Tampereen yliopisto

Petri Peitsalo

AIKASARJA-MOMENTUM SUOMEN OSAKEMARKKINOILLA

Johtamisen ja talouden tiedekunta

Pro gradu -tutkielma

Syyskuu 2019

Peitsalo, Petri Tomi Juhani: Aikasarja-momentum Suomen osakemarkkinoilla (55 sivua)

Pro gradu -tutkielma

Tampereen yliopisto, Johtamisen ja talouden tiedekunta, Kauppatieteiden tutkinto-ohjelma

Syyskuu 2019

Tämän tutkimuksen ideana on tutkia, että onko Suomen osakemarkkinoilla havaittavissa aikasarja-momentumiksi kutsuttua ilmiötä ja johtuuko se rationaalisista vai behavioraalisista tekijöistä. Aikasarja-momentum -ilmiössä on kyse siitä, että edeltävien periodien tuotoilla on ennustevoimaa koskien tulevien periodien tuottoja. Lisäksi tässä tutkimuksessa pyritään selvittämään, että voiko sijoittaja tuottaa lisäarvoa kyseiseen ilmiöön perustuvia sijoitussääntöjä noudattamalla verrattuna tavalliseen osta-ja-pidä -sijoitussääntöön.

Tehokkaiden markkinoiden hypoteesin mukaan milläkään sijoitussäännöllä ei pitäisi tuottaa parempaa tuotto-riski -suhdetta kuin osta-ja-pidä -sijoitussäännöllä. Täten tutkimusta voidaan pitää lisäksi tehokkaiden markkinoiden hypoteesin testinä.

Itse ilmiön tutkimiseen tässä työssä käytetään ARMA-mallia, AR-GARCH-in-Mean-mallia, probit-mallia sekä regiimin muutosmallia. Sijoitussääntöjen kannattavuuden arviointiin puolestaan käytetään riippumattomien otosten t-testiä, GARCH-mallia, Sharpen lukua, CAPM-mallia ja APT-mallia.

Edellä mainittuja malleja käyttämällä käy ilmi, että kyseinen ilmiö on havaittavissa Suomen osakemarkkinoilla. Lisäksi saadaan indikaatiota siitä, että se aiheutuisi behavioraalisten tekijöiden seurauksena. Työssä testataan kahta kyseistä ilmiötä hyödyntävää sijoitussääntöä. Kyseisen ilmiön mukaisten sijoitussääntöjen noudattaminen ei näytä tuottavan sijoittajalle odotusarvoisesti suurempaa tuottoa kuin osta-ja-pidä -sijoitussääntö, mutta kyseiset sijoitussäännöt näyttävät sisältävän pienemmän riskin kuin osta-ja-pidä -sijoitussääntö. Saadaan indikaatiota lisäksi siitä, että yksi sijoitussääntö, joka hyödyntää kyseistä ilmiötä, tuottaisi parempaa tuotto-riski -suhdetta kuin pelkkä osta-ja-pidä -sijoitussääntö. Toinen kyseistä ilmiötä hyödyntävä sijoitussääntö taas ei tuottanut parempaa tuotto-riski -suhdetta.

Avainsanat: taloustiede, rahoitus, tehokkaiden markkinoiden hypoteesi, behavioraalinen rahoitus, markkinapsykologia, markkinatehottomuus, Suomen osakemarkkinat, aikasarja-momentum, momentum, anomalia, empiirinen tutkimus, ekonometrinen tutkimus

Sisällysluettelo

1 Johdanto

Tämän työn tarkoituksena on tutkia aikasarja-momentum -ilmiötä Suomen osakemarkkinoilla. Aikasarja-momentum -ilmiössä on kyse siitä, että edeltävien periodien tuotoilla on yhteys tämän ja seuraavien periodien tuottoihin. Tämä työ sisältää kolme tutkimuskysymystä. Ensimmäinen tutkimuskysymys on, että voidaanko aikasarja-momentum -ilmiötä havaita Suomen osakemarkkinoilla. Toinen tutkimuskysymys on, että syntyykö aikasarja-momentum -ilmiö rationaalisten vai behavioraalisten tekijöiden seurauksena. Kolmas tutkimuskysymys on, että voiko sijoittaja hyödyntää kyseistä ilmiötä sijoitustoiminnassaan. Lisäksi kyseinen työ toimii tietynlaisena tehokkaiden markkinoiden hypoteesin testinä.

Jos siis tutkittavaa ilmiötä havaitaan, niin kyseessä on markkinatehottomuus. Tässä työssä muodostetaan kaksi sijoitussääntöä, jotka hyödyntävät tämän ilmiön mahdollista olemassaoloa. Kyseiset sijoitussäännöt määritetään pohjautuen ilmiötä tutkivien mallien tuloksiin koskien ajanjaksoa 01/1987-12/2008. Lisäksi saadaan vastattua tutkimuskysymyksiin yksi ja kaksi. Sitten kyseisien sijoitussääntöjen kannattavuus arvioidaan käyttäen tunnuslukuja ja malleja koskien ajanjaksoa 01/2009-4/2019. Näin saadaan siis vastattua tutkimuskysymykseen kolme.

Tämän työn tulokset osoittavat, että aikasarja-momentum -ilmiö on havaittavissa Suomen osakemarkkinoilla ja indikaatio siitä, että se ei synny rationaalisten tekijöiden seurauksena. Työssä käytetyksi sijoitussäännöiksi valikoitui yhden viiveen aikasarja-momentum -sijoitussääntö ja viiden viiveen liukuvan keskiarvon sijoitussääntö. Havaitaan, että kyseiset sijoitussäännöt eivät saavuta odotusarvoisesti suurempaa tuottoa kuin osta-ja-pidä -sijoitussääntö. Toisaalta saatujen tulokset indikoivat, että kyseiset sijoitussäännöt omaavat pienemmän riskin kuin osta-ja-pidä -sijoitussääntö. Tuottojakauman huomioon ottava Sharpen luku indikoi, että yhden viiveen aikasarja-momentum -sijoitussääntö tuottaa parempaa tuotto-riski -suhdetta kuin osta-ja-pidä -sijoitussääntö. Viiden viiveen liukuvan keskiarvon sijoitussääntö taas näyttää tuottavan huonompaa tuotto-riski -suhdetta kuin osta-ja-pidä -sijoitussääntö. Mikään käsiteltävistä sijoitussäännöistä ei näytä tuottavan Jensenin alfaa tilastollisesti merkitsevästi kummassakaan käytetyistä riskimalleissa.

Yhden viiveen aikasarja-momentum -sijoitussäännöstä saatuja tuloksia voidaan tulkita indikaationa markkinatehottomuudesta. Toisaalta on otettava huomioon yhteishypoteesiongelman olemassaolo. Tämä tarkoittaa siis sitä, että saatujen tuloksien takana voi myös olla puutteellinen malli tai tunnusluku.

Seuraavassa luvussa esitetään ilmiöön liittyvät teoriat ja käytetyt sijoitussäännöt. Kolmannessa luvussa esitetään tässä työssä käytetty aineisto, metodit sekä mallit. Neljännessä luvussa esitetään kyseisistä malleista saadut tulokset ja vastataan asetettuihin tutkimuskysymyksiin. Viides ja viimeinen luku sisältää yhteenvedon ja johtopäätelmät sekä mahdolliset jatkotutkimusaiheet.

