308.

Jyväskylän yliopisto

Taloustieteiden tiedekunta

ÄÄRIMMÄISET MARKKINAMUUTOKSET JA VALUE AT RISK

Kansantaloustieteen
pro gradu –tutkielma
Kesäkuu 2000
Tutkielman laatijat:
Jyri Koskela
Jari Ritvanen
Tutkielman ohjaaja:
KTL Juhani Raatikainen

SISÄLLYS

Jyväskylän yliopisto **Tutkielmatiivistelmä**

Aine: Kansantaloustiede

Tekijät: Jyri Koskela ja Jari Ritvanen

Otsikko: Äärimmäiset markkinamuutokset ja Value at Risk

Ohjaaja: KTL Juhani Raatikainen

Sivumäärä: 53

Yhteenveto

Äärimmäisten markkinamuutosten tarkka ennustaminen on tärkeää monissa rahoituksen sovelluksissa ja erityisesti Value at Risk –analyysissa. Tässä tutkielmassa selvitetään, voidaanko suuret muutokset ottaa tarkemmin huomioon käyttämällä tilastollista extreme value –teoriaa. Tutkielmassa käytetyn puoliparametrisen menetelmän ovat kehittäneet Danielsson ja de Vries (1997b). Menetelmän oletukset pyritään huomioimaan paremmin muokkaamalla havaintoaineisto pääkomponenttianalyysillä ja GARCH-mallituksella. Tutkielman lähteet ovat pääasiassa uusimpia riskienhallintaan liittyviä artikkeleita.

Tutkielma jakautuu kolmeen osaan. Alussa käsitellään tavallisimpien Value at Risk –mallien heikkouksia suurten markkinamuutosten kannalta. Kolmas luku sisältää teoreettisen viitekehyksen ja neljännessä luvussa suoritetaan extreme value –estimointi ja verrataan saatuja tuloksia normaalijakaumaan.

Tutkielmassa osoitetaan, että normaalijakaumaan perustuvat Value at Risk –menetelmät aliarvioivat markkinariskiä. Käytetty menetelmä on tuloksiltaan lupaava, mutta vaatisi kehittämistä käytännön soveltamisen helpottamiseksi.

Avainsanat: Value at Risk, extreme value –estimointi, markkinariski.

1. JOHDANTO

1.1. Rahoitusmarkkinat ja riskienhallinta

Taloustieteessä rahoitusmarkkinat määritellään yhteisnimitykseksi tavoille myydä rahoitusylijäämät niille, joilla on tarve investoida tai kuluttaa enemmän kuin mihin heidän samaan aikaan käytettävissä olevat tulonsa riittävät. Rahoitusmarkkinoilla myydään ja ostetaan rahoitusvaateita, jotka ovat tulevia maksuja koskevia sitoumuksia. Tulevaisuuden odotuksiin liittyy aina epävarmuutta, joten kaupankäyntiin rahoitusvaateilla sisältyy riski.

Talousteorian näkemyksen mukaan riski ei ole ainoastaan tappion vaan myös voiton mahdollisuus, joten riskin voidaan sanoa olevan epävarmuutta sijoituksen tulevasta arvosta (Jorion 1997, 63). Käytännössä ei kuitenkaan ole tarkoituksenmukaista puhua voiton mahdollisuudesta riskinä, sillä odottamattomalta voitolta suojautuminen ei ole tarpeellista. Rahoitusriskityypit pankkitoiminnassa voidaan epävarmuuden lähteen perusteella jakaa viiteen luokkaan (Simons 1996). *Luottoriskillä* tarkoitetaan vaaraa siitä, että vastapuoli ei täytä sopimusvelvoitteitaan. *Operationaalisella riskillä* puolestaan tarkoitetaan mahdollisesti sijoittajan järjestelmissä, toimintatavoissa tai valvontarutiinissa piileviä vaaroja. *Likviditeettiriski* merkitsee vaaraa siitä, että sijoittaja ei pysty suoriutumaan maksuistaan, kun taas *markkinariski* aiheutuu markkinahintojen muutoksiin liittyvästä epävarmuudesta. Rahoituslaitosten suorittamien oikeustoimien sisältämiä riskejä kutsutaan *juridisiksi riskeiksi.*

Rahoitusmarkkinoilla on nykyisin tarjolla aiempaa enemmän monimutkaisia sijoitusinstrumentteja. Uusien sijoituskohteiden hinnoittelun vaikeudesta johtuen niiden sisältämien riskien arvioiminen on ollut ongelmallista. Useat suuret rahoituslaitokset ovatkin kärsineet valtavia tappioita rahoitusmarkkinoilla. Useissa tapauksissa riskien realisoituminen on aiheutunut puutteellisesta valvonnasta, minkä seurauksena Value at Risk[1] (VaR) –laskennasta on tullut rahoitusmarkkinoilla toimivien yritysten riskienhallinnan tärkein työkalu. Osaltaan VaR-laskennan yleistymiseen on vaikuttanut myös Baselin komitean[2] mietintö, jonka mukaan liikepankit voivat käyttää VaR-laskentaa pääomavelvoitteidensa määräämisessä. VaR-analyysimenetelmät jaetaan yleensä kolmeen osaan: delta-normaaliin-lähestymistapaan, historialliseen simulointiin ja Monte Carlo –simulointiin. Kaikkiin näihin lähestymistapoihin liittyy ongelmia. Monte Carlo –simulointi ja delta-normaali-lähestymistapa perustuvat yleensä normaalijakaumaan, jolloin harvinaiset, mutta kuitenkin empiirisesti havaittavissa olevat tapahtumat saavat liian pienen esiintymistodennäköisyyden. Tämä tarkoittaa käytännössä sitä, että todennäköisyysjakauman hännät ovat liian ohuet. Historiallisen simuloinnin ongelma taas on se, että se ottaa huomioon vain aikaisemmin havaitut markkinamuutokset. Historiallinen simulointi olettaa myös, että rahoitusmarkkinoiden toiminta pysyy ajassa muuttumattomana, ja että valitun ajanjakson havainnot kuvaavat markkinoiden toimintaa täydellisesti.

Mielenkiintoinen ongelma on se, kuinka suurten tappioiden mahdollisuus voidaan ottaa paremmin huomioon. Yhtenä ratkaisuna tähän on pidetty stressitestejä, joissa määritellään katastrofiskenaario ja tutkitaan sen vaikutus salkun arvoon. Stressitestejä ei voi kuitenkaan pitää kovin hyvänä ratkaisuna useasta eri syystä: skenaariot asetetaan ad hoc –pohjalta, jolloin arviot salkun arvon kehityksestä perustuvat subjektiiviseen oletukseen markkinamuutoksesta. Stressitestit käsittelevät huonosti korrelaatioita, jotka kuitenkin ovat olennainen osa portfolion riskiä. Ongelmallista on

---

[1] Value at Risk mittaa suurinta mahdollista tappiota annetulla aikavälillä halutulla luottamustasolla (Beder 1995).

[2] Baselin komitea, the Basle Committee on Banking Supervision, on pankkivalvontaviranomaisista koostuva elin, joka säätelee pankkien pääomavelvoitteita. Baselin komitea perustettiin vuonna 1975 kymmenen keskuspankin toimesta. Nykyään sen toimintaan osallistuu keskuspankkien ja rahoitusvalvontaviranomaisten edustajia kahdestatoista eri maasta, jotka ovat Belgia, Kanada, Ranska, Saksa, Italia, Japani, Luxemburg, Hollanti, Ruotsi, Sveitsi, Englanti ja Yhdysvallat.

myös arvioida, mitä tapahtuu salkun arvolle monen instrumentin arvon ja yleisen markkinatilanteen muuttuessa. Lisäksi stressitestejä käytettäessä portfolion arvon muutokseen ei voida liittää arviota toteutumistodennäköisyydestä, jolloin menetetään tärkeä osa informaatiosta verrattuna Value at Risk –analyysiin. Stressitesteillä voidaan ainoastaan täydentää VaR-analyysia, mutta ei korvata sitä.

Viime vuosina stressitestejä parempana menetelmänä ottaa huomioon äärimmäiset markkinamuutokset on pidetty niin sanottua extreme value –estimointia, jota voidaan pitää normaalin Value at Risk –analyysin laajennuksena. Kyseisessä lähestymistavassa voidaan ottaa paremmin huomioon markkinoiden poikkeuksellisen suuret muutokset. Menetelmässä estimoidaan suurten tappioiden todennäköisyysjakaumaa jollain paksuhäntäiseksi tunnetulla jakaumalla, esimerkiksi Pareto-jakaumalla. Onnistuneen estimoinnin tulosten perusteella voidaan ennustaa tarkasti suurten tappioiden todennäköisyyttä. Yhdistämällä extreme value –estimointi esimerkiksi historialliseen simulointiin, saadaan muodostettua todennäköisyysjakauma, jonka keskiosan havainnot perustuvat empiiriseen jakaumaan ja häntien havainnot extreme value –tekniikalla laskettuihin tuloksiin.

Tässä tutkielmassa tavoitteena on selvittää, toimiiko extreme value –estimointi normaalijakaumaan perustuvaa Value at Risk –mallia luotettavammin. Tutkielman aineistona käytetään seitsemästätoista sijoitusinstrumentista muodostettua portfoliota. Normaalijakauman arvoja verrataan extreme value –estimoinnilla laskettaviin kynnysarvoihin, jolloin voidaan tehdä johtopäätökset normaalijakaumiin perustuvien VaR-mallien luotettavuudesta. Lisäksi tutkielmassa käsitellään tavanomaisia VaR-malleja ja sitä, kuinka ne ottavat huomioon suuret markkinamuutokset.

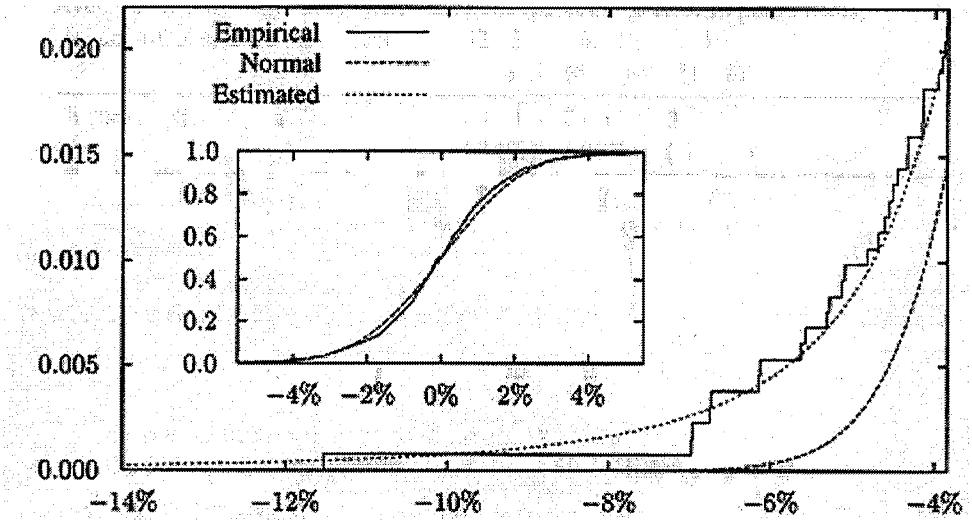
1.2. Aikaisempien tutkimusten tuloksia

Danielsson ja de Vries (1998) ovat selvittäneet extreme value –estimoinnin tuloksia vertaamalla niitä perinteisten Value at Risk –menetelmien tuloksiin. Vertailu toteutettiin käyttämällä aineistona päivittäisiä öljyn hintoja vuodesta 1986 vuoteen 1997 saakka. Danielsson et al vertasivat kehittämäänsä menetelmää (Danielsson & de Vries 1997b) historialliseen simulointiin ja varianssi-kovarianssi-menetelmän antamiin tuloksiin. Tutkimustulokset osoittavat, että extreme value –estimointi "pehmentää" jakauman häntää, jolloin saatu VaR-luku on luotettavampi verrattuna perinteisten menetelmien tuottamiin tuloksiin. Sen avulla voidaan myös laskea VaR-lukuja otoksen ulkopuolelta (vertaa historiallinen simulointi). Extreme value –estimointi mahdollistaa siis sellaistenkin suurten markkinamuutosten todennäköisyyksien arvioimisen, joita käsiteltävä aineisto ei sisällä.

Tuloksia voidaan havainnollistaa vertaamalla extreme value –estimoinnilla muodostettua käyrää varianssi-kovarianssi-metodilla laskettuun estimaattiin. Kuviosta 1 nähdään selvästi, että delta-normaaliin-menetelmään perustuva käyrä on kahden muun menetelmän tuottamien käyrien alapuolella. Voidaan siis selkeästi havaita, että normaalisuusoletus johtaa merkittävään VaR:n aliarviointiin. Esimerkiksi normaalijakaumaan perustuva VaR-luku ennustaa 13 kertaa pienemmän todennäköisyyden 6%:n pudotukselle kuin extreme value –estimoinnilla saatu estimaatti. Danielsson et al ennustavat, että kerran 15 vuodessa voidaan odottaa 28%:n päiväpudotusta öljyn hinnassa ja 44%:n päivittäistä pudotusta kerran 70 vuodessa. Lisäksi kerran vuodessa voidaan öljyn hinnan odottaa nousevan noin kymmenen prosenttia.

Kuvio 1.

Päivittäisten öljyn hintojen tappiohäntä 1992-1997



Lähde: Danielsson et al (1998).

Danielsson ja de Vries (1997a) vertailivat extreme value –estimoinnilla saatuja tuloksia historialliseen simulointiin ja varianssi-kovarianssi-menetelmään käyttäen osakemarkkina-aineistoa. Aineisto koostui yhdysvaltalaisista osakeportfolioista neljän vuoden ajalta. Menetelmiä testattiin vertaamalla niiden tuottamia virheitä odotettuihin VaR-virheisiin. Esimerkiksi Baselin komitean suosittelemalla menetelmällä arvioituna 99%:n luottamustasolla odotettu rikkeiden määrä on kymmenen. Varianssi-kovarianssi-menetelmä tuotti 16 virheellistä arviota kun taas extreme value –estimoinnilla vääriä arvioita tuli vain kahdeksan. Taulukosta 1 voidaan havaita, että extreme value –estimointi toimii menetelmistä luotettavimmin silloin kun luottamustaso on suuri. Alemmilla luottamustasoilla menetelmien väliset erot muuttuvat jonkin verran ja esimerkiksi 95%:n luottamustasolla varianssi-kovarianssi-menetelmä toimii parhaiten.

Taulukko 1.

Eri VaR-mallien toimivuus (1000 päivän ajanjaksolla).

Luottamustaso	**95 %**	**99 %**	**99,95 %**
Odotettu VaR rikkomusten lkm.	50	10	0,5
Toteutuneiden VaR-rikkomusten keskimääräinen lukumäärä. (prosentuaalinen virhe suluissa)			
Variassi-kovarianssi-menetelmä	52,45	16,28	3,55
	(4 %)	(63 %)	(610 %)
Historiallinen simulointi	43,24	7,66	0
	(-14 %)	(-23 %)	(-100 %)
Häntäestimaattori	43,14	8,19	0,59
	(-14 %)	(-18 %)	(18 %)

Mallit testattiin 500 satunnaisella portfoliolla, jotka koottiin seitsemästä yhdysvaltalaisesta osakkeesta periodilla 1993-1996.

Lähde: Danielsson et al (1998).

Danielsson ja de Vries (1997a) ovat selvittäneet, kuinka hyvin eri menetelmillä voidaan ennustaa suurten markkinamuutosten todennäköisyyksiä. Tässä selvityksessä käytettiin aineistona päivittäistä SP-500 indeksiä vuodesta 1990 vuoteen 1996. Vertailtavana olleet menetelmät olivat GARCH, Student-t GARCH ja extreme value –estimointi. Vertaamalla eri menetelmillä saatuja tuloksia empiirisesti todettuihin tuloksiin saatiin selville, että extreme value –estimointi toimi luotettavimmin (taulukko 2).

Taulukko 2.

Suuret päivittäiset markkinamuutokset SP-500 indeksistä aikaväliltä 1990-1996 ja niiden esiintymistodennäköisyydet estimoituna eri menetelmillä.

Havaittu tuotto	**Tuottojen todennäköisyydet**			
	Normaali	Student-t	EV-estim.	Empiirinen
-3,72 %	0,0000	0,0002	0,0007	0,0006
-3,13 %	0,0000	0,0010	0,0015	0,0011
-3,07 %	0,0002	0,0021	0,0016	0,0017
-3,04 %	0,0032	0,0071	0,0016	0,0023
-2,71 %	0,0098	0,0146	0,0026	0,0028
-2,62 %	0,0015	0,0073	0,0029	0,0034
3,66 %	0,0000	0,0011	0,0004	0,0006
3,13 %	0,0060	0,0096	0,0009	0,0011
2,89 %	0,0002	0,0022	0,0013	0,0017
2,86 %	0,0069	0,0117	0,0014	0,0023
2,53 %	0,0059	0,0109	0,0025	0,0028
2,50 %	0,0007	0,0038	0,0026	0,0034

Normaali GARCH, Student-t GARCH ja extreme value –menetelmä vs. empiiriset tulokset.
Lähde: Danielsson ja de Vries (1997a).

McNeil (1999) on esittänyt vaihtoehtoisen, täysin parametrisen, lähestymistavan äärimmäisten markkinamuutosten estimoinnille. Jälkitestauksen perusteella hän toteaa, että normaalijakaumaan perustuvat mallit aliarvioivat VaR-estimaattia. McNeilin johtopäätös on, että äärimmäisten tuottojen estimointi on ainoa tieteellisesti toimivaksi todistetuttu menetelmä poikkeavien havaintojen ennustamiseen.

Frey ja McNeil (1999) ovat lähestyneet äärimmäisten havaintojen mallittamista täysin parametrisesta näkökulmasta. Tutkimuksessaan he esittävät menetelmän, jolla heteroskedastisten tuottosarjojen häntää voidaan mallittaa normaaleja menetelmiä tarkemmin. Frey et al (1999) ratkaisevat ehdollisen volatiliteetin ja GARCH-mallinnuksen suurimman uskottavuuden (maximum likelihood) estimoinnin perusteella. Tutkimuksessa osoitetaan eri tilastollisten testien perusteella, että saadut residuaalit noudattavat i.i.d. (identically and independently distributed) –muotoista sarjaa jakauman häntien osalta. Frey et al käyttävät jakauman keskiosalle historiallisen simuloinnin perusteella saatua jakaumaosaa ja jakauman hännät on estimoitu mallittamalla aikaisemmin saatuja residuaaleja. Frey et al (1999) toteavat jälkitestauksen (backtesting) perusteella, että VaR-estimaatit ovat parempia kuin tavanomaisten menetelmien tuottamat tulokset.

Tämän tutkielman rakenne on seuraava. Kappaleessa kaksi käsitellään tavallisia Value at Risk –malleja ja niiden heikkouksia äärimmäisten markkinamuutosten kannalta. Tämän jälkeen kappaleessa kolme esitellään extreme value –estimoinnin suorittamiseen tarvittava teoria ja neljännessä luvussa selvitetään tutkielman empiirinen osuus. Luku viisi sisältää johtopäätökset.

