

VALUE AT RISK: EKSPONENTIAALISESTI PAINOTETTU HISTORIAALLINEN SIMULAATIO

Jyväskylän Yliopiston Kauppakorkeakoulu
Taloustiede
Pro Gradu-tutkielma
12.8.2014
Tekijä: Lauri Kallio
Ohjaaja: Juha Junntila

JYVÄSKYLÄN YLIOPISTON KAUPPAKORKEAKOULU

Tekijä Lauri Leo Kallio	
Työn nimi Value at Risk: Eksponentiaalisesti painotettu historiallinen simulaatio	
Oppiaine Taloustiede	Työn laji Pro Gradu-tutkielma
Aika 12.8.2014	Sivumäärä 51
Tiivistelmä – Abstract <p>Tämän tutkimuksen tarkoituksena oli toteuttaa Value at Risk-malleja vertaileva backtesting-tutkimus. Malleiksi valikoituivat perinteinen painottamaton historiallinen simulaatio ja eksponentiaalisesti painotettu historiallinen simulaatio. Tavoitteena oli tutkia kannattaako havaintoja painottaa koska molemmat mallit olivat siinä mielessä vertailukelpoisia, että painotettu malli toteutettuna painotuskertoimella 1, eli kaikki havainnot saavat saman painon vastaa painottamatonta variaatiota. Lähtökohtana oli, että painottaminen on kannattavaa koska painottamattomassa versiossa pienillä havaintomäärillä havaintojoukko ei ole edustava ja suurilla havaintomäärillä taas malli ei kykene huomioimaan markkinoiden dynamiikkaa etenkin volatilitietin suhteen. Tutkimus toteutettiin niin, että molemmille malleille laskettiin yhden päivän VaR-estimaatit 99 % luottamustasolla. Käytetyt havaintomäärät olivat 250, 750 ja 1500 ja painotuskertoimena käytettiin arvoa 0,99. Aineistoina toimi neljä osakeindeksiä: S&P 500, OMX Helsinki, Hang Seng ja MSCI Emerging Markets. Päivittäisiä havaintoja oli 20 vuodelta ja VaR-estimaatteja laskettiin kullekin havaintomäärälle ja mallille 5000 kpl. VaR-estimaattien tarkkuutta mitattiin tämän jälkeen Christoffersenin kaksiosaisella testillä. Tämän jälkeen toteutettiin painotuskertoimen optimointi edellä mainitun Christoffersenin testin suhteen. Eksponentiaalisesti painotettu historiallinen simulaatio on tulosten valossa tarkempi malli ja tämän perusteella näyttäisi siltä, että painottaminen on kannattavaa ja parantaa VaR-estimaattien tarkkuutta.</p>	
Asiasanat Value at Risk, historiallinen simulaatio, eksponentiaalinen painotus, backtesting, Christoffersenin testi	
Säilytyspaikka Jyväskylän yliopiston kauppakorkeakoulu	

KUVIOT

KUVIO 1 - S&P 500-indeksi.....	34
KUVIO 2 - OMXH-indeksi	34
KUVIO 3 - Hang Seng-indeksi.....	35
KUVIO 4 - MSCI Emerging Markets-indeksi	35
KUVIO 5 - S&P 500 250 havaintoa.....	39
KUVIO 6 - S&P 500 1500 havaintoa.....	40
KUVIO 7 - S&P 500 750 havaintoa.....	40
KUVIO 8 - OMXH 250 havaintoa	48
KUVIO 9 - OMXH 750 havaintoa	48
KUVIO 10 - OMXH 1500 havaintoa	49
KUVIO 11 - Hang Seng 250 havaintoa	49
KUVIO 12 - Hang Seng 750 havaintoa	49
KUVIO 13 - Hang Seng 1500 havaintoa	50
KUVIO 14 - MSCI Emerging Markets 250 havaintoa.....	50
KUVIO 15 - MSCI Emerging Markets 750 havaintoa.....	50
KUVIO 16 - MSCI Emerging Markets 1500 havaintoa.....	51

TAULUKOT

TAULUKKO 1 - VaR-estimaatin ylitykset.....	37
TAULUKKO 2 - Christoffersenin ehdottoman kattavuuden testi	37
TAULUKKO 3 - Christoffersenin riippumattomuustesti	38
TAULUKKO 4 - Christoffersenin ehdollisen kattavuuden testi	38
TAULUKKO 5 - Optimaalinen painotuskerroin	42

SISÄLLYS

TIIVISTELMÄ
KUVIOT JA TAULUKOT
SISÄLLYS

1	JOHDANTO	7
2	VALUE AT RISK	10
	2.1 Value at Risk yleisesti	10
	2.2 RiskMetrics-lähestymistapa	11
	2.3 Monte Carlo-simulaatiot	12
	2.4 Ääriarvoteoria	12
	2.5 Historiallinen simulaatio	13
3	HISTORIALLINEN SIMULAATIO	14
	3.1 Perinteinen historiallinen simulaatio	14
	3.2 Eksponentiaalinen painotus	15
	3.3 Bootstrap	16
4	BACKTESTING	19
	4.1 Baselin testi	19
	4.2 Kupiec-testi	19
	4.3 Christoffersenin testi	19
	4.4 Rosenblattin transformaatio	20
	4.5 Tappiofunktio	22
5	AIKAISEMPI TUTKIMUS	24
	5.1 Perinteisen historiallisen simulaation estimointitarkkuus	24
	5.2 Ajan perusteella painotetun historiallisen simulaation estimointitarkkuus	28
	5.3 Vertailevaa tutkimusta ja johtopäätöksiä mallien eroista	30
6	EMPIIRINEN OSIO	33
	6.1 Aineisto ja menetelmät	33
	6.2 Vertaileva osio	36
	6.3 Painotuskertoimen optimointi	41
7	LOPPUPÄÄTELMÄT	43
	LÄHTEET	44
	LIITTEET	48

1 JOHDANTO

Viimeisin talouskriisi eli Yhdysvaltojen asuntomarkkinoilta maailmalle levinnyt rahoituskriisi osoitti jälleen rahoitusalan riskienhallinnan riittämättömyyden. Alan riskienhallinnan standardiksi muodostunut Value at Risk-menetelmä on koko historiansa ajan ollut kiistelty ja kritisoitu lähestymistapa riskien mittaamiseen sekä niiden hallintaan. VaR-menetelmän soveltamisalasta ja hyvydestä voidaan olla eri mieltä, mutta riskien kvantifiointi on välttämätöntä varsinkin rahoitusosalalla ja lyhyellä tähtämellä riskienhallinnan kehitys on ainakin toistaiseksi sidottu VaR-mallien kehitykseen.

Hedgerahasto Long-Term Capital Managementin (LTCM) ajautuminen maksukyvyttömäksi vuonna 1998 käänsi katseet sekä kritiikin Value at Risk-malleja kohtaan koska rahaston tiedettiin käyttäneen VaR-menetelmää sijoitustensa riskin estimoimiseen. Jorion (2000, 299) ei kuitenkaan näe, että ongelma olisi pelkästään tai edes suuressa määrin VaR-menetelmässä. Hän argumentoi, että muiden toimijoiden VaR-mallit estimoivat riskiä vuonna 1998 verrattain hyvin ja hän kääntäisikin katseen riskienhallintaan yleisemmin. Esimerkiksi LTCM:n tapauksessa syyt olivat kymmenkertainen vivutus ja optimointiharha, joka syntyi kun rahasto optimoi tuottoensa VaR-estimaatin perusteella. Lisäksi hän huomauttaa, että VaR ei huomio likviditeettiriskiä. Nykyisen vuonna 2008 alkaneen rahoituskriisin osalta Campbell (2009) tuo esiin miten pankkien ja rahoituslaitosten VaR-estimaatit reagoivat oikealla tavalla vuoden 2008 tapahtumiin, etenkin Lehman Brothersin konkurssiin. Tutkija tuo kuitenkin esiin myös ongelmia, joita pankeissa on havaittu näihin menetelmiin liittyen. Liian vanhojen havaintojen käyttäminen laskelmien pohjana tarkoittaa sitä, että mallin dynamiikka katoaa lähes kokonaan. Tätä onkin moni toimija jo korjannut siirtymällä pienempiin havaintomääriin tai siirtymällä havaintojen painotukseen. Campbellin (2009, 46) esiin nostamista asioista todella mielenkiintoinen on Standard Chartered Bankin edustajan kommentti, jonka mukaan VaR toimi hyvin kaikilla muilla markkinoilla paitsi strukturoiduilla lainamarkkinoilla, jotka olivat silloin äärimmäisen epälikvidit. Näiden kahden eri kriisistä kertovan artikkelin perusteella näyttäisi siltä, että likviditeettiriski, jota VaR ei siis huomioi, on äärimmäisen keskeinen riskienhallinnan kannalta eikä riskienhallintaa voi rakentaa pelkästään VaR-menetelmän varaan. Muista menetelmistä hyvä esimerkki ovat stressitestit ja Jorion (2001, 231) muistuttaakin, että todella epätodennäköisten tapahtumien ennustaminen VaR-menetelmän avulla on lähes mahdotonta ja stressitestejä tulisi aina käyttää VaR-analyysin ohessa ja sen tukena.

McKinsey & Companyn (Mehta et al. 2012) julkaisemassa artikkelissa osoitetaan, miten VaR-menetelmä on nykyisin noussut riskienhallinnan standardiksi. Eikä tämä koske vain pankkeja ja rahoituslaitoksia vaan myös pankkivalvojia. Artikkelista ilmenee, että kaikki rahoitusalan toimijat käyttävät VaR-menetelmää ja jopa viranomaisten kehittämät Basel-komitean

vakavaraisuusvaatimukset pohjautuvat näihin laskelmiin. Samassa artikkelissa tuodaan esiin miten painottamaton historiallinen simulaatio on nykyisin ylivoimaisesti suosituin VaR-malli pankkien keskuudessa. Toiseksi suosituimmat ovat Monte Carlo-simulaatiot, joiden suosio on kuitenkin vähentynyt viimeisen kymmenen vuoden aikana. Perinteisesti MC-simulaatioiden suurimpana ongelmana on pidetty niiden suurta laskentatehotarvetta. Vähentynyt käyttö ei kuitenkaan näyttäisi tukevan tätä väitettä, koska tietokoneiden laskuteho on lisääntynyt viimeisen kymmenen vuoden aikana. Toinen todennäköinen syy siihen, että pankit valitsevat painottamattoman historiallisen simulaation MC-simulaation sijasta on robustisuus. Näin ollen tämän tutkimuksen lähtökohtana on se, että historiallinen simulaatio on alan standardi, johon muita malleja verrataan ja mallin suuren suosion syy piilee sen robustisuudessa. Tilastollisten oletusten vähäinen määrä ei kuitenkaan tule ilmaiseksi ja kuten seuraavaksi esitellään, painottamattomaan versioon liittyy tunnettu ongelma. Lisäksi on muistettava, että ”paholainen asuu hännissä” mikä kuvaa sitä kuinka hankalaa todella epätodennäköisten tapahtumien tilastollinen mallintaminen on.

Boudoukh, Richardson ja Whitelaw (1998) osoittavat miten painottamattomaan historialliseen simulaatioon liittyy kaksi ongelmaa. Ensinnäkin epätodennäköisten tapahtumien estimoiminen on hankalaa ilman suuria havaintomääriä. Toinen ongelma on oletus riippumattomasti ja identtisesti jakautuneista havainnoista, jonka seurauksena joudutaan olettamaan, että tuottojen volatilitteetti on ajan suhteen vakio. Ensimmäinen ongelma ratkeaa lisäämällä havaintojen määrää mutta samalla luovutaan ainoasta keinosta välttää toinen ongelma, koska vähäisellä määrällä havaintoja ajassa muuttuva volatilitteetti ei aiheuta yhtä suuria ongelmia. Tämän dilemman ratkaisuksi tutkijat esittävät havaintojen painotusta ajan suhteen. Näin kyetään käyttämään suurta havaintomäärää luopumatta nopeasta reagointikyvystä markkinoiden muutoksiin. Toinen mahdollinen hyöty syntyy vähäisillä havaintomäärillä, jos suuren tappion omaavat havainnot saavat suuremman painon ja helpottavat epätodennäköisten tapahtumien estimoimista. Tämän tutkielman tarkoituksena on vertailla eksponentiaalisesti painotetun historiallisen simulaation ja painottamattoman historiallisen simulaation estimointitarkkuutta. Suurin ero tämän tutkielman ja aikaisemman tutkimuksen välillä on käytettyjen aikasarjojen pituus ja viimeisimpien havaintojen tuoreus. Tässä tutkielmassa käytetyt aikasarjat ovat selkeästi pidempiä sekä viimeisimmät havainnot uudempia. Kolmas keskeinen ero aikaisempaan tutkimukseen löytyy kehittyvien markkinoiden painosta, joka tässä tutkielmassa on normaalia suurempi ja kattaa puolet käytetyistä aikasarjoista.

Seuraavassa luvussa esitellään Value at Risk-menetelmä yleisellä tasolla, sekä yleisimmät mallit sen arvon laskemiseen ja mallien paremmuuden arviointiin. Samassa yhteydessä esitellään finanssivalvonnan tuomia vaatimuksia malleille, niiden käytölle ja rajoituksille. Historiallinen simulaatio ja sen variaatiot esitellään tarkemmin sitä seuraavassa luvussa, jonka jälkeen esitellään backtesting-mallit, joilla testataan VaR-mallien hyvyttä. Tämän

jälkeen esitellään aikaisempien tutkimusten tuloksia kiinnostuksen kohteena olevien mallien osalta. Lopuksi toteutetaan vertaileva backtesting-osio vertailemalla mallien hyvyyttä Christoffersenin testillä ja estimoidaan painotetun historiallisen simulaation optimaalinen painotuskerroin.

2 VALUE AT RISK

Tässä luvussa esitellään Value at Risk-menetelmä yleisellä tasolla. Luku pohjautuu pitkälti Philippe Jorionin (2001) oppikirjaan. Erilaiset VaR-mallit esitellään pääpiirteissään ja historiallisiin simulaatioihin perehdytään tarkemmin seuraavassa luvussa. Tämän luvun yhteydessä tuodaan myös esiin Value at Riskin yhteys pankkien valvontaan ja pääomavaatimuksiin. Tältä osin keskitytään rahoitusmaailman kannalta keskeisiin Basel-komitean vaatimuksiin.

2.1 Value at Risk yleisesti

Jorion (2001, 22) pitää Value at Risk-termin keksijänä Till Guldin, joka työskenteli J.P. Morganilla 1980-luvun lopussa. Käsite Value at Risk syntyi, kun edellä mainitun investointipankin riskienhallintaosasto pyrki määrittelemään täydellisesti suojatun position. Arvonmuutokseen vaikuttavaa riskiä päädyttiin pitämään merkittävämpänä kuin tuloihin vaikuttavaa riskiä ja käteistä täydellisesti suojatun position määritelmänä.

Yksinkertaisesti ilmaistuna voidaan VaR-analyysin sanoa vastaavan seuraavaan kysymykseen: mikä on suurin mahdollinen tappio joka voidaan kokea valitulla aikaperiodilla ja halutulla luottamustasolla? Aikaperiodi valitaan usein yhden ja kymmenen päivän väliltä ja luottamustaso on tavallisesti joko 95 tai 99 %. VaR-luku löytyy täten tuottojakauman tappiohännästä halutulla merkitsevyydellä. Käytännössä VaR toimii niin, että esim. yhden päivän VaR-estimaatti 99 % luottamustasolla ylittyy kerran 100 päivässä. Vuodessa, olettaen 250 pankkipäivää, VaR-luvun pitäisi siis ylittyä 2,5 kertaa. Näin ollen VaR-mallin estimointitarkkuuden testaaminen vaatii sitä enemmän havaintoja mitä korkeampi luottamustaso tai pidempi ennustusperiodi on. Backtestauksen eli mallien estimointitarkkuuden testaamisen osalta Jorion (2001, 119) huomauttaa, että vaikka Basel-komitea vaatii laskemaan VaR-estimaatit kymmenelle päivälle 99 % luottamustasolla, voi estimaatit laskea Baselin ohjeistuksen mukaisesti yhdelle päivälle ja sen jälkeen muuntaa ne kymmenen päivän estimaateiksi kertomalla kymmenen neliöjuurella. Tämä helpottaa mallien backtestausta koska tarvittu havaintomäärä on kymmenen kertaa pienempi. Jorion huomauttaa samassa yhteydessä, että mallit kannattaa testata 95 % luottamustasolla, koska näin samalla havaintomäärällä pystytään paremmin arvioimaan mallin hyvyttä. Tässä tutkimuksessa tullaan noudattamaan Jorionin huomautusta estimointiperiodin suhteen, muttei luottamustason. Näkemykseni mukaan kaikki mallit eivät aina käyttydy samalla tavalla molemmilla luottamustasoilla

ja 95 % luottamustason tulokset eivät aina kuvaa 99 % luottamustason tuloksia. Tämä tulee esiin aikaisempaa tutkimusta käsittelevässä luvussa.

VaR-mallit voidaan jakaa kahteen eri ryhmään. Ensimmäisen ryhmän mallit ovat lokaaliin arvostukseen perustuvia malleja. Tämä tarkoittaa sitä, että riski arvioidaan nykyisellä tilanteella nykyiselle portfoliolle ja mahdollisia muutoksia mitataan lokaaleilla approksimaatioilla. Toisen ryhmän mallit perustuvat kokonaisarvostukseen ja portfolion riski määritellään simuloimalla arvonmuutokset monelle eri skenaariolle. (Jorion, 2001, 205)

Lokaaliin arvostukseen perustuvia malleja kutsutaan myös analyttisiksi tai parametrisiksi malleiksi. Kokonaisarvostukseen perustuvat mallit ovat simulaatioita, jotka voidaan jakaa historiallisiin ja Monte Carlo-simulaatioihin. Näiden lisäksi on vielä tullut käyttöön ääriarvoteoriaan perustuvia malleja.

2.2 RiskMetrics-lähestymistapa

Analyttisistä tai lokaaliin arvostukseen perustuvista malleista yleisin on delta normaali-malli. Sama malli tunnetaan myös varianssi-kovarianssi- tai RiskMetrics-mallina. Mallissa riskitekijöiden oletetaan olevan normaalijakautuneita ja arvonmuutosten olevan lineaarisia. Mallin pohjana on varianssi-kovarianssi-matriisi, jonka perusteella lasketaan portfolion volatilitteetti, joka on samalla portfolion riski. RiskMetrics-variaatioon tosin liitetään usein havaintojen painotus. Näistä kehittyneempi menetelmä on delta gamma-malli, jossa riskin muutoksia mitataan toisen asteen Taylor-polynomilla. Tässä mallissa joudutaan tosin luopumaan normaalijakaumasta. Jorion (2001, 220-221) kuvailee, että delta-normaali malli on laskennallisesti yksinkertainen ja helppo malli, mutta normalisuusoletus voi tuottaa ongelmia ja lineaarisuusoletus ei toimi epälineaarisesti käyttäytyvillä instrumenteilla. Johdannossa esittelystä McKinsey & Companyn (Mehta et al. 2012) tutkimuksesta tulee ilmi, ettei analyttisiä malleja enää käytetä pankki- ja rahoitusalailla.