2 Teoriatausta ja käsiteltävät sijoitussäännöt

2.1 Teoriat finanssimarkkinoista

Fama (1970) esitti työssään tehokkaiden markkinoiden hypoteesin, jonka mukaan kaikkien sijoituskohteiden hintojen pitäisi reflektoida kaikkea hallussa olevaa informaatiota oikein ja tarkasti. Fama (1970, 383) jakaa markkinatehokkuuden kolmeen tasoon, jotka ovat heikko, keskivahva sekä vahva. Heikko tehokkuusehto täyttyy, kun hinnat reflektoivat kaikkea historiatietoa. Keskivahva tehokkuusehto täyttyy, kun hinnat reflektoivat kaiken historiatiedon lisäksi myös kaikkea muutakin hallussa olevaa informaatiota, kuten tilinpäätöstietoa. Vahva tehokkuusehto täyttyy, kun hinnat reflektoivat kaiken olemassa olevan informaation, mukaan lukien jopa sisäpiiritiedon. Samuelson (1973) argumentoi työssään, että tehokkailla markkinoilla osakkeiden hinnat määräytyvät niiden osinko-odotusten nykyarvojen perusteella. Tehokkailla markkinoilla osakkeiden hinnat täten määräytyvät alla esitetyn kaavan mukaisesti (Gordon & Shapiro 1956, 104).

$$P_t = V_t = \sum_{s=1}^{\infty} \frac{D_{t+s}}{(1+r_f+\varphi)^s}, \qquad (1)$$

jossa P_t on hinta ajankohtana t, D_t on osinko ajankohtana t, r_f on riskittömän sijoituskohteen tuotto, φ on riskipreemio ja V_t on osakkeen todellinen arvo ajankohtana t. Mitä suurempi riski sijoituskohteeseen sisältyy, sitä suurempi on sijoittajan vaatima riskipreemio. Sijoituskohteen hinta määräytyy siis tehokkailla markkinoilla pelkästään kahden tekijän perusteella, jotka ovat tuotto ja riski. Sijoituskohteiden tuotot saadaan suhteellisen helposti selvitettyä tuottoja kuvaavista tilastoista, kuten osakekursseista tai indeksiarvoista. Niiden riskien määrittäminen onkin paljon haastavampaa. Volatiliteetti, eli sijoituskohteen tuottojen keskihajonta, on yleisesti käytetty kokonaisriskin mitta. Kokonaisriskit voidaan taas jakaa systemaattisiin ja epäsystemaattisiin riskeihin. (Markowitz 1952.)

Epäsystemaattisissa riskeissä on kyse riskeistä, jotka eivät korreloi muiden yritysten kanssa, eli ne koskettavat vain kyseisiä yrityksiä. Systemaattisissa riskeissä taas on kyse riskeistä, jotka korreloivat yritysten kesken eli ne koskettavat koko markkinaa. Ostamalla sijoituskohteita, jotka eivät korreloi ainakaan täydellisesti keskenään, epäsystemaattiset riskit voidaan hajauttamalla muuttaa suuruudeltaan suhteellisen pieniksi. Koska epäsystemaattiset riskit voidaan hajauttamalla muuttaa epäolennaisiksi, voidaan olettaa informoitujen sijoittajien tekevän näin. Täten tehokkailla markkinoilla informoitujen sijoittajien tuottovaatimus määräytyy pelkästään sijoituskohteen sisältämän systemaattisen riskin mukaisesti. (Markowitz 1952.)

Tehokkaat markkinat perustuvat ajatukseen siitä, että jos markkinat ovat tehottomat, niin syntyy mahdollisuus arbitraasiin. Arbitraasilla tarkoitetaan väärinhinnoittelun hyödyntämistä varman voiton saavuttamiseksi. Toisin sanoen aina kun on havaittavissa tehottomuutta, niin syntyy mahdollisuus ”ilmaiseen lounaaseen”. Kun ”ilmainen lounas” on havaittavissa hyvin kilpailuilla markkinoilla, voidaan olettaa, että se hyödynnetään heti kun se havaitaan. (Samuelson 1965.)

Tehokkaiden markkinoiden hypoteesia ei voida kunnolla testata yhteishypoteesiongelman takia (Fama 1991, 1575–1576). Yhteishypoteesiongelma tarkoittaa sitä, että aina kun testataan markkinoiden tehokkuutta, vääjäämättä testattavaksi tulee myös testaamisessa käytetty malli. Täten ei voida olla varmoja siitä, että onko saatujen tuloksien syynä puutteellinen malli vai markkinoiden todellinen tilanne.

Tehokkailla markkinoilla ylituottojen saavuttaminen olisi millä tahansa sijoitusstrategialla mahdotonta. Ylituotot tarkoittavat tuottoja, jotka suuruudeltaan ylittävät sen määrän, mitä kyseisen sijoituskohteen pitäisi tuottaa sen sisältämät riskit huomioon ottaen. Tehokkailla markkinoilla kaikista paras sijoitusstrategia olisi täten sijoittaa indeksiin tai valita osakkeet satunnaisesti, koska tällöin voitaisiin minimoida kaupankäynti- ja analysointikustannukset. Tehokkaat markkinat eivät kuitenkaan tarkoita sitä, että markkinahinnat eivät ikinä ole väärässä. Tehokkailla markkinoilla sijoituskohteiden markkinahintojen poikkeamat niiden todellisista arvoista olisivat vain aina täysin sattumanvaraisia. (Malkiel 2003.)

Anomaliat ovat poikkeamia tehokkaista markkinoista. Anomaliat ovat sijoitusstrategioita, joilla voidaan tuottaa systemaattisesti ylituottoa. Anomaliat olisivat siis mahdottomuus tehokkailla markkinoilla. Tämän vuoksi anomalioiden tutkiminen on yksi tapa testata markkinoiden tehokkuutta. Jos anomalioita havaitaan, niin markkinat ovat ainakin jonkin verran tehottomat. Jos taas yhtäkään anomaliaa ei ole havaittavissa, niin markkinat ovat tehokkaat. (Knüpfer & Puttonen 2014, 171–172.)

Yksi pääongelmista tehokkaiden markkinoiden hypoteesin paikkansapitävyydelle on arbitraasin rajoitteiden olemassaolo. Arbitraasin rajoitteet tarkoittavat tekijöitä, jotka estävät rationaalisia sijoittajia hyödyntämästä markkinatehottomuutta. Tällaisia ovat esimerkiksi korkeat kaupankäyntikustannukset ja lyhyeksi myyntiin liittyvät rajoitukset. (Brealey, Myers & Allen 2014, 334.)