2. VALUE AT RISK -MENETELMÄT

2.1. VaR-mallien lähtökohtia

Sijoitussalkun markkinariski aiheutuu markkinahinnoissa, kuten esimerkiksi koroissa ja valuuttakursseissa (tausta- eli riskifaktoreissa) tapahtuvista muutoksista (Beckström 1995). Markkinariski voidaan nähdä sijoitussalkun tulevan arvon todennäköisyysjakaumana, jonka muoto on riippuvainen salkun rakenteesta ja aikahorisontista (Hakuni 1997, 9). Value at Risk –mallien tarkoitus on muodostaa tietylle portfoliolle tulevan arvon vaihtelua kuvaava todennäköisyysjakauma. Yleensä itse jakaumaa ei esitetä, vaan poimitaan siitä valitun riskitason mukainen prosenttipiste VaR-luvuksi. Voidaan esimerkiksi laskea, että päivän VaR-taso 1 %:n todennäköisyydellä on miljoona euroa tai sitä suurempi. Tuolloin tappio ei 99%:n todennäköisyydellä ylitä miljoonaa euroa.

Value at Risk -luku lasketaan rahamääräisenä, mikä antaa johdolle ja omistajille mahdollisuuden arvioida onko otettu riski tietyllä aikavälillä heidän omasta näkökulmastaan siedettävän suuruinen. Lisäksi VaR auttaa tuotto/riski-suhteeltaan parhaimman sijoituskohteen valinnassa (Pritsker, 1997). Näin ollen se mahdollistaa esimerkiksi salkunhoitajan tai sijoitusrahaston saavuttaman tuoton arvioimisen.

VaR-luvun määrittämiseen vaikuttavat valittu mittayksikkö, aikahorisontti ja luottamustaso sekä volatiliteetti (Simons, 1996). Aikahorisontilla tarkoitetaan sitä ajanjaksoa, jolle Value at Risk lasketaan. Aikahorisontti määritetään portfolion instrumenttien likvidiyden perusteella siten, että aikahorisontti on sitä lyhyempi mitä nopeammin instrumentit ovat muutettavissa rahaksi. Näin ollen optimaalisin aikahorisontti vastaa sitä aikaväliä, jolla on mahdollista likvidoida tarkasteltava positio (Jorion 1997, 20). Luottamustason valintaan vaikuttaa yrityksen suhtautuminen riskiin sekä mahdollisten tappioiden sietokyky. Riskinkarttajilla luottamustaso on suurempi kuin riskiin neutraalimmin suhtautuvilla yrityksillä. Yleensä luottamustaso vaihtelee välillä 95-99% (Jorion 1997, 87).

2.2. VaR-mallin muodostaminen

Käytännössä VaR-mallien muodostaminen koostuu kahdesta osasta: sijoitussalkun hinnan mallintamisesta sekä taustatekijöiden vaihtelevuuden estimoinnista.

Hinnoittelumalli kertoo, millainen on taustatekijöiden ja arvopaperin hinnan välinen riippuvuus (Beckström 1995). Portfolion arvopapereiden hinnoittelumalleja ovat esimerkiksi CAP-malli (lineaarinen) ja Black & Scholes optionhinnoittelumalli (epälineaarinen). Ero lineaarisen ja epälineaarisen hinnoittelumallin välillä on oleellinen VaR-metodiikan puitteissa. Lineaarisessa hinnoittelussa oletetaan, että instrumenttien arvot voidaan määrittää riittävän tarkasti lineaarisena funktiona riskifaktoreiden arvoista. VaR-käsitteistössä tätä kutsutaan delta valuation –menetelmäksi. Delta-menetelmässä käytetään Taylorin ensimmäisen asteen sarjakehitelmää, joka tuottaa lineaariapproksimaation arvopaperin hinnan muutoksesta. Täydellisen hinnoittelun menetelmällä (full valuation) lasketaan portfolion arvopaperien arvot niiden oikeiden hinnoittelukaavojen mukaan.

Riskien mittaamiseen vaikuttavat hinnoittelumallien lisäksi arviot siitä, kuinka eri taustatekijät tulevaisuudessa muuttuvat. Tämän vuoksi joudutaan muodostamaan arvio taustatekijöiden yhteistodennäköisyysjakaumasta. Taustatekijöiden todennäköisyysjakaumaa voidaan mallintaa myös historialliseen aineistoon perustuen. Yleensä kuitenkin oletetaan, että taustatekijöiden muutokset ovat peräisin normaalijakaumasta, jonka parametrit ovat vakioita ajassa. Lisäksi oletetaan, että taustatekijöissä ei esiinny autokorrelaatiota. Taustatekijöiden todennäköisyysjakauman estimointiin voidaan käyttää historiatietojen sijaan myös optioiden hintoja. Koska option hinnan yksi tekijä on alla olevan kohde-etuuden keskihajonta, voidaan volatiliteetin estimaatti laskea option hinnan perusteella. Tätä niin sanottua implisiittistä volatiliteettiestimaattia voidaan puolestaan käyttää todennäköisyysjakauman parametrinä. On myös mahdollista, että todennäköisyysjakauma muodostetaan täysin subjektiivisen harkinnan perusteella. Kyseessä on tällöin skenaariolähestymistapa, jossa ei käytetä mitään kvantitatiivista metodia vaan menetelmä perustuu tutkijan esittämään arvaukseen.

2.3. VaR-mallien luokittelua

Perusajatus VaR-malleissa on se, että portfolion instrumenttien hinnat riippuvat muutamasta tausta- eli riskifaktorista. Riskitekijöiden estimointitavat voidaan karkealla tasolla jaotella parametriseen ja ei-parametriseen lähestymistapaan. Toinen tapa jaotella VaR-malleja on lähteä liikkeelle siitä, mistä Value at Risk –mallien rakentaminenkin alkaa eli hinnoittelumallin muodostamisesta. Hinnoittelumalli voidaan toteuttaa joko full valuation –menetelmällä tai delta valuation –menetelmällä. Kolme tavallisinta tapaa laskea Value at Risk –luku ovat delta-normaali-lähestymistapa ja Monte Carlo –simulointi, jotka ovat parametrisia menetelmiä, sekä historiallinen simulointi, joka on ei-parametrinen menetelmä (kts. taulukko 3). Lisäksi ovat olemassa stressitestit eli ns. skenaarioanalyysit, joita on perinteisesti käytetty yhdessä VaR-mallin kanssa arvioimaan erityisen suuria markkinaliikkeitä.

Taulukko 3.

VaR-mallien luokittelu

Estimointimenetelmä	**Hintamuutoksen laskeminen**	
	Delta valuation	Full valuation
Parametrinen	Varianssi-kovarianssi-menetelmä	Monte Carlo –simulointi
Ei parametrinen		Historiallinen simulointi
		Stressitestit

Lähde: Hakuni 1997

2.3.1. Delta-normaali-menetelmä

Delta-normaali-malli perustuu parametriseen lähestymistapaan, jossa hintamuutoksia arvioidaan delta valuation –menetelmällä. Delta-menetelmä olettaa, että pääomatuotot ovat multinormaalisti jakautuneet ja että i.i.d.-oletus on voimassa (Jamshidian & Zhu 1997). Matemaattisesti tämä esitetään muodossa

(1) $(\Delta X1,...,\Delta XN) \sim N(\mu, \sigma)$.

Menetelmässä historiallisesta aineistosta lasketaan sekä taustatekijöiden tuottojen varianssiestimaatit että korrelaatiot eri taustatekijöiden tuottojen välille (Jorion 1997, 186). Mallissa voidaan käyttää myös optiomarkkinadataa implisiittisen volatiliteetin ja korrelaatioiden estimoimiseksi. Instrumenttien tuottojen muutoksia mitataan lineaari- eli delta-approksimaatiolla. VaR-luku lasketaan käyttämällä ennusteita varianssi-kovarianssi-matriisista ja delta-approksimaatioista. Delta-normaali-menetelmä on laskennallisesti helppo ja nopea. Se ei vaadi tietokoneilta suurta laskentakapasiteettia, eikä sen käyttö ole työlästä suurenkaan sijoitussalkun yhteydessä (Pritsker, 1997).

Tarkastellaan esimerkiksi sijoitusinstrumenttia i ja oletetaan sen hinnan P_i olevan riippuvuussuhteessa taustatekijöihinsä $X_1,\dots,X_N$ seuraavasti

(2) $P_i = f_i(X_1,\dots,X_N,u_i)$.

Virhetermille u oletetaan, että u~(0,σ) ja että virhetermi ei korreloi taustatekijöiden X muutosten kanssa. Kuten aiemmin mainittiin, hintamuutokset lasketaan delta valuation –menetelmällä, jolloin yksittäisen sijoitusinstrumentin i hinnanmuutos ΔP_i saadaan yhtälöstä

$$(3)\Delta P_i = \frac{\partial f_i}{\partial X_1}\Delta X_1 + \frac{\partial f_i}{\partial X_2}\Delta X_2 + \dots + \frac{\partial f_i}{\partial X_N}\Delta X_N,$$

joka on Taylorin sarjakehitelmän ensimmäisen asteen termi. Mikäli mukaan otettaisiin sarjakehitelmän toisen asteen termi, olisi kyseessä delta-gamma-menetelmä (Pritsker, 1997). Delta-gamma-menetelmässä pyritään saamaan tarkempi arvio hintamuutoksista ottamalla mahdolliset sijoitusinstrumenttien epälineaarisuudet paremmin huomioon.

Itse Value at Risk –luku lasketaan kaavalla

$$(4) \qquad VaR = W_0(\alpha\sigma\sqrt{t} - \mu t),$$

jossa W_0 = position arvo hetkellä t_0

α = valitun riskitason perusteella määräytyvä kerroin

σ = volatiliteetti

μ = odotettu tuotto

Delta-normaaliin-menetelmään sisältyy useita heikkouksia. Se soveltuu huonosti sellaisen portfolion riskien mittaamiseen, joka sisältää paljon epälineaarisia instrumentteja, kuten esimerkiksi optioita (El Jahel, 1998). Lisäksi on havaittu empiirisesti, että rahoitustuotot eivät noudata normaalijakaumaa (Engle & Rotschild 1992). Jakaumat saattavat olla epäsymmetrisiä ja niiden hännät voivat olla normaalijakauman häntiä paksumpia.

Myös peräkkäiset havainnot ovat usein toisistaan riippuvia. Delta-menetelmän kannalta jakaumien paksuhäntäisyys on kuitenkin suurin ongelma (Danielsson & de Vries 1997a). Malli arvioi riskiä sitä heikommin mitä kauempana odotusarvosta tappio sijaitsee, jolloin mallista käytännössä saatava hyöty jää vähäiseksi.

2.3.2. Historiallinen simulointi

Historiallisessa simuloinnissa sovelletaan full valuation –hinnoittelua. Menetelmässä käytetään tuottojen historiallisia havaintoja ja määritetään aikasarjojen havainnoille niiden nykyarvot. Matemaattisesti malli voidaan esittää muodossa

$$(5)\ V^{t} = \sum_{k=1}^{M} A_k P_k^t, t = 1,\dots,T,$$

jossa V^t on sijoitussalkun arvo hetkellä t ja $P^t{}_k$ on arvopaperin k hinta hetkellä t. Portfolio koostuu kaiken kaikkiaan M:stä eri arvopaperista, joiden kappalemääriä sijoitussalkussa merkitään A_k:lla. Tarkasteluperiodin pituutta kuvataan T:llä. Yllä olevalla kaavalla saadaan laskettua portfolion arvo periodin t eri hetkille. Peräkkäisten hetkien arvojen muutokset puolestaan saadaan kaavasta

$$(6)\ \Delta V^t = V^t - V^{t-1}, t = 2,\dots,T.$$

Tuolloin saadaan T-1 kappaletta muutoksia, joiden perusteella voidaan muodostaa portfolion tulevien arvojen todennäköisyysjakauma. Historiallisen simuloinnin keskeinen oletus on, että menneet tuotot ennustavat parhaiten tulevia tuottoja. Erillistä oletusta todennäköisyysjakaumasta ei tehdä (Simons, 1996). Menetelmästä on myös olemassa sovellus, jossa lähempänä nykyhetkeä olevat arvot saavat suuremman painoarvon (EWMA).

Historiallinen simulointi ei vaadi rajoittavia oletuksia normaalisuudesta eikä hinnoittelumallin lineaarisuudesta. Sillä on kuitenkin omat heikkoutensa. Mallissa oletetaan, että historia toistaa itseään ja mallin parametrit pysyvät näin ollen muuttumattomina. Menetelmässä taustatekijöiden tiheysfunktion muodostaa vain yksi ainoa tuottojen realisaatiopolku (Beckström, 1995). On selvää, ettei mikään sijoitusperiodi kuvaa tyhjentävästi kaikkia mahdollisia markkinamuutoksia.

Oletetaan, että VaR-luku on laskettu historiallisella simuloinnilla vuoden pituiselta ajanjaksolta.[3] Se voi joko yli- tai aliarvioida riskejä riippuen siitä, millaisia havaintoja otos sisältää (Beckström, 1995). Harhaisuus voi johtua esimerkiksi otosperiodiin sisältyvistä positiivisista tulevaisuuden näkymistä, jotka ennusteperiodilla ovat vähemmän todennäköisiä. Tällöin jakaumaotos ei ole harhaton ennuste oikealle tuottojakaumalle. Menetelmä on siis erittäin herkkä valitun otosperiodin suhteen. Vaikka historiallisen simuloinnin otosperiodi sisältäisikin muutamia suuria markkinamuutoksia, tulisi jakauman hännistä diskreettejä (Danielsson & Hartman, 1998). Tällöin VaR-lukujen variansseista tulisi suuria, eivätkä ne enää palvelisi tarkoituksenmukaisella tavalla. Lisäksi historiallinen simulointi ei huomioi sellaisia markkinamuutoksia, joita ei ole vielä empiirisesti havaittu, mutta jotka ovat kuitenkin mahdollisia. Ne saavat esiintymistodennäköisyyden nolla ja jäävät näin ollen kokonaan huomioimatta. Historiallisen simuloinnin käyttö vaatii myös suuret tietokannat menneestä markkina-aineistosta, joiden ylläpito ja hankkiminen voi olla hyvin kallista tai jopa mahdotonta (Simons, 1996).

[3] Baselin komitea on asettanut yhden vuoden otosperiodin vähimmäispituudeksi VaR-lukua historiallisella simuloinnilla laskettaessa (Overview of the Amendment to the Capital Accord to Incorporate Market Risks 1996, 4).

2.3.3. Monte Carlo –simulointi

Monte Carlo –simulointi (stokastinen simulointi) on VaR-malleista kehittynein. Sillä päästään yleensä parempiin tuloksiin kuin muilla VaR-menetelmillä (Pritsker, 1997). Monte Carlo –simuloinnilla pystytään käsittelemään vaativaa matematiikkaa, joten se soveltuu hyvin myös monimutkaisempien rahoitusinstrumenttien tuottojakaumien muodostamiseen (Vose 1996, 11). Monte Carlo –simuloinnin lähestymistapa pohjautuu delta-normaalin-menetelmän kaltaisesti taustatekijöihin, mutta arvopapereiden hinnat lasketaan full valuation –menetelmällä (Jorion 1997, 200-201). Simuloinnissa oletetaan, että arvopaperin i hinta P_i on riippuvainen taustatekijöistä $X_1,\ldots,X_N$ noudattaen funktiota

(7) $P_i = f_i(X_1,\ldots,X_N, u_i)$

Monte Carlo –simuloinnissa taustatekijöille estimoidaan parametrinen yhteistodennäköisyysjakauma. Se muodostetaan joko optioiden hintojen avulla tai taustatekijöiden historiatietojen perusteella käyttämällä niin sanottua bootstrap-tekniikkaa[4]. Saadusta yhteistodennäköisyysjakaumasta poimitaan useita, yleensä tuhansia toisistaan riippumattomia otoksia erityisen satunnaislukugeneraattorin avulla (Pritsker, 1997). Jokaisen otoksen perusteella lasketaan arvopaperille hinta, jolloin saadaan otoksien lukumäärää vastaava määrä arvopapereiden hintoja. Monte Carlo –malli voidaan esittää matemaattisesti

(8) $P_i^k = f_i(X_1^k,\ldots,X_N^k),\ k = 1,\ldots,K,$

jossa otosten määrää merkitään K:lla ja otoksen järjestysnumeroa k:lla. Edellä kuvattu menettely toistetaan jokaiselle portfolion instrumentille M ja yhdistetään saadut tulokset, jolloin saadaan K kappaletta estimaatteja portfolion arvolle. Formaalisti esitettynä menettely on tällöin muotoa

---

[4] Bootstrap-menetelmä on suurten lukujen lakiin perustuva tilastollinen tekniikka, jolla aineistosta voidaan muodostaa uusia jakaumia, jotka voivat tuoda uusia ominaisuuksia esiin aineistosta.

$$(9)\, V^k = \sum_{j=1}^{M} A_j P_j^k,\ k = 1,\ldots,K.$$

Saaduista estimaateista muodostetaan histogrammi, jota voidaan käyttää jakaumana portfolion tulevalle arvolle. Jos portfolion tulevien arvojen sijasta halutaan käsitellä portfolion arvon muutoksia, on sekin mahdollista. Tämä tapahtuu yksinkertaisesti vähentämällä portfolion tulevasta arvosta sen nykyinen markkina-arvo, jota nyt merkitään V:llä eli formaalisti

$$(10)\, \Delta V^k = V^k - V,\ k = 1,\ldots K.$$

Kaavan (10) perusteella lasketuista muutoksista voidaan jälleen muodostaa histogrammi, joka kuvaa portfolion tulevan arvon ennustettuja muutoksia. Tätä histogrammia käytetään VaR-luvun muodostamisen perustana. Monte Carlo –simuloinnin pohjana voidaan myös käyttää jotain parametristä todennäköisyysjakaumaa, kuten esimerkiksi Pareto-jakaumaa.

Monte Carlo –lähestymistavan heikkous on sen tietokoneelta vaatima suuri laskentakapasiteetti (Beckström, 1995). Simuloinnissa saatetaan joutua laskemaan esimerkiksi 10000 hintapolkua salkun tuotteille ja määrittämään salkun arvo 10000 kertaa, jolloin saatetaan joutua suorittamaan yhteensä esimerkiksi 100 miljoonaa laskutoimitusta. Toinen Monte Carlo –menetelmän heikkous on sen monimutkaisuudesta johtuva malliriski (Jorion 1997, 200). Malli sallii kaikki jakaumat sekä epälineaarisuuden arvopapereissa.

Monte Carlo –simuloinnin käyttö voi johtaa riskien aliarvioimiseen, jos VaR-luku muodostetaan normaalijakauman pohjalta. Empiirisesti on havaittu, että rahoitustuottojen muodostamat jakaumat ovat paksuhäntäisempiä kuin normaalijakauma (Engle et al, 1992).

2.3.4. Stressitestit

VaR-laskennassa on tavallista, että käytetyt aikasarjat sisältävät vähän poikkeuksellisen suuria markkinamuutoksia. Tällöin syntyy riski, että suurten markkinaromahdusten, kuten "musta maanantai" (lokakuussa 1987) tai Aasian kriisi, esiintymistodennäköisyys aliarvioidaan (Mori & Ohsawa & Shimizu, 1996). Siitä huolimatta, että edellä mainitun tyyppiset poikkeukselliset tapahtumat otettaisiin huomioon, ne eivät vaikuttaisi VaR-lukuun juuri lainkaan, koska niiden esiintymistodennäköisyys on ainakin normaalijakaumaan perustuvassa päättelyssä lähellä nollaa.

Stressitestauksessa katastrofiskenaariota muodostamalla pyritään selvittämään portfolion arvon muutoksia. Stressitestit kuuluvat ei-parametriseen lähestymistapaan ja niissä käytetään full valuation –menetelmää hintamuutosten laskemiseksi. Tässä niin sanotussa skenaariolähestymistavassa taustatekijöiden muutoksia ei mallinneta tilastollisen estimoinnin avulla vaan niille muodostetaan arvio subjektiivisen harkinnan perusteella (Kupiec, 1998). Taustatekijöille määritetään erilaisia vaihtoehtoisia skenaarioita, joiden avulla tarkastellaan portfolion markkinariskiä. Mikäli skenaarioihin liitetään mukaan subjektiiviset todennäköisyydet voidaan laskea myös VaR-luvut.