Delta-normaalissa menetelmässä oletetaan, että on olemassa markkinariskitekijöitä, joiden logaritmiset hinnanmuutokset ovat yhteisesti normaalijakautuneita odotusarvolle nolla. Koska r_0 on logaritminen tuotto riskitekijästä 0 niin $r \sim (0, Q^*)$, missä Q^* on kovarianssimatriisi riskitekijöiden tuotoista. Näin ollen kaavaksi muodostuu:

$$VAR = m(x)^t \times Q \times m(x) \quad (1)$$

missä $m(x)$ on vektori, joka kuvaa portfolion altistumista riskitekijöille ja $m(x)^t$ on luonnollisesti tämän vektorin transpoosi. Q on normeerattu kovarianssimatriisi niin, että $Q = Z_\alpha^2 \times Q^*$, jossa Z_α on standardinormaali z-arvo ja α kuvaa haluttua merkitsevyytstasoa. (Mausser & Rosen, 1998, 6)

RiskMetrics-lähestymistavan (RiskMetrics Group, 1996, 83) eksponentiaalipainotus toteutetaan seuraavalla tavalla:

$$\sigma_{1,t+1|t}^2 = (1 - \lambda) \sum_{i=0}^{\infty} \lambda^i r_{1,t-1} \times r_{2,t-1} \quad (2)$$

missä r_1 ja r_2 edustavat haluttuja aikasarjoja tai varianssin ollessa kyseessä $r_1 = r_2$ ja λ on painotuskerroin. Painotuskertoimen osalta käytettäväksi ehdotetaan päivittäisille havainnoille 0,94 ja kuukausittaisille 0,97 (RiskMetrics Group, 1996, 97).

2.3 Monte Carlo-simulaatiot

Yleisesti käytetyistä VaR-malleista edistyneimmät ovat erilaiset Monte Carlo-simulaatiot. MC-simulaatioiden toteutus koostuu kahdesta osasta. Ensimmäisessä osassa määritetään stokastiset prosessit ja prosessien parametrit. Tämän jälkeen simuloidaan satunnaisesti monta tuhatta eri polkua näille muuttujille. Näistä tuloksista muodostetaan jakauma, josta voidaan laskea VaR-luku halutulla luottamustasolla. Jorion (2001, 225) nostaa MC-simulaatioiden suurimmaksi eduksi niiden tehokkuuden ja joustavuuden estimointitarkkuuden osalta. Heikkouksista Jorion (2001, 226) nostaa esiin suuren vaatimuksen laskentateholle ja malliriskin, joka käytännössä on seurausta MC-simulaatioiden joustavuuden tuomasta robustisuuden puutteesta. Mainituista heikkouksista ja vahvuuksista huomataan, että valittaessa MC- ja historiallisten simulaatioiden väliltä päädytään valitsemaan hitaan ja vaikean mutta tarkan, ja nopean ja helpon mutta hieman vähemmän tarkan väliltä. Historiallisten simulaatioiden suosio osoittaa, että laskentanopeus ja yksinkertaisuus saavat suuren painoarvon valittaessa sopivaa VaR-mallia.

2.4 Ääriarvoteoria

Ääriarvoteoriaa on jo pitkään käytetty laajalti VaR-metodologian ulkopuolella erittäin epätodennäköisten tapahtumien ennustamiseen. Teorian tarkoituksena on laajentaa keskeisestä raja-arvolausesta häntien jakauman mallintamiseen. Toisin sanoen tarkoituksena on esittää kertymäfunktion muoto katkaisupiste u :n jälkeen olevassa pisteessä x . (Jorion, 2001, 249-250)

Jakauman hännän kertymäfunktio on siis seuraavanlainen:

$$\hat{F}(y) = 1 - (N_u/N) \left[1 + \xi \frac{(x-u)}{\beta} \right]^{-1/\xi} \quad (3)$$

missä parametri $\beta > 0$ ja parametri ξ kuvaa hännän paksuutta, $\xi > 0$ viittaa paksuihin häntiin. N_u edustaa pisteen u jälkeen ilmeneviä havaintoja ja N kaikkia havaintoja. N_u/N tarkoituksena on varmistaa, että hännän todennäköisyyksien summa on yksi. VaR-estimaatti luottamustasolla c saadaan kun $\hat{F}(y) = c$, jolloin laskukaava muodostuu seuraavanlaiseksi:

$$VAR = u + \frac{\hat{\beta}}{\hat{\xi}} \left\{ [(N/N_u)(1 - c)]^{-\hat{\xi}} - 1 \right\} \quad (4)$$

(Jorion, 2001, 251)

Jorionin (2001, 252-253) mukaan perinteiset VaR-mallit toimivat verrattain hyvin vielä esimerkiksi 99 % luottamustasolla, mutta siirryttäessä vielä pidemmälle häntään muuttuu ääriarvoteoria hyödyllisimmäksi VaR-malliksi. Tietyn pisteen jälkeen normaalijakauma aliarvioi todennäköisyyksiä raskaasti ja jopa historialliset simulaatiot aliarvioivat riskiä koska käytetyt frekvenssijakaumat eivät välttämättä sisällä riittävän monta poikkeavaa havaintoa. Jorion kuitenkin muistuttaa, ettei ääriarvoteoriallakaan kyetä ennustamaan kaikkia äärimmäisen epätodennäköisiä tapahtumia, ja kuten aina, stressitestien kuuluisi täydentää VaR-analyysia.

2.5 Historiallinen simulaatio

Historiallisten simulaatioiden ideana on käyttää historiallista frekvenssijakaumaa tulevan tuottojakauman ennusteena. Eli tutkitaan kuinka paljon haluttu portfolio olisi tuottanut päivittäisiä tuottoja tai tappioita jos se olisi ollut olemassa x päivää. VaR-estimaatti saadaan simuloidun jakauman tappiohännästä halutulla merkitsevyytasolla eli kohdasta, jossa merkitsevyytason verran havaituista päivittäisistä tappioista ja tuotoista jää tappiohännässä VaR-luvun vasemmalle puolella eli merkitsevyytason verran toteutuneista tappiosta on suurempia kuin VaR-estimaatti. Jorion (2001, 222-223) pitää mallia erittäin intuitiivisena, robustina ja sen suurimpana etuna sitä, ettei mallissa jouduta juurikaan tekemään oletuksia jakauman muodosta, markkinoiden stokastisesta rakenteesta tai rahoitusinstrumenttien käyttäytymisestä. Malliin liittyy kuitenkin valintatilanne, joka voi muodostaa ongelmia. Käytettävien havaintojen määrän suhteen syntyy dilemma, jolloin vähän havaintoja tarkoittaa nopeaa reagointia, relevantteja havaintoja mutta kattamatonta havaintojoukkoa ja suuri määrä havaintoja vastakkaisia lähtökohtia (Jorion, 2001, 223-224). Kattamattomalla havaintojoukolla viitataan havaintojoukkoon, joka ei riitä kuvaamaan halutun kohteen historiallista kehitystä, esimerkiksi pieni havaintojoukko ei nousukausina sisällä suuria päivittäisiä tappioita ja näin ollen otetusta jakaumasta tulee yksipuolinen. Lisäksi on huomioitava, että havaintojen oletetaan olevan riippumattomia ja identtisesti jakautuneita. Käytetyistä havaintomääristä voidaan todeta, että Basel-komitean sallima minimimäärä on nykyisin 250 havaintoa ja harvemmin näkee malleja, joissa on käytetty enempää kuin 1500 havaintoa. Historiallisiin simulaatioihin perehdytään tarkemmin seuraavassa luvussa.

3 HISTORIALLINEN SIMULAATIO

3.1 Perinteinen historiallinen simulaatio

Perinteinen tai painottamaton historiallinen simulaatio on edelleen yleisimmin käytetty variaatio erilaisista historiallisista simulaatioista ja kuten jo johdannossa tuli ilmi historialliset simulaatiot ovat ylipäätään yleisempiä kuin muut lähestymistavat.

Painottamaton historiallinen simulaatio voidaan ilmaista seuraavasti:

$$r_{w,t} = \sum_{i=1}^N w_{i,t} r_{i,t} \equiv W_T' R_t, \quad t = 1, 2, \dots, T \quad (5)$$

$$VAR_{T+1|T}^\alpha \equiv r_w((T+1)\alpha) \quad (6)$$

missä $r_w((T+1)\alpha)$ saadaan suuruusjärjestyksessä olevasta simuloitujen tappioiden ja tuottojen jakaumasta $\{r_w(1), r_w(2), \dots, r_w(T)\}$. Jos $(T+1)\alpha$ ei ole kokonaisluku käytetään lineaarista interpolointia. Muuttuja $r_{i,t}$ kuvaa toteutunutta tappiota tai voittoa ja $w_{i,t}$ kyseisen havainnon saamaa painotusta, joka tässä tapauksessa on yhtä suuri kaikille havainnoille. Parametri α kuvaa haluttua merkitsevyytensä. (Zikovic ja Aktan, 2011)

Edellisessä luvussa esiteltiin historiallisiin simulaatioihin liittyviä ongelmia mutta kaksi näistä esitellään vielä tarkemmin koska muut variaatiot tästä simulaatiosta pohjautuvat pitkälti näiden ongelmien eliminoimiseen säilyttäen samalla painottamattoman simulaation tulosten robustisuuden. Jorion (2001, 224) osoittaa, että historiallisessa simulaatiossa joudutaan valitsemaan pitkän ja kattavan havaintojoukon tai lyhyen ja ei niin kattavan havaintojoukon väliltä. Käytännössä tämä tarkoittaa sitä, että kattava joukko kuten 1500 havaintoa sisältää havaintoja kuudelta vuodelta, jolloin mukana on erittäin todennäköisesti nousu- ja laskukausilta olevia havaintoja. Tällöin havaintojoukon voidaan olettaa kuvaavan kyseisen sijoitustuotteen historiallista kehitystä hyvinkin tarkasti. Tällöin kuitenkin voidaan joutua tilanteeseen, jossa monta vuotta vanhat nousukauden havainnot estävät mallia reagoimasta markkinoiden muutokseen, esimerkiksi finanssikriisiin. Toisaalta pienellä havaintojoukolla kuten 250 havaintoa voidaan olla tilanteessa, jossa kaikki viimeiseltä vuodelta olevat havainnot ajoittuvat nousukauden ajalle eikä havaintojoukossa ole juurikaan tappiollisia päiviä. Tämän seurauksena markkinoiden romahtaessa VaR-estimaatti aliarvioi riskiä hyvin rankasti, koska vääristyneen havaintojoukon perusteella mallissa oletetaan nousukauden olevan normaalitila. Toisaalta 250 havaintoon ei tarvitse lisätä monta suuren tappion omaavaa havaintoa, jotta VaR-estimaatti reagoi markkinoiden muutoksiin.

Toinen ongelma, joka vahvasti liittyy nimenomaan painottamattomaan historialliseen simulaatioon, on haamuilmiö. Boudoukh et al. (1998, 65)

esittämässä esimerkissä laskettaessa esimerkiksi 250 havainnolla 99 % luottamustason VaR-estimaattia voidaan päätyä tilanteeseen, jossa tällä viikolla koettujen kolmen todella pahasti tappiollisen päivän jälkeen VaR-estimaatti on vakio koko loppu vuoden, koska 1 % merkitsevyytaso löytyy toiseksi ja kolmanneksi tappiollisimman päivän välistä. Tämä siis oletuksella ettei aikaperiodilla tule vielä isompia tappioita. Näin saavutetaan tilanne, jossa VaR-estimaatin perusteella voisi olettaa, että markkinoilla on siirrytty pysyvään korkeamman riskin tilanteeseen vaikka todellisuudessa minkäänasteista muutosta ei ole tapahtunut. Isommilla havaintomäärillä kuten 1500 havainnolla voi käydä niin, että samalle kuuden vuoden periodille sattuu kaksi markkinaromahdusta, jolloin pahasti tappiollisia päiviä voi kertyä niin paljon, että VaR-estimaatti pysyy samassa äärimmäisen korkeassa lukemassa jopa monta vuotta putkeen. Päinvastainen haamuilmiö taas tapahtuu kun edellisen kaltaiset pahasti tappiolliset päivät putoavat pois liukuvasta ikkunasta. Tällöin VaR-estimaatti voi pudota päivässä tai viikossa merkittävästikin vaikka todellisessa riskissä ei tapahtuisi minkäänäköistä muutosta (Jorion, 2001, 224).

3.2 Eksponentiaalinen painotus

Jorionin mainitsemaa dilemmaa havaintojen määrään liittyen voidaan korjata painottamalla havaintoja kuten seuraavaksi esiteltävässä Boudoukhin, Richardsonin ja Whitelawn (1998) mallissa tehdään. Edellä mainittujen kehittämän hybridimallin tarkoituksena on yhdistää RiskMetrics-mallin mukainen eksponentiaalinen painotus perinteiseen historialliseen simulaatioon. Tämä toteutetaan seuraavalla tavalla (Boudoukh et al. 1998):

1. askel: $R(t)$ kuvaa toteutunutta tuottoa/tappiota hetkestä $t-1$ hetkeen t . Edellisiä K -määrä havaintoja: $R(t), R(t-1), \dots, R(t-K+1)$ painotetaan kertoimilla $[(1-\lambda)/(1-\lambda^K)], [(1-\lambda)/(1-\lambda^K)]\lambda, \dots, [(1-\lambda)/(1-\lambda^K)]\lambda^{K-1}$. Huomioitavaa on, että parametrisoinnilla $[(1-\lambda)/(1-\lambda^K)]$ on tarkoitus taata painojen yhteissumman olevan tasan 1.

2. askel: järjestä havainnot suuruusjärjestykseen.

3. askel: saadaksesi portfolion VaR-luvun x % luottamustasolla, tulee aloittaa suurimmasta tappiosta ja jatkaa painojen yhteen laskemista kunnes niiden kumulatiivinen määrä on x %. Jos jakauman x % kohdassa ei ole havaintoa, lasketaan VaR-luku kahden lähimmän väliltä lineaarisesti interpoloimalla.

Formaalimpi esitys mallista on seuraavanlainen:

$$VaR_{t+1,\alpha} = \sum_{j=t-N+1}^t r_j \times (\sum_{i=1}^N f_i(\lambda; N) I(r_{t+1-i} \leq r_j) = \alpha) \quad (7)$$

missä $f_i(\lambda; N)$ edustaa havainnon r_i painotuskerrointa ja $I(\cdot)$ on indikaattorifunktio (Aktan ja Zikovic, 2011).

Edellisen kaltaisella havaintojen painottamisella voidaan ainakin teoriassa vähentää tai jopa kokonaan päästä eroon havaintojen määrään liittyvästä dilemmasta. Reagointikyvyn pitäisi parantua suurilla havaintomäärillä ja pienillä havaintomäärillä havaintojoukkojen pitäisi muuttua hieman kattavimmiksi, jos tappiohännän relevantit havainnot saavat suuremman painon. Painotuksen pitäisi empiirisessä osiossa näkyä nimenomaan kohta esiteltävän Christoffersenin testin riippumattomuustestissä. Lisäksi painottamattoman historiallisen simulaation oletuksen havaintojen riippumattomuudesta ja identtisestä jakautuneisuudesta, joka osittain on epärealistinen, ei pitäisi vaikuttaa tuloksiin niin paljon koska kaikkien havaintojen ei enää oleteta olevan yhtä informatiivisia. Lisäksi painotuksen pitäisi poistaa haamuilmiöiden ilmeneminen. Koska vanhemmat havainnot saavat pienemmän painon poistuu edellisen alaluvun kaltaisen tilanteen mahdollisuus, jossa siis samat kolme havaintoa pitää VaR-estimaatin todella korkeana ja muuttumattomana pidemmän aikaa, jonka seurauksena näyttää siltä, että markkinoiden riskitaso on pysyvästi kohonnut. Tämä siksi, että ennen pitkää näiden kolmen vanhan havainnon yhteinen paino putoaa alle yhden prosentin merkitsevyytason, jolloin ne eivät enää määrää esimerkin mukaisen 99 % luottamustason VaR-estimaattia. Päinvastaisten haamuilmiöiden mahdollisuus katoaa myös, eli tilanteiden jolloin suuren päivittäisen tappion omaava havainto putoaa pois liukuvasta ikkunasta ja VaR-estimaatti osoittaa merkittävää riskin laskua vaikka markkinoiden riskitasossa ei ole tapahtunut mitään muutosta. Havaintoja painottamalla saavutetaan tilanne, jolloin pois putoavan havainnon paino on niin vähäinen, ettei VaR-estimaatissa tapahdu painottamattoman kaltaisia hyppäyksiä.

3.3 Bootstrap

Bootstrap-metodin tarkoituksena on soveltaa uudelleenotantaa, jolloin VaR-luku saadaan menetelmällä approksimoidusta otantajakaumasta. Tässä osiossa bootstrap-metodi esitellään sellaisena kuin se on tunnetussa Barone-Adesi et al. (2002) kehittämässä filtteroidussa bootstrap-mallissa. Artikkelia täydennetään filtteroinnin osalta saman tutkijaryhmän (Barone-Adesi et al. 1999) aikaisemmalla artikkelilla, jossa kyseinen menetelmä on esitelty tarkemmin. Samassa yhteydessä sivutaan myös ARMA-malleja ja GARCH-prosesseja.

Perinteisen historiallisen simulaation tarkoituksena on ottaa historiallinen frekvenssijakauma ja käyttää jakaumaa sellaisenaan VaR-luvun laskennan pohjana. Tästä edistyneempi menetelmä on bootstrap, jossa idea viedään pidemmälle. Yksinkertainen esimerkki metodin käytöstä on tilanne, jossa halutaan luoda 100 havainnon jakauma VaR-luvun laskemiseksi. Tällöin voidaan ottaa esimerkiksi 500 viimeisintä havaintoa ja valita satunnaisotannalla 100 havaintoa, joista muodostetaan haluttu uusi jakauma. Jo valittuja havaintoja

ei poisteta otantajakaumasta, joten sama havainto voidaan valita useampaan kertaan. Otannalla saatu uusi jakauma on edustava, kunhan alkuperäinen jakauma on sitä. (Jorion, 2001, 296-297)

Barone-Adesi et al. (2002, 34) tuovat artikkelissaan esiin miten perinteisen historiallisen simulaation oletus riippumattomista ja identtisesti jakautuneista havainnoista on virheellinen ja miten tämän seurauksena malli aliarvioi riskiä korkean volatilitteen aikana. Ongelman ratkaisuksi esitetään filterointia, joka yksinkertaisuudessaan tarkoittaa sitä, että historiallisesta frekvenssijakaumasta saadut havainnot muutetaan vastaamaan tämän hetken näkemystä sijoitusinstrumentin nykyisestä riskistä. Toisin sanoen havainnot standardisoidaan jakamalla nykyisellä volatilitteilla. Seuraavaksi esitellään metodi yksityiskohtaisemmin ja se miten tutkijaryhmä on toteuttanut oman versionsa bootstrap-simulaatiosta.