Täten markkinatasapainossa hintojen muodostumiseen vaikuttavat myös osakkeen todellisen arvon lisäksi markkinapsykologiset tekijät. Shiller (2014, 1496–1499) esittää tutkimuksessaan kyseisen markkinatasapainon alla esitellyllä kaavalla.

$$P_t = E_t \sum_{s=1}^{\infty} \frac{D_{t+s} + \varphi A_{t+s}}{(1 + r_f + \varphi)^s}, \tag{2}$$

jossa E_t on odotus ajankohtana t, A_t on "animal spirits" eli ei-informoitujen sijoittajien kysyntä osakkeelle ajankohtana t. Koska markkinapsykologialla on vaikutusta markkinatasapainossa, niin on tärkeää pohtia, että minkälaiset psykologiset tekijät vaikuttavat sijoittajien tekemiin päätöksiin. On havaittu, että sijoittajat omaavat monia psykologisia vääristymiä, jotka vaikuttavat heidän päätöksentekoonsa. Tällaisia psykologisia vaikuttimia on muun muassa konservatismi, yli-itsevarmuus ja edustavuusharha. Konservatismissa on kyse siitä, että sijoittajat uudistavat näkemyksiään liian hitaasti ja täten osakkeen hinta heijastaa osakkeen todellista arvoa vasta viiveellä. Yli-itsevarmuudella tarkoitetaan sitä, että sijoittajat yliarvioivat omia kykyjään, jonka seurauksena he käyvät liian paljon kauppaa, josta heille on keskimäärin vain haittaa. Edustavuusharhassa on kyse siitä, että sijoittajat tekevät päätelmät liian nopeasti sekä liian pienellä otoksella. Edustavuusharhan takia sijoittajat ekstrapoloivat trendejä liian pitkäksi ajaksi ja täten ylireagoivat yritystä koskeviin hyviin uutisiin. (Bodie, Kane & Marcus 2014, 388–394.)

Sijoittaminen on lisäksi sosiaalista toimintaa. Sijoittajat keskustelevat paljon keskenään siitä, että kuinka sijoitustoiminta on sujunut ja mihin kannattaisi seuraavaksi sijoittaa. Sijoittajat saattavat siis levittää ideoita toisilleen, jotka voivat perustua hyvinkin epämääräisiin todisteisiin tai puhtaaseen spekulointiin. Tällaisilla keskusteluilla voi taas olla suurikin vaikutus sijoittajien tekemiin sijoituspäätöksiin. Näiden keskustelujen seurauksena syntyy siis markkinatrendejä, koska sijoittajat toimivat toisiltaan saadun idean mukaisesti. (Shiller 1984, 457–458.)

Jos markkinat ovat tehokkaat, niin sijoituskohteiden hinnat noudattavat yksikköjuuriprosessia. Syynä tähän on se, että tehokkailla markkinoilla hinnoissa tapahtuvat muutokset ovat puhtaasti satunnaisia. Engle & Granger (1987, 252–253) argumentoivat työssään, että jos jokin aikasarja noudattaa yksikköjuuriprosessia, niin kyseiseen aikasarjaan todennäköisesti syntyy nousu- ja laskutrendejä. Jos vastaavasti markkinat ovat tehottomat ja sijoituskohteiden markkina-arvot lähestyvät niiden todellisia arvoja vasta viiveellä, niin syntyy myös nousu- ja laskutrendejä (Bodie, Kane & Marcus 2014, 400 & 401–404). Täten molemmissa edellä kuvatuista tapauksista, syntyy osakkeiden hintojen aikasarjoihin stokastisia trendejä. Toisaalta, jos markkinat ovat tehottomat, niin kyseiset trendit ovat helpommin havaittavissa ja pysyvämpiä kuin tilanteessa, jossa markkinat ovat tehokkaat.

2.2 Teoriat aikasarja-momentum -ilmiöstä

Moskowitz, Ooi & Pedersen (2012) esittävät tutkimuksessaan ilmiön, jota he kutsuvat aikasarja-momentumiksi. Momentumista on puhuttu ilmiönä jo pitkään akateemisessa kirjallisuudessa, mutta sillä on yleensä tarkoitettu poikkileikkauksellista-momentumia (Jegadeesh & Titman 1993). Kun poikkileikkauksellisessa-momentumissa ollaan kiinnostuneita osakkeen tuotoista suhteessa muiden osakkeiden tuottoihin, niin aikasarja-momentumissa ollaan taas kiinnostuneita pelkästään osakkeen tuotoista suhteessa sen omiin aiempiin tuottoihin. (Moskowitz, Ooi & Perdersen 2012, 228–230.) Moskowitz, Ooi & Pedersen (2012, 244–245) argumentoivat tutkimuksessaan, että aikasarja-momentumissa spekuloijat keskimäärin rikastuvat suojautujien kustannuksella. He siis argumentoivat, että kyseessä on suojautujien maksama preemio spekuloijien tarjoamasta likviditeetistä ja suojautumisesta. Moskowitz, Ooi & Pedersen (2012, 241–244) raportoivat lisäksi tutkimukseen, että aikasarja-momentumin tuottojen pääkomponenttina on tuottojen autokovarianssi. Hong & Satchell (2015, 1471) toteavat tutkimuksessaan, että liukuvan keskiarvon sijoitussääntö on tapa hyödyntää sijoituskohteiden hintojen autokorrelaatiorakenteita, ilman että tarvitsee tuntea niiden tarkkaa rakennetta.

Lim, Wang & Yao (2018, 284–286) toteavat tutkimukseen, että aikasarja-momentum -portfolio tuottaa ylituottoa tilastollisesti merkitsevästi standardeissa riskimalleissa. Kyseisessä työssä aikasarja-momentum -portfolio muodostettiin jakamalla käytetyt osakkeet voittaja ja häviäjä osakkeisiin niiden edeltävien periodien tuottojen etumerkkien mukaan. Voittaja osakkeet sitten ostettiin portfolioon ja häviäjä osakkeet lyhyeksi myytiin. Käytetyt riskimallit kyseisessä tutkimuksessa olivat CAPM, Fama-French kolmen tekijän malli sekä Fama-French viiden tekijän malli. Kyseisistä malleista kaksi ensimmäistä on avattu tässä työssä luvussa 3.6. Lim, Wang & Yao (2018, 286–287) lisäksi muodostavat erilliset portfoliot pienen, keskisuuren ja suuren markkina-arvon yrityksistä ja testasivat, että tuottavatko ne kaikki tilastollisesti merkitsevää ylituottoa eri riskimalleissa ja eri painoilla. He havaitsevat, että melkein kaikilla painoilla kaikki kyseisistä portfolioista tuottavat tilastollisesti merkitsevästi ylituottoa. Lim, Wang & Yao (2018, 287–289) myös tutkivat, että miten aikasarja-momentum -portfolio pärjää eri markkinatilanteissa. He huomaavat, että kyseinen portfolio pärjää parhaiten huonoina aikoina ja huonoiten hyvinä aikoina. Lim, Wang & Yao (2018, 289) tutkivat lisäksi sitä, että onko informaation saannin frekvenssillä vaikutusta aikasarja-momentum -portfolion tuottoihin. He havaitsevat, että mitä yleisemmin uutta informaatiota saadaan, sitä suurempia ovat aikasarja-momentum -portfolion tuotot ja ylituotot.

Frazzini (2006.) havaitsee tutkimuksessaan, että jos osake on ollut paperilla tappiolla ennen uuden negatiivisen uutisen tulemista, niin syntyy ”earnings drift”-ilmiö uutisen tulemisen jälkeen. ”Earnings drift” -ilmiössä on siis kyse siitä, että tuotot tai tappiot tulevat ikään kuin viiveellä. Vastaavasti, jos osake on ollut paperilla voitolla ja uutena tietona tulee jokin positiivinen uutinen, syntyy vastaava ilmiö. Toisaalta kun osake on ollut paperilla tappiolla ennen uutisen tulemista ja sitten tulee positiivinen uutinen, vastaavaa ilmiötä ei synny. Efekti ei myöskään synny silloin, kun osake on ollut aiemmin paperilla voitolla, ja sitten siitä tulee jokin negatiivinen uutinen. He & Li (2015) lisäksi havaitsevat, että momentum strategioiden kannattavuuteen vaikuttaa positiivisesti momentum-kaupankäyjien aktiivisuus ja negatiivisesti takaisinkatsomisajan pituus.

