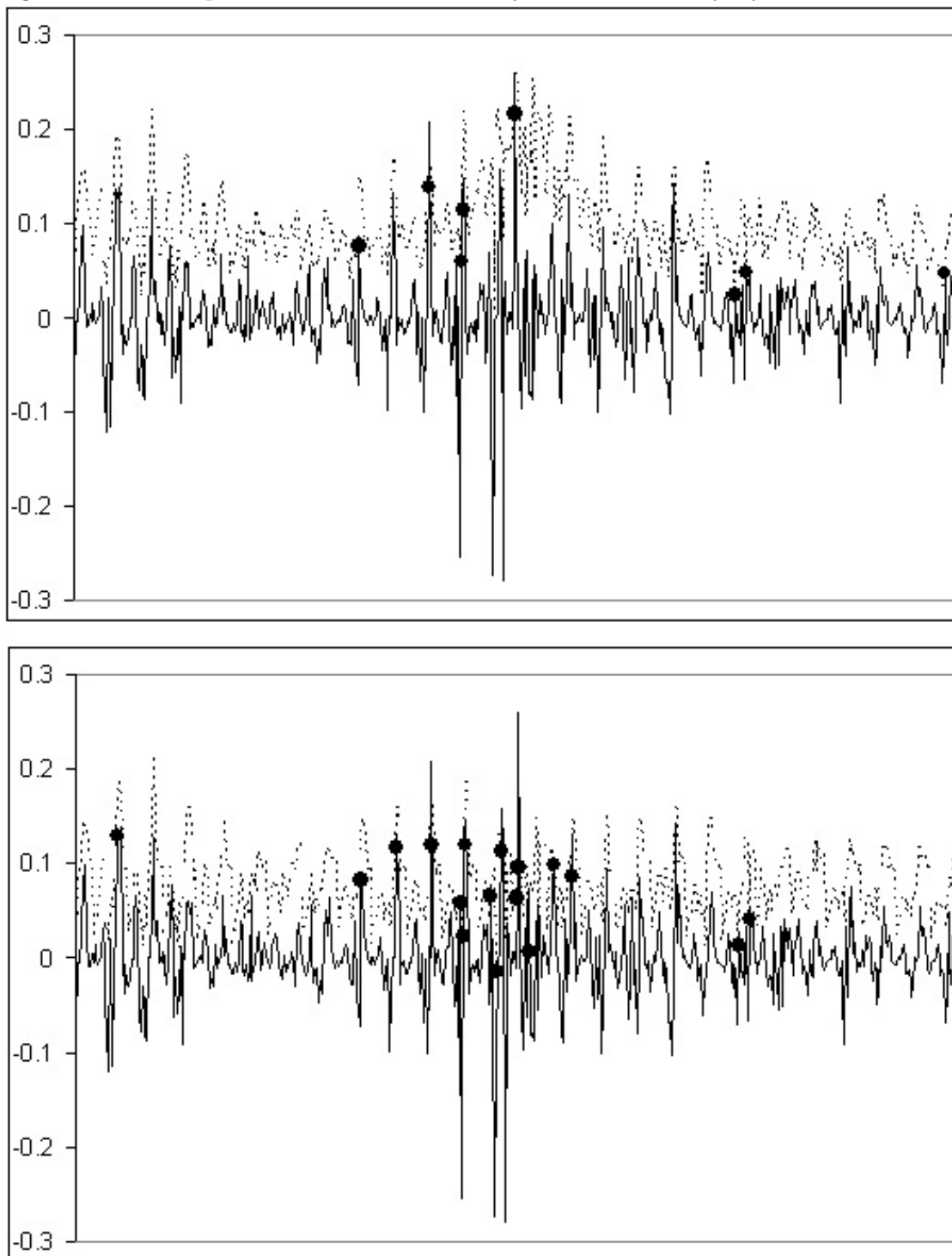


Liite 3: Ennustejakson VaR:t ja ylitykset

Kuva 15: Ehdollisen ja ehdollistamattoman ääriarvomenetelmän 95 % VaR ja ylitykset ajalla 15.10.–9.11.2007. Yhtenäinen viiva on hinnan logaritmimuutos, pisteviiva ehdollinen VaR, katkoviiva ehdollistamaton VaR, ympyrä ehdollisen VaR:n (20 kpl) ja kolmio ehdollistamattoman VaR:n ylitys (13 kpl)