Lähtökohtana on, ettei aineistoon sovelleta minkäänlaista teoreettista jakaumaa ja täten käytetään pelkkää todellista empiiristä jakaumaa. Havainnoista pitää kuitenkin poistaa mahdollinen autokorrelaatio ja mahdolliset volatilitteen klusteroitumiset. Autokorrelaatio voidaan eliminoida käyttämällä liukuvaa keskiarvo. Volatilitteen klusteroitumisen poistamiseksi pitää mallintaa itse prosessi, jonka seurauksena nämä muodostuvat. Volatilitteen klusteroituminen voidaan havaita mallintamalla havainnot GARCH-prosesseina. Näin ollen tarvittava ARMA-GARCH(1,1) voidaan ilmaista seuraavalla tavalla:

$$r_t = \mu r_{t-1} + \theta \varepsilon_{t-1} + \varepsilon_t \quad \varepsilon_t \sim N(0, h_t) \quad (8)$$

$$h_t = \omega + \alpha(\varepsilon_{t-1} + \gamma)^2 + \beta h_{t-1} \quad (9)$$

missä μ on autoregressiivinen AR(1)-termin parametri, θ on liukuvan keskiarvon MA-termi, ω on vakio ja ε_t on virhetermi. GARCH(1,1)-yhtälössä virhetermin ε_t volatilitteettia selitetään vakiotermin ω ja edellisen periodin virhetermin ε_{t-1} ja edellisen periodin volatilitteen h_{t-1} avulla. Termit α ja γ määrittävät edellisen periodin virhetermin painon ja asymmetrian. Termi β taas määrittää edellisen periodin volatilitteen painon. Virhetermien standardisointi toteutetaan jakamalla estimoitu virhetermi $\hat{\varepsilon}_t$ vastaavalla volatilitteettiennusteella $\sqrt{\hat{h}_t}$. Näin ollen standardisoitu virhetermi saadaan seuraavasti:

$$e_t = \hat{\varepsilon}_t / \sqrt{\hat{h}_t} \quad (10)$$

missä e_t on GARCH-hypoteesin mukaisesti riippumaton ja identtisesti jakautunut, jonka seurauksena se soveltuu historiallisen simulaatioon. (Barone-Adesi et al. 1999, 585-586)

Tämän jälkeen voidaan suorittaa itse bootstrap-osio, joka tässä tapauksessa toteutetaan niin, että simuloidaan 5000 eri polkua seuraavalle

kymmenelle päivälle ja satunnaisotannalla valitut luvut palautetaan takaisin sarjaan niin, että sama luku voidaan käyttää moneen kertaan. Laskemisen pohjana käytetään kaavan 10 mukaisella kaavalla standardisoitujen lukujen e_t sarjaa. Näiden satunnaisotannalla valittujen lukujen pohjalta voidaan laskea varianssien polut, joiden perusteella saadaan se lopullinen jakauma, josta VaR-estimaatti saadaan. VaR-estimaatti saadaan lopullisesta simuloidusta jakaumasta samalla tavalla kuin perinteisessä historiallisessa simulaatiossa. Ensimmäiseksi valittu luku skaalataan huomiseksi ennustetulla volatilitteetilla seuraavalla tavalla:

$$z_{t+1} = e_1 \sqrt{h_{t+1}} \quad (11)$$

jonka tuloksena saadaan yhden päivän ennuste arvopaperin hinnanmuutoksesta p seuraavalle päivälle:

$$p_{t+1} = p_t + p_t(\mu r_t + \theta z_t + z_{t+1}) \quad (12)$$

jolloin kyseessä on siis kaavan 8 mukainen ARMA-yhtälö niin, että ε on standardisoitu ja skaalattu kaavojen 10 ja 11 mukaisesti. Tämä riittää jos tehdään VaR-ennuste vain yhdelle päivälle ja tämän jälkeen seuraava askel on uusia simulointi haluttu määrä kertoja. Kuitenkin esimerkin mukaisen kymmenen päivän VaR-estimaatin saamiseksi sama toimenpide toistetaan vielä yhdeksän kertaa niin, että toisena arvotulla luvulla ennustetaan riskiä kahden päivän päähän ensimmäisen päivän pohjalta, kolmantena arvotulla luvulla kolmen päivän päähän toisen päivän pohjalta ja niin edelleen kunnes saavutetaan haluttu kymmenen päivän taso. Yleisesti ottaen volatilitteetin h polku mallinnetaan seuraavasti:

$$h_{t+1} = (\omega + \alpha(z_{t+i-1} - \gamma)^2 + \beta h_{t+i-1}) \quad i \geq 2 \quad (13)$$

joka on siis kaavan 9 mukainen GARCH-prosessi, jossa ε on standardisoitu ja skaalattu samalla tavalla kuin edellä. Tämän kaavan perusteella pystytään laskemaan tarvittava volatilitteettiennuste koko periodille, tässä tapauksessa 10 päivälle. Useamman instrumentin portfolion ollessa kyseessä edellä mainittu toimenpide toistetaan kaikille instrumenteille käyttäen havaintoja samoilta satunnaisesti valituilta päiviltä. Tällöin instrumenttien väliset korrelaatiot ilmenevät implisiittisesti kuten perinteisessä historiallisessa simulaatiossa. (Barone-Adesi et al. 2002, 45-36)

4 BACKTESTING

4.1 Baselin testi

Basel-komitean (1996) testissä lasketaan VaR-estimaatin ylitysten määrä vuoden eli 250 päivän ajalta ja mallin hyvyys arvioidaan ylitysten määrän mukaan. Esimerkiksi käytettäessä 99 % luottamustasoa ja yhden päivän estimointiperiodia voi malli saada kolme erilaista arvosanaa riippuen siitä kuinka monta kertaa VaR-estimaatti ylittyy vuoden aikana. Alle viisi ylitystä antaa arvosanan vihreä ja malli hyväksytään. 5-9 ylitystä saa arvosanan keltainen ja mallia pitää parantaa. 10 tai enemmän ylitystä tarkoittaa punaista ja malli hylätään ja se pitää vaihtaa. Baselin testi antaa hyvin samankaltaisia tuloksia kuin seuraavaksi esiteltävä Kupiec-in testi, mutta Baselin testissä malleja ei rangaista riskin yliarvioimisesta.

4.2 Kupiec-testi

Kupiec-in (1995) testin lähtökohtana on se, että VaR-estimaatin ylitykset voidaan mallintaa binomijakaumalla arvon 1 edustaessa ylitystä ja 0 tilannetta, jolloin VaR-estimaatti ei ylity. Käytännössä tämä toimii niin, että $a^*=x/250$, jossa a^* on haluttu merkitsevyystaso prosenteissa ja x edustaa toteutuneita VaR-estimaatin ylityksiä. Testi, joka noudattaa asympotoottisesti $\chi^2(1)$ -jakaumaa ilmaistaan $(1-p)$ % luottamustasolla seuraavalla tavalla:

$$LR_{uc} = 2[\log(\alpha^{*x}(1 - \alpha^*)^{250-x}) - \log(p^x \times (1 - p)^{250-x})] \quad (14)$$

jossa uc viittaa ehdottomaan kattavuuteen. Ehdottoman kattavuuden malleissa keskitytään ainoastaan VaR-estimaatin ylitysten määrään ja oletetaan, etteivät VaR-estimaatin ylitykset tule ryppäissä.

4.3 Christoffersenin testi

Christoffersenin (1998) tavoitteena oli kehittää backtesting-malli, jossa huomioidaan se mahdollisuus, että VaR-estimaatin ylitykset tulevat ryppäissä. Toisin sanoen sen lisäksi, että huomioidaan ykkösten ja nollien määrä, eli alitusten ja ylitysten määrä, tarkistetaan myös se missä järjestyksessä ne ilmenevät. Näin ollen voidaan välttää sellaisen mallin hyväksyminen, joka välillä aliarvioi ja välillä yliarvioi selkeästi riskiä, mutta näyttäisi keskimäärin olevan tarkka. Tämänkaltaisen lähestymistavan mukaista mallia kutsutaan ehdollisen kattavuuden malliksi. Christoffersenin testin molemmat osatellit

noudattavat asympotoottisesti $\chi^2(1)$ -jakaumaa ja koko testi $\chi^2(2)$ -jakaumaa ja ne muotoillaan seuraavasti:

$$LR_{cc} = LR_{uc} + LR_{ind} \quad (15)$$

jossa cc viittaa ehdollisen kattavuuteen ja ind riippumattomuuteen.

$$LR_{uc} = -2[\log((1-p)^{n_0}p^{n_1}) - \log((1-\pi_1)^{n_0}\pi_1^{n_1})] \quad (16)$$

$$\pi_1 = n_1/(n_0 + n_1)$$

jossa n_1 kuvaa päiviä jolloin VaR-estimaatti ylittyy, n_0 päiviä jolloin se ei ylitä ja $(1-p)$ haluttua luottamustasoa.

$$LR_{ind} = -2[\log((1-\pi_2)^{(n_{00}+n_{10})}\pi_2^{(n_{01}+n_{11})}) - \log((1-\pi_{01})^{n_{00}}\pi_{01}^{n_{01}}(1-\pi_{11})^{n_{10}}\pi_{11}^{n_{11}})] \quad (17)$$

$$\pi_2 = (n_{01} + n_{11})/(n_{00} + n_{10} + n_{01} + n_{11})$$

$$\pi_{01} = n_{01}/(n_{01} + n_{00})$$

$$\pi_{11} = n_{11}/(n_{11} + n_{10})$$

jossa n_{ij} :n alaindeksi j kuvaa ylittyykö VaR-estimaatti tänään (1) vai ei (0) ja i edustaa samaa tilannetta edelliseltä päivältä.

4.4 Rosenblattin transformaatio

Yksi Christoffersenin testin suurista ongelmista on binomimuuttujien käyttäminen, koska ne voivat saada ainoastaan kaksi eri arvoa 1 tai 0 ja näistäkin 1 esiintyy VaR-mallien backtestauksen yhteydessä äärimmäisen harvoin. Toinen tapa on muuttaa kaikki toteumat sarjaksi riippumattomia ja identtisesti jakautuneita muuttujia. Tätä kutsutaan Rosenblattin transformaatioksi. Oletetaan, että meillä on stokastinen prosessi y_t , jonka todennäköisyyden tiheysfunktio on $f(y_t)$ ja siihen liittyvä kertymäfunktio $F(y_t) = \int_{-\infty}^{y_t} f(u)du$. VaR-mallit perustuvat käänteiseen kertymäfunktioon kuten:

$$\bar{y}_t = F^{-1}(\alpha) \quad (18)$$

missä α on haluttu merkitsevyytystaso. Esimerkiksi 99 % luottamustason VaR-luku on yhtä kuin \bar{y} niin, että todennäköisyys sille että $y < \bar{y}$ on 0,01. Näiden perusteella voidaan Rosenblattin transformaatio esittää seuraavalla tavalla:

$$x_t = \int_{-\infty}^{y_t} \hat{f}(u)du = \hat{F}(y_t) \quad (19)$$

missä y_t on toteutunut tappio tai voitto, $\hat{f}(\cdot)$ on tappion ennustettu esiintymistiheys ja x_t on riippumaton ja identtisesti jakautunut sekä noudattaa tasaisesti jakaumaa $(0,1)$. $\hat{F}(\cdot)$ on jakauma tehdyistä ennusteista, johon pankkivalvojat voivat soveltaa edellä esitettyä transformaatiota ja tämän jälkeen testata löytyykö rikkomuksia iid- tai tasajakaumaoletusta vastaan. Mahdollisten rikkomusten testaamiseen löytyy useita erilaisia menetelmiä ja rikkomukset tarkoittavat sitä, että VaR-malli hylätään. Seuraavaksi esitellään esimerkkitestit molemmille oletuksille. Lisäksi on huomioitava, ettei transformaation toimivuus muutu vaikka käytetyn portfolion koostumus tai VaR-malli muuttuisi. (Berkowitz, 2001, 466)

Oletetaan annetulle mallille generoitu jono $z_t = \Phi^{-1}(\hat{F}(y_t))$. Edellä tuotiin esiin miten z_t kuuluisi olla riippumaton havaintojen välillä ja standardi normaalisti jakautunut, jonka seurauksena sen ominaisuuksia voidaan testata monella eri tavalla. Yksi esimerkki on testata nollahypoteesia eli normalisuutta ensimmäisen asteen autoregressiivistä vaihtoehtoa vastaan, jonka odotusarvo eroaa nollassa tai varianssi yhdestä. Tämä voidaan ilmaista näin:

$$z_t - \mu = \rho(z_{t-1} - \mu) + \epsilon_t \quad (20)$$

missä edellä esitetyn nollahypoteesin mukaisesti $\mu = 0$, $\rho = 0$ ja $\sigma_{\epsilon_t}^2 = 1$. Kaavan 20 mukainen malli voidaan ilmaista myös log-uskottavuutena tuntemattomien parametrien osalta $L(\mu, \sigma^2, \rho)$. Havaintojen riippumattomuuden testi voidaan ilmaista seuraavalla tavalla:

$$LR_{ind} = -2(L(\hat{\mu}, \hat{\sigma}^2, 0) - L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})) \quad (21)$$

niin, että nollahypoteesin mukainen jakauma on $\chi^2(1)$. Tämä tarkoittaa sitä, että nollahypoteesi hylätään ja havainnot eivät ole riippumattomia jos testin tulos ylittää jakauman ja vapausasteen mukaisen kriittisen arvon. Testiä voidaan kuitenkin laajentaa kattamaan myös ensimmäisenä mainittu nollahypoteesi standardi normaalista jakaumasta. Tällöin odotusarvo ja varianssi muutetaan vastaamaan standardi normalisuusoletusta $(0,1)$ ja testi voidaan muotoilla molemmat hypoteesit kattavaksi seuraavasti:

$$LR = -2(L(0,1,0) - L(\hat{\mu}, \hat{\sigma}^2, \hat{\rho})) \quad (22)$$

jolloin jakauman pitäisi uuden yhdistetyn nollahypoteesin perusteella noudattaa $\chi^2(3)$ -jakaumaa, josta saadaan kriittinen arvo ja sen perusteella voidaan määrittää hylätäänkö nollahypoteesi ja ovatko havainnot riippumattomia ja noudattavat standardi normaalia jakaumaa. (Berkowitz, 2001, 468)

4.5 Tappiofunktio

Tappiofunktion tarkoituksena on luoda funktio, jonka minimointi tuottaa halutunlaisen eli toisin sanoen optimaalisen lopputuloksen. Testattaessa VaR-malleja tavoitteena ei ole kuitenkaan funktion lopputuloksen minimointi vaan tavoitteena on päästä mahdollisimman lähelle haluttua merkitsevyystasoa. Aluksi esitellään yksinkertainen malli, jossa oletetaan, että VaR-estimaatin ylitykset voidaan mallintaa binomijakaumalla. Tämän jälkeen esitellään Lopezin esittämä edistyneempi tappiofunktio, jossa huomioidaan myös VaR-estimaatin ylityksen suuruus.

Yleisesti VaR-mallien backtesting-tarkoituksiin sopivat tappiofunktiot voidaan esittää seuraavalla tavalla:

$$C_{mt+1} = \begin{cases} f(\varepsilon_{t+1}, VaR_{mt}) & \text{jos } \varepsilon_{t+1} < VaR_{mt} \\ g(\varepsilon_{t+1}, VaR_{mt}) & \text{jos } \varepsilon_{t+1} \geq VaR_{mt} \end{cases} \quad (23)$$

missä C_{mt+1} on itse tappiofunktio, $\varepsilon_{t+1} < VaR_{mt}$ kuvaa tilannetta, jolloin todellinen tappio on suurempi kuin VaR-estimaatti ja $f(x, y)$ ja $g(x, y)$ ovat funktioita, joille pätee annetulla y :lla $f(x, y) \geq g(x, y)$. Tämän jälkeen lasketaan yhteinen lopputulos koko otokselle seuraavalla tavalla, tässä tapauksessa yhdelle vuodelle eli 250 päivälle:

$$C_m = \sum_{i=1}^{250} C_{mt+1} \quad (24)$$

jonka jälkeen voidaan määrittää haluttu vertailulukku ja arvioida VaR-estimaatin hyvyttä. (Lopez, 1998, 121)

Binomijakaumaan perustuvassa tappiofunktiossa VaR-estimaatin ylitys saa arvon 1 ja muissa tapauksissa arvon 0. Tällöin tappiofunktio muodostuu seuraavanlaiseksi:

$$C_{mt+1} = \begin{cases} 1 & \text{jos } \varepsilon_{t+1} < VaR_{mt} \\ 0 & \text{jos } \varepsilon_{t+1} \geq VaR_{mt} \end{cases} \quad (25)$$

tällöin oikea C_{mt+1} :n vertailulukku 1 % merkitsevyystasolla on $E[C_{mt+1}] = 0,01$ ja koko otokselle $E[C_{mt}] = 2,5$, oletuksena 250 päivää. (Lopez, 1998, 121)

Edellisessä esimerkissä kiinnostuksen kohteena oli pelkästään VaR-estimaatin ylitysten määrä, mutta myös ylitysten suuruudella on merkitystä. Esimerkiksi Lopezin (1998, 122) esittelemä ylityksen suuruuden neliöintiin perustuva testi on seuraavanlainen:

$$C_{mt+1} = \begin{cases} 1 + (\varepsilon_{t+1} - VaR_{mt})^2 & \text{jos } \varepsilon_{t+1} < VaR_{mt} \\ 0 & \text{jos } \varepsilon_{t+1} \geq VaR_{mt} \end{cases} \quad (26)$$

Malli toimii kuten edellä mainittu binomimalli, mutta nyt suuremmat VaR-estimaatin ylitykset saavat suuremman painon kuin pienet ylitykset. C_m

saadaan kaavan 24 mukaisesti, kuten aikaisemminkin. Koska tappiofunktiossa huomioidaan myös ylitysten suuruus, voidaan tämän avulla saada tarkempi kuva siitä miten hyvin VaR-malli toimii taustalla olevan jakauman f_{t+1} hännässä. Vertailuluvun määrittäminen muodostuu kuitenkin hankalaksi koska jakauman f_{t+1} muoto on tuntematon. (Lopez, 1998, 122)

Lopez (1998, 124) kuitenkin huomauttaa, etteivät tappiofunktioon perustuvat mallit korvaa hypoteesitestaukseen perustuvia malleja kuten Kupiec ja Christoffersen, vaan ne tulevat näiden mallien tueksi antamaan lisäinformaatiota mallien toimintakyvystä.