Käytännössä stressitestit toteutetaan siten, että portfolion jokaisen instrumentin tuotto arvioidaan valitun skenaarion mukaan, jonka jälkeen lasketaan portfoliolle uusi arvo. Matemaattisesti tämä voidaan esittää muodossa

$$(11)\ R_{p,s} = \sum_{i=1}^{N} w_{i,t} R_{i,s},$$

jossa

R = tuotto

p = portfolio

w = painokerroin

s = skenaario.

Skenaariolähestymistapaan liittyy kuitenkin eräitä heikkouksia. Ensinnäkin, on hankalaa määritellä skenaarioita siten, että ne kuvaisivat tarkasti harvinaisia, mutta joskus toteutuvia markkinaliikkeitä. Perinteinen menettelytapa, joka asettaa skenaariot täysin ad hoc –pohjalta ei ole kelvollinen. Ongelmallista on päättää, rajoitetaanko riskinottoa sellaisen skenaarion avulla, joka perustuu hyvin suureen markkinaromahdukseen. Toinen ongelma liittyy siihen, että skenaarioissa rahassa ilmaistuun tappiolukuun ei yleensä liitetä sen toteutumisen todennäköisyyttä. Jos skenaarioille asetetaan todennäköisyydet ovat ne ainoastaan subjektiivisia "arvauksia" tulevaisuudesta, eikä niitä siksi voida pitää luotettavina. Kaikista suurin puute stressitesteissä on se, että ne käsittelevät portfolion eri instrumenttien välisiä korrelaatioita huonosti (Jorion 1997, 198). Korrelaatiot ovat kuitenkin olennainen osa portfolion riskiä. Pitäisi olla mahdollista ottaa huomioon, kuinka muut markkinat reagoivat yhden markkinan romahtaessa. Jos skenaarion mukaan osakemarkkinat romahtavat kaksikymmentä prosenttia yhden päivän aikana, tulisi voida ottaa huomioon se, kuinka suuria muutoksia tapahtuu esimerkiksi johdannaismarkkinoilla.

3. EXTREME VALUE -ESTIMOINTI

3.1. Estimoinnin lähtökohdat

Äärimmäisiksi havainnoiksi sanotaan havaintoja, jotka osuvat jakauman ääripäihin. Äärimmäisten muutosten mallittamiseen on olemassa karkeasti sanottuna kaksi eri lähestymistapaa. Vanhempaa tyyppiä edustavat ns. maksimilohkomallit (block maxima). Nämä ovat malleja, joissa poikkeavat havainnot on poimittu useista suurista otoksista. Maksimilohkomallilla voidaan esimerkiksi arvioida muutoksia vuoden ajalle keräämällä tietystä instrumentista muutostiedot päivän sisältä tai useiden päivien ajalta. (McNeil, 1998.)

Toinen ja samalla modernimpi lähestymistapa ovat POT-mallit (Peaks-Over-Threshold). Nämä mallit sopivat kaikille poikkeaville havainnoille, jotka ylittävät tietyn kynnyksen. POT-mallit ovat parempia käytännön sovelluksissa kuin "blokkimallit", sillä ne toimivat paremmin aineistolla, joka sisältää vähän poikkeavia havaintoja. POT-mallit voidaan edelleen jakaa kahteen eri suuntaukseen, puoliparametrisiin malleihin ja täysin parametrisiin malleihin. Puoliparametriset mallit pohjautuvat Hillin estimaattori –tyyppisiin estimaattoreihin[5], kun taas parametriset mallit perustuvat yleistettyyn Pareto-jakaumaan. Molemmat tavat ovat empiirisesti toimivia ja teoreettisesti hyvin perusteltuja. (McNeil, 1998.)

Rahoitusteoriassa lähtökohtana on usein oletus siitä, että peräkkäiset tuotot ovat identtisesti ja itsenäisesti jakautuneita (i.i.d). Tämä tarkoittaa sitä, että tuotot ovat riippumattomia ajasta ja että ne ovat jakautuneet identtisesti ajan suhteen. Lisäksi tuottojen oletetaan usein noudattavan normaalijakaumaa. Käytännössä on havaittu, että jakaumat voivat olla normaalijakaumaa paksuhäntäisempiä, huipukkaampia (leptokurtosis-ilmiö) ja vinoja (Lim & Lye & Martin & Martin 1998). Jakaumien paksuhäntäisyys on seurausta suurista markkinaliikkeistä, joita esiintyy epäsäännöllisesti. Normaalijakaumaa oletuksenaan käyttävä VaR-malli aliarvioi suurten tappioiden mahdollisuutta siksi, että käytännössä tuotot muodostavat paksuhäntäisen jakauman (Danielsson et al 1997a).

Yleensä VaR-analyysit pohjautuvat itsenäiseen ja identtiseen jakaumaan. Tuolloin on olemassa hyvin suuri mahdollisuus virheeseen, koska usein käytännössä peräkkäiset tuotot korreloivat keskenään. Danielsson ja de Vries (1997b) lähtevät tutkimuksessaan liikkeelle ottamatta huomioon edellä mainittua ristiriitaa teorian ja käytännön välillä. Ristiriidan poistamiseen on olemassa kolme eri vaihtoehtoa. Yksinkertaisin keino on soveltaa ajatusta siitä, että kunkin periodin maksimiarvot ovat yleensä vähemmän klusteroituneita (clustered) verrattuna alkuperäiseen aineistoon, josta maksimiarvot on poimittu (Dowd, 1999). Klusteroituminen vähenee sitä

---

[5] Katso tarkemmin kpl 3.4. Danielsson ja de Vries-menetelmän soveltaminen ε_t-muuttujiin.

enemmän mitä kauemmin maksimien poiminnasta on aikaa. Kuitenkaan ei ole olemassa täyttä varmuutta siitä, kuinka pitkiä näiden "blokkiperiodien" pitäisi olla. Toinen tapa perustuu teoriaan, jonka mukaan estimaattoreita muutetaan siten, että ne ottavat klusteroinnin huomioon (Embrechts & Kluppelberg & Mikosch, 1997).

Tässä tutkielmassa sovelletaan tapaa, jossa pyritään minimoimaan virhe teorian ja käytännön välillä muuttamalla käytettävä aineisto sellaiseen muotoon, jossa Danielssonin ja de Vriesin (1997b) tekemät oletukset käytettävän aineiston ominaisuuksista pitävät tarkemmin paikkansa. Havaintoaineistoa muokataan Alexanderin (1996) suosittelemalla menetelmällä, jossa aineistolle ensin suoritetaan pääkomponenttianalyysi ja tämän jälkeen GARCH-proseduurilla ratkaistaan standardoidut residuaalit. Menetelmässä on useita etuja: korrelaatio- ja kovarianssimatriisin laskemisessa ei ole merkittäviä laskennallisia ongelmia, vaikka dimensio olisi suuri. Toinen etu on, että tarvittaessa aineiston dimensiota voidaan pienentää poistamalla merkitykseltään vähäisimmät pääkomponentit. Pääkomponenttianalyysin tuloksena muuttujien korrelaatiot ja volatiliteetit ovat sellaisia, joihin GARCH-mallitus onnistuu usein helposti. Menettelyn tuloksena saadut korrelaatio- ja varianssimatriisit ovat positiivisesti definiittejä.

Pääkomponenttianalyysissa syntyy todennäköisesti jonkin verran virhettä, koska pääkomponenttianalyysin oletus on, että muuttujien aikasarjat ovat stationaarisia. Kuitenkin rahoituksen aikasarjat ovat vain harvoin stationaarisia, sillä ainakin varianssi ja odotusarvo muuttuvat ajassa. Tämä menettelytapa on kuitenkin tällä hetkellä yleisesti hyväksytty ja sitä käytetään johtavien riskienhallinnan ohjelmistovalmistajien tuotteissa. Tässä tutkielmassa käytetään GARCH-proseduurin perusteella saatuja standardoituja residuaaleja extreme value –estimoinnin pohjana.

3.2. Pääkomponenttianalyysi

Tutkielmassa käytetty menetelmä on yleistys Englen ja Leen (1993) faktori-GARCH-mallista (Alexander & Chibumba, 1995). Erona on kuitenkin, että tuottoja ei malliteta yhdellä riskifaktorilla, vaan mallissa käytetään useita ortogonaalisia faktoreita. Ortogonaalinen faktorimalli on muotoa

$$(12)\quad Y_{i,t} = \mu_i + \sum_{j=1}^{m} \varphi_{ij} P_{j,t} + \varepsilon_{i,t},$$

jossa $i=1,\dots,k$, $t=1,\dots,T$ ja $m \le k$. Täten mallissa on yhteensä k tuottosarjaa ja jokaisessa tuottosarjassa on T havaintoa. Jatkossa nämä havainnot esitetään Txk matriisina Y.

Ortogonaaliset faktorit $P_1,\dots,P_m$ saadaan pääkomponenttianalyysin perusteella. Aluksi jokaisen sarakkeen aineisto standardoidaan vähentämällä jokaisesta havainnosta sarakkeen otoskeskiarvo ja jakamalla tämä otoksen keskihajonnalla. Tällä tavoin saadaan alkuperäisestä matriisista Y normalisoitu tuottomatriisi X. Tässä W:llä merkitään ominaisarvojen matriisia X'X, joka on järjestetty ominaisarvojen koon mukaan. Tuolloin pääkomponenttien muodostama matriisi on

$$(13)\quad P=XW.$$

Mikäli $W=(w_{ij})$, kun $i,j=1, \dots,k$, niin tuolloin $\lambda_1>\lambda_2> \dots>\lambda_k$, kun λ_i:llä merkitään matriisin W sarakkeen i ominaisarvovektorin ominaisarvoa. Koska $X'XW=W\Lambda$ (Λ=diagonaalimatriisi X'X:n ominaisarvoista), niin

$$(14)\quad P'P=W'X'XW=W'W\Lambda=\Lambda$$

Pääkomponenttien matriisin P sarakkeet eivät siis korreloi keskenään.

Huomionarvoista on, että $X=PW'$, eli tarkemmin

(15) $x_i = w_{i1}P_1 + w_{i2}P2 + w_{i3}P_3 + \ldots + w_{ik}P_k$

Tässä X_i ja P_i vastaavat X:n ja P:n sarakkeita i. Termit w_{ij} ($j=1, \ldots ,k$) ovat kunkin faktorin painokertoimia. Tarpeen vaatiessa dimensiota voidaan pienentää ottamalla huomioon vain m ensimmäistä pääkomponenttia, jolloin pääkomponenttien matriisi P on Txm-matriisi ($m \leq k$) ja faktoreiden painojen matriisi on kxm. Koska $X_i = (Y_i - \mu_i)/\sigma_i$ johtaa pääkomponenttiesitys linearisoituun faktorimalliin (12), jossa faktoripainot on normalisoitu $\varphi_{ij} = w_{ij}\sigma_i$. Mallin virhetermi kertoo virheen, joka syntyy kun pääkomponenteista otetaan huomioon vain m kpl. Tulkintojen etsiminen pääkomponenteille ei tässä tutkielmassa ole tarpeellista, sillä ne eivät ole tutkimusongelman kannalta olennaisia.

Mallin (12) perusteella saadaan ratkaistua varianssit

(16) $V = \phi D \phi' + V_\varepsilon$,

jossa $V = V(Y)$, $\phi = (\varphi_{ij})$, $D = \mathrm{diag}(V(P_1), \ldots V(P_m))$ ja V_ε on virheiden varianssi-kovarianssi matriisi. Tässä ei lasketa residuaalien varianssia, vaan ehdollinen varianssi-kovarianssi-matriisi V_t alkuperäisestä systeemistä käyttämällä ehdollista diagonaali varianssi-kovarianssi-matriisia $D_t = \mathrm{diag}(V_t(P_1), \ldots V_t(P_m)$. $V_t(P_i)$:llä merkitään pääkomponentin P_i ehdollista GARCH-varianssia. Tuolloin

(17) $V_t \approx \phi D_t \phi'$.

Kaavalla (17) saadaan estimoitua V_t. Faktoripainomatriisi ϕ saadaan pääkomponenttianalyysista ja diagonaalimatriisin D_t GARCH-estimaatit saadaan tavallisilla menetelmillä.

Tässä tutkielmassa ei ole tarvetta vähentää muuttujien lukumäärää, sillä tarkoitus oli ainoastaan poistaa multikollineaarisuus. Lisäksi faktoreita on melko vähän, joten kaikkien faktoreiden käsittely on perusteltua. Mikäli pääkomponenttianalyysilla tavoitellaan muuttujien lukumäärän vähentämistä, voidaan se tehdä kolmen vaihtoehtoisen kriteerin perusteella. Nämä kriteerit eivät ole tärkeitä tämän tutkielman kannalta, joten niitä ei lähemmin tarkastella[6].

3.3. GARCH-esitys pääkomponenteille

Pääkomponenttianalyysin jälkeen muodostetaan pääkomponenteille GARCH-esitys. GARCH-mallituksen perusteella saadaan laskettua standardoidut residuaalit, jotka noudattavat i.i.d.-oletusta. Pääkomponenttianalyysin ja GARCH-esityksen perusteella aineisto saadaan sellaiseen muotoon, jossa sitä voidaan hyväksyttävästi käyttää Danielssonin ja de Vriesin esityksen pohjana. Standardoidut residuaalit ovat tällöin muotoa

(18) $\varepsilon_t = r_t/\sigma \quad \text{i.i.d.},$

jossa r_t on toteutunut tuotto.

Eri rahoitusinstrumenttien tuotot näyttävät usein olevan sarjakorreloituneita. Tämä tarkoittaa käytännössä sitä, että suuria tuottoja seuraavat suuret tuotot ja negatiivisia tuottoja seuraavat usein negatiiviset tuotot. Yksi tapa ottaa volatiilisuuden sarjakorrelaatio huomioon on ARCH-mallien (Autoregressive Conditional Heteroscedasticy) käyttäminen. Tässä tutkielmassa pääkomponentteihin pyrittiin sovittamaan joko ARCH-, GARCH(p,q)- tai EGARCH-malli. Monimuuttuja-GARCH-mallit jätettiin tutkielman ulkopuolelle vakiintuneen menettelytavan seuraamisen vuoksi.

Engle (1982) määrittelee diskreetin stokastisen prosessin (ε_t) ARCH-mallissa muodossa

---

[6] Tarkemmin pääkomponenttien vähentämiskriteereistä esimerkiksi Koutsoyiannis (1977, 433-434).

(19) $\varepsilon_t = z_t h_t^{1/2}$, z_t i.i.d ja $var(z_t)=1$

jossa h_t on ajassa muuttuva positiivinen mitattavissa oleva funktio hetken t-1 informaatiosta. ε_t:n ehdollinen varianssi on yhtä suuri kuin h_t ja voi muuttua ajassa. Ajasta riippuva parametrisoitu muoto h_t:lle on

$$(20)\ h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2,\ \alpha_0 > 0\, ja\ \alpha_i > 0$$

Rajoiteparametrit ovat positiivisia, koska ehdollinen varianssi h_t on aina positiivinen. Varianssi h_t ilmaistaan menneiden ARCH(q)-mallien asteiden q neliöityjen arvojen lineaarisena kombinaationa. Usein asteena q käytetään varsin suurta arvoa, mutta sen kasvattaminen muuttaa helposti parametrin α_i negatiiviseksi. Parametri α_i mittaa shokkien pysyvyyttä.

Mikäli muutokset volatiliteetissa ovat jatkuvia, sopii Bollerslevin (1986) yleistetty ARCH-malli eli GARCH (Generalized Autoregressive Conditional Heteroskedasticity) paremmin volatiliteetin estimointiin. GARCH-malli yleisessä muodossa h_t on formaalisti

$$(21)\ h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-1}^2 + \sum_{i=1}^{p} \beta_i h_{t-1}\ ,\alpha_0 > 0, \alpha_i > 0\ ja\ \beta_i > 0.$$

Ehdollinen varianssi h_t riippuu lineaarisesti autoregressiivisen AR(q)-prosessin neliöityjen arvojen ja itse ehdollisen varianssin liukuvan keskiarvon MA(p)-prosessin arvosta. Erilaisten shokkien pysyvyys saadaan parametrien α_i ja β_i summana.

Nelsonin (1990) kehittämä EGARCH-malli (Exponential GARCH) ottaa paremmin huomioon shokkien epäsymmetriset vaikutukset. Yleensä tätä epäsymmetrisyyttä esiintyy osakkeiden hinnoissa ja negatiiviset muutokset lisäävät volatiliteettia enemmän kuin positiiviset muutokset. EGARCH-mallissa vipuvaikutus otetaan huomioon ehdollisessa varianssissa edellisten ε_t:n epäsymmetrisenä funktiona. EGARCH-mallin matemaattinen spesifiointi on seuraava:

$$(22)\quad \log h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i (\phi z_{t-i} + \gamma [\,| z_{t-i} | - E | z_{t-i} |\,]) + \sum_{i=1}^{p} \beta_i \log h_{t-i}.$$

Mallissa varianssi riippuu sekä viimeaikaisten muutosten suuruudesta että niiden suunnasta.

Sovittamalla pääkomponentteihin niihin sopiva GARCH-malli saadaan ratkaistua ehdollinen varianssi, jota käyttämällä voidaan ratkaista standardoidut residuaalit ε_t. Etukäteen vanhojen tutkimusten perusteella voidaan tehdä oletus, että tavallinen GARCH(1,1)-malli sopii useisiin pääkomponentteihin melko hyvin (kts. Ahlstedt 1998 ja taulukko 4).

Taulukko 4.
Volatiliteetin estimointimenetelmien tutkimustuloksia.

Tutkijat	Menetelmät	Estimointi-periodi	Aineisto	Johtopäätökset
C.O. Alexander & C.T. Leigh (1997)	Yhtäläisesti painotettu lka (BIS-menetelmä), EWMA, GARCH (1,1)	1, 5, 10 ja 25 päivää.	FTSE 100, S&P 500, CAC 40, DAX, NIKKEI 225 päätöskurssit ja näitä indeksejä vastaavat dollarikurssit Testaus tapahtui 1.1.96-6.10.96 päivittäisellä aineistolla.	EWMA-malli toimi huonoiten lukuunottamatta USA:n osakemarkkinoita. GARCH (1,1) ja yhtäläisesti painotettu lka toimivat paremmin; ne sijoittuivat vihreälle alueelle BIS:n jälkitestauksessa.
Jacob Boudoukh & Matthew Richardson & Robert F. Whitelaw (1997)	Historiallinen keskihajonta, Riskmetrics^{TM}-malli (λ=0.94), GARCH (1,1), MDE (Multivariate Density Estimation)	Estimoinnin vyörytysperiodina käytettiin 150 päivää.	Kolmen kuukauden USA:n valtion obligaation päivittäinen korko. Testaus suoritettiin vuosien 1983-1992 aineistolla.	Tulosten yleistys hankalaa, koska testattu vain 150 päivän periodilla, tällöin GARCH (1,1)-malli toimii epäluotettavasti. Malleista MDE toimi parhaiten ja GARCH (1,1) huonoiten.
Philippe Jorion (1995)	Implisiittinen keskihajonta (ISD), lka(20), GARCH (1,1).	1 päivä.	Valuuttafutuurit Saksan (1/85-2/92), Japanin (7/86-2/92) ja Sveitsin (3/85-2/92) valuutoille.	GARCH-malli oli vähemmän harhainen kuin 20 päivän perusteella laskettu liukuva keskiarvo, mutta selitysaste oli vastaavasti hieman heikompi. ISD-malli tuotti tarkat estimaatit, mutta sen tuottamaa ennustetta volatiliteetista ei välttämättä kannata soveltaa seuraavalle päivälle, vaan koko option jäljellä olevalle ajalle.
Kenneth D. West, Dongchul Cho (1995)	GARCH (1,1) IGARCH (1,1), homoskedastinen malli, kaksi AR-mallia ja epäparametrinen Gaussian kernel-malli.	1, 12 ja 24 viikkoa.	Viikottaiset USD:n vaihtokurssit Kanadan, Ranskan, Saksan, Japanin ja Iso-Britannian valuuttojen suhteen Testaus tapahtui vuosien 1973-1989 aineistolla.	GARCH-mallit toimivat viikon periodilla parempia kuin muut mallit, mutta pidemmällä ajanjaksolla yhdenkään mallin toimivuus ei yksiselitteisesti ollut muita parempi. Yksikään malleista ei toiminut hyvin, kun testattiin ennustustehokkuutta.