Kuva 16: Normaalijakaumaan (ylhällä) ja t-jakaumaan (alhaalla) perustuvat 95 %:n ehdolliset VaR:t ja ylitykset ajalla 15.10.–9.11.2007. Yhtenäinen viiva on hinnan logaritmimuutos, pisteviiva ehdollinen VaR ja kolmio VaR:n ylitys



VATT TUTKIMUKSET -SARJASSA ILMESTYNEITÄ

PUBLISHED VATT RESEARCH REPORTS

129. Kiander Jaakko: Julkisen talouden liikkumavara vuoteen 2030 mennessä. Helsinki 2007.
130. Lintunen Jussi: Tuloerojen ja taloudellisen eriarvoisuuden mittaamisesta: Sovellus Suomen kulutustutkimuksilla. Helsinki 2007.
131. Kirjavainen Tanja: Nuorten lukiokoulutuksen tehokkuus 2000–2004. Helsinki 2007.
132. Ollikainen Virve: Ammatillisen peruskoulutuksen kustannustehokkuus 2001–2003. Helsinki 2007.
133. Kyyrä Tomi: Studies on Wage Differentials and Labour Market Transitions. Helsinki 2007.
134. Mannermaa Kauko: Ohjailusta kilpailuun – Suomen hallitusten kasvu- ja rakennepolitiikka vuosina 1962–1999. Helsinki 2007.
135. Aaltonen Juhon – Kirjavainen Tanja – Moisio Antti – Ollikainen Virve: Perusopetuksen, lukioiden ja ammatillisen peruskoulutuksen tuottavuus ja tehokkuus – Loppuraportti. Helsinki 2007.
136. Parkkinen Pekka: Väestön ikääntymisen vaikutukset kuntatalouteen. Helsinki 2007.
137. Rätty Tarmo – Aaltonen Juhon – Kirjavainen Tanja: Tuloksellisuuden ja tuottavuuden mittaaminen ammattikorkeakouluissa. Helsinki 2008.
138. Rätty Tarmo – Harava Maiju: Kokonaistuottavuuden kehitys yliopistoissa. Helsinki 2008.
139. Honkatukia Juha – Forsström Juha: Ilmasto- ja energiapolitiittisten toimenpiteiden vaikutukset energiajärjestelmään ja kansantalouteen. Helsinki 2008.
140. Lyytikäinen Teemu: Studies on the Effects of Property Taxation, Rent Control and Housing Allowances. Helsinki 2008.
141. Lehtonen Sanna – Lyytikäinen Teemu – Moisio Antti: Kuntien rahoitus- ja valtionosuusjärjestelmä: Vaihtoehtoja uudistuksen toteuttamiseksi. Helsinki 2008.
142. Pekkala Johanna: Tuloerojen ja terveyden välinen yhteys Suomessa. Helsinki 2009.
- 143.1. Perrels Adriaan – Hongisto Mikko – Hyvönen Kaarina – Katajajuuri Juha-Matti – Nissinen Ari: A quick scan of climate policy services and of underlying data system approaches. Climate Bonus project report (WP1). Helsinki 2009.
144. Karvinen Anni-Mari: Maahanmuuttajien poismuutto Suomesta. Helsinki 2009.
145. Anna Sahari – Adriaan Perrels: Ekotehokkuutta parantavat investoinnit kesämökeillä. Helsinki 2009.



VALTION TALOUDELLINEN TUTKIMUSKESKUS
STATENS EKONOMISKA FORSKNINGSCENTRAL
GOVERNMENT INSTITUTE FOR ECONOMIC RESEARCH

Valtion taloudellinen tutkimuskeskus
Government Institute for Economic Research
P.O.Box 1279
FI-00101 Helsinki
Finland

ISBN 978-951-561-852-8
ISSN 0788-5008