5 AIKAISEMPI TUTKIMUS

Seuraavaksi esitellään aikaisempaa tutkimusta perinteisen ja ajan perusteella painotetun historiallisen simulaation osalta. Tässä luvussa esiteltävä tutkimus voidaan jakaa kahteen osaan, kuten käytännössä kaikki muukin näitä malleja koskeva tutkimus. Joko tutkimuksissa on ollut mukana perinteinen versio tai molemmat. Tämä heijastelee perinteisen asemaa alan standardina. Edeltävä tutkimus esitetään kolmessa osassa. Ensimmäisessä osassa esitellään painottamatonta historiallista simulaatiota koskevaa tutkimusta, toisessa osassa painotetun historiallisen simulaation tutkimusta ja kolmannessa kahden ensimmäisen osan johtopäätöksiä sekä vertailevaa tutkimusta mallien eroista. Käytännössä tämä tarkoittaa sitä, että osa vertailevasta tutkimuksesta esitellään kolmeen kertaan, tosin eri perspektiiveistä. Yksinkertainen syy tähän on se, että tässä tutkimuksessa kiinnostuksen kohteena olevat mallit ovat harvoin tutkimusten keskiössä. Tästä syystä monissa tutkimuksissa painottamattomassa ja ajan perusteella painotetussa menetelmässä on käytetty eroavia havaintomääriä. Tämän tutkimuksen kysymyksenasettelu on sen suuntainen, että eroavilla havaintomäärillä lasketut perinteinen ja ajan perusteella painotettu historiallinen simulaatio eivät ole suoraan vertailukelpoisia. Näin ollen osa aikaisemmasta tutkimuksesta antaa enemmänkin tietoa kyseisten mallien toimivuudesta kuin mallien välisistä eroista ja vertailevan tutkimuksen tulosten käsittelyn osalta keskitytään tutkimuksiin, joissa käytetyt havaintomäärät ovat vertailukelpoisia. Painottamaton historiallinen simulaatio, jossa on käytetty X -määrää havaintoja tullaan tässä osiossa esittelemään muodossa $HS(X)$ ja ajan perusteella painotettu, jossa on Y -painotuskerroin ja X määrä havaintoja tullaan ilmaisemaan sen kehittäjien nimikirjainten mukaan muodossa $BRW(Y; X)$. Tässä luvussa esiteltävien tutkimusten ennusteperiodi on yksi päivä, ellei toisin mainita.

5.1 Perinteisen historiallisen simulaation estimointitarkkuus

Seuraavaksi esitellään aikaisempaa tutkimusta perinteisen historiallisen simulaation osalta. Alkuun tutkimukset esitellään järjestyksessä käytetyn backtesting-menetelmän mukaisesti ja lopussa on tarkoitus esittää mahdollisia säännönmukaisuuksia.

Basel-komitean mukaista liikennevalotestiä ovat omissa tutkimuksissaan käyttäneet Soni (2005) ja Mutu, Balogh ja Moldovan (2011). Molemmissa tutkimuksissa käytetty malli oli $HS(250)$ ja riskiä ennustettiin 99 % luottamustasolla. Malli hyväksyttiin Sonin (2005) yhden vuoden periodilta tarkastelluilla intialaisilla korkoswapsalkuilla, vaikkakin malli hieman yliarvioi riskiä. Mutun et al. (2011) tutkimusryhmän käyttämällä itäeurooppalaisilla kuuden vuoden periodilta olevilla viidellä osakeindeksillä saatiin vastakkaisia

tuloksia. Tulosten mukaan malli hylätään jokaisella indeksillä, koska se aliarvioi riskiä.

Seuraavissa kehittyneiden maiden havaintoaineistoilla toteutetuissa tutkimuksissa käytettiin Kupiecien testiä. Historiallinen simulaatio menestyy heikoimmin Changchienin, Linin ja Kaon (2012) sekä Angelidisin, Benosin ja Degiannakis (2007) tutkimuksissa. Ensin mainitussa (Changchien et al. 2012) tutkittu HS(500) aliarvioi riskiä tutkimuksessa käytetyillä viidellä eri futuurilla ja neljällä eri luottamustasolla, väliltä 90 ja 99,5 %. Huomioitavaa on, että tutkimuksessa on käytetty havaintoja hyvin pitkältä ajalta, lähes 20 vuodelta. Angelidisin et al. (2007) ryhmän tuloksissa HS(1750)-mallin hylkäys johtuu riskin yliarvioinnista. Käytössä olivat 97,5 ja 99 % luottamustasot ja käytetty aineisto kattoi kaksi DJ Euro Stoxx-indeksiä yhdeksältä vuodelta. Engelin ja Gizyckin (1999) tutkimuksessa käytettyjen kymmeneltä vuodelta olevien valuuttakurssisarjojen osalta esiin nousee trendi, jonka mukaan malli aliarvioi riskiä pienillä havaintomäärillä ja yliarvioi suurilla. Tutkimuksessa testatut mallit olivat HS(125, 250, 500, 750 ja 1250). 95 % luottamustasolla tämä tarkoittaa sitä, että ainoastaan HS(250 ja 500) hyväksytään ja 99 % osalta ainut hyväksytty on HS(750). Luottamustasojen ero näyttää, kuinka riskin aliarviointi tässä tapauksessa lisääntyy luottamustason kasvaessa. Perignon ja Smith (2008) käyttivät viiden ison pankin oikeaa sijoitussalkkua kolmelta vuodelta ja HS(250) hyväksyttiin jokaiselle salkulle 99 % luottamustasolla, mutta estimointitarkkuus heikkeni hieman luottamustason laskiessa. Gustafssonin ja Lundbergin (2009) tutkima HS(2000) hylättiin 95 % luottamustasolla Tukholman pörssin indeksillä ja kolmen kuukauden Ruotsin valtion liikkeelle laskemilla rahamarkkinainstrumenteilla. 99 % luottamustasolla mallin ennustekyky parani mutta malli aliarvioi riskiä osakeindeksin osalta.

Kehittyneiden markkinoiden jälkeen siirrytään kehittyville markkinoille ja esiteltävissä tutkimuksissa käytetty backtesting-menetelmä on edelleen Kupiecien testi. Ensin esitellään kehittyvien markkinoiden osakeindeksejä käyttävistä tutkimuksista seuraavien tutkijoiden tutkimukset: Andjelic, Djakovic ja Radisic (2010) ja Hsieh ja Chou (2008). Viimeisenä mainitut (Hsieh & Chou, 2008) testasivat mallia HS(242) neljälle kiinalaiselle osakeindeksille, joista käytettiin vuoden tutkimusperiodia. Malli aliarvioi hieman riskiä 95 % luottamustasolla mutta se hylättiin vain yhdellä indeksillä. Andjelicin et al. (2010) tarkastelussa olivat HS(50, 100, 200 ja 250) ja tutkimusaineistona toimi neljä eurooppalaista kehittyvien markkinoiden osakeindeksiä, joista jokainen oli jaettu kolmeen vuoden mittaiseen periodiin. Tutkimuksessa käytetyt 95 ja 99 % luottamustasot eivät vaikuttaneet merkittävästi lopputulokseen, kuten ei myöskään valittu havaintomäärä. Tässä tapauksessa käytännössä kaikki mallit hylättiin kahdella ensimmäisellä aikaperiodilla mutta hyväksyttiin viimeisellä. Skiadopoulos, Lambadiaris, Papadopoulou ja Zoulis (2003) tutkivat HS(100 ja 252)-mallia kreikkalaisilla osake- ja korkopaperiportfolioilla. Molemmille aineistoille laskettiin estimaatit vuoden mittaiselle periodille. Osakeportfolion osalta molemmille havaintomäärille saadaan hyväksyviä tuloksia niin 95 % kuin myös 99 % luottamustasolla vaikka molemmat mallit hieman yliarvioivat

riskiä. Korkopapereiden osalta mallit hyväksytään mutta pientä riskin aliarvioimista ilmenee. Mielenkiintoista on, että HS(252) yliarvioi osakkeilla ja aliarvioi korkopapereilla enemmän kuin mitä HS(100) yli- ja aliarvioi kyseisillä aineistoilla. Yawalkar (2004) havaitsi tutkiessaan historiallisen simulaation estimointitarkkuutta intialaisilla osakkeilla ja joukkovelkakirjalainoilla, että malli aliarvioi riskiä niin 95 kuin 99 % luottamustasolla.

Tutkimukset, joissa on käytetty Christoffersenin testiä, esitellään kolmessa osassa. Ensimmäisenä vuorossa ovat tutkimukset, joiden tutkimusaineistona on käytetty kehittyneiden markkinoiden rahoitusinstrumentteja ja osakeindeksejä. Sen jälkeen esitellään tutkimusta kehittyvien markkinoiden osalta ja lopuksi esiin tuodaan tutkimusta, jossa on käytetty molempien markkinoiden aineistoa. Ottink (2009) käytti tutkimuksessaan mallia HS(260), Fierli (2004) vuorostaan HS(500) ja Billinger ja Eriksson (2009) käyttivät HS(250 ja 1000). Ottinkin (2009) tutkimuksessa käytetyille kolmelle eri maturiteetin omaavalle Italian liikkeelle laskemalle valtionlainalle saatiin seuraavanlaisia tuloksia: 95 ja 99 % luottamustasoilla malli estimoiti tarkasti lyhimmän maturiteetin riskiä, mutta mallin estimointitarkkuus heikkeni maturiteetin pidentyessä ja malli hylättiin muilla kuin lyhyimmällä maturiteetilla. Tutkimus toteutettiin kolmen vuoden aineistolle. Fierlin (2004) aineistona toimi kaksi yhdysvaltalaisista osakeportfoliota kahden vuoden ajalta. Tutkimuksessa käytetty malli on tarkka 99 % luottamustasolla mutta hylätään 95 % tasolla. Billinger ja Eriksson (2009) toteuttivat tutkimuksensa S&P 500-osakeindeksille ja oikeaa pankin sijoitussalkkua jäljittelevälle portfoliolle ja malleja testattiin molemmilla aineistoilla neljän vuoden mittaisella periodilla. HS(250) on tarkempi molemmilla aineistoilla ja hyväksytään Christoffersenin riippumattomuustestissä, mutta paremminkaan pärjännyt malli ei läpäise Christoffersenin ehdottoman tai ehdollisen kattavuuden testiä.

Kehittyvillä markkinoilla tehdyistä tutkimuksista, joiden tuloksia arvioitiin Christoffersenin testillä, esitellään seuraavien tutkijoiden ja tutkimusryhmien tuloksia ja testattuja malleja: Angelidis ja Benos (2005) HS(1000), Zikovic (2008) HS(50, 100, 250 ja 500) ja Unal (2011), jonka tarkastelussa oli useita eri havaintomääriä käyttäviä perinteisiä historiallisia simulaatioita. Angelidisin ja Benosin (2005) tutkimuksessa käytettiin 15 vuoden aineistoa Kreikan osakemarkkinoilta. Malli läpäisi 99 % luottamustasolla Christoffersenin ehdottoman kattavuuden testin kaikilla paitsi yhdellä aineistolla. Riippumattomuustestin osalta tulos oli päinvastainen ja malli hyväksyttiin ainoastaan yhdellä aineistolla. Tämä johti mallien hylkäämiseen ehdollisen kattavuuden testissä. Zikovicin (2008) tutkimuksen aineistona toimi neljä itäeurooppalaista pörssi-indeksiä kahdelta vuodelta. 95 % luottamustasolla HS(250) on ainoa vaihtoehto. Siirryttäessä 99 % luottamustasolle huomataan, että estimointitarkkuus paranee mitä enemmän mallissa käytetään havaintoja. Unal (2011) käytti tutkimuksessaan turkkilaisia osakkeita ja indeksejä vajaalta 15 vuodelta. Eri aineistoista, luottamustasoista ja havaintomääristä saatiin 32 erilaista kombinaatiota. Huomattavaa on, että

malleista 30 läpäisee ehdottoman kattavuuden testin mutta ainoastaan kuusi läpäisee riippumattomuustestin.

Christoffersenin testiä hyödyntävistä tutkimuksista esitellään vielä kolme tutkimusta, joiden aineisto on kerätty niin kehittyneiltä kuin kehittyviltä markkinoilta. Näistä ensimmäisenä esitellään Zikovicin ja Filerin (2009) tutkimus. Testattuja malleja olivat HS(250 ja 500) ja näitä testattiin 99 % luottamustasolla kahdeksalle eri kehittyvien markkinoiden ja kahdeksalle kehittyneiden markkinoiden osakeindeksille neljän vuoden periodilla. Tulokset indikoivat, että estimointitarkkuus kasvaa havaintojen lisääntyessä. Esiin on myös nostettava tutkimuksen tulos, jonka mukaan mallit toimivat paremmin kehittyvillä markkinoilla. Brandolini ja Colucci (2012) tutkivat HS(500) estimointitarkkuutta yhdeksällä indeksillä, joiden joukossa oli mm. S&P 500 ja MSCI Emerging Markets. Tutkimuksessa testattiin niin 95 % kuin 99 % luottamustasoa ja tutkimus toteutettiin yhdentoista vuoden periodille. Mallia ei hyväksytty millekään indeksille kummallakin luottamustasolla. Yleisesti ottaen estimointitarkkuus oli parempi 95 % luottamustasolla. S&P 500-indeksin osalta malli hyväksyttiin riippumattomuustestissä mutta ehdottoman kattavuuden testi hylkää mallin. MSCI Emerging Markets-indeksin osalta havaittiin mielenkiintoinen tulos. Käytetty malli sai 99 % luottamustasolla tarkimmat estimaatit juuri tältä indeksiltä, kun taas 95 % luottamustasolla malli mittasi riskiä kaikkein heikoimmin kyseiselle indeksille. Ero syntyy riippumattomuustestissä, jossa tulos vaihtui loistavasta surkeaksi. Sitä miksi näin tapahtui, ei tutkimuksessa osattu selittää. Terzic, Milojevic ja Dzamic (2013) tutkivat samaa mallia HS(500) kuin edellisessä tutkimuksessa. Aineistoina toimi viiden vuoden osakeindeksihavainnot kahdelta kehittyneeltä ja kolmelta kehittyvältä markkinalta. Ehdottoman kattavuuden osalta malli hylättiin ainoastaan molemmilla kehittyneillä markkinoilla 99 % luottamustasolla. Riippumattomuustestin osalta malli hyväksyttiin niin 95 kuin 99 % luottamustasolla ainoastaan S&P 500-indeksillä. Näin ollen toinen kahdesta tapauksesta, jossa malli hyväksytään ehdollisen kattavuuden testissä, oli S&P 500-indeksille 95 % luottamustasolla.

Lopuksi esitellään tutkimusta, joissa on käytetty muita backtesting-metodeja. Osa seuraavista testeistä muistuttaa esim. Kupiecin testiä mutta koska seuraavissa tutkimuksissa ei sitä eksplisiittisesti ole tuotu esiin, käytössä olleisiin backtesting-malleihin ei tulla sillä nimellä viittaamaan. Ensimmäisenä esitellään kaksi tutkimusta, joissa estimointitarkkuutta on mitattu yksinkertaisella binomijakaumatestillä. Ensimmäinen näistä on Debin ja Banerjeen (2009) tutkimus. Siinä on muista tutkimuksista poiketen käytetty viikon ennusteperiodia. Mallissa käytettiin 200 viikkohavaintoa, luottamustasoina olivat 95 ja 99 % ja aineisto koostui monesta eri intialaisesta rahastosta seitsemän vuoden periodilta. Malli pärjasi hyvin koko aineistossa vaikka lievää riskin yliarviointia ilmenikin. Toinen tutkimus, jossa sovelletaan samanlaista backtesting-menetelmää, on Cabedon ja Moyan (2003). Testattu malli on HS(1750) 99 % luottamustasolla vuoden raakaöljydatalle. Malli aliarvioi hinnan muutoksia ja hylättiin. VaR-estimaatin ylitysten

prosentuaalisen osuuden perusteella mallien tarkkuutta mittaavia tutkimuksia ovat seuraavien tutkijoiden tutkimukset: Gencay, Selcuk ja Ulugulyagi (2003), Vlaar (2000) ja Beder (1995). Gencay (et al. 2003) tutkimusryhmineen tutki HS(1000)-mallia viidellä eri luottamustasolla, väliltä 95 ja 99,9 %. Tutkimus suoritettiin kahdelle osakeindeksille, Istanbulin pörssi ja S&P 500, tutkimusperiodin ollessa 15 vuotta. Malli aliarvioi riskiä Istanbulin osalta jokaisella luottamustasolla ja S&P 500-indeksillä kaikilla muilla paitsi 95 % luottamustasolla. Kokonaisuutena malli antaa tarkempia estimaatteja S&P 500-indeksille. Vlaar (2000) käyttää tutkimuksessaan 99 % luottamustasoa mutta muista poiketen 10 päivän estimointiperiodia. Testattavina malleina ovat HS(250, 550, 750, 1250 ja 2250). Malleja testattiin 25 hollantilaisella joukkovelkakirjalainasalkulla 15 vuoden ajalta. Tulosten perusteella malli on tarkka 750 havainnolla, kun taas sitä pienemmällä havaintomäärillä malli aliarvioi ja suuremmilla yliarvioi riskiä. Tutkija antaa mahdollisena selityksenä sen, että korkojen volatiliteetti Euroopassa on laskenut siirryttäessä 1980-luvulta 1990-luvulle. Ensimmäiset tutkimuksessa käytetyt havainnot ovat vuodelta 1985. Beder (1995) testasi malleja HS(100 ja 250) 95 ja 99 % luottamustasolla, käyttäen niin yhden kuin kymmenen päivän estimaatteja. Aineistoina toimi yksi korkopaperisalkku, yksi optiosalkku ja yksi näiden sekoitus. Tutkimuksessa havaittiin, että estimaattien tarkkuus vaihtelee hetkittäin aika paljonkin, kun havaintomäärä, ennusteperiodi tai luottamustaso muuttuu. Hendricks (1996) testasi tutkimuksessaan seuraavia malleja HS(125, 250, 500 ja 1250). Tavoitteena oli tutkia historiallisen simulaation estimointitarkkuutta eri havaintomäärillä, 95 ja 99 % luottamustasolla ja 1000 satunnaisesti valitulla valuuttaportfoliolla 12 vuodelta. Tutkimuksen perusteella historiallinen simulaatio on havaintomääristä riippumatta tarkka 95 % luottamustasolla mutta 99 % luottamustasolla ainoastaan HS(1250) voidaan laskea tarkaksi.