3.4. Danielsson ja de Vries-menetelmän soveltaminen ε_t-muuttujiin

Danielssonin ja de Vriesin menetelmän päämääränä on ottaa paremmin huomioon jakaumien empiirisesti todettavissa oleva paksuhäntäisyys. Lisäksi tarkoituksena heidän menetelmässään on, että malli toimisi hyvin myös ennen empiirisesti havaitsemattomien tuottojen (tappioiden) tapauksissa.

Jakauman paksuhäntäisyys voidaan määritellä säännöllisen vaihtelun perusteella (Danielsson & de Vries 1997b). Jakauma funktio F(x) vaihtelee normaalisti äärettömyydessä häntäindeksillä α [7], jos

$$(23)\quad \lim_{t\to\infty}\frac{1-F(tx)}{1-F(t)}=x^{-\alpha},\alpha>0,x>0.$$

Säännöllisen vaihtelun ominaisuudella tarkoitetaan sitä, että α:aa suuremmat jakauman X ei-ehdolliset tunnusluvut (momentit) ovat rajoittamattomia. Tämän vuoksi säännöllisesti vaihtelevien jakaumien joukko on paksuhäntäinen. Käytännössä säännöllinen vaihtelu on ainoa rajoittava oletus jakauman X paksuhäntäisyyden analysoimisessa. Teoriaosuudessa keskitytään jakauman ylempään häntään, mutta tappiohännän osalta analyysi on täysin analoginen.

Yleinen parametrinen muoto paksuhäntäiselle jakaumalle on

$$(24) f(x)=a\alpha x^{-\alpha-1}+ab(\alpha+\beta)x^{-\alpha-\beta-1}+o(x^{-\alpha-\beta-1}).$$

Yllä olevan funktion viimeinen termi $o(x^{-\alpha-\beta-1})$ on tiheysfunktion virhetermi. Funktion erikoistapauksina saadaan mm. Studentin t-jakauma ja Fréchet-jakauma. Menetelmän lähestymistapa perustuu ns. Hill-estimaattorin yleistykseen ja valintakriteerinä käytetään estimaattorin asymptoottista

[7] Häntäindeksi α kuvaa jakaumahännän muodon.

keskineliövirhettä[8]. Tarkemmin sanottuna lähestymistapa perustuu estimaattorien tarkentuvuuteen konvergenssin todennäköisyysmielessä, joka on jakaumakonvergenssia voimakkaampi ominaisuus.

Paksuhäntäisen jakauman yleisen muodon (24) parametri α saadaan ratkaistua häntäindeksin käänteisluvun eli k-momentin estimaattorista (25). Funktion (25) erikoistapaus, jossa $k = 1$ ja $w_1(s_n) = u_1(s_n)$ tunnetaan Hillin estimaattorina (Hill 1975). Hillin estimaattori on eräs funktion (25) momentti estimaattoreista. Koko k-momentin estimaattoreiden joukon tarkastelu on mielekästä siksi, että ne antavat tiettyjen olosuhteiden vallitessa harhattomammat tulokset keskineliövirheelle. Optimaalisen kynnysarvon s_n laskemiseksi joudutaan käyttämään vähintään kahta momenttiestimaattoria. Häntäindeksin käänteisluvun estimaattori on muotoa

$$(25)\quad w_k(s_n) \equiv 1/\hat{\alpha} = \frac{u_k(s_n)}{k u_{k-1}(s_n)},$$

jossa $k = 1,2,\ldots$ on kokonaisluku, $u_0(s_n) = 1$ ja

$$(26)\quad u_k(s_n) \equiv \frac{1}{M}\sum_{i=1}^{M}(\log\frac{X_i}{s_n})^k \Big| X_i > s_n,$$

Funktiossa (26) X_i on rajan (kynnysarvon) s_n ylittävä havainto ja M näiden havaintojen lukumäärä. Funktio (19) on k:nnen asteen empiirisen jakauman logaritminen momentti otoksesta $X_1, \ldots, X_n$, joka on muodostettu poimimalla n kertaa (bootstrap-menetelmällä) i.i.d jakaumasta $F(x)$.

Niille satunnaismuuttujille, jotka toteuttavat funktion (24), voidaan k:nnen momentin estimaattorin $w_k(s_n)$ asymptoottinen virhe määritellä seuraavasti:

---

[8] Keskineliövirhe (MSE=Mean Squared Error) on estimaattorin tuottaman estimaatin ja todellisen arvon erotuksen neliö. MSE(θ^) = E(θ^ - θ)^2. Mikäli harha on nolla, niin MSE(θ^) = Var(θ^).

$$(27)\ E\left[w_k(s_n)-\frac{1}{\alpha}\right]=-\frac{b\beta\alpha^{k-2}}{(\alpha+\beta)^k}s_n^{-\beta}+o(s_n^{-\beta}).$$

Mikäli kynnysarvo s_n valitaan siten, että todennäköisyys $Ms_n^{\alpha}/an \rightarrow 1$, kun $n \rightarrow \infty$ ja jos yhtälö (24) pätee niin

$$(28)\ Var\left[w_k(s_n)-\frac{1}{\alpha}\right]=\frac{\kappa(k)}{M\alpha^2}+o\left(\frac{1}{M}\right),$$

jossa

$$\kappa(k)=\frac{(2k)!}{(k!)^2}+\frac{(2k-2)!}{((k-1)!)^2}-2\frac{(2k-1)!}{k!(k-1)}.$$

$\kappa(k)$:n arvo kasvaa räjähdysmäisesti k:n kasvaessa. Esimerkki funktio $\kappa(k)$:n käyttäytymisestä on taulukossa 5.

Taulukko 5.

$\kappa(k)$:n käyttäytyminen

K	1	2	3	4	5	6	7
$\kappa(k)$	1	2	6	20	70	252	924

Yhdessä (27) ja (28) merkitsevät sitä, että w_k:t ovat konsistentteja[9] estimaattoreita tietyille s_n:lle, jos otokset X_i ovat i.i.d ja toteuttavat funktion (24). w_k estimaattorit ovat konsistentteja myös eri tyyppisten riippuvuuksien vallitessa.

[9] Konsistentilla estimaattorilla tarkoitetaan estimaattorin tarkentuvuutta otoskoon kasvaessa. Kun otoskoko lähenee ääretöntä, niin estimaattorin tuottama estimaatti vastaa tarkalleen parametrin todellista arvoa ($\lim_{n\to\infty} E(\hat{\theta}_n)=\theta$). Tällöin sekä harha että varianssi ovat yhtä suuria kuin nolla ($\lim_{n\to\infty} Var(\hat{\theta}_n)=0$).

3.4.1. Optimaalisen kynnysarvon määrittely

Extreme value –teorian soveltamisessa keskeistä on määritellä kynnysarvo, jota pienemmät arvot (tappiot) katsotaan äärimmäisiksi havainnoiksi. Näin ollen tarkastellaan koko jakaumahännän kokoa, koska se vaikuttaa itse VaR-estimaattiin välillisesti häntäindeksin α kautta. Jakaumahännän koon ja siten varsinaisen kynnysarvon (s_n) määrittäminen on kompromissi harhan ja varianssin suhteen. Jos häntää kasvatetaan, kynnysarvo siirtyy kohti jakauman keskiosaa ja poikkeavia havaintoja saadaan enemmän. Tämä lisää häntäestimaattorin tarkkuutta ja vähentää sen varianssia, mutta toisaalta harha kasvaa, koska jakauman keskiosalle annetaan suurempi paino kuin äärimmäisille havainnoille. Vastaavasti, jos kynnysarvoa siirrettäisiin jakauman keskiosasta poispäin, häntä lyhenisi ja häntäindeksin estimaattorin harhaisuus vähenisi, mutta vastaavasti sen varianssi kasvaisi vähentyneiden havaintojen takia.

Danielssonin ja de Vriesin ratkaisu edellä mainittuun ongelmaan on määrittää kynnysarvo, joka minimoi tappiohännän (asymptoottisen) keskineliövirheen (MSE) tuottamalla tulokseksi kompromissin, jossa harha ja varianssi kumoavat toisensa. Asymptoottinen keskineliövirhe (AMSE) häntäindeksiestimaattorille $w_k(s_n)$ saadaan kaavojen (27) ja (28) perusteella.

$$(29)\ AMSE(w_k(s_n)) \approx \frac{\kappa(k)}{a\alpha^2}\frac{s_n^{\alpha}}{n} + \frac{b^2\beta^2\alpha^{2k-4}}{(\alpha+\beta)^{2k}} s_n^{-2\beta}.$$

Kaavan oikeanpuoleisista termeistä se, joka on asymptoottisesti dominoiva määritetään sen asteen mukaan, jossa $s_n \to \infty$, kun $n \to \infty$. Lisäksi on olemassa tietty s_n, joka asymptoottisesti tasapainottaa kaavan (29) oikeanpuoleiset termit. s_n saadaan ottamalla osittaisderivaatta asymptoottisesta keskineliövirheestä s_n:n suhteen ja asettamalla se nollaksi. Optimissa toisen asteen osittaisderivaatan tulee olla positiivinen.

Mikäli todennäköisyys $(Ms_n^{\alpha})/an \to 1$ ja yhtälö (24) pätee, niin asymptoottista kynnysarvoa $\bar{s}_n$ minimoiva AMSE on muotoa

$$(30)\ \bar{s}_n(w_k) = \left[\frac{2ab^2\beta^3\alpha^{2k-3}}{(\alpha+\beta)^{2k}\kappa(k)}\right]^{\frac{1}{\alpha+2\beta}} n^{\frac{1}{\alpha+2\beta}}.$$

Tällöin asymptoottisesti minimi $w_k(s_n)$:n keskineliövirhe voidaan esittää muodossa

$$(31)\ \overline{MSE}[w_k(\bar{s}_n)] = \frac{\kappa(k)}{a\alpha}\left[\frac{1}{\alpha}+\frac{1}{2\beta}\right]\left[\frac{2ab^2\beta^3\alpha^{2k-3}}{(\alpha+\beta)^{2k}\kappa(k)}\right]^{\frac{\alpha}{2\beta+\alpha}} n^{-\frac{2\beta}{2\beta+\alpha}}.$$

Asymptoottisen keskineliövirheen minimoimiseksi täytyy harhan ja varianssin kumota toisensa. Poikkeavien arvojen asymptoottinen lukumäärä $\bar{m}_n$, jossa harha- ja varianssiosat ovat tasapainossa, saadaan laskettua lisäämällä kaavaan (30) oletus, jonka mukaan $(Ms_n^{\alpha})/an \to 1$ (Danielsson & de Vries 1997b). Näin saadaan

$$(32)\ \overline{m}_n(w_k) = a\left[\frac{2ab^2\beta^3\alpha^{2k-3}}{(\alpha+\beta)^{2k}\kappa(k)}\right]^{-\frac{\alpha}{2\beta+\alpha}} n^{\frac{2\beta}{2\beta+\alpha}}.$$

Kaavojen (29-31) perusteella selviää se, että s_n:n lähestyessä ääretöntä hitaammin kuin $n^{1/(2\beta+\alpha)}$ MSE:n harhaosa dominoi varianssia. Mikäli s_n lähestyy ääretöntä nopeammin kuin $n^{1/(2\beta+\alpha)}$ niin vastaavasti MSE:n varianssiosa dominoi harhaosaa. AMSE:n minimoimiseksi s_n:n tulee lähestyä ääretöntä täsmälleen nopeudella $n^{1/(2\beta+\alpha)}$. Estimaattoreille $w_k(s_n)$ ainoat AMSE-kriteerin perusteella relevantit estimaattorit ovat w_1 ja w_2[10].

[10] Todistus Danielsson et al 1997b.

3.4.2. Kynnysarvon s_n estimointi

Kynnysarvon s_n estimointi on ongelmallista ja aiheesta ei ole juuri lainkaan julkaistua tietoa. Ennen Danielssonin ja de Vriesin (1997b) tutkimusta ainoastaan Hall (1990) on käsitellyt aihetta. Hallin tutkimuksessa on kuitenkin tiettyjä puutteita, joiden vuoksi sitä ei esimerkiksi voi soveltaa kaikkien jakaumien kanssa[11].

Danielsson et al käyttävät seuraavaa suuretta s_n:n estimoinnin pohjana

(33) $z(s_n) \equiv w_2(s_n) - w_1(s_n),\ \lim_{n\to\infty} z(\bar{s}_n) = 0.$

Kynnysarvo s_n saadaan minimoimalla kaavaa

(34) $\frac{1}{R}\sum_{r}^{R}\left[z_r(\hat{s}_{n1}, n_1)\right]^2,$

jossa

$Z_r(s_n) \equiv w_2(s_n) - w_1(s_n).$

r viittaa r:nnen bootstrap-otoksen perusteella laskettuun suureeseen ja bootstrap-otosten lukumäärää merkitään R:llä. Yleisen paksuhäntäisen jakaumafunktion (24) parametri β saadaan ratkaistua estimoimalla se β:n ja α:n konsistentin estimaattorin

(35) $\frac{\hat{\beta}}{\hat{\alpha}} = \frac{\log \hat{m}_{n1}(z)}{2\log n_1 - 2\log \hat{m}_{n1}(z)}$

perusteella, jossa n_1 on bootstrap-otoksen koko ja m_{n1} on rajaa s_n järjestystunnusluvuilla vastaava raja.

---

[11] Lisää Hallin (1990) tutkimuksen ongelmista Danielsson et al (1997 b).

3.5. Äärimmäisten tuottojen ennustaminen

Keskeistä äärimmäisten tuottojen ennustamisessa on määrittää sellainen estimaattori, joka toimii sekä otokseen kuuluvien että sen ulkopuolisten tuottojen yhteydessä. Estimaatin $1/\alpha$ perusteella määräytyvät ominaisuudet sekä kvantiilien että niiden todennäköisyyksien estimaattoreille. Otoksen ulkopuoliset (P,Q) estimaatit saadaan ns. *Bahadur-Kiefer*-tuloksen perusteella.

Oletetaan todennäköisyys rajat (p ja t), joille $p < 1/n < t$ ja jossa n on otoskoko. Todennäköisyyksiin liittyvät kvantiilit ovat x_p ja x_t, missä x_p: $1-F(x_p) = p$ ja x_t: $1-F(x_t) = t$. Koska p<1, niin oletettavasti $x_p > \max\{X_1,\ldots,X_n\}$. Kvantiili x_p voidaan estimoida ekstrapoloimalla empiiristä jakaumafunktiota $F_n(x)$ hyödyntämällä säännöllisen vaihtelun ominaisuutta (Danielsson & de Vries 1997b). Soveltamalla funktiota

$$(36)\quad F(x) = 1 - \alpha x^{-\alpha}\left[1 + bx^{-\beta} + o(x^{-\beta})\right],\ \beta > 0,\ kun\, x \to \infty,$$

jonka muunnos on funktio (24), saadaan

$$(37)\ \frac{t}{p} = \left(\frac{x_p}{x_t}\right)^{\alpha} \frac{1 + bx_t^{-\beta} + o(x_t^{-\beta})}{1 + bx_p^{-\beta} + o(x_p^{-\beta})}$$

ja ehdolla β>0 saadaan vastaavasti

$$(38)\ x_p = x_t\left(\frac{t}{p}\right)^{\frac{1}{\alpha}}\left(\frac{1 + bx_p^{-\beta} + o(x_p^{-\beta})}{1 + \beta x_t^{-\beta} + o(x_t^{-\beta})}\right)^{\frac{1}{\alpha}}.$$

Estimaattori x_p:lle saadaan hylkäämällä kaavan (38) korkeamman asteen termi ja korvaamalla t m/n:llä sekä x_t laskevalla (m+1):ä vastaavalla järjestystunnusluvulla ja korvaamalla $1/\alpha$ estimaattorilla w_k

$$(39)\hat{x}_p = X_{(m+1)}\left(\frac{m}{np}\right)^{w_k}.$$

Vaihtoehtoisesti on mahdollista hyödyntää w_k:n tulkintaa kynnykseksi siten, että todennäköisyys t korvataan satunnaismuuttujalla M/n kun x_t on vakio s_n:ssä. Tällöin estimaattori on muotoa

$$(40)\hat{x}_p = x_t\left(\frac{M}{np}\right)^{w_k}.$$

Estimaattori x_p voidaan osoittaa asymptoottisesti normaalisti jakautuneeksi.

Käänteisessä tapauksessa voidaan yhtälö kirjoittaa muokkaamalla kaavaa (38), jolloin saadaan

$$(41)p = t\left(\frac{x_t}{x_p}\right)^{\alpha}\left(\frac{1+bx_p^{-\beta}+o(x_p^{-\beta})}{1+\beta x_t^{-\beta}+o(x_t^{-\beta})}\right),$$

josta edelleen

$$(42)\hat{p} = \frac{M}{n}\left(\frac{x_t}{x_p}\right)^{\hat{\alpha}}.$$

Poikkeavien havaintojen estimaattori p on asymptoottisesti normaalisti jakautunut. Määritettyjen kvantiilien asymptoottiset jakaumat ja todennäköisyydet poikkeavat kertoimella $-\alpha^2$. Tämä on Bahadur-Kiefer-tyyppinen tulos otoksen ulkopuolisille (P,Q)-kombinaatioille. Otoksen ulkopuolisten havaintojen tapauksessa empiirinen jakauma korvataan (p,^x) tai (^p,x_p) –käyrillä. (Danielsson & de Vries 1997b.)

3.6. Danielsson ja de Vries –menetelmän testaus

Yleensä VaR-mallien tarkkuutta arvioidaan ns. jälkitestaus (backtesting) menettelyllä. Tuolloin riskinmittaamiseen käytetyn mallin ennustamia tuloksia verrataan toteutuneisiin tuloksiin. Tässä tutkielmassa ei suoriteta varsinaista jälkitestausta, koska tarkoitus on selvittää itse extreme value –estimoinnin toimivuutta ja verrata sitä normaalijakaumaan perustuviin VaR-malleihin. Vertaamalla EV-tekniikalla laskettua kynnysarvoa normaalijakauman yhden ja viiden prosentin pisteisiin, voidaan tehdä päätelmät menetelmien suorituskyvyistä. Menettely on riittävä, koska käytetty portfolio on kuvitteellinen siitä huolimatta, että se perustuu todelliseen aineistoon. Valitulla testaustavalla vältetään myös mahdollinen virhe riskifaktoreiden hinnoittelussa.