2.3 Käsiteltävät sijoitussäännöt

Sijoitussäännöt ovat tapoja harjoittaa sijoitustoimintaa. Sijoitussääntöjä noudattamalla voidaan helpottaa sijoittamiseen liittyvää päätöksentekoa sekä siihen liittyvää tuloksien seurantaa. Usein sijoitussäännöillä pyritään hyödyntämään jotakin markkinailmiötä. Sijoitussääntö voi esimerkiksi olla pienen markkina-arvon yrityksiin sijoittaminen tai markkina-arvon ja kirjanpito arvon alhaisen suhteen omaaviin yrityksiin sijoittaminen. (Knüpfer & Puttonen 2014, 172–174.)

Tässä työssä verrataan kolmesta eri sijoitussäännöistä saatuja tuottoja keskenään. Ensimmäinen sijoitussäännöistä on osta-ja-pidä -sijoitussääntö, jossa vain yksinkertaisesti pidetään osakkeet ikuisesti. Tämän sijoitussäännön tuotot toimivat siis vertailu arvoina vaihtoehtoisten sijoitussääntöjen tuotoille. Toinen sijoitussäännöistä on aikasarja-momentum -sijoitusääntö. Aikasarja-momentum -sijoitussääntö on esitelty alla. (Marshall, Nguyen & Visaltanachoti 2017, 407–408.)

$$TSMOM_{t,l} = P_t - P_{t-l}, \tag{3}$$

jossa $TSMOM_{t,l}$ on aikasarja-momentum ajankohtana t viiveellä l. Mikäli $TSMOM_{t,l}$ on positiivinen, kyseessä on ostosignaali. Jos se taas on negatiivinen, kyseessä on myyntisignaali. Aikasarja-momentum -sijoitussäännön kohdalla optimaalisen takaisinkatsomisajan määrittämiseen käytettyjä malleja ovat ARMA-, AR-GARCH-in-Mean- sekä probit-malli. Nämä mallit on esitetty luvussa 3.5. Työn kolmas sijoitussääntö on liukuvan keskiarvon sijoitussääntö. Liukuvan keskiarvon sijoitussääntö poikkeaa aikasarja-momentum -sijoitussäännöstä siten, että nykyistä markkinahintaa verrataan entisen markkinahinnan sijaan markkinahinnan liukuvaan keskiarvoon. Liukuvan keskiarvon sijoitussääntö on esitelty alla. (Marshall, Nguyen & Visaltanachoti 2017, 407–408.)

$$MAR_{t,l} = P_t - \frac{P_{t-l+1}+P_{t-l+2}+\cdots+P_t}{l}, \tag{4}$$

jossa $MAR_{t,l}$ on liukuva keskiarvon sijoitussääntö ajankohtana t viiveellä l. Mikäli MAR on positiivinen, kyseessä on ostosignaali. Kun taas jos MAR on negatiivinen, kyseessä on myyntisignaali. Liukuvan keskiarvon sijoitussäännössä on myös tarpeellista määrittää optimaalinen takaisinkatsomisaika. Tämän määrittäminen tapahtuu käyttämällä regiimin muutosmallia, joka on esitetty luvussa 3.5.

Aikasarja-momentum -sijoitussäännöllä ja liukuvan keskiarvon sijoitussäännöllä on yhteys toisiinsa. Aikasarja-momentum -sijoitussäännön osto- ja myyntisignaalit perustuvat hinnassa tapahtuvaan muutokseen tietyllä aikavälillä. Kyseinen muutos puolestaan vaikuttaa hinnan liukuvan keskiarvon suuntaan ja täten liukuvan keskiarvon sijoitussäännön antamiin signaaleihin. (Marshall, Nguyen & Visaltanachoti 2017, 407–408.)

Aikasarja-momentum -sijoitussääntö ei anna ostosignaalia ennen kuin hinnan liukuva keskiarvo vaihtaa suuntaan, kun taas liukuvan keskiarvon sijoitussäännössä ostosignaaliin riittää se, että nykyinen hinta ylittää liukuvan keskiarvon. Tästä syystä liukuvan keskiarvon sijoitussääntö antaa signaalit todennäköisesti aikaisemmin kuin aikasarja-momentum -sijoitussääntö. Sijoitussääntöjen välinen yhteys on esitelty alla. (Marshall, Nguyen & Visaltanachoti 2017, 407–408.)

$$TSMOM_{t,l} = l\left(\frac{P_{t-l+1}+P_{t-l+2}+\cdots+P_t}{l} - \frac{P_{t-l}+P_{t-l+1}+\cdots+P_{t-1}}{l}\right) \tag{5}$$

3 Tutkimusmenetelmät ja -aineisto

3.1 Käytetty aineisto ja siitä määritetyt tunnusluvut

Tämä työn aineistona käytetään Suomen osakemarkkinoiden yleisindeksiä, joka on laskettu euroissa, ja havaintoyksikkö on logaritmoitu kuukausituotto, joka on laskettu käyttämällä kuukauden viimeisen päivän arvoja. Kyseinen aineisto jaetaan kahteen osaan. Ensimmäinen osa on 01/1987 – 12/2008 ja toinen osa on 01/2009 – 04/2019. Ensimmäistä osaa käytetään ilmiötä tutkivien mallien estimointiin ja täten on tämän työn in-sample-osa. Toista osaa käytetään taas kannattavuus-mallien estimointiin ja on täten out-of-sample-osa.

Itse ilmiön tutkimiseen käytetään hintaindeksejä, kun taas sen kannattavuuden analysointiin käytetään tuottoindeksejä. Kyseinen aineisto on ladattu Thomson Reutersin päätteeltä. Riskittömän koron edustajana käytän tässä työssä yhden kuukauden euribor-korkoa. Euribor-koron aikasarja on otettu Eurostat:n sivuilta (https://ec.europa.eu/eurostat/data/database). Lisäksi kannattavuuden analysoinnissa käytetään kahta eri riskitekijää, jotka ovat maailman osakemarkkinoiden tuottoindeksi ja euron ja Yhdysvaltojen dollarin valuuttakurssi. Ne on myös ladattu Thomson Reuters päätteeltä. Logaritmoitu kuukausituotto ja riskipreemio on laskettu käyttämällä alla olevia kaavoja.

$$R_t = \ln(P_t) - \ln(P_{t-1}) \tag{6}$$

$$RP_t = R_t - \ln(1 + r_{f,t}), \tag{7}$$

jossa P_t on indeksiluku ajankohtana t, R_t on logaritmoitu kuukausituotto ajankohtana t, $r_{f,t}$ on yhden kuukauden euribor ajankohtana t ja RP_t on logaritmoitu kuukausiriskipreemio ajankohtana t. Sijoitussäännöt on toteutettu seuraavalla tavalla. Ensimmäiseksi on laskettu, että pitäisikö sijoitussäännön mukaan ostaa vai myydä kyseistä sijoituskohdetta sen hetkisellä tiedolla. Vaihtoehtoiset sijoitussäännöt on esitetty luvussa 2.3. Ideana on siis se, että kuvitteellinen sijoittaja tarkastaa tilanteen jokaisen kuukauden viimeisenä päivänä. Sitten määritetään kyseisen sijoituspäätöksen tuottama tuotto. Jos sijoitussäännön mukaan kyseessä on ostosignaali, niin tämän periodin tuotoksi muodostuu tämän ja seuraavan kuukauden viimeisen päivän välillä tapahtunut logaritmoitu muutos. Jos taas sijoitussäännön mukaan kyseessä on myyntisignaali, niin periodin tuotoksi muodostuu joko seuraavan kuukauden logaritmoitu yhden kuukauden euribor-korko tai nolla, riippuen kumpi näistä on suurempi. Ideana tässä on se, että kuvitteellinen sijoittaja sijoittaa rahat euribor-korkoon tai käyttötililleen siksi aikaa, kun myyntisignaali on voimassa. Kaupankäyntikustannuksiksi tässä työssä on määritetty 0,1 prosenttia kauppahinnasta kauppaa kohden.