Kuten edellä esitellyistä tutkimuksista huomataan, historiallisen simulaation estimointitarkkuus vaihtelee hyvin paljon luottamustason, aineiston, backtesting-menetelmän tai tutkimusperiodin muuttuessa. Edellä esitelyjen tutkimusten pohjalta näyttäisi kuitenkin muodostuvan pari säännönmukaisuutta, jotka tulevat esiin useammassa tutkimuksessa. Ensinnäkin havaintomäärän kasvun myötä VaR-estimaatti muuttuu konservatiivisemmaksi. Toinen huomio liittyy Christoffersenin riippumattomuustestiin, jossa perinteinen historiallinen simulaatio näyttää pärjäävän huonosti.

5.2 Ajan perusteella painotetun historiallisen simulaation estimointitarkkuus

Tässä osiossa esitellään aikaisempaa tutkimusta ajan perusteella painotetun historiallisen simulaation osalta.

Fierli (2004), Zikovic (2008) ja Zikovic ja Filer (2009) tutkivat kaikki malleja BRW(0,97 ja 0,99; 250). VaR-estimaatit laskettiin 95 ja 99 % luottamustasolla yhdelle päivälle ja backtesting-menetelmänä käytettiin Christoffersenin testiä, lukuun ottamatta Zikovicin ja Filerin (2009) tutkimusta, jossa käytettiin ainoastaan 99 % luottamustasoa. Fierli (2004) toteutti testinsä kahdelle yhdysvaltalaisista osakkeista koostuvalle portfoliolle ja noin kahden vuoden periodille. Tulosten mukaan malli on molemmilla painotuskertoimilla heikko 99 % luottamustasolla mutta tarkka 95 % luottamustasolla. Lisäksi painotuskerroin 0,97 on tarkempi 95 % luottamustasolla, kun taas 99 % luottamustasolla eroa ei synny. Zikovic (2008) käytti tutkimuksessaan viittä itäeurooppalaista kehittyvien maiden osakeindeksiä ja mallien testaus toteutettiin kahden vuoden periodille. Molemmilla painotuskertoimilla varustetut mallit hyväksytään 95 % luottamustasolla kaikilla indekseillä ja Christoffersenin testin molemmilla osilla. 99 % luottamustasolla estimointikyky heikkenee ja painotuskerroin 0,99 näyttäisi olevan tarkempi kuin 0,97. Zikovicin ja Filerin (2009) tutkimuksen aineistona toimi 16 osakeindeksiä, joista puolet kehittyneiden ja puolet kehittyvien markkinoiden. Malleja testattiin neljän vuoden periodille. Kehittyneiden ja kehittyvien markkinoiden tulosten välillä ei näyttäisi olevan eroa, mutta painotuskerroin 0,99 näyttäisi olevan tarkempi. Edellisten kaltaista asetelmaa mallien osalta on käyttänyt myös Ottink (2009). Hän tutki mallia BRW(0,97; 260) 95 ja 99 % luottamustasolla yhden päivän estimaateille käyttäen Christoffersenin testiä. Tutkimus toteutettiin kolmelle eripituiselle joukkovelkakirjalle, joista hyviä tuloksia saatiin ainoastaan lyhimmän maturiteetin omaavalle. Yleisesti ottaen malli on tarkempi 95 % luottamustasolla.

Osassa tutkimuksista on ajan perusteella painotettua historiallista simulaatiota käytetty perinteisen historiallisen simulaation niin kutsutun haamuvaikutuksen eliminoimiseksi. Haamuvaikutus syntyy kun havainto, joka on päivältä jolloin kärsittiin suuri tappio, tippuu pois historiallisen simulaation liukuvasta ikkunasta. Tällöin mallin ennusteessa voi syntyä suurikin muutos vaikka todellisuudessa markkinoilla ei ole tapahtunut mitään siihen viittaavaa. Seuraavaksi esiteltävissä tutkimuksissa on luovuttu liukuvasta ikkunasta ja otettu käyttöön malli, jossa vanhoja havaintoja ei pudoteta pois laskelmista vaan malliin käytetty havaintomäärä lisääntyy joka päivä. Tämä on mahdollista koska painotuskerroin häivyttää vanhemmat havainnot hienovaraisesti. Tämän tyyllisistä tutkimuksista esitellään Billingerin ja Erikssonin (2009) ja Perignonin ja Smithin (2008) tutkimukset. Ensin mainitut (Billinger & Eriksson, 2009) käyttivät hyvin korkeaa 0,9999-painotuskerrointa ja jälkimmäiset (Perignon & Smith, 2008) eivät mainitse käyttämänsä painotuskerrointa. Billinger ja Eriksson (2009) aloittivat tutkimuksensa 1000 havainnolla ja viimeinen estimaatti laskettiin 2000 havaintoa käyttäen, eli tutkimus toteutettiin neljän vuoden periodille. Tutkimus toteutettiin S&P 500-indeksille ja oikeaa pankin portfoliota jäljittelevälle salkulle. Malli pärjää heikosti ja hylätään molemmilla Christoffersenin testeillä niin 95 kuin 99 % luottamustasoilla. Perignonin ja Smithin (2008) tutkimuksessa aineistona toimi viiden ison pankin portfolio

kolmelta vuodelta. Käytetyn mallin lähtöhavaintomäärä oli 250 ja viimeinen estimaatti laskettiin 1000 havainnolla. Tulosten perusteella mallin estimointitarkkuus heikkenee luottamustason kasvaessa, tässä tapauksessa 95, 97,5 ja 99 %. Malli hylätään Kupiecin testin perusteella jokaisella aineistolla korkeimmalla luottamustasolla ja malli pärjää kokonaisuudessaan heikosti.

Kiohos ja Dimopoulos (2004) tutkivat mallia käyttäen verrattain pitkää tutkimusaineistoa ja testasivat mallia 15 vuoden periodilla. Aineistona toimi kolme isoa osakeindeksiä ja yhdysvaltalaisia ja brittiläisiä joukkovelkakirjalainoja. Christoffersenin testi hylkää mallin kaikilla paitsi yhdellä indeksillä 95 % luottamustasolla. Malli on hieman parempi 95 % luottamustasolla kuin 99 % luottamustasolla. Tutkijat eivät kuitenkaan ilmoita käyttämänsä havaintomäärää tai painotuskerrointa. Tämä vaikeuttaa tulosten tulkintaa ja mahdollisia johtopäätöksiä, mutta tutkimuksen tulokset noudattavat verrattain hyvin muiden tutkimusten tuloksia, joiden mukaan malli näyttäisi pärjäävän paremmin matalammilla luottamustasoilla.

Lopuksi esitellään Zikovicin ja Aktanin (2011) mielenkiintoinen tutkimus liittyen painotuskertoimeen. Tutkimus toteutettiin monelle eri aineistolle ja rahoitusinstrumentille 99 % luottamustasolla ja yhdelle päivälle. Yleinen tendenssi näyttäisi olevan, että raaka-aineiden osalta painotuskertoimen pitäisi olla yli 0,99, mutta osakeindeksien osalta syntyy enemmän hajontaa. Optimaalinen kerroin vaihtelee ajassa, mutta keskimäärin suurempi määrä havaintoja näyttäisi toimivan paremmin korkeammalla kertoimella. Optimoidut painotuskertoimet jäävät osakeindeksien osalta välille 0,973 ja 0,998. Huomioitavaa on, että perinteinen historiallinen simulaatio on sama malli kuin ajan perusteella painotettu käyttäen painotuskerrointa 1.

Ajan perusteella painotettua historiallista simulaatiota on harvemmin tutkittu samassa tutkimuksessa monella eri havaintomäärällä, joten varmojen johtopäätösten tekeminen on hankalaa. Lisäksi pitkälle vietyjen johtopäätösten tekeminen verrattain pienellä otannalla on hankalaa. Kuitenkin näyttäisi siltä, että malli on tarkempi 95 % luottamustasolla verrattaessa korkeampiin luottamustasoihin kuten 99 %:iin. Lisäksi voitaisiin todeta, että painotuskerroin 0,99 näyttäisi olevan hieman parempi kuin 0,97. Tätä tukee myös edellä esitetty tutkimus liittyen optimaaliseen painotuskertoimeen. Painotuskerrointutkimuksen tulokset implikoivat painotetun version paremmuutta painottomaan nähden, koska painotuskerroin 1 olisi antanut tukea painottamattomalle variaatiolle. Lisäksi voidaan todeta, että painotetulla versiolla ei nähdä samantyyllisiä eroja ehdottoman ja ehdollisen testin tuloksissa.

5.3 Vertailevaa tutkimusta ja johtopäätöksiä mallien eroista

Tämän luvun viimeisessä osiossa esitellään vertailevaa tutkimusta perinteisen ja ajan perusteella painotetun historiallisen simulaation osalta. Osion lopuksi vedetään johtopäätöksiä tämän ja kahden edeltävän osion tulosten perusteella.

Aloitan esittelemällä kolme tutkimusta, joissa tutkimusasetelma on havaintomäärien osalta vertailukelpoinen mallien välillä. Dimitrakopoulos, Kavussanos ja Spyrou (2010) vertasivat HS- ja BRW(0,98)-malleja käyttäen 50, 250, 500 ja 1000 havaintoa. Aineisto koostui 16:sta kehittyvien markkinoiden ja neljästä kehittyneiden osakeindeksistä ja estimaatteja testattiin neljän vuoden periodilla 95 ja 99 % luottamustasolla. Tulosten perusteella BRW on parempi, suurimmaksi osaksi todella tarkka ja hylätään Christoffersenin testissä vain muutaman kerran. Kriisien aikana BRW on selkeästi parempi ja kriisienkin jälkeen hieman parempi. Tämä on pitkälti painotuksen ansiota, koska estimaatit muuttuvat nopeammin kriisien alkaessa kuten myös niiden ollessa ohi. Ottink (2009) tutkimuksessaan, joka jo osittain esiteltiin edellisissä osioissa, sai tuloksia joiden mukaan BRW on kokonaisuudessaan hieman tarkempi mutta häviää niukasti 99 % luottamustasolla. Hullin ja Whiten (1998) tutkimuksessa verrattiin malleja HS(500) ja BRW(0,98; 500) 99 ja 99,5 % luottamustasoilla yhdelle ja kymmenelle päivälle. Mallien paremmuutta mitattiin testillä, jonka tarkoituksena on optimoida tarvittavan pääoman määrä. Aineistona toimi 12 kehittyneiden maiden valuuttaa ja 5 kehittyneiden maiden osakeindeksiä ja estimaatteja testattiin kahdeksan vuoden periodilla. BRW pärjää paremmin valuutoilla mutta heikommin osakeindekseillä.

Tutkimuksista, joissa käytettiin painotuskerrointa 0,99 esitellään Zikovicin (2008) ja Zikovicin ja Filerin (2009) tutkimukset. Molemmissa tutkimuksissa oli mukana BRW(0,99; 250) ja vastaava painottamattoman malli HS(250). Zikovic (2008) vertaili malleja viidellä itäeurooppalaisella osakeindeksillä, joista estimaatit laskettiin kahden vuoden periodille. Painotettu malli on tarkempi molemmilla luottamustasoilla ja läpäisee riippumattomuustestin joka kerta. 99 % luottamustasolla tulee kaksi hylkäystä ehdottoman kattavuuden testissä. Painottamaton malli läpäisee 95 % luottamustasolla riippumattomuustestin ja hylätään kahdella indeksillä ehdottoman kattavuuden testissä. 99 % luottamustasolla vastaavat hylkäysmäärät ovat riippumattomuustestissä 2 ja ehdottoman kattavuuden testissä 3. Zikovic ja Filer (2009) vertailivat samoja malleja kuin edellisessä tutkimuksessa, aineistona käytettiin 16 osakeindeksiä, joista puolet oli kehittyneiltä markkinoilta ja puolet kehittyviltä, ja estimaatit laskettiin neljän vuoden periodille. Painotettu malli läpäisee riippumattomuustestit jokaisella aineistolla. Painottoman historiallinen simulaatio on riippumattomuustestin perusteella melkein yhtä hyvä ja hylkäyksiä tulee neljä, jotka jakautuvat tasaisesti kehittyvien ja kehittyneiden markkinoiden välillä. Ehdottoman kattavuuden osalta painotettu malli saa samanlaisia tuloksia molemmilla markkinoilla, 6/8 hyväksytään kehittyneillä ja kehittyvillä 5/8. Painottoman malli pärjää heikosti ja hylätään kaikilla kehittyneiden markkinoiden indekseillä ja hyväksytään vain kahdella kehittyvien markkinoiden indekseillä.

Vertailevan tutkimuksen osalta esitellään vielä kolme tutkimusta, joissa tutkimusasetelma ei ole samalla tavalla vertailukelpoinen, viitaten vastaaviin havaintomääriin, kuin edellisissä kolmessa tutkimuksessa. Mohammadi ja Shaker (2012) vertaillessaan HS- ja BRW-mallia Iranin Realilla havaitsivat, että

BRW on tarkempi niin 95 kuin 99 % luottamustasolla. Pritsker (2006) vertaili tutkimuksessaan perinteistä ja ajan perusteella painotettua historiallista simulaatiota useilla eri testeillä. Saatujen tulosten perusteella voidaan huomata, että molemmat mallit voivat hetkittäin olla todella paljon väärässä. Tutkija myös huomauttaa molempien mallien omaavan saman ongelman, eli kumpikin huomioi vain yhdensuuntaista volatilitteettia. McKinsey & Companyn (Mehta et al. 2012) julkaisemassa tutkimuspaperissa huomautetaan, että molempien mallien estimaatit ylittyvät kriisin alkaessa mutta BRW reagoi muutokseen nopeammin. Tämä pätee myös kriisien jälkeiseen periodiin, jolloin BRW:n estimaatit laskevat nopeammin kohti oikeaa riskitasoa.

Tämän osion ja samalla luvun lopuksi vedetään yhteen kahden ensimmäisen osan tulokset ja yhdistetään tämän osion tuloksiin. Näin pyritään luomaan jonkinasteinen kokonaiskuva mallien välisistä eroista. Edellä olevista tutkimuksista huomataan, että molemmilla malleilla on hetkittäin ongelmia ja luottamustason, aineiston ja tutkimusperiodin valinta vaikuttaa lopputuloksiin todella paljon. Painottamattomalle versiolle näyttäisi olevan tyypillistä estimaattien muuttuminen konservatiivisemmiksi. Painotetun osalta tutkimus on sen suuntaista, että vastaavan pohtiminen on mahdotonta, mutta vanhojen havaintojen pienempi paino voi vähentää VaR-estimaatin konservatiivisuutta havaintomäärän kasvaessa. Toinen näkemys voisi olla painottamattomalle havaitut erot ehdottoman ja ehdollisen kattavuuden testeissä. Samaa ei näyttäisi ilmenevän painotetulla ja tämä käytännössä tarkoittaa sitä, että perinteisessä versiossa VaR-estimaattien ylitykset näyttäisivät tulevan ryppäissä. Vähäisen vertailevan tutkimuksen pohjalta voitaneen todeta, että painotettu näyttää olevan parempi ainakin valuutoilla mutta osakkeilla ero ei ole ollenkaan selkeä. Zikovicin ja Aktanin (2011) tutkimus liittyen painotuskertoimeen antaa tukea nimenomaan painotetulle versiolle. Yhdellekään mallille tai aineistolle ei optimaalinen painotuskerroin ollut 1, joka implikoi painottamisen tarpeellisuutta. Lopuksi voitaneen todeta, että painotuksesta näyttäisi olevan hyötyä nimenomaan kriisien alkaessa ja niiden ollessa ohitse.

6 EMPIIRINEN OSIO

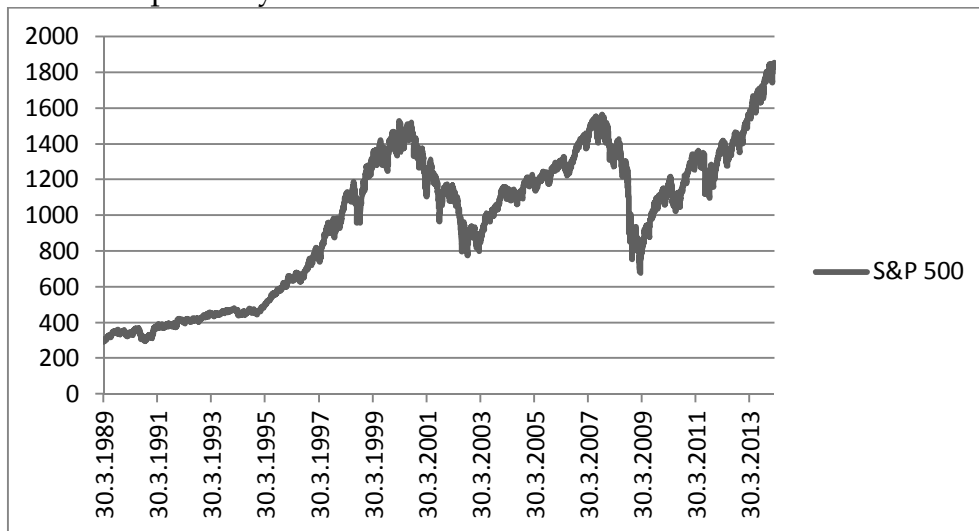
Tässä osiossa esitellään tutkimuksen empiirinen osa. Luvun alussa esitellään ja perustellaan valittu aineisto, jonka jälkeen esitellään käytettävät menetelmät ja niiden suhde aikaisempaan tutkimukseen. Tämän jälkeen esitellään vertailevan osion tulokset ja luku päätetään painotetun historiallisen simulaation painotuskertoimen optimointiin.

6.1 Aineisto ja menetelmät

Kuten johdannossa jo tuli ilmi, tämän backtesting-tutkimuksen tarkoituksena oli tutkia paraneeko historiallisen simulaation estimointitarkkuus jos havaintoja painotetaan Boudoukhin, Richardsonin ja Whitelawn (1998) ehdottaman mallin mukaisesti.