4. EXTREME VALUE –ESTIMOINNIN TOTEUTUS JA TESTAUS

4.1. Käytetty aineisto ja sen käsittely

Aineisto valittiin siten, että se edustaisi mahdollisimman hyvin suomalaisen liikepankin portfoliota. Yleensä pankkien sijoitusstrategiana on hankkia tuottoa kohtuullisella riskillä. Riskin vähentämiseksi ne hajauttavat portfolioitaan. Usein salkku hajautetaan alueellisesti eri markkinoille ja kullakin alueella eri sijoitusinstrumentteihin.

Tässä tutkielmassa portfolioon on valittu neljä eri markkina-aluetta, jotka ovat suomalaisten pankkien kannalta merkittävimmät. Valitut markkina-alueet ovat Suomi, Ruotsi, Saksa ja Yhdysvallat. Portfolioon valittiin jokaiselta alueelta aineistoa korko- ja osakemarkkinoilta. Osakemarkkinoita edustavat kunkin alueen tärkeimmät pörssi-indeksit. Indeksien käyttöä yksittäisten osakkeiden sijaan voidaan perustella sillä, että koska pankit joka tapauksessa hajauttavat sijoituksiaan, niin hajautetun portfolion tuotto lähenee indeksin kehitystä. Havaintoaineisto kattaa jokaisen muuttujan osalta saman ajanjakson, joka alkaa 16.2.1994 ja päättyy 2.7.1999. Havaintoaineiston lähteenä on ETLA:n tietokanta.

Kaikilta markkina-alueilta valittiin portfolioon viiden ja kymmenen vuoden korot. Muita portfoliossa mukana olevia korkoja ovat Ruotsin kuuden kuukauden korko ja kolmen kuukauden ECU/Euro-korko. Valitsemalla sekä lyhyitä, että pitkiä korkoja saadaaan portfolioon konveksisuutta. Osakeindeksejä portfoliossa on mukana siten, että Suomesta on mukana HEX-indeksi, Ruotsista OMX-indeksi, Saksasta DAX-indeksi ja Yhdysvalloista Dow Jones-indeksi. Lisäksi mukana ovat valuuttakurssit, jotta aineisto saadaan muunnettua yhteen valuuttaan. Luonnollisesti valuutoista myös muodostuu osa portfolion riski- ja tuottomahdollisuuksista. Havaintoaineistoa muokattiin muuntamalla aluksi valuuttamuuttujat euromääräisiksi. Tämän jälkeen kaikista osakeindekseistä ja valuutoista otettiin log-prosenttimuutokset, jotta ne saataisiin tuottomuotoon. Korot ovat valmiiksi tuottomuodossa, joten niille ei ollut tarpeellista tehdä vastaavia muutoksia. Lopuksi kaikille muuttujille suoritettiin ei-ehdollinen standardointi, jonka jälkeen muuttujille voidaan suorittaa pääkomponenttianalyysi.

Pääkomponenttianalyysi toteutettiin standardia menettelytapaa noudattaen. Pääkomponentit laskettiin Estiman Rats-ohjelmalle tarjoamaa proseduuria apuna käyttäen (http://www.estima.com/procindx.htm).

4.2. GARCH-estimointi pääkomponenteille

Ratkaistuihin pääkomponentteihin sovitettiin niihin parhaiten sopivat GARCH-mallit. Aikaisempien tutkimusten perusteella (Ahlstedt 1998 ja taulukko 4) voitiin olettaa, että tavallinen GARCH(1,1)-malli toimisi melko hyvin pääkomponenttien volatiliteetin estimoinnissa. Oletus piti paikkansa, mutta kaikkiin pääkomponentteihin GARCH(1,1)-mallia ei kuitenkaan saatu sovitettua. Tavoite oli, että tuloksissa ei esiintyisi sarjakorrelaatiota.

GARCH-sovitteet estimoitiin soveltamalla Estiman www-sivuillaan tarjoamaa ohjelmaa GARCH-sovitteiden laskemiseen[12]. Parametrit estimoitiin suurimman uskottavuuden menetelmällä siten, että SIMPLEX-algoritmillä estimoitiin BHHH-algoritmille parametrien alkuarvot.

Useimpiin pääkomponentteihin sopi parhaiten GARCH(1,1)-malli. Esimerkiksi ensimmäisen pääkomponentin kaikki testisuureet ovat kunnossa, mikä indikoi, että ongelmia GARCH-mallin oletusten suhteen ei ole. GARCH(1,1) mallittaa parhaiten myös toisen, kolmannen, kahdeksannen, yhdeksännen, kymmenennen, yhdennentoista, kahdennentoista, kolmannentoista, neljännentoista ja viidennentoista pääkomponentin ei-ehdollista volatiliteettia. Pieniä ongelmia esiintyy joissakin pääkomponenteissa. Esimerkiksi pääkomponentissa yhdeksän esiintyy lievää sarjakorrelaatiota mikäli viiverakenne on lyhyt, mutta pidemmällä viiverakenteella autokorrelaatiota ei voida havaita. Useimpien pääkomponenttien GARCH-proseduurin jälkeiset standardoidut residuaalit eivät noudata normaalijakaumaa ja esimerkiksi pääkomponentin kolme Jarque-Bera testisuure saa arvon 17344.5572 ja on siten erittäin merkitsevä. Tämä indikoi voimakasta poikkeamaa normaalijakaumasta. Aikaisempien tutkimusten perusteella (Engle & Rotshild (1992), Friedman & Vandersteel (1982), Hsieh (1989), Boothe & Glassman (1987)) voidaan olettaa, että poikkeama normaalijakaumasta johtuu jakaumien paksuhäntäisyydestä. Olennaista on se, että poikkeama normaalijakaumasta vahvistaa extreme value –analyysin jatkamisen tarpeellisuuden. Toisaalta ongelmana voidaan pitää käytettyjen testisuureiden luotettavuutta, sillä ne kaikki perustuvat oletukseen normaalijakaumasta.

Muutamien pääkomponenttien GARCH-estimoinnissa on suurempia ongelmia. Esimerkiksi pääkomponentissa viisi GARCH(1,1)-malli ei toiminut riittävän hyvin, koska muuttujat eivät olleet merkitseviä. EGARCH-mallilla päästiin kuitenkin jo hieman parempiin tuloksiin. Pääkomponenteissa kuusi ja seitsemän esiintyi erilaisia ongelmia kaikilla kokeilluilla malleilla ja

[12]Ohjelma on ladattavissa osoitteessa http://www.estima.com/procindx.htm.

viiverakenteilla. Ljung-Box-testin tuloksista voidaan nähdä, että residuaaleissa esiintyi sarjakorrelaatiota. Periaatteessa olisi ollut mahdollista lisätä viiverakenteiden lukumäärää, mutta koska sarjakorrelaatiota esiintyy myös silloin, kun viiveitä on 24, on todennäköistä, että pidemmätkään viiverakenteet eivät olisi merkittävästi parantaneet tilannetta. Toinen vaihtoehto olisi ollut käyttää monimuuttuja-GARCH-malleja tai esimerkiksi FIGARCH-malleja. Nämä tekniikat vaativat kuitenkin niin paljon laskentatehoa, että niiden hyödyntäminen käytännössä on ainakin tällä hetkellä mahdotonta. Lisäksi tässä tutkielmassa pyritään seuraamaan mahdollisimman tarkasti standardin mukaista menettelytapaa (Alexander 1999). Pääkomponenteille käytetyt mallit ja niiden kertoimet ovat taulukossa 6 ja Jarque-Bera-testin sekä Ljung-Box-testien tulokset taulukossa 7. Tulokset on raportoitu kokonaisuudessaan liitteessä 1.

Taulukko 6.
Pääkomponenttien GARCH-estimoinnin tulokset 16.2.1994 -2.7.1999. T-testisuureet suluissa.

	Malli	C	P1	Q1	P2	Q2	P3	Q3
PK1	GARCH(1,1)	0.0005	0.8739	0.0824	-	-	-	-
	ARMA(1,0)	(4.21)	(43.74)	(6.78)				
PK2	GARCH(1,1)	0.0632	0.8579	0.1165	-	-	-	-
	ARMA(2,0)	(3.76)	(50.35)	7.94				
PK3	GARCH(1,1)	0.1045	0.8191	0.0820	-	-	-	-
	ARMA(0,0)	(5.74)	(36.38)	(8.68)				
PK4	GARCH(3,3)	0.0131	1.0295	0.1009	-1.1486	-0.0309	0.6207	0.1181
	ARMA(5,0)	(4.31)	(12.92)	(4.93)	(-16.43)	(-2.19)	(9.68)	(5.64)
PK5	EGARCH(2,2)	-0.1286	-0.1230	0.2039	0.8793	0.1856	-	-
	ARMA(4,0)	(-7.14)	(-21.24)	(6.16)	(185.34)	(6.12)		
PK6	GARCH(1,1)	0.0365	0.8418	0.1080	-	-	-	-
	ARMA(1,0)	(4.06)	(43.67)	(9.33)				
PK7	GARCH(1,1)	0.0162	0.8436	0.1345	-	-	-	-
	ARMA(0,0)	(2.88)	(47.03)	(9.34)				
PK8	GARCH(1,1)	0.0167	0.8724	0.092	-	-	-	-
	ARMA(1,0)	(3.35)	(41.38)	(6.29)				
PK9	GARCH(1,1)	0.005	0.9479	0.039	-	-	-	-
	ARMA(1,0)	(2.41)	(87.22)	(5.10)				
PK10	GARCH(1,1)	0.0002	0.8395	0.1224	-	-	-	-
	ARMA(3,0)	(3.63)	(43.57)	(8.48)				
PK11	GARCH(1,1)	0.0005	0.6013	0.079	-	-	-	-
	ARMA(3,0)	(3.03)	(4.86)	(3.13)				
PK12	GARCH(1,1)	0.0000	0.9513	0.038	-	-	-	-
	ARMA(2,0)	(2.88)	(95.80)	(5.15)				
PK13	GARCH(1,1)	0.0000	0.9683	0.0245	-	-	-	-
	ARMA(2,0)	(3.79)	(174.0)	(5.11)				
PK14	GARCH(1,1)	0.0000	0.8994	0.062	-	-	-	-
	ARMA(2,0)	(5.31)	(70.72)	(7.52)				
PK15	GARCH(1,1)	0.0000	0.9262	0.0604	-	-	-	-
	ARMA(3,0)	(3.58)	(89.72)	(6.29)				
PK16	GARCH(1,1)	2.5190	0.5966	0.5305	-	-	-	-
	ARMA(4,0)	e-005	(31.21)	(15.63)				
		(6.50)						
PK17	GARCH(1,1)	2.2495	0.8565	0.1166	-	-	-	-
	ARMA(3,0)	e-005	(62.03)	(7.87)				
		6.64						

Taulukosta 6 nähdään, että GARCH(1,1)-malli sopii hyvin useimpiin pääkomponentteihin. Muutamissa pääkomponenteissa esiintyi suurempia ongelmia, joita yritettiin ratkaista käyttämällä pitkiä viiverakenteita ja eri GARCH-malleja. Mikäli kunnolla toimivaa mallia ei löytynyt, käytettiin tavallista GARCH(1,1)-mallia.

Taulukko 7.
Ljung-Box- ja Jarque-Bera-testien tulokset pääkomponenttien GARCH-estimaateille 16.2.1994 – 2.7.1999.

	Jarque-Bera-testi	Ljung-Box-testisuureet		
		LB(8)	LB(16)	LB(24)
PK1	131.2400	6.94	19.48	29.94
PK2	98.9613	9.56	12.42	33.48
PK3	17344.5572	3.52	16.69	22.78
PK4	81.5835	9.53	21.59	24.12
PK5	3000.9178	7.03	30.47	43.16
PK6	67.4497	361.04	772.21	1156.90
PK7	5.5213	2654.48	5418.93	8057.94
PK8	24.8584	12.71	25.47	33.30
PK9	39.0034	13.16	20.05	27.86
PK10	998.9293	8.86	18.33	24.25
PK11	1093.6546	8.54	18.77	31.27
PK12	150.8616	7.42	18.24	20.62
PK13	799.3307	3.13	9.81	19.37
PK14	1092.0833	9.07	16.76	27.67
PK15	296.3008	8.54	12.29	25.09
PK16	10178.5342	12.24	31.55	42.28
PK17	628.86	11.19	21.20	29.36

Taulukosta 7 voidaan nähdä, että useimpien pääkomponenttien GARCH-estimaattien Jarque-Bera-testisuure on huomattavan suuri, mikä indikoi voimakasta poikkeamaa normaalijakaumasta. Suurimmat ongelmat sarjakorrelaation suhteen olivat pääkomponenteissa kuusi ja seitsemän, joiden Ljung-Box-testisuureet saavat hyvin korkean arvon. Muissa pääkomponenteissa ei esiintynyt näin suuria ongelmia.

4.3. Parametrien α ja β estimointi

Extreme value –estimoinnin pohjana käytettävä jakaumaestimointi suoritettiin soveltamalla Danielsson ja de Vriesin (1997b) menetelmää (estimaattorien lausekkeet on esitetty kaavoissa (25), (26), (34) ja (35)). Toisen asteen approksimaatio suoritettiin siksi, että menetelmää haluttiin noudattaa tarkoin. Toisen asteen approksimaatiossa ratkaistaan sekä parametri α että parametri β. Menetelmää olisi mahdollista yksinkertaistaa ensimmäisen asteen approksimaatiolla, jolloin estimoitaisiin vain parametri α. Tarkoitus oli kuitenkin pyrkiä sekä tuloksissa että noudatetuissa menetelmissä mahdollisimman suureen tarkkuuteen, joten molemmat jakaumaparametrit estimoitiin.

Parametreja α ja β estimoitaessa bootstrap-otoksen koon valinta on olennaista. Mikäli otoskoko on liian suuri, on bootstrap-otosten poiminta turhaa, koska suuret otokset lähenevät ominaisuuksiltaan alkuperäisten havaintojen muodostamaa jakaumaa. Toisaalta, poimimalla liian pieniä bootstrap-otoksia ei välttämättä saada oikeaa kuvaa havaintoaineiston käyttäytymisestä. Danielsson et al (1997b) esittävät, että optimaalinen bootstrap-otos voidaan valita kaavalla

(43) $N_1=n^{1-\varepsilon}, 0<\varepsilon<½$

Kaavan mukaan laskettu optimaalinen bootstrap-otoksen koko tässä tutkielmassa käytetylle aineistolle on 224, kun $\varepsilon = 0.75$.

Estimoitaessa parametrit α ja β on sitä parempi, mitä enemmän bootstrap-otoksia poimitaan, koska tuolloin konvergoituminen on varmempaa. Käytännössä joudutaan kuitenkin tekemään kompromissi laskenta-ajan ja parametrien tarkkuuden suhteen. Parametrit estimoitiin ottamalla sata bootstrap-otosta. Estimointi toteutettiin Juhani Raatikaisen kirjoittamalla rutiinilla (Raatikainen 2000). Taulukossa 8 on esitetty estimoitujen parametrien α ja β arvot.

Taulukko 8.

Extreme value –estimoinnilla saadut parametrit α ja β.

Pääkomponentti	α	β
PK1	-92.9883	-97.09424
PK2	33.49406	15.69327
PK3	-23.22506	-172.90198
PK4	-12.97281	-218.89301
PK5	8.26264	6.8406
PK6	-68.50808	-109.88457
PK7	68.93991	33.97785
PK8	315.70957	653.88176
PK9	-352.47958	-487.74685
PK10	-92.05408	-120.91403
PK11	-39.07255	-105.82577
PK12	-44.10874	-99.5467
PK13	107.08904	90.96527
PK14	-50.515	-101.6337
PK15	-44.60573	-129.5255
PK16	21.87715	21.88988
PK17	-161.74042	-320.70273

Parametrien α ja β arvojen tulkinta on ongelmallista, koska kyseessä ovat eri pääkomponenteille lasketut arvot. Tämän vuoksi tuloksia ei voida yleistää suoraan eri riskifaktoreille. Muutenkaan parametrien estimointi on ei ole aivan yksinkertaista, sillä niiden funktioiden käyttäytyminen ei ole säännönmukaista. Ehdottomasti oikean arvon löytäminen on vaikeaa.

Tässä kohdassa havaitaan ongelma myös Danielsson & de Vries –teorian oletuksissa. Danielsson et al (1997b) johtivat teoriansa oletukset sillä perusteella, että havaintojen lukumäärä lähenee ääretöntä. Käytännössä oletus on mahdoton, koska rahoituksen aikasarjoja on olemassa hyvin rajallinen määrä. Tässä tutkielmassa eri riskifaktoreista on havaintoja 1347 ja koska käytetty malli ei mahdollisesti toimi luotettavasti pienillä otoksilla, tulee parametrien α ja β arvoihin suhtautua varauksella. Mikäli aikomus olisi käyttää menetelmää tuotantokäytössä, tulisi aineistoa olla käytettävissä mahdollisimman paljon ja silloinkin sen riittävyys tulisi varmistaa.

Parametrien α ja β estimoinnin jälkeen voitiin laskea kynnysarvo ja ylitysten lukumäärä jokaiselle pääkomponentille. Kynnysarvo on luku, jonka ylittävät havainnot lasketaan äärimmäisiksi havainnoiksi. Kynnysarvot ja ylitysten lukumäärä ovat taulukossa 9.

Taulukko 9.

Pääkomponenteille lasketut kynnysarvot ja ylitysten lukumäärä.

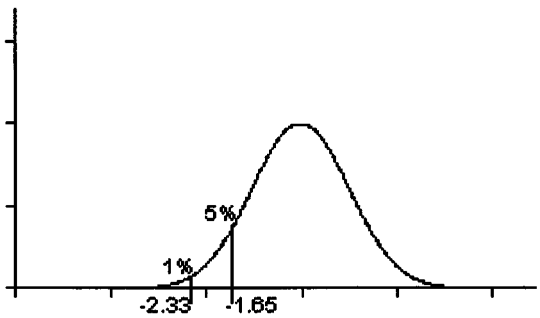
Pääkomponentti	Kynnysarvo	Ylitysten lukumäärä
PK1	-1.94961	729
PK2	-2.26648	250
PK3	-1.61206	907
PK4	-1.93419	807
PK5	-2.08684	510
PK6	-1.67538	1142
PK7	-2.67929	269
PK8	-1.54194	1433
PK9	-1.80318	919
PK10	-1.79151	909
PK11	-1.41671	1672
PK12	-1.54319	1508
PK13	-2.06277	532
PK14	-1.54899	1348
PK15	-1.33204	1870
PK16	-1.98388	630
PK17	-1.53624	1348

Ylitysten lukumäärästä voidaan selvästi nähdä, että todennäköisyysjakauman paksuhäntäisyydestä on selviä viitteitä. Ylitysten lukumäärä ei kuitenkaan ole kovin suuri, kun otetaan huomioon se, että aineiston lukumääräksi bootstrap-otosten poimimisen jälkeen tulee 22400.

Lasketun kynnysarvon perusteella voidaan päätellä olisiko extreme value –estimointi tuottanut Value at Risk –analyysissa luotettavammat tulokset, verrattuna normaalijakaumaan pohjautuviin malleihin. Normaalijakaumassa 1%:n luottamustasoa vastaa luku 2.33. Luonnollisesti luku tappiohännän puolella on –2.33. Vastaavasti 5%:n luottamustasoa vastaavat luvut 1.65 ja –1.65. Mikäli pääkomponentin kynnysarvo on suurempi kuin tavallisen normaalijakauman tappiohännän 5%:n piste, niin tavallinen normaalijakaumaan perustuva VaR-malli aliarvioi riskiä sekä yhden että viiden prosentin luottamustasoilla. Vastaavasti jos kynnysarvo on 1%:a ja 5%:a vastaavien lukujen välissä, aliarvioi 1%:n VaR-luku riskiä. Edelleen, jos kynnysarvo on pienempi kuin 1%:a vastaava arvo –2.33, niin normaalijakaumaan perustuva Value at Risk –analyysi antaisi oikeat tulokset sekä viiden että yhden prosentin luottamustasoilla. Kuvio 2 selventää asiaa.