Tunnuslukuja, joita kyseisen aineiston kohdalla esitetään, ovat keskiarvo, keskihajonta, jakauman vinous ja jakauman huipukkuus. Näiden neljän tunnusluvun tarkoituksena on esittää eri sijoitussääntöjen tuottojakaumat kunkin aineiston kohdalla. Näistä tunnusluvuista keskiarvo kuvaa, mikä on jakauman keskipiste, ja keskihajonta kertoo, kuinka paljon saadut arvot keskimäärin poikkeavat keskiarvosta. Kyseisien tunnuslukujen kaavat on esitetty alla.

$$\bar{R} = \frac{1}{N}\sum_{t=1}^{N} R_t \tag{8}$$

$$s = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}(R_t - \bar{R})^2}, \tag{9}$$

jossa $\bar{R}$ on tuottojen otoskeskiarvo, s on tuottojen otoskeskihajonta ja N on havaintojen lukumäärä. Vinous ja huipukkuus taas kuvaavat jakauman muotoa. Jos tuottojakauma on positiivisesti vino, niin tuottojen keskihajonta yliarvioi riskin suuruuden. Jos taas se on negatiivisesti vino, niin keskihajonta aliarvioi riskin suuruuden. Huipukkuus taas vaikuttaa todennäköisyyteen saada suuria poikkeamia odotusarvosta. Mitä suurempi on jakauman huipukkuus, sitä suurempi on todennäköisyys saada suuria poikkeamia. Jos siis jakaumalla on suuri huipukkuus, niin keskihajonta aliarvio todennäköisyyden saada suuria poikkeamia. Vinouden ja huipukkuuden kaavat on esitelty alla. (Bodie, Kane & Marcus 2014, 137–139.)

$$Vinous = \frac{1}{N}\sum_{t=1}^{N}\left[\frac{(R_t-\bar{R})^3}{s^3}\right] \tag{10}$$

$$Huipukkuus = \frac{1}{N}\sum_{t=1}^{N}\left[\frac{(R_t-\bar{R})^4}{s^4}\right] \tag{11}$$

Lisäksi mitataan sijoitussääntöihin sisältyvää riskiä eri aineistojen kohdilla ”Value-at-Risk” ja ”Expected Shortfall” -tunnusluvuilla. VaR eli ”Value-at-Risk” -tunnusluku on tuottojakauman jokin persentiili. Esimerkiksi, jos kyseessä on 95% VaR, niin 95% todennäköisyydellä tappio on pienempi kuin kyseinen VaR-arvo ja 5% todennäköisyydellä tappio on suurempi kuin kyseinen VaR-arvo. VaR-arvo on siis tässä tapauksessa kyseisen tuottojakauman 5% persentiili. Käytetyimmät prosentit kyseistä tunnuslukua koskien ovat 95% ja 99%. Tässä työssä määritetään 95% VaR-tunnusluku. ES eli ”Expected Shortfall” kuvaa taas häntäriskin suuruuden. Se siis kertoo, että kuinka suuri on tappion odotusarvo, kun tappiot ovat VaR-tasoa suuremmat. Kuten VaR:n kohalla, myös ES esitetään tässä työssä 95% luottamusvälillä. Tässä työssä sekä VaR että ES määritetään kahdella eri tavalla. Ensimmäisenä ne määritetään perustuen oletukseen, että tuotot ovat normaalistijakautuneet. Tällöin 95% VaR ja ES voidaan määrittää käyttäen alla esitettyjä kaavoja (Bodie, Kane & Marcus 2014, 139–140).

$$VaR = \bar{R} - 1{,}65s \tag{12}$$

$$ES = \frac{1}{0.05} e^{\bar{R}} \Phi[-s - \Psi(0.95)] - 1, \tag{13}$$

jossa Φ on standardi normaalijakauman kertymäfunktio ja Ψ on standardi normaalijakauman kertymäfunktion käänteisarvo. Lisäksi VaR ja ES arvioidaan perustuen otosjakaumiin. Tällöin voidaan luopua tuottojen normaalijakauma-oletuksesta. Ne siis arvioidaan laittamalla otoksen havainnot suuruusjärjestykseen ja valitsemalla niistä käytetty persentiili. Koska kyseistä persentiiliä ei välttämättä löydy kokonaislukuna otosjakaumasta, niin VaR täytyy estimoida interpoloimalla. ES estimoidaan taas katsomalla kyseisestä otosjakaumasta, että mitkä arvot alittavat kyseisen VaR:n tason ja määrittämällä niiden keskiarvon. Otosjakaumasta määritetyt VaR:n ja ES:n estimaatit ovat harhattomat, mutta ne ovat alttiita estimointivirheisiin, koska niiden arviot perustuvat yleensä suhteellisen pieniin otoksiin. (Bodie, Kane & Marcus 2014, 139–140.)

Kyseisen aineiston kuvaamiseen käytetään myös kahta eri kuviota. Ensimmäinen näistä kuvioista on tuottojen aikasarjakuvio ja toinen on autokorrelaatiokuvio, joka sisältää käsiteltävän aikasarjan eri viiveiden autokorrelaatiokertoimet ja osittaisautokorrelaatiokertoimet. Kyseisien kertoimien määrittämiseen käytetään alla olevia kaavoja (Gujarati & Porter 2009, 749).

$$\hat{\rho}_l = \frac{\frac{\sum (R_t - \bar{R})(R_{t+l} - \bar{R})}{N}}{\frac{\sum (R_t - \bar{R})^2}{N}}, \tag{14}$$

jossa $\hat{\rho}_l$ on tuottojen otosautokorrelaatiokerroin viiveellä l. Osittaisautokorrelaatio tarkoittaa ehdollista autokorrelaatiota ehdolla viivästetyt arvot. Osittaisautokorrelaatiokertoimet estimoidaan pienemmin neliösumman menetelmällä käyttämällä seuraavalla sivulla esiteltyä kaavaa. (Lütkepohl, Krätzig & Phillips 2004, 13.)

$$a_h = Corr(y_t, y_{t-h} | y_{t-1}, \dots, y_{t-h+1}) \tag{15}$$

$$y_t = \nu + \sum_{i=1}^{h} a_i y_{t-i} + u_t \tag{16}$$

$$u_t \sim WN, \tag{17}$$

jossa y_t on aikasarjamuuttuja y ajankohtana t ja a_h on osittaisautokorrelaatiokerroin viiveellä h.