Aineiston osalta päätin keskittyä osakemarkkinoihin, koska kuten edellisestä luvusta ilmenee, painotuskerroin näyttäisi vaihtelevan eri omaisuusluokkien välillä. Aikaisempien tutkimusten perusteella näyttäisi siltä, että vakiokerroin 0,99 on osakkeiden osalta verrattain hyvä approksimaatio optimaalisesta painotuskertoimesta. Samanlaisia johtopäätöksiä ei vielä voida tehdä muiden omaisuusluokkien osalta. Toinen syy osakeaineiston käyttöön on vertailukelpoisuus edellisiin tutkimuksiin. Osakeaineistoksi valittiin indeksejä, joiden on tarkoitus edustaa erilaisia mutta hajautettuja osakesalkkuja kuten aikaisemmissakin tutkimuksissa on tehty. Aineistoksi valikoitui neljä erilaista indeksia. Kehittyneiltä markkinoilta valikoituivat yleisesti käytetty S&P 500 sekä OMX Helsinki. Ensin mainittu on yksi tunnetuimmista ja vakaimmista osakeindekseistä kun taas Helsingin pörssi sijaitsee kehittyneiden markkinoiden periferiassa. Jälkimmäisessä näkyy myös Nokian suuren painon tuoma nousu ja lasku, joka heijastuu indeksin komposition muutoksena ajassa. Kehittyviltä markkinoilta valittiin Hang Seng ja MSCI Emerging Markets. Ensin mainittu heijastelee vahvasti manner-Kiinan kehitystä ja jälkimmäinen kattaa hyvin koko kehittyvät markkinat sekä niiden markkinaheilahdukset. Jokaisen indeksin osalta valittiin hintaindeksi ja käytettyjä havaintoja oli jokaisesta 6501 kpl väliltä 30.3.1989 - 27.2.2014, joista laskettiin 6500 päivittäistä muutosta. Aineisto hankittiin Thomson Reutersin Datastreamista. Kuvioissa 1-4 on esitelty indeksin kehitys edellä mainitulta ajalta. S&P 500-indeksin historiallinen kehitys on odotetunlaisesti kaikkein vakain valitusta neljästä indeksistä. Vuosituhannen vaihteen IT-kupla ja vuonna 2008 alkanut finanssikriisi ovat ainoat suuremmat laskut indeksin arvossa. OMX Helsinki noudattelee aika pitkälti samaa kehitysuraa kuin edellä mainittu, vaikka tiettyjä kansallisia piirteitä löytyy esimerkiksi 90-luvun alun laman kohdalta. Silmiinpistäväntä on kuitenkin Nokian vanavedessä syntynyt hurja arvonmuutos, joka ajoittuu IT-kuplan aikoihin vuosituhannen vaihteeseen. Kehittyvien markkinoiden indeksit Hang Seng ja MSCI Emerging Markets omaavat hyvin samankaltaisen

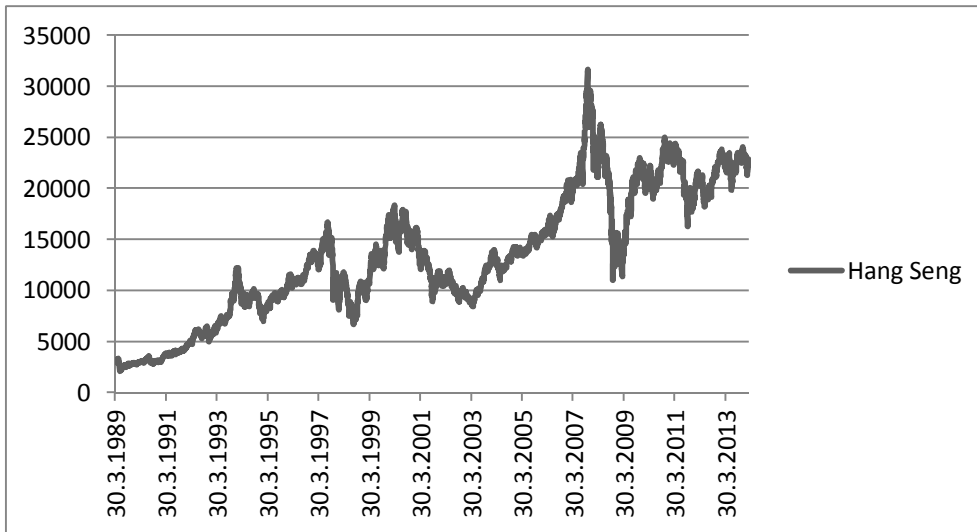
hintakehityksen. Verrattaessa esimerkiksi S&P 500-indeksiin havaitaan, että kehittyvillä markkinoilla on ollut epävakaita koko 90-luvun, kun taas esimerkiksi S&P 500-indeksin hintakehitys on ollut vakaata ja heilahtelut vähäisiä. IT-kupla näkyy selkeästi heikommin kehittyvillä markkinoilla mutta siitä eteenpäin erot kehittyneiden ja kehittyvien markkinoiden välillä ovat kaventuneet ja 2008 finanssikriisi iski kaikille markkinoille yhtä suuressa määrin. Kuitenkin S&P 500 hintakehitys näyttäisi olevan aika selkeää ja trendinomaista kun taas kehittyvillä markkinoilla hinnan laskut ja nousut tulevat nopealla syklillä.



KUVIO 1 - S&P 500-indeksi



KUVIO 2 - OMXH-indeksi



KUVIO 3 - Hang Seng-indeksi



KUVIO 4 - MSCI Emerging Markets-indeksi

Tutkimuksen ensimmäinen osio eli vertaileva osio toteutettiin niin, että painottamattomalle ja painotetulle historialliselle simulaatiolle laskettiin molemmille VaR-estimaatit käyttäen 250, 750 ja 1500 havaintoa. Havaintomäärien osalta oli tarkoitus valita määrät, joilla kyetään tutkimaan mallien eroja pienillä ja isoilla havaintomäärillä. Näitä edustavat Basel-komitean asettama minimi 250 ja hyvin kattava 1500 havaintoa. 750 havaintoa on mukana, jotta nähdään miten mallit vertautuvat keskiuurilla määrillä. Painotetun mallin painotuskertoimena käytettiin 0,99. VaR-estimaatit laskettiin liukuvilla ikkunoilla päivittäisistä muutoksista ja niin, että jokaiselle mallille saatiin 5000 estimaattia. Estimaatit laskettiin kaikille havaintomäärille samalle periodille, joten 1500 havainnon ikkuna alkaa ennen 750 havaintoa ja täten esim. 250 havainnon osalta ensimmäisiä 1250 havaintoa ei käytetä ollenkaan.

Estimaatit laskettiin 99 % luottamustasolla yhdelle päivälle Basel-komitean vaatimusten mukaisesti. VaR-estimaattien tarkkuutta mitattiin aikaisemmin esitellyllä kaavan 15 mukaisella Christoffersenin testillä, johon sisältyy kaavan 16 mukainen ehdottoman kattavuuden testi ja kaavan 17 mukainen riippumattomuustesti. Ehdottoman kattavuuden testi mittaa kuinka lähellä VaR-estimaattien ylitysten prosentuaalinen määrä on haluttua 1 % päivistä ja riippumattomuustesti testaa tulevatko ylitykset ryppäissä. Päädyin valitsemaan Christoffersenin testin koska sen osana olevaa riippumattomuustestiä ei muista yleisesti käytetyistä backtesting-malleista, eli Baselin tai Kupiecien testistä löydy. Toisaalta esimerkiksi Rosenblattin transformaatiot ja Lopezin testi eivät ole niin yleisesti käytettyjä tämän tutkimuksen kaltaisilla aineistoilla tai malleilla kuin edellä mainitut ja Christoffersenin testi antaa hyvän sekoituksen tarkkuutta ja toisaalta tuloksia, jotka ovat vertailukelpoisia aikaisemman tutkimuksen kanssa.

Kuten edellisen luvun johdannossa tuli esiin, niin aikaisemmissa tutkimuksissa on harvemmin verrattu tämän tutkimuksen kiinnostuksen kohteena olevia malleja vertailukelpoisilla havaintomäärillä. Dimitrakopoulos et al. (2010) tekemä tutkimus on aikaisemman tutkimuksen osalta kaikkein lähimpänä tätä tutkimusta. Havaintomäärien, painotuskertoimen ja valittujen osakeindeksien osalta heidän tutkimuksensa eroaa hieman tästä tutkimuksesta, mutta merkittävin ero syntyy aineiston rakenteen suhteen. Heidän tutkimuksessaan malleja on testattu 20 osakeindeksillä neljältä vuodelta, kun taas tässä tutkimuksessa käytettiin neljää osakeindeksiä 20 vuodelta. Päädyin tähän valintaan koska näyttäisi siltä, että mallien hyvyys vaihtelee enemmän ajassa kuin eri aineistojen välillä tai ainakin vaihtelu on vähäisempää esimerkiksi eri kehittyneiden maiden indeksien välillä. Vertailevan osion jälkeen optimoidaan painotetun historiallisen simulaation painotuskerroin mukailien Zikovicin ja Aktanin (2011) tutkimusta. Näin ollen suurimmaksi eroksi aikaisempaan tutkimukseen nähden muodostuu perinteisen backtesting-tutkimuksen yhdistäminen painotuskertoimen optimointiin.

6.2 Vertaileva osio

Tässä osiossa esitellään painotetun ja painottamattoman historiallisen simulaation vertailun tulokset. Tulosten esittely aloitetaan ylitysten määrästä, jonka jälkeen vuorossa on Christoffersenin testin tulokset. Taulukoissa painottamaton historiallinen simulaatio on lyhennetty HS, painotettu BRW ja perässä oleva luku edustaa malliin käytettyjen havaintojen määrää. Taulukoiden tulokset on järjestetty niin, että saman havaintomäärän omaavat mallit ovat peräkkäin. Eli lähtökohtaisena tarkoituksena ei ole pohtia mikä malli on paras vaan kumpi malli on tarkempi käyttäen x-määrää havaintoja ja aineistoa y.

TAULUKKO 1 - VaR-estimaatin ylitykset

	S&P 500	OMX Helsinki	Hang Seng	MSCI Emerging Markets
HS250	98.72	98.64	98.80	98.80
BRW250	98.78	98.92	98.86	98.88
HS750	98.44	98.76	98.74	98.76
BRW750	98.78	98.84	98.90	98.94
HS1500	98.44	98.48	98.68	98.60
BRW1500	98.72	98.82	98.82	98.84

Kaikki luvut prosenteissa (tavoite 99.00)

Taulukossa 1 on ilmaistu VaR-estimaattien ylitysten prosentuaalinen määrä. Tulosten paremman tulkittavuuden vuoksi luvut esitetään käänteisinä, eli luvut kuvaavat kuinka monena päivänä (prosentteissa ilmaistuna) toteutunut tappio ei ylitä VaR-estimaattia. Tuloksista ilmenee, ettei riskin yliarvioimista ilmene kummallakaan mallilla millekään havaintomäärälle tai aineistolle. Tutkimuksen kannalta kiinnostavinta on, että painotettu malli on lähempänä haluttua tasoa jokaisella vertailuparilla. Painottamattoman historiallisen simulaation osalta ei löydy selkeitä eroja aineistojen välillä mutta painotetun malli näyttäisi olevan tarkin MSCI Emerging Markets-indeksillä ja heikoin S&P 500-indeksillä. Engelin ja Gizyckin (1999) ja Vlaarin (2000) havaitsemaa ilmiötä, jossa painottamattoman historiallisen simulaation VaR-estimaatit muuttuvat konservatiivisemmiksi havaintomäärän kasvaessa, ei tällä kertaa havaittu. Enemminkin näyttäisi siltä, että trendi on päinvastainen jos jonkinnäköistä trendiä ylipäätään löytyy.

TAULUKKO 2 - Christoffersenin ehdottoman kattavuuden testi

	S&P 500	OMX Helsinki	Hang Seng	MSCI Emerging Markets
HS250	3.637723	5.8834532	1.8988025	1.8988025
BRW250	2.284267	0.3150256	0.9471256	0.7000884
HS750	13.529671	2.7029255	3.1542482	2.7029255
BRW750	2.284267	1.2296569	0.4891720	0.1783228
HS1500	13.529671	11.7807763	4.6991622	7.1870303
BRW1500	3.637723	1.5470773	1.5470773	1.2296569

Kriittinen arvo (tilastollisesti merkitsevät lihavoitu) $\chi_{0.010}^2(1) = 6.635$

Christoffersenin ehdottoman kattavuuden testin tulokset löytyvät taulukosta 2. Testi kertoo poikkeavatko taulukon 1 luvut merkitsevästi 99 %:sta, joka oli VaR-mallien haluttu luottamustaso. Testin perusteella painotettu malli hyväksytään jokaisella aineistolla ja havaintomäärällä ja kuten jo taulukon 1 esittelyssä ilmeni, se on jokaisella tarkasteluparilla tarkempi kuin painottamaton malli. Painottamaton malli pärjää hyvin 250 havainnolla mutta havaintomäärän kasvu näyttää heikentävän estimointitarkkuutta ja malli läpäisee testin kahdeksan kertaa ja hylätään neljästi. MSCI Emerging Markets-indeksi näyttäisi antavan tarkimmat tulokset S&P 500-indeksin jäädessä heikoimmaksi.

TAULUKKO 3 - Christoffersenin riippumattomuustesti

	S&P 500	OMX Helsinki	Hang Seng	MSCI Emerging Markets
HS250	0.03822233	9.2330227	20.667530	11.454884
BRW250	1.50712622	0.2512367	8.246104	12.725375
HS750	6.94621046	7.0793520	35.141673	15.117713
BRW750	1.50712622	4.5068281	8.754430	5.405483
HS1500	10.19147973	14.4755496	44.750735	25.824140
BRW1500	1.66002627	4.3414368	11.761616	4.506828

Kriittinen arvo (tilastollisesti merkitsevät lihavoitu) $\chi_{0,010}^2(1) = 6.635$

Taulukossa 3 esitellään Christoffersenin riippumattomuustestin tulokset, eli kuinka pahasti VaR-estimaattien ylitykset tulevat ryppäissä. Erot painottamattoman ja painotetun version välillä ovat hyvin selkeät. Tuloksissa näkyy myös selkeä ero kehittyneiden ja kehittyvien markkinoiden välillä ja painotettu malli hyväksytäänkin kaikilla havaintomäärillä molemmille kehittyneille markkinoille, kun taas kaikki variaatiot hylätään Hang Seng-indeksillä. Mielenkiintoista on, että MSCI Emerging Markets-indeksi antoi tarkimmat tulokset ehdottoman kattavuuden testissä ja S&P 500-indeksi heikoimmat, kun taas riippumattomuustestin osalta tilanne on lähes päinvastainen. Tämän perusteella voidaan olettaa, että kehittyvien markkinoiden epävakaus ilmenee nimenomaan riippumattomuustestissä. Mallien vertailun osalta tilanne ei ole yhtä selkeä kuin edellisessä testissä, mutta painotettu on heikompi ainoastaan kaksi kertaa eikä kertaakaan niin, että painottamaton hyväksyttäisiin ja painotettu hylättäisiin. Kuitenkin painotettu läpäisee testin kahdeksan kertaa ja painottamaton vain kerran, mikä on selkeä ero viime testin läpäisyihin, 12 vastaan kahdeksan. Havaintomäärän kasvu näyttäisi lisäävän VaR-estimaatin ylitysten ilmenemistä peräkkäisillä päivillä. Tämä todennäköisesti johtuu pienempien havaintomäärien mukanaan tuomasta nopeammasta reagoitakyvystä markkinoiden muutoksiin.

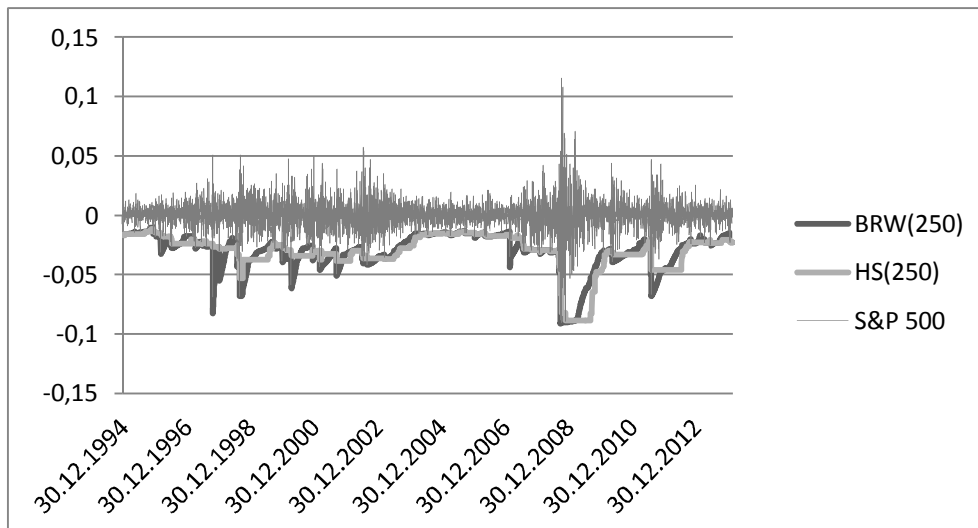
TAULUKKO 4 - Christoffersenin ehdollisen kattavuuden testi

	S&P 500	OMX Helsinki	Hang Seng	MSCI Emerging Markets
HS250	3.675946	15.1164759	22.566332	13.353686
BRW250	3.791394	0.5662624	9.193229	13.425463
HS750	20.475882	9.7822775	38.295921	17.820638
BRW750	3.791394	5.7364850	9.243602	5.583805
HS1500	23.721151	26.2563259	49.449897	33.011170
BRW1500	5.297750	5.8885141	13.308693	5.736485

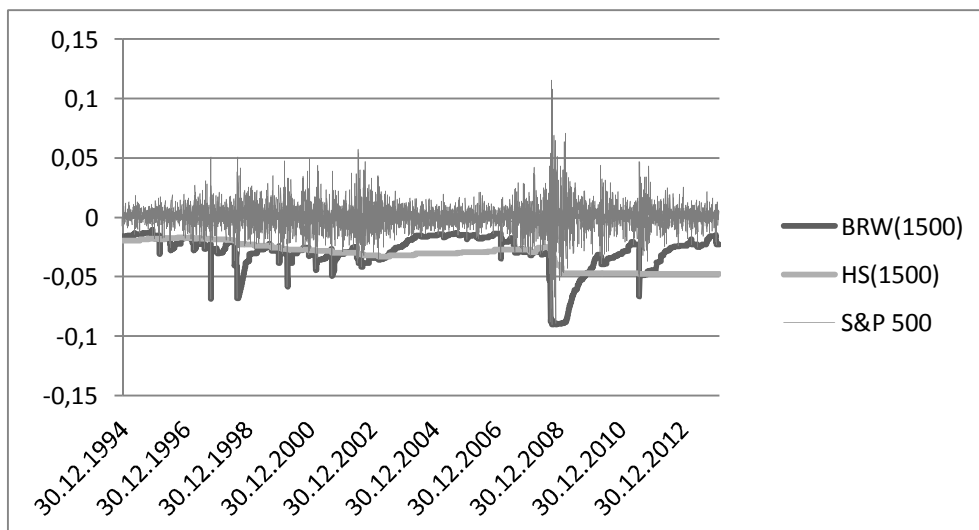
Kriittinen arvo (tilastollisesti merkitsevät lihavoitu) $\chi_{0,010}^2(2) = 9.210$

Taulukko 4 kertoo lopullisen Christoffersenin testin eli ehdollisen kattavuuden testin tulokset, jotka saadaan ehdottoman kattavuuden ja riippumattomuustestin summana. Toisin sanoen tulokset kuvaavat estimointitarkkuutta ottaen huomioon VaR-estimaattien ylitysten mahdollinen

ilmeneminen ryppäissä. Painotettu malli näyttäisi olevan tarkka kehittyneiden markkinoiden osakeindekseillä ja hieman heikompi kehittyvien osalta. Sama tendenssi näkyy myös painottamattoman osalta mutta malli hyväksytään ainoastaan S&P 500-indeksillä käyttäen 250 havaintoa. Riippumattomuustestissä ilmenneet erot kehittyneiden ja kehittyvien markkinoiden välillä näyttäytyvät myös ehdollisen kattavuuden testissä. Riippumattomuustestissä havaittu havaintomäärän kasvun myötä heikkenevä tarkkuus tulee esiin myös tässä testissä. Painotetun mallin osalta tämän trendin rikkoo ainoastaan MSCI Emerging Markets käyttäen 250 havaintoa. Tämä onkin ainut havaintomäärä millä tulee hylkäys kyseisellä aineistolla. Hang Seng-indeksillä saatiin vastakkaisia tuloksia ja 250 havaintoa käyttävä painottamaton malli oli ainut tarkka. Vertailevan osion osalta painotetulla saadaan selkeästi paremmat tulokset ja malli hyväksytään 9/12, kun painottamattoman tulos on 1/12. Painottamaton malli on marginaalisesti parempi kuin painotettu malli 2/12 kertaa mutta ero on niin olematon, että käytännössä voidaan todeta painotetun version olevan selkeästi parempi 10/12 kertaa ja yhtä hyvä kuin painottamaton malli 2/12 kertaa. Yleisesti aineiston osalta voidaan todeta, että kehittyvien markkinoiden ennustaminen näyttäisi olevan selkeästi hankalampaa. Painotettu malli näyttäisi pärjäävän todella hyvin kehittyneillä markkinoilla, tilanteen kehittyvillä markkinoilla ollessa heikompi. Painottamaton taas näyttäisi olevan heikko kaikilla aineistoilla. Käytettyjen testien osalta näyttäisi siltä, että riippumattomuustestin rooli ehdollisen kattavuuden testin tulosten kannalta on selkeästi merkittävämpi. Tässä tapauksessa syy voi olla käytettyjen mallien erot reagointikyvyssä markkinoiden muutoksiin. Toisin sanoen painottamattoman mallin heikko reagointikyky markkinoiden muutoksiin voi aiheuttaa tilanteen, jossa ennusteet ovat jatkuvasti joko riskiä yliarvioivia tai aliarvioivia. Tämän kaltainen ongelma ei välttämättä ilmene ehdottoman kattavuuden testissä.