Kuvio 2.

Normaalijakauman 1%:n ja 5%:n pisteet.



Taulukossa 10 on esitetty ne tapaukset, joissa normaalijakaumaan pohjautuva VaR-malli olisi käytetyllä aineistolla aliarvioinut tappion todennäköisyyttä.

Taulukko 10.

Normaalijakaumaan perustuvan VaR-mallin virheet käytetyllä aineistolla 5%:n ja 1%:n luottamustasolla. Tilannetta, jossa tavallinen VaR-malli olisi aliarvioinut tappiota merkitään x:llä.

Pääkomponentti	Virhe 5%:n tasolla	Virhe 1%:n tasolla
PK1		X
PK2		X
PK3	X	X
PK4		X
PK5		X
PK6		X
PK7		
PK8	X	X
PK9		X
PK10		X
PK11	X	X
PK12	X	X
PK13		X
PK14	X	X
PK15	X	X
PK16		X
PK17	X	X

Taulukosta 10 voidaan nähdä selvästi, että suurimmassa osassa pääkomponenteista 1%:n luottamustasolla riski aliarvioitaisiin, jos käytössä olisi normaalijakaumaan perustuva VaR-malli. Sen sijaan 5%:n luottamustasolla extreme value –estimoinnin hyöty ei olisi yhtä suurta. Tulos osoittaa, että riskienhallinnan kannalta normaalijakaumaan perustuvan Value at Risk –mallin käyttö etenkään 1%:n luottamustasolla ei ole luotettavaa.

Kynnysarvojen lisäksi voidaan laskea kunkin pääkomponentin kohdalta sen ylitysten lukumäärän suhde bootstrap-otosten jälkeisen aineiston kokonaismäärään. Tämän prosenttiluvun tulisi kuulua samaan alueeseen jakaumassa kuin kynnysarvo. Taulukossa 11 esitetään ylitysten prosenttiosuudet suhteessa kokonaisotokseen. Merkillä x tarkoitetaan taulukossa sitä, että prosenttiosuus tukee kynnysarvojen perusteella tehtyjä päätelmiä.

Taulukko 11.

Ylitysten lukumäärien prosenttiosuus koko aineistosta. X-merkki tarkoittaa, että prosenttiosuus tukee kynnysarvojen perusteella tehtyjä päätelmiä.

Pääkomponentti	Ylitysten lukumäärän osuus koko aineistosta	Prosenttiosuus tukee kynnysarvojen tuloksia
PK1	3,25%	X
PK2	1,12%	X
PK3	4,05%	
PK4	3,60%	X
PK5	2,28%	X
PK6	5,10%	
PK7	1,20%	
PK8	6,40%	X
PK9	4,10%	X
PK10	4,06%	X
PK11	7,46%	X
PK12	6,73%	X
PK13	2,38%	X
PK14	6,02%	X
PK15	8,35%	X
PK16	2,81%	X
PK17	6,02%	X

Taulukoiden 10 ja 11 perusteella näyttää selvältä, että extreme value –estimoinnilla voidaan parantaa VaR-analyysin luotettavuutta. Täytyy kuitenkin yhä muistaa menetelmän rajoitukset, koska tässä tutkielmassa käytetty aineisto on ollut melko pieni.

5. JOHTOPÄÄTÖKSET

Perinteisen Value at Risk –analyysin pohjana on usein normaalijakauma tai oletus siitä, että historialliset tuotot ennustavat tulevia tuottoja. Käytännössä nämä oletukset eivät aina pidä paikkaansa. Todellisuudessa rahoituksen aikasarjat muodostavat yleensä paksuhäntäisen todennäköisyysjakauman, joka voi olla huipukas ja vino. Jakauman oikean muodon huomioiminen on erittäin tärkeää, sillä väärän jakaumaoletuksen käyttäminen VaR-analyysissä johtaa riskien aliarvioimiseen.

Tässä tutkielmassa tavoite oli selvittää, voitaisiinko niin sanotulla extreme value –estimoinnilla ottaa normaalijakaumaan perustuvia VaR-malleja paremmin huomioon suuria markkinamuutoksia. Ongelmaa tarkasteltiin Danielssonin ja de Vriesin (1997b) kehittämän menetelmän pohjalta, jossa muodostetaan yhdistetty todennäköisyysjakauma. Yhdistetyssä todennäköisyysjakaumassa jakauman keskiosa tulee esimerkiksi historiallisesta simuloinnista ja jakauman hännät extreme value –estimoinnista. Estimoitujen parametrien perusteella tehtiin tulkinnat extreme value –estimoinnin tarpeellisuudesta tutkielman aineistolla.

Danielsson ja de Vries (1997b) tutkimus pohjautuu oletukseen instrumenttien identtisestä ja itsenäisestä jakautumisesta (i.i.d.). Useissa empiirisissä tutkimuksissa on kuitenkin todettu, että tuotot eivät käytännössä noudata tätä oletusta. Teorian ja empiiristen tulosten ristiriidasta seuraa, että tuloksiin syntyy virhettä. Danielsson et al (1997b) eivät huomioi tätä virhettä. Tässä tutkielmassa tämä virhe pyritään välttämään suorittamalla ensin riskifaktoreille pääkomponenttianalyysi ja sovittamalla ratkaistuihin pääkomponentteihin parhaiten sopivat GARCH-mallit. Näin saadut jäännökset noudattavat i.i.d. –oletusta ja niitä voidaan hyväksyttävästi käyttää extreme value –estimoinnin aineistona. Tutkielman empiirisessä osuudessa voitiin GARCH-mallituksen jälkeen nähdä, että riskifaktorit eivät noudattaneet normaalijakaumaa. Tämä vahvisti extreme value –analyysin tarpeellisuuden.

Eri pääkomponenteille laskettiin omat kynnysarvonsa, joita verrattiin normaalijakauman 1%:n ja 5%:n pisteisiin. Käytetyllä aineistolla näyttää siltä, että 1%:n luottamustasolla normaalijakaumaan perustuvat VaR-mallit selvästi aliarvioivat tappioita, mutta 5%:n luottamustasolla extreme value –analyysi ei ole enää kovin paljon parempi. Tulokset ovat linjassa Daniessonin ja de Vriesin (1997b) tulosten kanssa. On kuitenkin otettava huomioon, että tutkielman aineiston määrä ei täytä alkuperäisen teorian oletusta aineiston määrän äärettömyydestä. Tämä on voinut johtaa epätarkkuuteen etenkin parametrien α ja β estimoinnissa.

Tutkielmassa sovelletun menettelyn toimivuus on kiistatonta. Menetelmän soveltaminen riskienhallinnan välineenä vaatii kuitenkin vielä kehittämistä. Tässä muodossa menetelmä vaatii liian monta työvaihetta ja aineistoa tulisi olla käytettävissä huomattavasti enemmän. Käytännössä tulosten tulisi olla saatavissa nopeasti ja helposti, jotta niistä olisi hyötyä riskienhallinnan kannalta. Lisäksi GARCH-sovitteiden liittäminen pääkomponentteihin on hieman ongelmallista. GARCH-sovitteiden ongelmat olisi mahdollisesti ratkaistavissa monimuuttuja GARCH-malleilla. Nämä mallit ovat kuitenkin niin monimutkaisia, että niiden soveltaminen käytännössä on hankalaa.

Perusmuodossaan Danielsson & de Vries (1997b) –menetelmä on lupaava, mutta vielä kehittämistä vaativa tekniikka parantaa Value at Risk –analyysin luotettavuutta. Jatkossa kannattaisikin selvittää olisiko menetelmää mahdollista muokata siten, että tulokset olisivat helpommin saatavissa. Lisäksi olisi tarpeellista saada lisätietoa menetelmän käyttäytymisestä pienten otosten yhteydessä.

LÄHTEET

Alexander, C.O. & Chibumba, A.M. (1996) Multivariate Orthogonal Factor GARCH. School of Mathematical Sciences. University of Sussex.

Alexander, C.O. & Leigh, C.T. (1997) On the Covariance Matrices Used in Value at Risk Models. The Journal of Derivatives. Spring 1997.

Basle Committee on Banking supervision (1996) Overview of the Amendment to the Capital Accord to Incorporate Market Risk. Basle.

Beckström, Rod (1995) Value at Risk; Theoretical Foundations. Capital Market Strategies vol. 7 November.

Beder, Tanya (1995) VaR: Seductive but dangerous. Financial Analysts Journal September/Oktober.

Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. Journal of Econometrics, Vol. 31.

Boothe, P. & Glassman, D. (1987) The Statistical Distribution of Exchange Rates: Empirical Evidence and Economic Implications. Journal of International Economics, 22.

Boudoukh, Jacob & Richardson, Matthew & Whitelaw, Robert F. (1997) Investigation of a Class of Volatility Estimators. The Journal of Derivatives, Spring 1997.

Danielsson, Jon & de Vries, Casper G. (1997a) Value-at-Risk and Extreme Returns. Institute of Economic studies at University of Iceland. Working Paper Series September 1997.

Danielsson, Jon & de Vries, Casper G. (1997b) Tail Index and Quantile Estimation with Very High Frequency Data. Journal of Empirical Finance No 4.

Danielsson, Jon & Hartman, Philipp (1998) The Cost of Conservatism: Extreme Returns, Value-at-Risk, and the Basle "Multiplication Factor". Risk Magazine vol 11 No 1, Risk Publications Ltd.

Dowd, Kevin (1999) Financial Engineering News. http://fenews.com 26.11.1999.

Duffie, Darrel & Pan, Jun (1997) An Overview of Value at Risk. The Journal of Derivatives, vol. 3, No 4.

El Jahel, Lina (1998) Value at Risk. Sveriges Riksbank. Quarterly Review, vol. 2/1998.

Embrechts, P & Kluppelberg, C. & Mikosch, T. (1997) Modelling Extremal Events for Insurance and Finance.

Engle, R.E. (1982) Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation. Econometrica, Vol 50.

Engle, F. & Rothschild, M. (1992) Arch Models in Finance. Journal of Econometrics, 52.

Frey, Rüdiger & McNeil, Alexander J. (1999) Estimation of Tail-Related Risk Measures for Heteroscedastic Financial Time Series: an Extreme Value Approach.

Friedman, D. & Vandersteel, S. (1982) Short Run Fluctuations in Foreign Exchange Rates: Evidence from Data 1973-79. Journal of International Economics, 13.

Estiman www-sivu. http://www.estima.com/procindx.htm, 10.01.2000.

Hakuni, Mika (1997) Value-at-risk –mallit – Lyhyt katsaus tilastollisiin perusteisiin. Suomen Pankin työpaperi No 2/97, Helsinki.

Hsieh, D (1989) Modelling Heteroscedasticity in Daily Foreign-Exchange Rates. Journal of Business & Economics, 7.

Jamshidian, Farshid (1997) Scenario Simulation; Theory and Methodology. Finance and Stochastics, vol 1 No 1 January 1997.

Jorion, Philippe (1995) Predicting Volatility in the Foreign Exchange Market. The Journal of Finance. Vol L , No 2. June 1995

Jorion, Philippe (1997) Value-at-risk. Irwin professional publishing, Chicago.

Koutsoyiannis, Anna. (1977) Theory of Econometrics. The MacMillan Press Ltd, London.

Kupiec, Paul H. (1998) Stress testing in a Value at Risk framework. The Journal of Derivatives, vol 4., No.3, Fall 1998.

Kupiec, Paul H. (1995) Techniques for Verifying the Accuracy of Risk Measurement Models. The Journal of Derivatives, Winter 1995.

Lim, G.C. & Lye, J.N & Martin, G.M. & Martin, V.L. (1998) The Distribution of Exchange Rate Returns and Pricing of Currency Options. Journal of International Economics 45.

McNeil, Alexander J. (1998) Calculating Quantile Risk Measures for Financial Return Series using Extreme Value Theory.

McNeil, Alexander J. (1999) Extreme Value Theory for Risk Managers.

Mori, Atutoshi & Ohsawa, Makoto & Shimizu, Tokiko (1996) A Framework for More Effective Stress Testing. Institute for Monetary and Economic Studies Bank of Japan, Discussion Paper 96-E-2.

Nelson, D.B. (1990) Conditional Heteroskedasticity in Asset Returns: A New Approach. Econometrica, 59.

Pritsker, Matthew (1997) Evaluating Value at Risk Methodologies: Accuracy versus Computational Time. Journal of Financial Services Research, vol 12, October/December 1997.

Raatikainen, Juhani (2000) Algoritmi extreme value –estimointiin. Jyväskylän yliopisto. Taloustieteiden tiedekunta.

Simons, Katerine (1996) Value-at-risk – New Approaches to Risk Management. Federal Reserve Bank of Boston New England Economic Review, Sep/Oct 1996.

Vose, David (1996) Quantitative Risk Analysis: A Guide to Monte Carlo Simulation Modelling. John Wiley & Sons Ltd, West Sussex, England.

West, Kenneth D. & Cho, Dongchul (1995) The Predictive Ability of Several Models of Exchange Rate Volatility. Journal of Econometrics 69 (1995).

LIITE 1.
Pääkomponenttien GARCH-estimoinnin tulokset.

Pääkomponentti 1.

```
**********************************************************************
*  Estimating an ARMA(1,0) for the mean of PK1                       *
*   w/ a GARCH(1,1) model for the conditional variance               *
**********************************************************************

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations  1355     Degrees of Freedom  1350
Function Value            1138.29684630

  Variable              Coeff
*******************************************
1. A1                -0.002966774
2. A2                 1.000088987
3. C                  0.005237353
4. Q1                 0.125477446
5. P1                 0.376859846

Estimation by BHHH
Iterations Taken   25
Usable Observations  1355     Degrees of Freedom  1350
Function Value            1169.99042027

  Variable              Coeff        Std Error      T-Stat     Signif
*********************************************************************************
1. A1                -0.006250453  0.002688117    -2.32522    0.02006040
2. A2                 0.999632328  0.000904629  1105.01964    0.00000000
3. C                  0.000498150  0.000118407     4.20709    0.00002587
4. Q1                 0.082427485  0.012160715     6.77818    0.00000000
5. P1                 0.873868458  0.019978793    43.73980    0.00000000
```

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-0.03512	4.52303

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:	4.4258	Significance Level:	0.21901
LB(8)	Test Statistic:	6.9370	Significance Level:	0.43547
LB(12)	Test Statistic:	9.0545	Significance Level:	0.61686
LB(16)	Test Statistic:	19.4810	Significance Level:	0.19276
LB(20)	Test Statistic:	24.5045	Significance Level:	0.17751
LB(24)	Test Statistic:	29.9440	Significance Level:	0.15102

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 131.2400 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.2407	Significance Level:	0.91531
ARCH(8)	Test Statistic:	0.6025	Significance Level:	0.77639
ARCH(12)	Test Statistic:	0.5774	Significance Level:	0.86174
ARCH(16)	Test Statistic:	0.5191	Significance Level:	0.93864
ARCH(20)	Test Statistic:	0.6259	Significance Level:	0.89610
ARCH(24)	Test Statistic:	0.6179	Significance Level:	0.92453

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	0.9816	Significance Level:	0.80571
LB(8)	Test Statistic:	4.9533	Significance Level:	0.66566
LB(12)	Test Statistic:	6.9509	Significance Level:	0.80304
LB(16)	Test Statistic:	8.1135	Significance Level:	0.91915
LB(20)	Test Statistic:	12.4010	Significance Level:	0.86778
LB(24)	Test Statistic:	14.6177	Significance Level:	0.90781

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -0.2123 Significance Level: 0.83189
The Negative Size Bias Test
Test Statistic: 0.5524 Significance Level: 0.58075
The Positive Size Bias Test
Test Statistic: 1.0132 Significance Level: 0.31114
The Joint Test for the Three Effects
Test Statistic: 1.1830 Significance Level: 0.31490

****

Pääkomponentti 2.

```
*********************************************************************
*  Estimating an ARMA(2,0) for the mean of PK2                      *
*   w/ a GARCH(1,1) model for the conditional variance              *
*********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1354 Degrees of Freedom 1348
Function Value -2392.33882282

Variable	Coeff
1. A1	-0.000464747
2. A2	0.253210305
3. A3	-0.123075585
4. C	0.217646804
5. Q1	0.373909251
6. P1	0.617698392

Estimation by BHHH
Iterations Taken 17
Usable Observations 1354 Degrees of Freedom 1348
Function Value -2354.06380569

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	-0.050435234	0.035309655	-1.42837	0.15318554
2. A2	0.203866992	0.030873282	6.60335	0.00000000
3. A3	-0.024062468	0.026440514	-0.91006	0.36279060
4. C	0.063231197	0.016838319	3.75520	0.00017321
5. Q1	0.116541922	0.014681707	7.93790	0.00000000
6. P1	0.857900118	0.017040330	50.34528	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.44396	3.98271

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4) Test Statistic: 0.4190 Significance Level: 0.81099
LB(8) Test Statistic: 9.5605 Significance Level: 0.14442
LB(12) Test Statistic: 12.4158 Significance Level: 0.25819
LB(16) Test Statistic: 23.9655 Significance Level: 0.04626
LB(20) Test Statistic: 27.9361 Significance Level: 0.06303
LB(24) Test Statistic: 33.4838 Significance Level: 0.05538

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 98.9613 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	3.0254	Significance Level:	0.01694
ARCH(8)	Test Statistic:	1.6283	Significance Level:	0.11212
ARCH(12)	Test Statistic:	1.4809	Significance Level:	0.12450
ARCH(16)	Test Statistic:	1.2597	Significance Level:	0.21537
ARCH(20)	Test Statistic:	1.0271	Significance Level:	0.42567
ARCH(24)	Test Statistic:	1.1082	Significance Level:	0.32586

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	13.2156	Significance Level:	0.00135
LB(8)	Test Statistic:	14.3335	Significance Level:	0.02613
LB(12)	Test Statistic:	19.2782	Significance Level:	0.03687
LB(16)	Test Statistic:	21.2996	Significance Level:	0.09420
LB(20)	Test Statistic:	21.8407	Significance Level:	0.23913
LB(24)	Test Statistic:	27.2336	Significance Level:	0.20249

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -1.1008 Significance Level: 0.27116
The Negative Size Bias Test
Test Statistic: 1.4058 Significance Level: 0.16001
The Positive Size Bias Test
Test Statistic: 0.2573 Significance Level: 0.79697
The Joint Test for the Three Effects
Test Statistic: 3.3334 Significance Level: 0.01885

****

Pääkomponentti 3.