3.2 Regressioanalyysi ja parametrien estimointimenetelmät

Tässä työssä käytettävien mallien estimointiin käytetään regressioanalyysia. Regressioanalyysissa pyritään estimoimaan selittävien muuttujien arvojen yhteys selitettävän muuttujan keskimääräisiin arvoihin. Regressioanalyysin pohjana käytetään empiiristä aineistoa. Regressioanalyysilla havaittu yhteys eri muuttujien välillä on siis tilastollinen, eikä funktionaalinen tai deterministinen. Tilastollinen yhteys sisältää aina jonkin verran sattumanvaraisuutta. Siksi ei voida ikinä todentaa täydellä varmuudella, että onko havaitun yhteyden syynä kausaliteetti vai puhdas sattuma. Tilastollisilla menetelmillä voidaan siis vain testata korrelaatiota. Teoriat ovat täten ainoa tapa osoittaa korrelaation yhteys kausaliteettiin. (Gujarati & Porter 2009, 15 & 19.) Esimerkki regressiomallista on esitelty alla.

$$y_t = \beta_0 + \sum_{j=1}^{K-1} \beta_j x_{j,t} + \varepsilon_t \tag{18}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{19}$$

jossa y_t on selitettävä muuttujat ajankohtana t, $x_{j,t}$ on selittävä muuttuja j ajankohtana t ja ε_t on virhetermi ajankohtana t. Yllä esitetyssä regressiomallissa parametreja ovat beetat. Parametrit siis kuvaavat selittävien muuttujien vaikutusta selitettävän muuttujan keskiarvoon. Residuaali on taas havaittujen arvojen poikkeama ennustetuista arvoista. Residuaali on täten estimoitu virhetermi. Jos virhetermi noudattaa yllä esitettyä jakaumaa ja mallista ei puutu merkittäviä muuttujia, niin malli on hyvä ennustaja. Siksi on tärkeää tutkia residuaalien käyttäytymistä tilastotieteellisillä testeillä.

Regressiomallin parametrien estimointiin voidaan käyttää eri menetelmiä ja eri tilanteissa eri menetelmät ovat sopivia. Tässä työssä siihen käytetään kahta eri menetelmää. Yksi näistä on pienimmän neliösumman menetelmä. Pienimmän neliösumman menetelmässä mallin parametrit valitaan siten, että mallin residuaaleista syntyvä neliösumma on mahdollisimman pieni (Gujarati & Porter 2009, 55–61). Tämän määrittämiseen käytetään seuraavalla sivulla olevaa tavoitefunktiota (Verbeek 2008, 8–11).

$$S = \sum_{t=1}^{N}(y_t - \beta_0 - \sum_{j=1}^{K}\beta_j x_{j,t})^2, \tag{20}$$

Jossa S on residuaalien neliösumma, y_t on selitettävä muuttuja ajankohtana t ja $x_{j,t}$ on selittävä muuttuja j ajankohtana t. Regressiomalleissa on tärkeää määrittää parametrien keskivirheet. Keskivirheet tarkoittavat parametrien estimoitua keskihajontaa ja kuvaavat täten estimaattien tarkkuutta. Keskivirheet saadaan pienimmän neliösumman menetelmässä määritettyä käyttämällä alla esiteltyjä kaavoja (Verbeek 2008, 17–20).

$$s_\varepsilon^2 = \frac{1}{N-K}\sum_{t=1}^{N}\hat{\varepsilon}_t^2 \tag{21}$$

$$\vec{x}_t = [1, x_{1,t} \dots x_{K-1,t}]^T \tag{22}$$

$$\hat{V}_{OLS} = s_\varepsilon^2 \left(\sum_{t=1}^{N}\vec{x}_t\vec{x}_t^T\right)^{-1} \tag{23}$$

$$\hat{V}_{OLS} = \begin{bmatrix} \sigma_{0,0}^2 & \cdots & \sigma_{K-1,0}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{0,K-1}^2 & \cdots & \sigma_{K-1,K-1}^2 \end{bmatrix} \tag{24}$$

$$se(\hat{\beta}_i) = \sqrt{\sigma_{i,i}^2}, \tag{25}$$

jossa s_ε^2 on virhetermin estimoitu varianssi, $\vec{x}_t$ on selittävien muuttujien vektori ajankohtana t, $\hat{V}_{OLS}$ on parametrien estimoitu kovarianssi-varianssi matriisi, $\sigma_{i,i}^2$ on β_i:n estimaatin estimoitu varianssi ja $se(\hat{\beta}_i)$ on β_i:n estimaatin keskivirhe. Jos kaikki Gauss-Markov -ehdot pitävät paikkansa, niin pienimmän neliösumman menetelmällä estimoidut parametrit ovat parhaat lineaariset ja harhattomat estimaattorit Gauss-Markov -teoreeman mukaisesti. Gauss-Markov -ehdot on lueteltu alla. (Verbeek 2008, 16–21.)

Ehto 1: $\forall t: E(\varepsilon_t) = 0$

Ehto 2: $[\varepsilon_1, \dots, \varepsilon_N]^T\ ja \begin{bmatrix} x_{1,1} & \cdots & x_{K-1,1} \\ \vdots & \ddots & \vdots \\ x_{1,N} & \cdots & x_{K-1,N} \end{bmatrix}\ ovat\ riippumattomat$

Ehto 3: $\forall t: Var(\varepsilon_t) = \sigma_\varepsilon^2$

Ehto 4: $\forall t, j: Cov(\varepsilon_t, \varepsilon_j) = 0, t \neq j$

Gauss-Markov -ehtoja on siis neljä. Ensimmäinen niistä on se, että virhetermin odotusarvo on nolla. Toinen niistä on, että virhetermi ei korreloi minkään selittävän muuttujan kanssa. Kolmas ehdoista on, että virhetermin varianssi on vakio. Neljäs ehto on, että virhetermi ei ole autokorreloitunut.

Pienimmän neliösumman menetelmän estimaattorit ovat lisäksi asymptoottisesti normaalistijakautuneita. Kun edellisellä sivulla esitetyt ehdot täyttyvät, niin kyseiset estimaattorit noudattavat likimain alla esiteltyä jakaumaa. (Verbeek 2008, 35.)

$$\hat{\beta}_i \sim \mathcal{N}\left(\beta_i, se(\hat{\beta}_i)\right), \tag{26}$$

jossa $\hat{\beta}_i$ on β_i:n pienimmän neliösumman menetelmällä estimoitu arvo. Toinen tässä työssä käytetty menetelmä on suurimman uskottavuuden menetelmä. Suurimman uskottavuuden menetelmässä mallin parametrit valitaan sillä perustella, että parametreilla olisi suurin todennäköisyys tuottaa havaitut arvot. Täten menetelmässä täytyy ensimmäisenä selvittää tai arvioida otoksen havaintojen todennäköisyysjakauma. Tästä jakaumasta käytetään tässä menetelmässä nimitystä uskottavuusfunktio. Sen jälkeen parametrit määritetään siten, että kyseisen uskottavuusfunktion arvo maksimoituu. Yleensä uskottavuusfunktiona käytetään normaalijakaumaa. Koska yleensä uskottavuusfunktio on vaikeasti käsiteltävässä muodossa, siitä otetaan luonnollinen logaritmi. Täten tavoitefunktiona yleensä käytetään log-uskottavuusfunktiota. Jos esimerkiksi jakaumana käytetään normaalijakaumaa, niin log-uskottavuusfunktio on alla esitettyä muotoa. (Verbeek 2008, 172–176.)