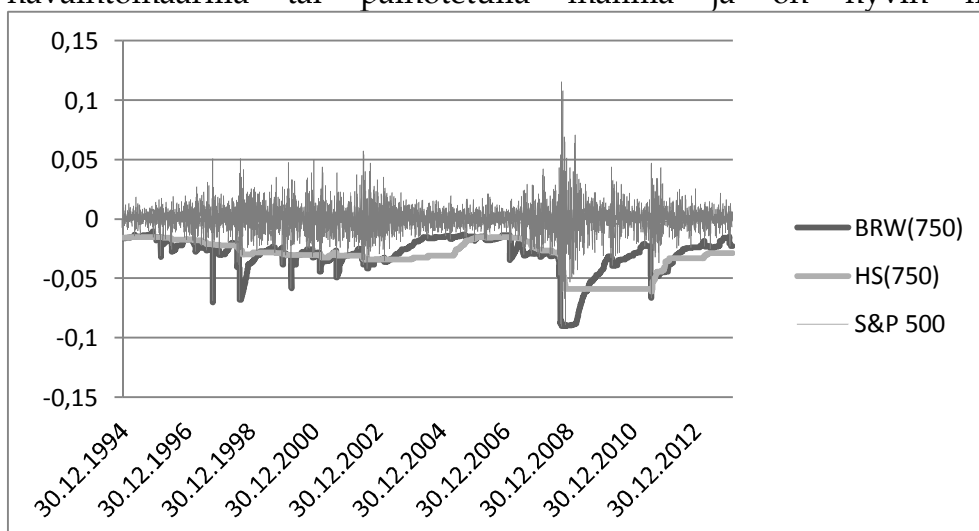


KUVIO 5 - S&P 500 250 havaintoa



KUVIO 6 - S&P 500 1500 havaintoa

Kuvioissa 5 ja 6 on esitelty molempien mallien VaR-estimaatit S&P 500-indeksille 250 ja 1500 havainnolla sekä kyseisen indeksin todelliset muutokset. Kuvioiden perusteella voidaan todeta, että painottamattoman mallin reagointikyky markkinoiden muutoksiin heikkenee havaintomäärän kasvaessa ja 250 havainnolla mallit antavat hyvin samankaltaisia tuloksia kuin painotetulla. Tätä tukee myös ehdollisen kattavuuden testin tulokset. 1500 havainnolla painotettu malli reagoi markkinoiden muutoksiin toisin kuin painottamaton, joka ei juuri vaihtele markkinoiden muutosten mukana. Painottamattoman mallin ainut suurempi reaktio markkinoiden muutoksiin on 2008 kriisin jälkeen. Reaktio kuitenkin tulee myöhässä ja aiheuttaa estimateissa haamureaktion ja malli yliarvioikin riskiä todella pitkään vielä kriisin jälkeen. Tämän kaltaista haamureaktiota ei näyttäisi ilmenevän pienillä havaintomäärillä tai painotetulla mallilla ja on hyvin mahdollista,



KUVIO 7 - S&P 500 750 havaintoa

että painottamattoman mallin heikko estimointikyky kaikilla indekseillä 1500 havainnolla johtuu edellisen kaltaisista haamuilmiöistä. Kuviossa 7 esitetään vastaavat 750 havainnon tulokset. Tulosten perusteella voidaan todeta, että 750 havainnolla painottamaton historiallinen simulaatio reagoi selkeästi vähemmän pieniin markkinamuutoksiin verrattaessa 250 havaintoon. Reagointi on kuitenkin selkeästi merkittävämpää verrattaessa 1500 havaintoon ja 750 havainnon tulokset löytyvät odotetunlaisesti 250 ja 1500 havainnon tulosten välistä. Eri havaintomäärien väliset erot ja ominaispiirteet tulevat esiin myös muilla aineistoilla, joiden osalta kuvat löytyvät liitteistä.

Yleisesti ottaen tulokset eroavat hieman edellisistä tutkimuksista mutta vertailevan osion osalta vastaavat aikaisemmat tutkimukset saavat tukea. Zikovicin (2008) ja Zikovicin ja Filerin (2009) mukaan havaintomäärän lisääminen parantaa painottamattoman historiallisen simulaation estimointitarkkuutta. Nämä tulokset eivät kuitenkaan vastaa tämän tutkimuksen tuloksia. Tulosten perusteella asiasasia näyttäisi päinvastaiselta. Mahdollinen selitys eroaville lopputuloksille lienee käytetyn aineiston pituus, aikaisemmissa oli käytetty aineistoa 2-4 vuoden periodilta ja käytetyt havaintomäärät, aikaisemmissa tutkimuksissa suurin käytetty havaintomäärä oli 500 havaintoa. Angelidis ja Benos (2005), Unal (2011) ja Terzic et al. (2013) havaitsivat painottamattoman historiallisen simulaation suurimmiksi ongelmiksi hylkäykset riippumattomuustestissä. Tulos vastaa hyvin pitkälti saamiani tuloksia. Hieman vastakkaisia tuloksia saivat Zikovic ja Filer (2009), joiden mukaan ero painotetun mallin eduksi syntyy ehdottoman kattavuuden testissä. Tätä voidaan kenties selittää lyhyemmällä aineistolla tai käytetyillä havaintomäärillä, jotka olivat pienemmät. Luvun alussa mainitussa Dimitrakopoulos et al. (2010), joka aikaisemmista tutkimuksista eniten muistuttaa tätä tutkimusta, saatiin samankaltaisia tuloksia kuin tässä eli painotettu historiallinen simulaatio on tarkempi.

6.3 Painotuskertoimen optimointi

Painotuskertoimen optimointi toteutettiin niin, että laskettiin se λ :n arvo kaavan 7 mukaiselle painotetulle historialliselle simulaatiolle, joka minimoi kaavan 15 mukaisen Christoffersenin ehdollisen kattavuuden testin tuloksen. Käytännössä tämä toteutettiin niin, että laskettiin Christoffersenin testin tulokset eri λ :n arvoille ja optimaaliseksi painotuskertoimeksi valittiin se luku, jolle Christoffersenin testin tulos oli pienin. Painotuskerroin laskettiin tässä tapauksessa kolmen desimaalin tarkkuudella. Optimointi toteutettiin siis vastaavalla tavalla kuin Zikovicin ja Aktanin (2011) tutkimuksessa sillä poikkeukselle, että Lopezin testi vaihdettiin edellä mainittuun Christoffersenin testiin.

TAULUKKO 5 - Optimaalinen painotuskerroin

	S&P 500	OMX Helsinki	Hang Seng	MSCI Emerging Markets
BRW250	0.992	0.990	0.982	0.984
BRW750	0.994	0.992	0.984	0.996
BRW1500	0.997	0.992	0.984	0.996

Taulukossa 5 esitellään painotuskertoimen optimoinnin tulokset. Tuloksissa havaitaan vastaava trendi kuin Zikovic ja Aktan (2011) tutkimuksessa, eli havaintomäärän kasvu kasvattaa optimaalista painotuskerrointa. Toinen asia joka nousee välittömästi esiin, ovat MSCI Emerging Markets-indeksin optimaaliset kertoimet. 250 ja 750 havainnon välille syntyy muista poikkeava ero ja optimaalinen painotuskerroin vaihtuu selkeästi alle 0,99 olevasta selkeästi yli 0,99 olevaksi. Painotuskerroin 0,99 näyttäisi olevan aika hyvä approksimaatio ja optimaalinen kerroin vaihtelee väliltä 0,982 ja 0,997. Tulos on samansuuntainen mutta vähäisemmän vaihtelun omaava kuin Zikovicin ja Aktanin (2011), joiden mukaan optimaalinen kerroin on välillä 0,973 ja 0,998. Alimman painotuskertoimen suhteen syntyvä ero selittynee käytetyn backtesting-mallin ja aineiston pituuden erolla. Vertailtaessa optimoituja kertoimia Christoffersenin ehdollisen kattavuuden testiin havaitaan, että painotuskertoimen vaikutus tuloksiin muodostaa joko vinon tai bimodaalisen jakauman tai molemmat. Tämä on ainakin mahdollista johtuen Christoffersenin testin kaksiosaisuudesta. Intuitiivinen tulkinta optimaalisesta painotuskertoimesta on seuraavanlainen: mitä lähempänä kerroin on lukua 1, sitä todennäköisempää on, että painottamaton malli on tarkempi. Kuitenkin optimaalisen painotuskertoimen vertailu Christoffersenin tuloksiin antaa erilaisen kuvan. MSCI Emerging Markets-indeksin osalta optimaalinen painotuskerroin indikoi, että 250 havaintoa käyttäen painottamaton on selkeästi heikompi, kun taas suuremmilla havaintomäärillä tulkinta on päinvastainen. Christoffersenin ehdollisen kattavuuden testi antaa täysin vastakkaiset tulokset. Toinen esimerkki on S&P 500-indeksi, jonka osalta voisi olettaa, että painotettu malli on tarkempi 250 havainnolla mutta heikompi 1500 havainnolla. Kuitenkin Christoffersenin testin tulosten tarkastelu paljastaa päinvastaisen tuloksen. Christoffersenin ehdollisen kattavuuden testin tulosten vertailu optimaalisen painotuskertoimen perusteella näyttää siltä, että liian pienen painotuskertoimen vaikutus tuloksiin on pienempi kuin liian ison painotuskertoimen.

7 LOPPUPÄÄTELMÄT

Tämän tutkimuksen tarkoituksena oli tutkia onko ajan perusteella painotettu historiallinen simulaatio tarkempi kuin painottamaton historiallinen simulaatio. Tutkimus oli siis luonteeltaan Value at Risk-malleja vertaileva backtesting-tutkimus. Backtesting-menetelmänä käytettiin Christoffersenin kaksiosaista testiä ja aineistona neljää pitkältä ajalta olevaa osakeindeksiä. Nämä indeksit olivat S&P 500, OMX Helsinki, Hang Seng ja MSCI Emerging Markets. Backtesting-osion jälkeen toteutettiin myös painotetun historiallisen simulaation painotuskertoimen optimointi.

Tulosten perusteella voidaan todeta, että painotettu historiallinen simulaatio on, ainakin osakkeiden osalta, tarkempi ja parempi malli kuin painottamaton historiallinen simulaatio. Painotuskerroin näyttäisi kasvavan havaintomäärän kasvaessa ja vaihtelevan eri aineistojen ja omaisuusluokkien välillä. Painotuskertoimen osalta tulokset vastaavat aika pitkälti aikaisempaa tutkimusta, jota tosin on todella vähän. Mallien vertailevan osion tulokset vahvistavat niiden samankaltaisten tutkimusten tulokset, joiden mukaan ajan perusteella painotettu historiallinen simulaatio on tarkempi malli. Lisäksi näyttäisi siltä, että suurin ero painotetun eduksi havaitaan nimenomaan Christoffersenin riippumattomuustestissä.

Tämän tutkimuksen tulosten perusteella uskaltaisin todeta, että havaintojen painottaminen ajan perusteella on kannattavaa. Aiheesta vaaditaan kuitenkin lisää tutkimusta, jotta voidaan selvittää onko 0,99 keskimäärin paras painotuskerroin osakkeille, miten painotuskerroin muuttuu siirryttäessä osakkeista muihin omaisuusluokkiin ja kuinka paljon painotuskerrointa pitäisi lisätä havaintomäärän kasvaessa? Lisäksi olisi mielenkiintoista tutkia mikä on painotuskertoimen ja Christoffersenin testin tulosten suhde, koska saamani tulokset indikoivat, että suhde ei ole lineaarinen. Tulevaisuudessa olisi myös hyödyllistä perehtyä johdannossa mainittuun ongelmaan, jonka mukaan VaR-mallit menestyvät hyvin kriisien aikana paitsi jos kyseiset markkinat muuttuvat erittäin epälikvideiksi.

LÄHTEET

Andjelic, G., Djakovic, V. & Radisic, S. 2010. Application of VaR in emerging markets: A case of selected Central and Eastern European Countries. *African Journal of Business Management* 4. 17, 3666-3680

Angelidis, T. & Benos, A. 2005. Value-at-Risk for Greek stocks. SSRN Working Paper Series Feb 2005. [Viitattu 30.1.2014] Saatavissa: papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID661242_code362451.pdf?abstractid=661242&mirid=1

Angelidis, T., Benos, A. & Degiannakis, S. 2007. A robust VaR model under different time periods and weighting schemes. *Review of Quantitative Finance and Accounting* 28. 2, 187-201

Barone-Adesi, G., Giannopoulos, K. & Vosper, L. 1999. VaR without correlations for portfolios of derivative securities. *The Journal of Futures Markets* 19. 5, 583-602

Barone-Adesi, G., Giannopoulos, K. & Vosper, L. 2002. Backtesting derivative portfolios with Filtered Historical Simulation (FHS). *European Financial Management* 8. 1, 31-58

Basel-komitea. 1996. Supervisory framework for the use of "backtesting" in conjunction with the internal models approach to market risk capital requirements. BIS, Basel, Switzerland

Beder, T. 1995. VAR: Seductive but dangerous. *Financial Analysts Journal* 51. 5, 12-24

Berkowitz, J. 2001. Testing density forecasts, with applications to risk management. *Journal of Business and Economic Statistics* 19. 4, 465-474

Billinger, O. & Eriksson, B. 2009. Star VaRs: Finding the optimal value-at-risk approach for the banking industry. Master thesis. Lund University [Viitattu 5.2.2014] Saatavissa: <http://lup.lub.lu.se/luur/download?func=downloadFile&recordId=1436974&fileId=1646978>

Boudoukh, J., Richardson, M. & Whitelaw, R. 1998. The best of both worlds: A hybrid approach to calculating Value at Risk. *Risk* 11. 5, 64-67

Brandolini, D. & Colucci, S. 2012. Backtesting value-at-risk: a comparison between filtered bootstrap and historical simulation. *Journal of Risk Model Validation* 6. 4, 3-16

Cabedo, J. & Moya, I. 2003. Estimating oil price "Value at Risk" using the historical simulation approach. *Energy Economics* 25. 3, 239-252

Campbell, A. 2009. The risk of value-at-risk. *Risk* 22. 4, 42-46.