```
************************************************************************
*   Estimating an ARMA(0,0) for the mean of PK3                        *
*    w/ a GARCH(1,1) model for the conditional variance                *
************************************************************************



A1 is the constant in the mean equation
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations  1356     Degrees of Freedom  1353
Function Value              -1902.81031817

  Variable                 Coeff
******************************************
1.  C                0.1736139617
2.  Q1                0.0945369133
3.  P1                0.7522126028

Estimation by BHHH
Iterations Taken   86
Usable Observations  1356     Degrees of Freedom  1353
Function Value              -1900.67148577

  Variable                 Coeff        Std Error      T-Stat     Signif
*******************************************************************************
1.  C                0.1045442889 0.0182283880     5.73525  0.00000001
2.  Q1                0.0819969938 0.0094420973     8.68419  0.00000000
3.  P1                0.8190992009 0.0225159995    36.37854  0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

        Skewness                Kurtosis
        -1.07195               20.38928

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4)      Test Statistic:  2.5145       Significance Level:  0.64203
LB(8)      Test Statistic:  3.5224       Significance Level:  0.89745
LB(12)     Test Statistic:  12.1176      Significance Level:  0.43629
LB(16)     Test Statistic:  16.6922      Significance Level:  0.40578
LB(20)     Test Statistic:  20.0743      Significance Level:  0.45329
LB(24)     Test Statistic:  22.7875      Significance Level:  0.53237

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
        Test Statistic:17344.5572    Significance Level:  0.00000

F-Test of no ARCH vs. ARCH in NRESIDS
ARCH(4)    Test Statistic:  2.0127       Significance Level:  0.09038
ARCH(8)    Test Statistic:  1.0729       Significance Level:  0.37940
ARCH(12)   Test Statistic:  0.7205       Significance Level:  0.73243
ARCH(16)   Test Statistic:  0.5579       Significance Level:  0.91572
ARCH(20)   Test Statistic:  0.4583       Significance Level:  0.98052
ARCH(24)   Test Statistic:  0.3939       Significance Level:  0.99632
```

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	7.8753	Significance Level:	0.09626
LB(8)	Test Statistic:	8.4733	Significance Level:	0.38865
LB(12)	Test Statistic:	8.6076	Significance Level:	0.73603
LB(16)	Test Statistic:	9.0227	Significance Level:	0.91248
LB(20)	Test Statistic:	9.3082	Significance Level:	0.97906
LB(24)	Test Statistic:	9.6299	Significance Level:	0.99591

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -1.7480 Significance Level: 0.08069
The Negative Size Bias Test
Test Statistic: -1.4777 Significance Level: 0.13971
The Positive Size Bias Test
Test Statistic: -0.8594 Significance Level: 0.39027
The Joint Test for the Three Effects
Test Statistic: 1.1987 Significance Level: 0.30895

****

Pääkomponentti 4.

```
**********************************************************************
*  Estimating an ARMA(5,0) for the mean of PK4                        *
*   w/ a GARCH(3,3) model for the conditional variance                *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1349 Degrees of Freedom 1337
Function Value 245.02017763

Variable	Coeff
1. A1	0.3306690404
2. A2	0.2302255649
3. A3	0.1801922859
4. A4	0.1323535056
5. A5	0.1047016067
6. C	0.0193297643
7. Q1	0.1419665751
8. Q2	0.1023062552
9. Q3	0.1028353583
10. P1	0.0373746022
11. P2	0.0828487651
12. P3	0.1197315357

Estimation by BHHH
Iterations Taken 57
Usable Observations 1349 Degrees of Freedom 1337
Function Value 254.78836787

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.330802051	0.028976971	11.41603	0.00000000
2. A2	0.236372568	0.030951214	7.63694	0.00000000
3. A3	0.192625861	0.031185389	6.17680	0.00000000
4. A4	0.120826061	0.031219433	3.87022	0.00010874
5. A5	0.098503704	0.028263490	3.48519	0.00049178
6. C	0.013116577	0.003045626	4.30669	0.00001657
7. Q1	0.100937350	0.020463007	4.93267	0.00000081
8. Q2	-0.030915157	0.014126499	-2.18845	0.02863673
9. Q3	0.118115639	0.020928477	5.64378	0.00000002
10. P1	1.029495092	0.079694993	12.91794	0.00000000
11. P2	-1.148572667	0.069897555	-16.43223	0.00000000
12. P3	0.620659567	0.064138269	9.67690	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.12840	4.17707

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(8) Test Statistic: 9.5337 Significance Level: 0.02298
LB(16) Test Statistic: 21.5894 Significance Level: 0.02775
LB(24) Test Statistic: 24.1158 Significance Level: 0.19175

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 81.5835 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS
ARCH(4) Test Statistic: 0.4722 Significance Level: 0.75619
ARCH(8) Test Statistic: 0.7471 Significance Level: 0.64981
ARCH(12) Test Statistic: 1.5243 Significance Level: 0.10873
ARCH(16) Test Statistic: 1.1805 Significance Level: 0.27637
ARCH(20) Test Statistic: 1.0925 Significance Level: 0.35048
ARCH(24) Test Statistic: 0.9936 Significance Level: 0.47133

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS
LB(8) Test Statistic: 6.1475 Significance Level: 0.10465
LB(16) Test Statistic: 20.5669 Significance Level: 0.03815
LB(24) Test Statistic: 24.4250 Significance Level: 0.18035

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: 0.3703 Significance Level: 0.71124
The Negative Size Bias Test
Test Statistic: -0.5941 Significance Level: 0.55256
The Positive Size Bias Test
Test Statistic: -0.6599 Significance Level: 0.50946
The Joint Test for the Three Effects
Test Statistic: 1.1005 Significance Level: 0.34781

****

```
**********************************************************************
*  Estimating an ARMA(4,0) for the mean of PK5                       *
*    w/ an EGARCH(2,2) model for the conditional variance *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag
D is the scale parameter for the generalized error distribution

```
****************************************
* The distribution used is a G.E.D.  *
****************************************
```

Estimation by Simplex
Usable Observations 1351 Degrees of Freedom 1341
Function Value -1695.42057953

Variable	Coeff
1. A1	-0.082708393
2. A2	0.042541135
3. A3	0.117014222
4. A4	0.063848916
5. C	-0.235575103
6. Q1	0.210314601
7. Q2	0.080724528
8. P1	0.133815964
9. P2	0.081574857
10. L1	-0.036170720

Estimation by BHHH
Iterations Taken 14
Usable Observations 1351 Degrees of Freedom 1340
Function Value -1622.52622868

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	-0.006284324	0.027405097	-0.22931	0.81862626
2. A2	0.070377217	0.026679082	2.63792	0.00834169
3. A3	0.101991655	0.027108833	3.76230	0.00016836
4. A4	0.044736310	0.026671339	1.67732	0.09348037
5. C	-0.128571935	0.017999218	-7.14320	0.00000000
6. Q1	0.203880014	0.033106707	6.15827	0.00000000
7. Q2	0.185625119	0.030304644	6.12530	0.00000000
8. P1	-0.123074871	0.005794941	-21.23833	0.00000000
9. P2	0.879301106	0.004744137	185.34478	0.00000000
10. L1	0.211132753	0.115222892	1.83239	0.06689401
11. D	1.575297297	0.056237682	28.01142	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-0.63937	10.18853

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(8)   Test Statistic: 7.0285   Significance Level: 0.13439
LB(16)   Test Statistic: 30.4734   Significance Level: 0.00237
LB(24)   Test Statistic: 43.1153   Significance Level: 0.00197

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic:3000.9178   Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS
ARCH(4) Test Statistic: 4.1460   Significance Level: 0.00242
ARCH(8) Test Statistic: 2.1100   Significance Level: 0.03213
ARCH(12) Test Statistic: 1.5231   Significance Level: 0.10914
ARCH(16) Test Statistic: 1.3731   Significance Level: 0.14633
ARCH(20) Test Statistic: 1.2248   Significance Level: 0.22390
ARCH(24) Test Statistic: 1.0432   Significance Level: 0.40539

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS
LB(8)   Test Statistic: 16.3862   Significance Level: 0.00254
LB(16)   Test Statistic: 20.9351   Significance Level: 0.05134
LB(24)   Test Statistic: 23.6024   Significance Level: 0.26019

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -2.4133   Significance Level: 0.01594
The Negative Size Bias Test
Test Statistic: -2.1183   Significance Level: 0.03434
The Positive Size Bias Test
Test Statistic: -0.8248   Significance Level: 0.40962
The Joint Test for the Three Effects
Test Statistic: 2.3131   Significance Level: 0.07436

****

Pääkomponentti 6.

```
**********************************************************************
*  Estimating an ARMA(1,0) for the mean of PK6                       *
*    w/ a GARCH(1,1) model for the conditional variance              *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1625.44104230

Variable	Coeff
1. A1	0.2694592334
2. C	0.4158912947
3. Q1	0.2268076906
4. P1	0.1743167590

Estimation by BHHH
Iterations Taken 64
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1612.07569384

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.2253271237	0.0280551542	8.03158	0.00000000
2. C	0.0365224239	0.0089930278	4.06119	0.00004882
3. Q1	0.1079718593	0.0115688348	9.33299	0.00000000
4. P1	0.8417516519	0.0192770491	43.66600	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.06409	4.08547

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic: 163.2863	Significance Level: 0.00000
LB(8)	Test Statistic: 361.0401	Significance Level: 0.00000
LB(12)	Test Statistic: 575.5740	Significance Level: 0.00000
LB(16)	Test Statistic: 772.2119	Significance Level: 0.00000
LB(20)	Test Statistic: 972.7862	Significance Level: 0.00000
LB(24)	Test Statistic:1156.8970	Significance Level: 0.00000

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 67.4497 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic: 6.5059	Significance Level: 0.00003
ARCH(8)	Test Statistic: 4.2603	Significance Level: 0.00005
ARCH(12)	Test Statistic: 2.9976	Significance Level: 0.00037
ARCH(16)	Test Statistic: 2.4325	Significance Level: 0.00126
ARCH(20)	Test Statistic: 2.1946	Significance Level: 0.00177
ARCH(24)	Test Statistic: 2.1826	Significance Level: 0.00083

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	27.3707	Significance Level:	0.00000
LB(8)	Test Statistic:	36.2279	Significance Level:	0.00001
LB(12)	Test Statistic:	37.8903	Significance Level:	0.00008
LB(16)	Test Statistic:	40.6711	Significance Level:	0.00036
LB(20)	Test Statistic:	44.8028	Significance Level:	0.00073
LB(24)	Test Statistic:	54.0495	Significance Level:	0.00026

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -3.1790 Significance Level: 0.00151
The Negative Size Bias Test
Test Statistic: -4.5142 Significance Level: 0.00001
The Positive Size Bias Test
Test Statistic: -1.4348 Significance Level: 0.15157
The Joint Test for the Three Effects
Test Statistic: 7.4795 Significance Level: 0.00006

****

Pääkomponentti 7.

```
**********************************************************************
*  Estimating an ARMA(0,0) for the mean of PK7                       *
*    w/ a GARCH(1,1) model for the conditional variance              *
**********************************************************************
```

A1 is the constant in the mean equation
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1356 Degrees of Freedom 1353
Function Value -1555.14754883

Variable	Coeff
1. C	0.1475921933
2. Q1	0.4822467864
3. P1	0.3520502432

Estimation by BHHH
Iterations Taken 25
Usable Observations 1356 Degrees of Freedom 1353
Function Value -1519.52797582

Variable	Coeff	Std Error	T-Stat	Signif
1. C	0.0161946980	0.0056258552	2.87862	0.00399420
2. Q1	0.1345471811	0.0143900971	9.34998	0.00000000
3. P1	0.8436320712	0.0179368069	47.03357	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.05061	3.29576

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:1265.3846	Significance Level: 0.00000
LB(8)	Test Statistic:2654.4835	Significance Level: 0.00000
LB(12)	Test Statistic:4045.1428	Significance Level: 0.00000
LB(16)	Test Statistic:5418.9257	Significance Level: 0.00000
LB(20)	Test Statistic:6760.7477	Significance Level: 0.00000
LB(24)	Test Statistic:8057.9436	Significance Level: 0.00000

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 5.5213 Significance Level: 0.06325

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic: 3.6601	Significance Level: 0.00568
ARCH(8)	Test Statistic: 2.4724	Significance Level: 0.01162
ARCH(12)	Test Statistic: 2.0933	Significance Level: 0.01493
ARCH(16)	Test Statistic: 1.6615	Significance Level: 0.04790
ARCH(20)	Test Statistic: 1.6085	Significance Level: 0.04331
ARCH(24)	Test Statistic: 1.3826	Significance Level: 0.10316

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	14.9946	Significance Level:	0.00471
LB(8)	Test Statistic:	20.2775	Significance Level:	0.00934
LB(12)	Test Statistic:	24.7087	Significance Level:	0.01627
LB(16)	Test Statistic:	26.6907	Significance Level:	0.04505
LB(20)	Test Statistic:	31.8907	Significance Level:	0.04448
LB(24)	Test Statistic:	33.4274	Significance Level:	0.09537

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -0.0343 Significance Level: 0.97263
The Negative Size Bias Test
Test Statistic: 0.7537 Significance Level: 0.45115
The Positive Size Bias Test
Test Statistic: 0.7507 Significance Level: 0.45294
The Joint Test for the Three Effects
Test Statistic: 1.1719 Significance Level: 0.31916

****

Pääkomponentti 8.

```
**********************************************************************
*  Estimating an ARMA(1,0) for the mean of PK8                       *
*   w/ a GARCH(1,1) model for the conditional variance               *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1347.89142488

Variable	Coeff
1. A1	-0.042362550
2. C	0.132217760
3. Q1	0.208138962
4. P1	0.527177616

Estimation by BHHH
Iterations Taken 19
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1325.53843866

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	-0.124428258	0.027990890	-4.44531	0.00000878
2. C	0.016685610	0.004978103	3.35180	0.00080288
3. Q1	0.091511912	0.014548560	6.29010	0.00000000
4. P1	0.872408029	0.021079613	41.38634	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.05217	3.65529

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4) Test Statistic: 2.1678 Significance Level: 0.53833
LB(8) Test Statistic: 12.7078 Significance Level: 0.07956
LB(12) Test Statistic: 16.4064 Significance Level: 0.12670
LB(16) Test Statistic: 25.4654 Significance Level: 0.04403
LB(20) Test Statistic: 29.7802 Significance Level: 0.05466
LB(24) Test Statistic: 33.3020 Significance Level: 0.07595

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 24.8584 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS
ARCH(4) Test Statistic: 1.4950 Significance Level: 0.20130
ARCH(8) Test Statistic: 1.2620 Significance Level: 0.25941
ARCH(12) Test Statistic: 0.9004 Significance Level: 0.54598
ARCH(16) Test Statistic: 0.7125 Significance Level: 0.78339
ARCH(20) Test Statistic: 0.8461 Significance Level: 0.65759
ARCH(24) Test Statistic: 0.9683 Significance Level: 0.50633

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	5.8398	Significance Level:	0.11967
LB(8)	Test Statistic:	10.3783	Significance Level:	0.16813
LB(12)	Test Statistic:	11.4136	Significance Level:	0.40929
LB(16)	Test Statistic:	11.7094	Significance Level:	0.70088
LB(20)	Test Statistic:	18.0344	Significance Level:	0.52014
LB(24)	Test Statistic:	23.5335	Significance Level:	0.43000

* Sign and Size Bias Tests *

The Sign Bias Test
    Test Statistic: -1.6402   Significance Level: 0.10120
The Negative Size Bias Test
    Test Statistic: -1.4723   Significance Level: 0.14116
The Positive Size Bias Test
    Test Statistic: -0.9715   Significance Level: 0.33148
The Joint Test for the Three Effects
    Test Statistic: 1.1237   Significance Level: 0.33829

****

Pääkomponentti 9.

```
**********************************************************************
*  Estimating an ARMA(1,0) for the mean of PK9                       *
*    w/ a GARCH(1,1) model for the conditional variance              *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1159.80136511

Variable	Coeff
1. A1	-0.083256403
2. C	0.022250781
3. Q1	0.112426210
4. P1	0.836002775

Estimation by BHHH
Iterations Taken 15
Usable Observations 1355 Degrees of Freedom 1351
Function Value -1148.03615898

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	-0.152839275	0.026731731	-5.71752	0.00000001
2. C	0.004570914	0.001894869	2.41226	0.01585404
3. Q1	0.038671907	0.007585905	5.09786	0.00000034
4. P1	0.947850679	0.010867256	87.22079	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-0.01878	3.83032

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4) Test Statistic: 9.2771 Significance Level: 0.02583
LB(8) Test Statistic: 13.1555 Significance Level: 0.06841
LB(12) Test Statistic: 14.6675 Significance Level: 0.19823
LB(16) Test Statistic: 20.0523 Significance Level: 0.16994
LB(20) Test Statistic: 24.1974 Significance Level: 0.18869
LB(24) Test Statistic: 27.8609 Significance Level: 0.22109

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 39.0034 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS
ARCH(4) Test Statistic: 0.9816 Significance Level: 0.41646
ARCH(8) Test Statistic: 0.6959 Significance Level: 0.69551
ARCH(12) Test Statistic: 0.5760 Significance Level: 0.86289
ARCH(16) Test Statistic: 0.6107 Significance Level: 0.87753
ARCH(20) Test Statistic: 0.6556 Significance Level: 0.87151
ARCH(24) Test Statistic: 0.5987 Significance Level: 0.93682

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	4.1331	Significance Level:	0.24745
LB(8)	Test Statistic:	5.7292	Significance Level:	0.57170
LB(12)	Test Statistic:	7.1280	Significance Level:	0.78863
LB(16)	Test Statistic:	9.6669	Significance Level:	0.84015
LB(20)	Test Statistic:	13.5234	Significance Level:	0.81073
LB(24)	Test Statistic:	15.2448	Significance Level:	0.88565

* Sign and Size Bias Tests *

The Sign Bias Test
    Test Statistic: -2.8795   Significance Level: 0.00405
The Negative Size Bias Test
    Test Statistic: -1.7578   Significance Level: 0.07900
The Positive Size Bias Test
    Test Statistic: -1.0466   Significance Level: 0.29548
The Joint Test for the Three Effects
    Test Statistic: 2.8646   Significance Level: 0.03559

****

Pääkomponentti 10.

```
**********************************************************************
*  Estimating an ARMA(3,0) for the mean of PK10                     *
*    w/ a GARCH(1,1) model for the conditional variance             *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1353 Degrees of Freedom 1347
Function Value 1656.53601403

Variable	Coeff
1. A1	0.5452832973
2. A2	0.4121708145
3. A3	0.0434346461
4. C	0.0011288425
5. Q1	0.2666011522
6. P1	0.5781630741

Estimation by BHHH
Iterations Taken 32
Usable Observations 1353 Degrees of Freedom 1347
Function Value 1683.95675050

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.4479389325	0.0288772658	15.51182	0.00000000
2. A2	0.3745422104	0.0285204204	13.13242	0.00000000
3. A3	0.1693254741	0.0276984110	6.11318	0.00000000
4. C	0.0002437497	0.0000670765	3.63391	0.00027916
5. Q1	0.1223886719	0.0144368712	8.47751	0.00000000
6. P1	0.8395463243	0.0192690302	43.56972	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.33926	7.15439

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:	7.2392	Significance Level:	0.00713
LB(8)	Test Statistic:	8.8622	Significance Level:	0.11469
LB(12)	Test Statistic:	17.7057	Significance Level:	0.03875
LB(16)	Test Statistic:	18.3294	Significance Level:	0.14541
LB(20)	Test Statistic:	19.9739	Significance Level:	0.27557
LB(24)	Test Statistic:	24.2465	Significance Level:	0.28127

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 998.9293 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.1661	Significance Level:	0.95560
ARCH(8)	Test Statistic:	0.3027	Significance Level:	0.96514
ARCH(12)	Test Statistic:	0.2873	Significance Level:	0.99138
ARCH(16)	Test Statistic:	0.2873	Significance Level:	0.99737
ARCH(20)	Test Statistic:	0.6358	Significance Level:	0.88822
ARCH(24)	Test Statistic:	0.5870	Significance Level:	0.94364

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	0.6819	Significance Level:	0.40893
LB(8)	Test Statistic:	2.5002	Significance Level:	0.77647
LB(12)	Test Statistic:	3.5817	Significance Level:	0.93673
LB(16)	Test Statistic:	4.6690	Significance Level:	0.98178
LB(20)	Test Statistic:	13.1792	Significance Level:	0.72411
LB(24)	Test Statistic:	14.9680	Significance Level:	0.82455

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -0.2368 Significance Level: 0.81286
The Negative Size Bias Test
Test Statistic: 0.7318 Significance Level: 0.46441
The Positive Size Bias Test
Test Statistic: 0.3323 Significance Level: 0.73973
The Joint Test for the Three Effects
Test Statistic: 0.6574 Significance Level: 0.57835

****

Pääkomponentti 11.