$$logL = -\frac{N}{2}\log(2\pi s_\varepsilon^2) - \frac{1}{2}\sum_{t=1}^{N}\frac{\left(y_t - \beta_0 - \sum_{j=1}^{K}\beta_j x_{j,t}\right)^2}{s_\varepsilon^2}, \tag{27}$$

jossa $logL$ on log-uskottavuusfunktio. Suurimman uskottavuuden menetelmässä keskivirheet määritetään käyttäen alla olevia kaavoja (Verbeek 2008, 176–178).

$$logL = \sum_{t=1}^{N} logL_t \tag{28}$$

$$\hat{V}_{ML} = \left(-\frac{1}{N}\sum_{t=1}^{N}\frac{\partial^2 logL_t}{\partial\vec{\theta}\partial\vec{\theta}^T}\bigg|\hat{\vec{\theta}}\right)^{-1} \tag{29}$$

$$\hat{V}_{ML} = \begin{bmatrix} \sigma_{0,0}^2 & \cdots & \sigma_{K-1,0}^2 \\ \vdots & \ddots & \vdots \\ \sigma_{0,K-1}^2 & \cdots & \sigma_{K-1,K-1}^2 \end{bmatrix} \tag{30}$$

$$se(\hat{\beta}_i) = \sqrt{\sigma_{i,i}^2}, \tag{31}$$

jossa $logL_t$ on log-uskottavuusfunktio koskien ajankohtaa t ja $\hat{V}_{ML}$ on parametrien estimoitu kovarianssi-varianssi matriisi. Jos uskottavuusfunktio on oikein määritetty, niin kyseiselle menetelmällä estimoidut parametrit ovat tarkentuvia, asymptoottisesti tehokkaita ja asymptoottisesti normaalistijakautuneita. Niiden suhde todellisiin arvoihin on siis alla esitettyä muotoa. (Verbeek 2008, 177.)

$$\sqrt{N}\left(\hat{\vec{\theta}} - \vec{\theta}\right) \rightarrow \mathcal{N}(0, V_{ML}), \tag{32}$$

jossa $\vec{\theta}$ on parametrien todellisten arvojen vektori ja $\hat{\vec{\theta}}$ on suurimman uskottavuuden menetelmällä estimoitujen parametrien vektori. Jos kyseisessä menetelmässä käytetty uskottavuusfunktio ei ole täysin oikein spesifioitu, niin siitä saaduista estimaateista käytetään nimitystä kvasi-suurimman uskottavuuden estimaattorit. Kyseiset estimaatit ovat tietynlaisten ehtojen vallitessa edelleen tarkentuvia eli lähestyvät todellisia parametriarvoja pitkällä aikavälillä, mutta niiden kovarianssi-varianssi matriisi on erilainen kuin tavallisten suurimman uskottavuuden estimaattien kovarianssi-varianssi matriisi. Täten on tarpeen tehdä tarvittavat korjaukset kovarianssi-varianssi matriisin määrittämisessä, jos epäillään, että uskottavuusfunktio ei ole täysin oikein spesifioitu. Kovarianssi-varianssi matriisin määrittäminen kvasi-suurimman uskottavuuden tapauksessa on esitetty alla. (Verbeek 2008, 176–194.)

$$\frac{\partial logL_t}{\partial \vec{\theta}} \equiv s_t \tag{33}$$

$$V_{QMLE} = \left(\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} E\left(-\frac{\partial s_t}{\partial \vec{\theta}^T}\right)\right)^{-1} \lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} E(s_t s_t{}^T)\left(\lim_{N\to\infty}\frac{1}{N}\sum_{t=1}^{N} E\left(-\frac{\partial s_t}{\partial \vec{\theta}^T}\right)\right)^{-1}, \tag{34}$$

jossa V_{QMLE} on parametrien todellinen kovarianssi-varianssi matriisi. Kvasi-suurimman uskottavuuden estimaattoreiden suhde todellisiin arvoihin lähestyy alla esitettyä jakaumaa (Verbeek 2008, 193).

$$\sqrt{N}\left(\hat{\vec{\theta}} - \vec{\theta}\right) \rightarrow \mathcal{N}(0, V_{QMLE}) \tag{35}$$

3.3 Mallien sopivuuden vertailu sekä parametrien ja selitysasteen testaaminen

Monesti käytetty regressiomalli täytyy valita erilaisista vaihtoehdoista. Tällöin on tarpeellista verrata niiden sopivuutta suhteessa käytettyyn aineistoon. Lisäksi on hyvä verrata erilaisten regressiomallien sopivuutta keskenään, kun yritetään tulkita niistä saatuja tuloksia. Tässä työssä sopivuuden vertailuun käytetään Akaiken informaatiokriteeriä eli AIC:ta, selitysastetta eli R^2-arvoa ja korjattua selitysastetta eli $\bar{R}^2$-arvoa. Selitysaste kuvaa, että kuinka monta prosenttia havainnoista malli pystyy selittämään. Selitysasteella ei sinällään ole olemassa yhtä oikeaa viitearvoa, koska selitysasteen keskiarvo voi olla hyvinkin erisuuruinen eri asiayhteyksissä. Toisissa tapauksissa hyvinkin pienet arvot voivat olla suhteellisen hyviä ja toisissa hyvinkin suuret arvot voivat olla suhteellisen huonoja. Selitysasteen kaava on esitetty alla. (Verbeek 2008, 21–23.)

$$R^2 = \frac{1/(N-1)\sum_{i=1}^{N}(\hat{y}_t-\bar{y})^2}{1/(N-1)\sum_{i=1}^{N}(y_t-\bar{y})^2}, \tag{36}$$

jossa $\hat{y}_t$ on käytetyllä mallilla estimoitu selitettävän muuttujan arvo ajankohtana t, $\bar{y}$ on selitettävän muuttujan keskiarvo ja K on parametrien lukumäärä. Selitysasteen ongelmana on se, että sitä voidaan kasvattaa lisäämällä jopa tilastollisesti merkityksettömiä muuttujia. Toisin sanoen mitä enemmän muuttujia lisätään malliin, sitä suuremmaksi mallin selitysaste muuttuu riippumatta siitä, että onko kyseisillä muuttujilla minkäänlaista selitysvoimaa. Täten yleensä eri mallien paremmuuden vertailussa sen sijaan käytetään korjattua selitysastetta eli $\bar{R}^2$ - arvoa, joka ottaa huomioon tämän vaikutuksen. Kyseisen tunnusluvun määritetään käyttämällä alla olevaa kaavaa. (Verbeek 2008, 21–23.)

$$\bar{R}^2 = 1 - \frac{1/(N-K)\sum_{i=1}^{N}\hat{\varepsilon}_t^2}{1/(N-1)\sum_{i=1}^{N}(y_t-\bar{y})^2}, \tag{37}$$

jossa $\hat{\varepsilon}_t$ on residuaali ajankohtana t. Akaiken informaatiokriteeri ottaa huomioon sen, että kuinka hyvin kyseinen malli sopii kyseiseen aineistoon sekä sen, että kuinka monta muuttujaa siinä on. Kyseistä tunnuslukua tulkitaan siten, että mitä pienempi sen arvo on, sitä parempi malli on. AIC:n kaava on esitetty alla (Verbeek 2008, 61).

$$AIC = log\frac{1}{N}\sum_{t=1}^{N}\hat{\varepsilon}_t^2 + \frac{2K}{N} \tag{38}$$

Regressiomallien kaikki estimoidut parametrit ja selitysasteet täytyy testata. Syy miksi parametrit ja selitysasteet täytyy testata, johtuu siitä, että hallussa on vain otoksesta estimoidut arvot, eikä populaatiosta määritetyt todelliset arvot. Otoksesta estimoituihin arvoihin taas sisältyy aina jonkin verran sattumanvaraisuutta, joten niitä tarkkailemalla ei voida yksistään tehdä päätelmiä, vaan ne täytyy myös testata.