Changchien, CC., Lin, CH. & Kao, WS. 2012. Capturing value-at-risk in futures markets: a revised filtered historical simulation approach. *Journal of Risk Model Validation* 6. 4, 67-93

Christoffersen, P. 1998. Evaluating interval forecasts. *International Economic Review* 39. 4, 841-862

Deb, S. & Banerjee, A. 2009. Downside risk analysis of Indian equity mutual funds: a Value at Risk approach. *International Research Journal of Finance and Economics* 1. 23, 216-230

Dimitrakopoulos, D., Kavussanos, M. & Spyrou, S. 2010. Value at Risk models for volatile emerging markets equity portfolio. *The Quarterly Review of Economics and Finance* 50. 4, 515-526

Engel, J. & Gizycki, M. 1999. Conservatism, accuracy and efficiency: comparing value-at-risk models. Working paper 2, Reserve bank of Australia [Viitattu 2.2.2014] Saatavissa: <http://www.apra.gov.au/AboutAPRA/Research/Documents/Conservatism-Accuracy-and-Efficiency-Comparing-VaR-models-Feb-1999.pdf>

Fierli, F. 2004. Hedging and risk measurement for option portfolios. Thesis. Universita della Svizzera Italiana [Viitattu 5.2.2014] Saatavissa: <http://doc.rero.ch/record/4780>

Gencay, R., Selcuk, F. & Ulugulyagci, A. 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance Mathematics and Economics* 33. 2, 337-356

Gustafsson, M. & Lundberg, C. 2009. An empirical evaluation of Value at Risk. Master thesis. University of Gothenburg [Viitattu 30.1.2014] Saatavissa: https://gupea.ub.gu.se/bitstream/2077/19299/1/gupea_2077_19299_1.pdf

Hendricks, D. 1996. Evaluation of value-at-risk models using historical data. *Economic Policy Review* 2. 1, 39-70

Hsieh, CS. & Chou, JH. 2008. Forecasting of Value at Risk (VAR) by cluster method in Chinese stock market. *Journal of Money, Investment and Banking*. 5, 57-66

- Hull, J. & White, A. 1998. Incorporating volatility updating into the historical simulation method for Value at Risk. *Journal of Risk*. Fall 1998, 5-19
- Jorion, P. 2000. Risk management lessons from Long-Term Capital Management. *European Financial Management* 6. 3, 277-300.
- Jorion, P. 2000. Risk management lessons from Long-Term Capital Management. *European Financial Management* 6. 3, 277-300
- Jorion, P. 2001. *Value at Risk: A new benchmark for managing financial risk*. New York: McGraw-Hill
- Kiohos, A. & Dimopoulos, A. 2004. Estimation portfolio VAR with three different methods: financial institution risk management approach. *Spoudai* 54. 20, 59-83
- Kupiec, P. 1995. Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives* 3, 73-84
- Lopez, J. 1998. Methods for evaluating Value-at-Risk estimates. *Federal Reserve Bank of New York, Economic Policy Review*. 2, 119-124
- Mausser, H. & Rosen, D. 1998. Beyond VaR: From measuring risk to managing risk. *Algo Research Quarterly* 1. 2, 5-20.
- Mehta, A., Neukirchen, M., Pfetsch, S. & Poppensieker, T. 2012. Managing market risk: Today and tomorrow. *McKinsey Working Papers on risk*, Number 32 [Viitattu 17.1.2014] Saatavissa: www.mckinsey.com/~media/mckinsey/dotcom/client_service/risk/working%20papers/working_papers_on_risk_32.ashx
- Mohammadi, P. & Shaker, S. 2012. Measuring Exchange Rate Fluctuations Using the Value-at-Risk. *Journal of Applied Finance & Banking* 2. 3, 65-79
- Mutu, S., Balogh, P. & Moldovan, D. 2011. The efficiency of value at risk models on Central and Eastern European stock markets. *International Journal of Mathematics and Computers in Simulation* 5. 2, 110-117
- Ottink, N. 2009. Historical Value at Risk models applied to Italian floating rate government bonds. Master thesis. University of Groningen [Viitattu 5.2.2014] Saatavissa: <http://gloria-mundi.com/UploadFile/2010-6/662010403000.pdf>
- Perignon, C. & Smith, D. 2008. A new approach to comparing VaR estimation methods. SSRN working paper series [Viitattu 5.2.2014] Saatavissa: <http://ssrn.com/abstract=981207>

Pritsker, M. 2006. Hidden dangers of historical simulation. *Journal of Banking and Finance* 30. 2, 561-582

RiskMetrics Group. 1996. RiskMetrics - Technical Document. Fourth edition. J.P.Morgan/Reuters. [Viitattu 25.6.2014] Saatavissa: <http://yats.free.fr/papers/td4e.pdf>

Skiaopoulos, G., Lambadiaris, G., Papadopoulou, L. & Zoullis, Y. 2003. VaR: History or simulation? *Risk* 16. 9, 122-127

Soni, V. 2005. A comparison of Value-at-Risk methods for portfolios consisting of interest rate swaps in the Indian market under the GARCH framework. Credence Analytics (I) Pvt. Ltd. [Viitattu 30.1.2014] Saatavissa: <http://gloria-mundi.com/UploadFile/2010-2/vs.pdf>

Terzic, I., Milojevic, M. & Dzamic, V. 2013. Value at risk performance in developed and emerging stock markets. *International Journal of Information Technology and Business Management* 16. 1, 129-136

Unal, G. 2011. Value-at-Risk forecasts: a comparison analysis of extreme-value versus classical approaches. *The Journal of Risk Model Validation* 5. 3, 59-76

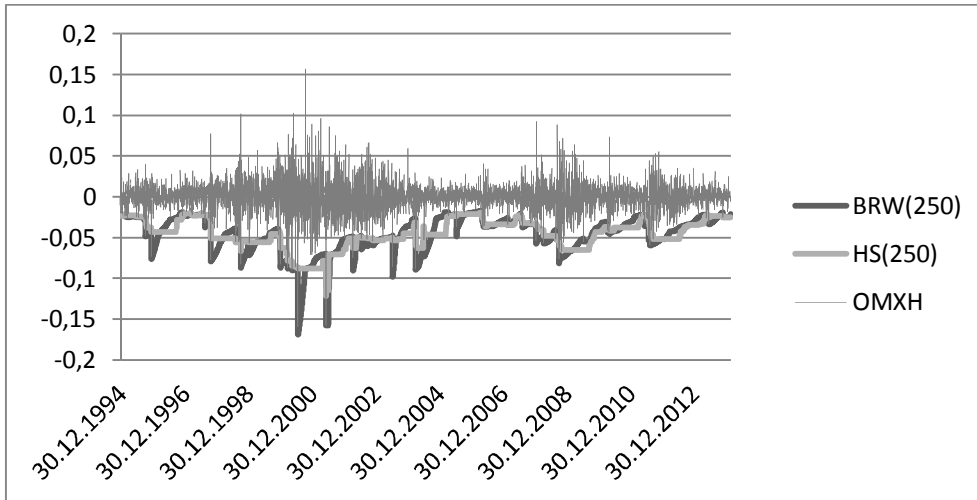
Vlaar, P. 2000. Value at risk models for Dutch bond portfolios. *Journal of Banking & Finance* 24. 7, 1131-1154

Yawalkar, P. 2004. Comparison of Value at Risk methods for FIS and equity portfolios in the Indian market. Credence Analytics (I) Pvt. Ltd. [Viitattu 30.1.2014] Saatavissa: http://gloria-mundi.com/UploadFile/2010-2/py_1.pdf

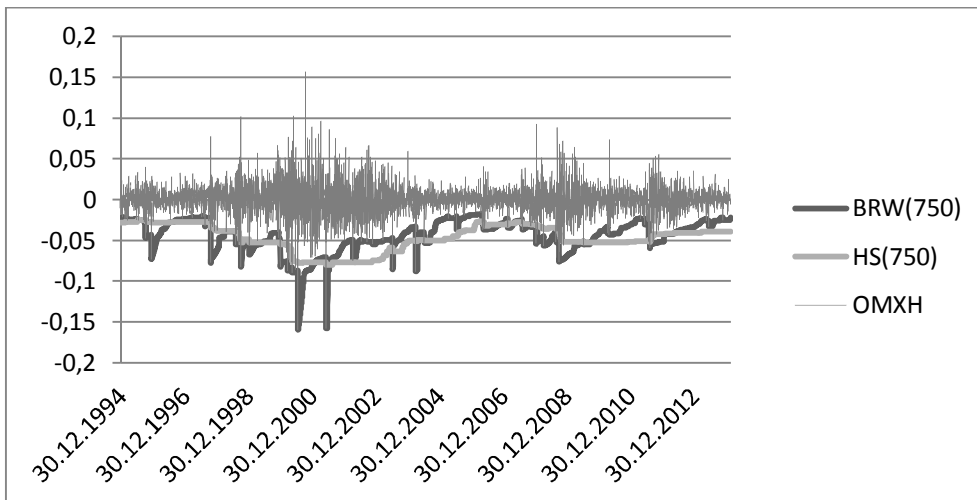
Zikovic, S. 2008. Calculating VaR in EU Candidate States. *South East European Journal of Economics and Business* 3. 1, 23-33

Zikovic, S. & Filer, R. 2009. Hybrid historical simulation VaR and ES: Performance in developed and emerging markets. CESifo working paper series no. 2820 [viitattu 5.2.2014] Saatavissa: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1491418

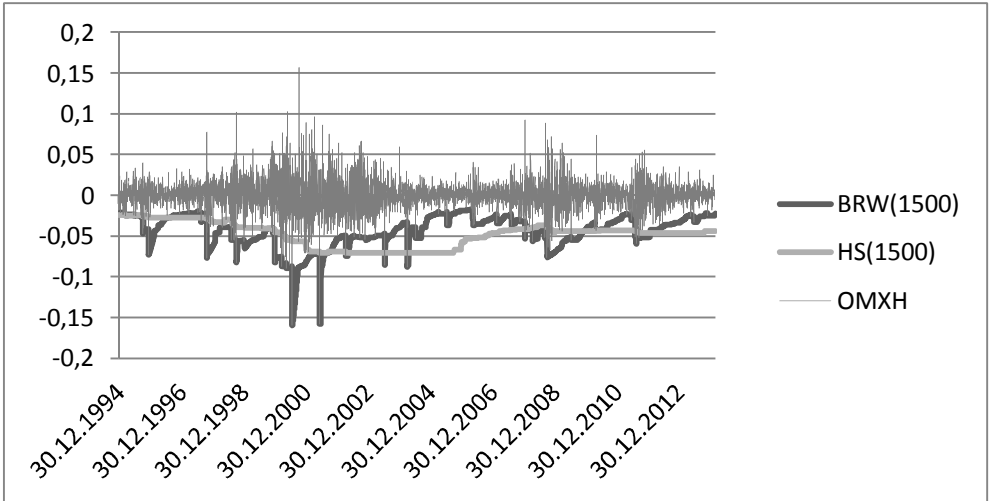
Zikovic, S. & Aktan, B. 2011. Decay factor optimisation in time weighted simulation - Evaluating VaR performance. *International Journal of Forecasting* 27. 4, 1147-115

LIITTEET

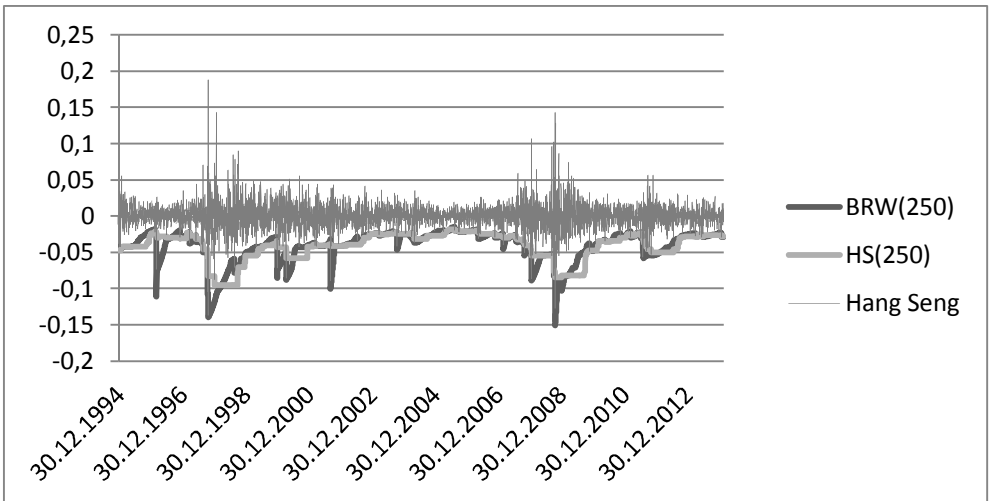
KUVIO 8 - OMXH 250 havaintoa



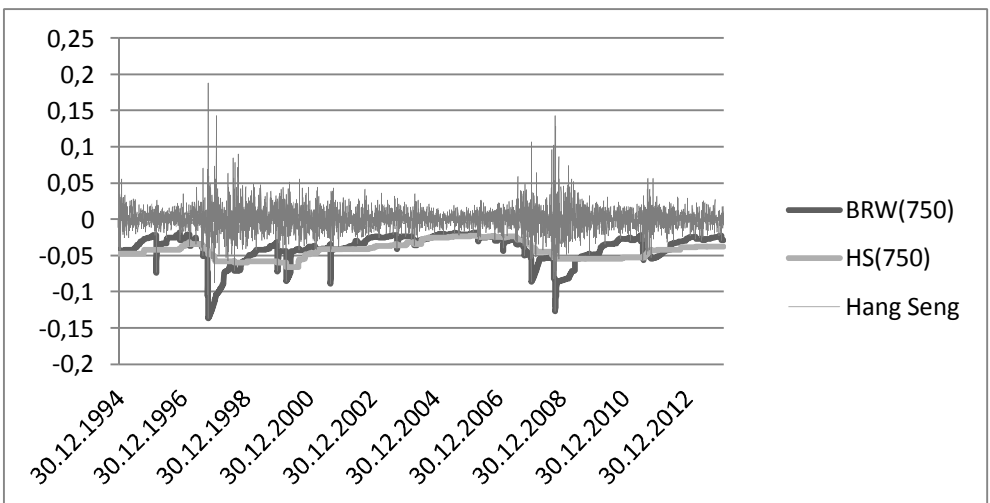
KUVIO 9 - OMXH 750 havaintoa



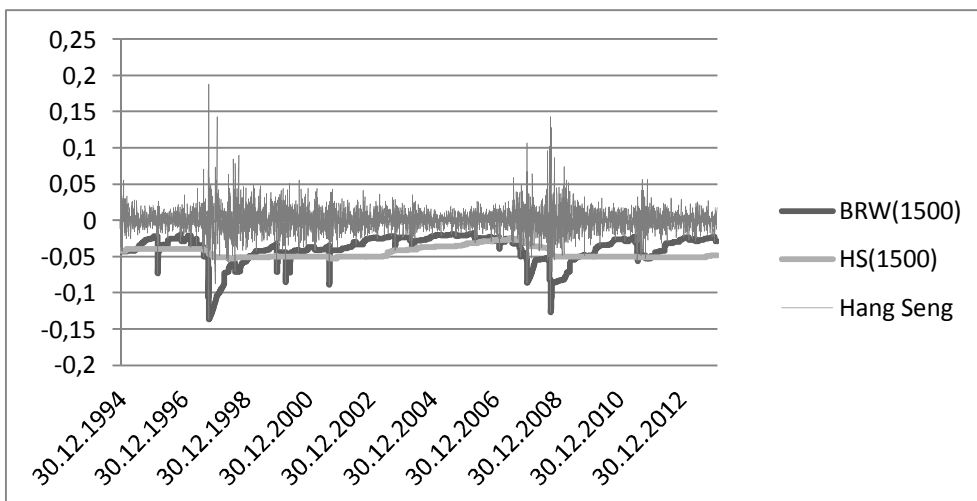
KUVIO 10 - OMXH 1500 havaintoa



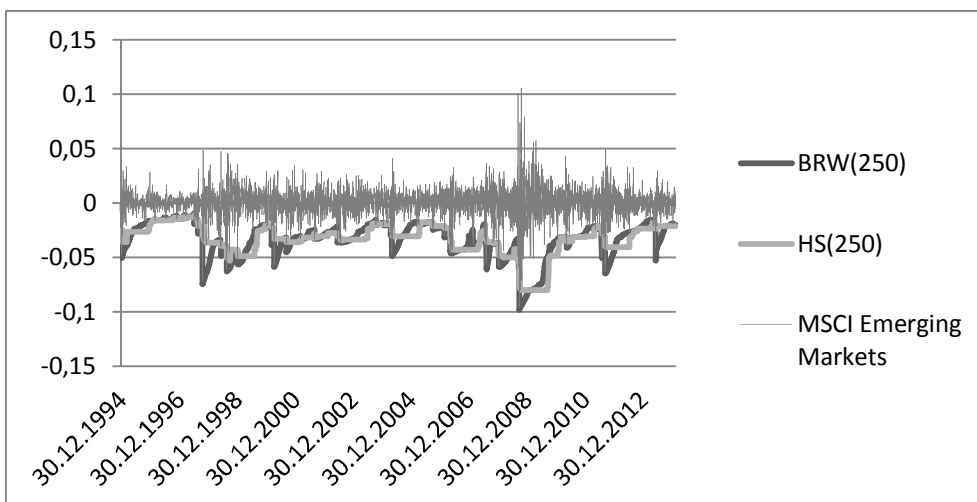
KUVIO 11 - Hang Seng 250 havaintoa



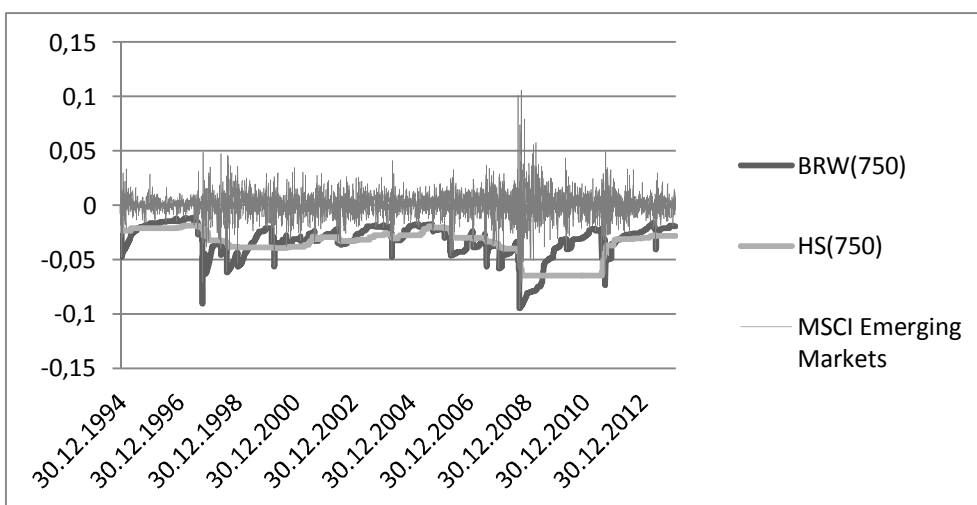
KUVIO 12 - Hang Seng 750 havaintoa



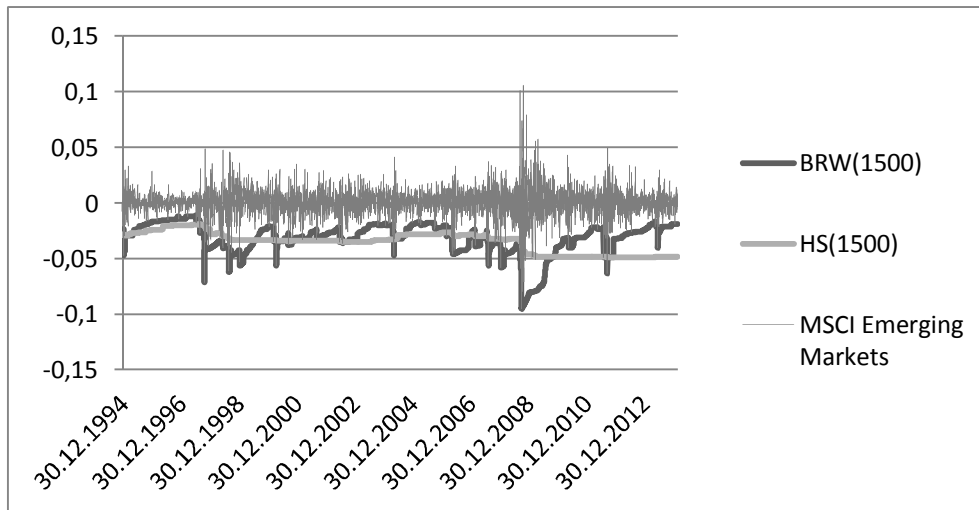
KUVIO 13 - Hang Seng 1500 havaintoa



KUVIO 14 - MSCI Emerging Markets 250 havaintoa



KUVIO 15 - MSCI Emerging Markets 750 havaintoa



KUVIO 16 - MSCI Emerging Markets 1500 havaintoa