```
*********************************************************************
*  Estimating an ARMA(3,0) for the mean of PK11                     *
*    w/ a GARCH(1,1) model for the conditional variance             *
*********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1353 Degrees of Freedom 1347
Function Value 2483.19198980

Variable	Coeff
1. A1	0.6936016609
2. A2	0.1847515115
3. A3	0.1116087554
4. C	0.0009757509
5. Q1	0.2007252153
6. P1	0.2060418375

Estimation by BHHH
Iterations Taken 24
Usable Observations 1353 Degrees of Freedom 1347
Function Value 2489.24524571

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.6942781128	0.0305266146	22.74337	0.00000000
2. A2	0.1838406262	0.0351563211	5.22923	0.00000017
3. A3	0.1119636207	0.0290068599	3.85990	0.00011343
4. C	0.0004825427	0.0001592819	3.02949	0.00244969
5. Q1	0.0789491486	0.0252178703	3.13068	0.00174401
6. P1	0.6013079871	0.1238282478	4.85598	0.00000120

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.23053	7.38030

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4) Test Statistic: 2.5436 Significance Level: 0.11074
LB(8) Test Statistic: 8.5480 Significance Level: 0.12851
LB(12) Test Statistic: 17.3749 Significance Level: 0.04316
LB(16) Test Statistic: 18.7738 Significance Level: 0.13029
LB(20) Test Statistic: 25.4720 Significance Level: 0.08463
LB(24) Test Statistic: 31.2664 Significance Level: 0.06934

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic:1093.6546 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.0867	Significance Level:	0.98659
ARCH(8)	Test Statistic:	0.1058	Significance Level:	0.99904
ARCH(12)	Test Statistic:	0.3261	Significance Level:	0.98481
ARCH(16)	Test Statistic:	0.3532	Significance Level:	0.99130
ARCH(20)	Test Statistic:	0.4154	Significance Level:	0.98937
ARCH(24)	Test Statistic:	0.5125	Significance Level:	0.97573

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	0.3542	Significance Level:	0.55174
LB(8)	Test Statistic:	0.8910	Significance Level:	0.97087
LB(12)	Test Statistic:	4.0741	Significance Level:	0.90647
LB(16)	Test Statistic:	5.6106	Significance Level:	0.95918
LB(20)	Test Statistic:	8.4608	Significance Level:	0.95568
LB(24)	Test Statistic:	12.8304	Significance Level:	0.91445

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -2.3741 Significance Level: 0.01773
The Negative Size Bias Test
Test Statistic: -1.2777 Significance Level: 0.20157
The Positive Size Bias Test
Test Statistic: -1.0719 Significance Level: 0.28395
The Joint Test for the Three Effects
Test Statistic: 1.8920 Significance Level: 0.12904

****

Pääkomponentti 12.

```
*********************************************************************
*  Estimating an ARMA(2,0) for the mean of PK12                     *
*    w/ a GARCH(1,1) model for the conditional variance             *
*********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2172.22218305

Variable	Coeff
1. A1	0.7621719768
2. A2	0.2042531557
3. C	0.0013259581
4. Q1	0.2794497896
5. P1	0.2340745404

Estimation by BHHH
Iterations Taken 24
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2221.90377358

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.7853272719	0.0252347010	31.12093	0.00000000
2. A2	0.1859406545	0.0260404035	7.14047	0.00000000
3. C	0.0000237469	0.0000082352	2.88359	0.00393174
4. Q1	0.0384104789	0.0074578225	5.15036	0.00000026
5. P1	0.9512639023	0.0099295489	95.80132	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-0.08013	4.62738

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:	1.4584	Significance Level:	0.48231
LB(8)	Test Statistic:	7.4155	Significance Level:	0.28412
LB(12)	Test Statistic:	15.5898	Significance Level:	0.11199
LB(16)	Test Statistic:	18.2417	Significance Level:	0.19600
LB(20)	Test Statistic:	19.6559	Significance Level:	0.35252
LB(24)	Test Statistic:	20.6244	Significance Level:	0.54404

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 150.8616 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.9968	Significance Level:	0.40813
ARCH(8)	Test Statistic:	0.9543	Significance Level:	0.47041
ARCH(12)	Test Statistic:	0.8962	Significance Level:	0.55033
ARCH(16)	Test Statistic:	1.1717	Significance Level:	0.28378
ARCH(20)	Test Statistic:	0.9528	Significance Level:	0.51874
ARCH(24)	Test Statistic:	1.1970	Significance Level:	0.23337

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	3.9737	Significance Level:	0.13713
LB(8)	Test Statistic:	7.5660	Significance Level:	0.27166
LB(12)	Test Statistic:	9.6126	Significance Level:	0.47511
LB(16)	Test Statistic:	15.3791	Significance Level:	0.35274
LB(20)	Test Statistic:	17.0617	Significance Level:	0.51887
LB(24)	Test Statistic:	24.2915	Significance Level:	0.33214

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: 0.6921 Significance Level: 0.48897
The Negative Size Bias Test
Test Statistic: -0.8770 Significance Level: 0.38065
The Positive Size Bias Test
Test Statistic: 1.6320 Significance Level: 0.10292
The Joint Test for the Three Effects
Test Statistic: 1.2272 Significance Level: 0.29839

****

Pääkomponentti 13.

```
**********************************************************************
*  Estimating an ARMA(2,0) for the mean of PK13                      *
*    w/ a GARCH(1,1) model for the conditional variance              *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2533.70473479

Variable	Coeff
1. A1	0.8744806435
2. A2	0.0991462830
3. C	0.0008360898
4. Q1	0.1941827312
5. P1	0.2322779283

Estimation by BHHH
Iterations Taken 90
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2557.44004828

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.8095714667	0.0254677203	31.78814	0.00000000
2. A2	0.1709210000	0.0262023968	6.52311	0.00000000
3. C	0.0000104873	0.0000027650	3.79294	0.00014888
4. Q1	0.0245335495	0.0048005470	5.11057	0.00000032
5. P1	0.9682555676	0.0055644667	174.00689	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.08592	6.76016

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:	1.3600	Significance Level:	0.50663
LB(8)	Test Statistic:	3.1260	Significance Level:	0.79288
LB(12)	Test Statistic:	8.2940	Significance Level:	0.60014
LB(16)	Test Statistic:	9.8052	Significance Level:	0.77628
LB(20)	Test Statistic:	15.0896	Significance Level:	0.65581
LB(24)	Test Statistic:	19.3683	Significance Level:	0.62249

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 799.3307 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	1.9925	Significance Level:	0.09334
ARCH(8)	Test Statistic:	1.3075	Significance Level:	0.23535
ARCH(12)	Test Statistic:	1.0692	Significance Level:	0.38250
ARCH(16)	Test Statistic:	0.8920	Significance Level:	0.57860
ARCH(20)	Test Statistic:	0.7248	Significance Level:	0.80350
ARCH(24)	Test Statistic:	0.9682	Significance Level:	0.50652

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	8.1958	Significance Level:	0.01661
LB(8)	Test Statistic:	10.7917	Significance Level:	0.09503
LB(12)	Test Statistic:	13.2693	Significance Level:	0.20900
LB(16)	Test Statistic:	14.3560	Significance Level:	0.42354
LB(20)	Test Statistic:	14.5185	Significance Level:	0.69472
LB(24)	Test Statistic:	22.2303	Significance Level:	0.44621

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -0.0635 Significance Level: 0.94939
The Negative Size Bias Test
Test Statistic: -0.3294 Significance Level: 0.74194
The Positive Size Bias Test
Test Statistic: 2.7208 Significance Level: 0.00660
The Joint Test for the Three Effects
Test Statistic: 3.6303 Significance Level: 0.01256

****

Pääkomponentti 14.

```
*********************************************************************
*  Estimating an ARMA(2,0) for the mean of PK14                    *
*   w/ a GARCH(1,1) model for the conditional variance             *
*********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2507.51881000

```
  Variable              Coeff
*****************************************
1.  A1                0.7405211966
2.  A2                0.2179017217
3.  C                 0.0007882668
4.  Q1                0.1631671965
5.  P1                0.3099716581
```

Estimation by BHHH
Iterations Taken 88
Usable Observations 1354 Degrees of Freedom 1349
Function Value 2514.46677630

```
  Variable              Coeff        Std Error      T-Stat     Signif
*******************************************************************************
1.  A1                0.7262458773 0.0259810084   27.95295  0.00000000
2.  A2                0.2340676481 0.0256308616    9.13226  0.00000000
3.  C                 0.0000613374 0.0000115452    5.31279  0.00000011
4.  Q1                0.0617924552 0.0082171611    7.51993  0.00000000
5.  P1                0.8994011141 0.0127185271   70.71582  0.00000000
```

* Analyzing the normalized residuals for serial correlation and normality *

```
      Skewness              Kurtosis
       0.14713              7.38986
```

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

```
LB(4)    Test Statistic:  4.5288    Significance Level:  0.10389
LB(8)    Test Statistic:  9.0747    Significance Level:  0.16942
LB(12)   Test Statistic: 14.2824    Significance Level:  0.16050
LB(16)   Test Statistic: 16.7558    Significance Level:  0.26942
LB(20)   Test Statistic: 19.2858    Significance Level:  0.37442
LB(24)   Test Statistic: 27.6711    Significance Level:  0.18684
```

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic:1092.0833 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	2.3527	Significance Level:	0.05217
ARCH(8)	Test Statistic:	1.4232	Significance Level:	0.18198
ARCH(12)	Test Statistic:	1.1195	Significance Level:	0.33953
ARCH(16)	Test Statistic:	0.9377	Significance Level:	0.52477
ARCH(20)	Test Statistic:	0.7593	Significance Level:	0.76472
ARCH(24)	Test Statistic:	0.6367	Significance Level:	0.91108

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	9.7025	Significance Level:	0.00782
LB(8)	Test Statistic:	11.9843	Significance Level:	0.06232
LB(12)	Test Statistic:	13.7366	Significance Level:	0.18535
LB(16)	Test Statistic:	15.4122	Significance Level:	0.35057
LB(20)	Test Statistic:	15.6846	Significance Level:	0.61455
LB(24)	Test Statistic:	16.0430	Significance Level:	0.81375

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: 0.8050 Significance Level: 0.42094
The Negative Size Bias Test
Test Statistic: -0.6136 Significance Level: 0.53958
The Positive Size Bias Test
Test Statistic: 2.0021 Significance Level: 0.04547
The Joint Test for the Three Effects
Test Statistic: 1.4731 Significance Level: 0.22011

****

Pääkomponentti 15.

```
***********************************************************************
*  Estimating an ARMA(3,0) for the mean of PK15                        *
*    w/ a GARCH(1,1) model for the conditional variance                *
***********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1353 Degrees of Freedom 1347
Function Value 2639.96142738

Variable	Coeff
1. A1	0.8130762596
2. A2	0.0658924712
3. A3	0.0618872022
4. C	0.0004916376
5. Q1	0.3390153922
6. P1	0.3508743709

Estimation by BHHH
Iterations Taken 55
Usable Observations 1353 Degrees of Freedom 1347
Function Value 2666.26560626

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.7730911936	0.0265794182	29.08608	0.00000000
2. A2	0.1129228537	0.0366168153	3.08391	0.00204301
3. A3	0.0588574755	0.0271638502	2.16676	0.03025335
4. C	0.0000201985	0.0000056385	3.58224	0.00034066
5. Q1	0.0604326334	0.0096115115	6.28753	0.00000000
6. P1	0.9261689551	0.0103233194	89.71620	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
0.19336	5.25972

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(4) Test Statistic: 1.6176 Significance Level: 0.20342
LB(8) Test Statistic: 8.5445 Significance Level: 0.12867
LB(12) Test Statistic: 10.1058 Significance Level: 0.34199
LB(16) Test Statistic: 12.2854 Significance Level: 0.50439
LB(20) Test Statistic: 17.9838 Significance Level: 0.38986
LB(24) Test Statistic: 25.0870 Significance Level: 0.24342

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 296.3008 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	1.0094	Significance Level:	0.40133
ARCH(8)	Test Statistic:	1.5567	Significance Level:	0.13329
ARCH(12)	Test Statistic:	1.0654	Significance Level:	0.38589
ARCH(16)	Test Statistic:	1.1064	Significance Level:	0.34312
ARCH(20)	Test Statistic:	1.0560	Significance Level:	0.39157
ARCH(24)	Test Statistic:	0.9816	Significance Level:	0.48791

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	4.1759	Significance Level:	0.04100
LB(8)	Test Statistic:	11.6326	Significance Level:	0.04018
LB(12)	Test Statistic:	12.2588	Significance Level:	0.19911
LB(16)	Test Statistic:	16.7134	Significance Level:	0.21274
LB(20)	Test Statistic:	20.2066	Significance Level:	0.26380
LB(24)	Test Statistic:	23.2007	Significance Level:	0.33337

* Sign and Size Bias Tests *

The Sign Bias Test
    Test Statistic: 1.1301   Significance Level: 0.25864
The Negative Size Bias Test
    Test Statistic: -0.8061   Significance Level: 0.42034
The Positive Size Bias Test
    Test Statistic: 1.4805   Significance Level: 0.13897
The Joint Test for the Three Effects
    Test Statistic: 1.4029   Significance Level: 0.24029

****

Pääkomponentti 16.

```
**********************************************************************
*  Estimating an ARMA(4,0) for the mean of PK16                      *
*   w/ a GARCH(1,1) model for the conditional variance               *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1352 Degrees of Freedom 1345
Function Value 3311.09093456

Variable	Coeff
1. A1	0.6614761863
2. A2	0.1132618679
3. A3	0.1280398773
4. A4	0.0652519138
5. C	0.0000244268
6. Q1	0.4938311021
7. P1	0.6279215576

Estimation by BHHH
Iterations Taken 236
Usable Observations 1352 Degrees of Freedom 1345
Function Value 3332.79189075

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	0.7901	0.0298	26.49945	0.00000000
2. A2	0.0943	0.0245	3.84734	0.00011941
3. A3	0.1606	0.0364	4.41298	0.00001020
4. A4	-0.0820	0.0270	-3.03115	0.00243626
5. C	2.5190e-005	3.8766e-006	6.49797	0.00000000
6. Q1	0.5305	0.0339	15.62733	0.00000000
7. P1	0.5966	0.0191	31.20898	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-1.30496	16.18606

The Ljung-Box Q-Test for Serial Correlation in NRESIDS
LB(8) Test Statistic: 12.2374 Significance Level: 0.01567
LB(16) Test Statistic: 31.5520 Significance Level: 0.00162
LB(24) Test Statistic: 42.2821 Significance Level: 0.00254

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic:10178.5342 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.2857	Significance Level:	0.88739
ARCH(8)	Test Statistic:	0.2355	Significance Level:	0.98428
ARCH(12)	Test Statistic:	0.2101	Significance Level:	0.99806
ARCH(16)	Test Statistic:	0.1639	Significance Level:	0.99993
ARCH(20)	Test Statistic:	0.1537	Significance Level:	0.99999
ARCH(24)	Test Statistic:	0.1737	Significance Level:	1.00000

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(8)	Test Statistic:	1.7994	Significance Level:	0.77259
LB(16)	Test Statistic:	2.5604	Significance Level:	0.99793
LB(24)	Test Statistic:	4.2880	Significance Level:	0.99992

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: -0.2755 Significance Level: 0.78301
The Negative Size Bias Test
Test Statistic: 0.1249 Significance Level: 0.90061
The Positive Size Bias Test
Test Statistic: -0.4574 Significance Level: 0.64747
The Joint Test for the Three Effects
Test Statistic: 0.0843 Significance Level: 0.96862

****

Pääkomponentti 17.

```
**********************************************************************
*  Estimating an ARMA(3,0) for the mean of PK17                     *
*    w/ a GARCH(1,1) model for the conditional variance             *
**********************************************************************
```

A1 is the constant in the mean equation
The A2->An coefficients refer to the AR equation for the mean
C is the constant in the conditional variance equation
The Q coefficients refer to the lagged squared residuals
The P coefficients refer to the lagged conditional variance
The numbers in the A,MA,Q,P coefficients refer to the lag

Estimation by Simplex
Usable Observations 1353 Degrees of Freedom 1346
Function Value 3082.32783279

Variable	Coeff
1. A1	-0.000464347
2. A2	0.657737900
3. A3	0.037558550
4. A4	0.136597206
5. C	0.000086519
6. Q1	0.356843038
7. P1	0.598999578

Estimation by BHHH
Iterations Taken 60
Usable Observations 1353 Degrees of Freedom 1346
Function Value 3105.13386850

Variable	Coeff	Std Error	T-Stat	Signif
1. A1	-1.1330e-006	5.6583e-004	-0.00200	0.99840240
2. A2	0.5732	0.0285	20.12676	0.00000000
3. A3	0.1558	0.0352	4.42650	0.00000958
4. A4	0.0673	0.0302	2.22450	0.02611465
5. C	2.2495e-005	3.3888e-006	6.63791	0.00000000
6. Q1	0.1166	0.0148	7.86861	0.00000000
7. P1	0.8565	0.0138	62.02951	0.00000000

* Analyzing the normalized residuals for serial correlation and normality *

Skewness	Kurtosis
-0.52075	6.17336

The Ljung-Box Q-Test for Serial Correlation in NRESIDS

LB(4)	Test Statistic:	8.9427	Significance Level:	0.00279
LB(8)	Test Statistic:	11.1883	Significance Level:	0.04777
LB(12)	Test Statistic:	16.3604	Significance Level:	0.05973
LB(16)	Test Statistic:	21.2017	Significance Level:	0.06903
LB(20)	Test Statistic:	27.3714	Significance Level:	0.05284
LB(24)	Test Statistic:	29.3613	Significance Level:	0.10559

The Jarque-Bera Normality Test, ChiSqr(2), for NRESIDS
Test Statistic: 628.8600 Significance Level: 0.00000

F-Test of no ARCH vs. ARCH in NRESIDS

ARCH(4)	Test Statistic:	0.2581	Significance Level:	0.90480
ARCH(8)	Test Statistic:	1.5987	Significance Level:	0.12047
ARCH(12)	Test Statistic:	1.2307	Significance Level:	0.25585
ARCH(16)	Test Statistic:	1.1299	Significance Level:	0.32092
ARCH(20)	Test Statistic:	0.9555	Significance Level:	0.51517
ARCH(24)	Test Statistic:	0.8442	Significance Level:	0.68110

* Analyzing the squared normalized residuals for serial correlation *

The Ljung-Box Q-Test for Serial Correlation in SQNRESIDS

LB(4)	Test Statistic:	0.9742	Significance Level:	0.32363
LB(8)	Test Statistic:	12.2647	Significance Level:	0.03134
LB(12)	Test Statistic:	14.9982	Significance Level:	0.09099
LB(16)	Test Statistic:	17.9385	Significance Level:	0.15987
LB(20)	Test Statistic:	19.3800	Significance Level:	0.30716
LB(24)	Test Statistic:	20.1837	Significance Level:	0.50965

* Sign and Size Bias Tests *

The Sign Bias Test
Test Statistic: 0.9155 Significance Level: 0.36011
The Negative Size Bias Test
Test Statistic: -1.2132 Significance Level: 0.22526
The Positive Size Bias Test
Test Statistic: 0.3387 Significance Level: 0.73492
The Joint Test for the Three Effects
Test Statistic: 1.5755 Significance Level: 0.19349

****