Ideana tilastotieteellisissä testeissä on se, että jos jokin testisuure noudattaa nollahypoteesin mukaista jakaumaa, niin silloin nollahypoteesi on tosi. Täten testaaminen lähtee liikkeelle nollahypoteesin asettamisesta. Sitten määritetään, mitä jakaumaa mikäkin testisuure noudattaa, jos nollahypoteesi on tosi. Lopuksi yritetään päätellä, noudattaako se sitä vai ei. Päättely siitä noudattaako testisuure jotakin jakaumaa, tehdään kyseisen jakauman kertymäfunktion arvon avulla.

Jos havaittu arvo on joko arvoltaan tai itseisarvoltaan, riippuen testistä, suurempi kuin kertymäfunktion arvo, niin silloin nollahypoteesi hylätään. Jos taas havaittu arvo on arvoltaan tai itseisarvoltaan, riippuen testistä, pienempi kuin kertymäfunktion arvo, nollahypoteesi hyväksytään. Kertymäfunktion arvo taas määräytyy täysin testaajan asettaman riskitason mukaisesti. Täten nollahypoteesin hylkääminen ja hyväksyminen määräytyy loppupeleissä asetetun riskitason mukaisesti. Täten on mielekästä määrittää kyseiseen testisuureeseen liittyvä p-arvo ja tehdä päättely p-arvon perusteella. P-arvo tarkoittaa alhaisinta riskitasoa millä nollahypoteesi voidaan hylätä (Gujarati & Porter 2009, 835). Mitä alhaisempi p-arvo taas on, sitä enemmän on todisteita sen puolesta, että nollahypoteesi on epätosi, ja päinvastoin. Tässä työssä kaikkien tilastotieteellisten testien kohdalla päättelyn tehdään siten, että jos p-arvo < 0.05, niin nollahypoteesi hylätään ja jos taas p-arvo ≥ 0.05, niin nollahypoteesi hyväksytään.

Parametrien kohdalla testataan, että onko vakio tai kyseisen selittävän muuttujan parametrin populaatioarvo nollasta poikkeava vai nolla. Jos parametrin populaatioarvo on nolla, niin silloin parametria koskevalla selittävällä muuttujalla ei ole minkäänlaista vaikutusta selitettävään muuttujaan. Toisin sanoen näiden välillä ei ole mitään yhteyttä. Jos taas parametrin populaatioarvo on nollasta poikkeava, niin silloin näiden kahden välillä on ainakin jonkinlainen yhteys. Parametrit testataan käyttäen t-testisuuretta. Kyseisen testisuureen kaava sekä nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla (Verbeek 2008, 23–25).

$H_0: \beta_i = 0$

$H_1: \beta_i \neq 0$

$$t_{havaittu} = \frac{\widehat{\beta_i}}{se(\widehat{\beta_i})} \sim t_{N-K}, \tag{39}$$

jossa $\widehat{\beta_i}$ on β_i:n estimoitu arvo ja $se(\widehat{\beta_i})$ on β_i:n estimaatin keskivirhe. Selitysasteen kohdalla taas testataan, että onko selitysaste nolla vai suurempi kuin nolla. Jos selitysaste on nolla, niin silloin kyseinen malli ei pysty tilastollisesti merkitsevästi selittämään enemmän vaihtelua selitettävässä muuttujassa kuin pelkällä vakiotermimallilla. Jos taas selitysaste on suurempi kuin nolla, niin silloin kyseinen malli pystyy tilastollisesti merkitsevästi selittämään paremmin vaihtelua selitettävässä muuttujassa kuin pelkkä vakiotermimalli. (Verbeek 2008, 21–23.) Selitysastetta testataan käyttäen F-testisuuretta. Kyseisen testin kaava sekä nollahypoteesi ja vaihtoehtoinen hypoteesi esitetään seuraavalla sivulla (Verbeek 2008, 27–29).

$H_0: R^2 = 0$

$H_1: R^2 > 0$

$$F_{havaittu} = \frac{R^2/(K-1)}{(1-R^2)/(N-K)} \sim F_{K-1,N-K} \tag{40}$$

3.4 Taustaoletukset, diagnostiset testit ja robustit keskivirheet

Regressiomalleihin sisältyy taustaoletuksia, jotka täytyy testata. Yksi näistä on aikasarjan stationaarisuus. Muuttuja Y on heikosti stationaarinen, jos alla esitetyt ehdot täyttyvät (Gujarati & Porter 2009, 737–741).

Ehto 1: $\forall t$: $\mathrm{E}(Y_t) = \mu$

Ehto 2: $\forall t$: $\mathrm{Var}(Y_t) = \mathrm{E}[(Y_t - \mu)^2] = \sigma^2$

Ehto 3: $\forall t, k$: $\mathrm{Cov}(Y_t, Y_{t+k}) = \mathrm{E}[(Y_t - \mu)(Y_{t+k} - \mu)] = \gamma_k$

Heikon stationaarisuuden ehtoja on siis yhteensä kolme. Ensimmäinen niistä on, että aikasarjan odotusarvo ei riipu ajankohdasta. Toinen niistä on, että aikasarjan varianssi ei riipu ajankohdasta. Kolmas ehto on, että aikasarjan kovarianssi ei riipu mistään muusta kuin viivepituudesta. (Gujarati & Porter 2009, 741.) Aikasarjan stationaarisuutta voidaan testata tilastollisilla testeillä. Tässä työssä testataan aikasarjojen stationaarisuutta laajennetulla Dickey-Fuller testillä eli ADF-testillä. ADF-testin kaava, nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla (Verbeek 2008, 286–287).

H_0: *Aikasarjassa on yksikköjuuri eli se on epästationaarinen* $(\delta = 0)$

H_1: *Aikasarja on stationaarinen* $(\delta < 0)$

$$\Delta Y_t = \propto + \delta Y_{t-1} + \sum_{i=1}^{\xi-1} \gamma_i \Delta Y_{t-i} + \varepsilon_t \tag{41}$$

$$\varepsilon_t \sim WN \tag{42}$$

$$\frac{\hat{\delta}}{se(\hat{\delta})} \sim DF_\tau \tag{43}$$

Yllä esitetyssä testisuureessa differenssiviivetermien lukumäärä määritetään siten, että niitä lisätään, kunnes mallin virhetermi noudattaa asymptoottisesti valkoisen kohinan prosessia. Monesti tähän tarvittavien termien lukumäärää ei tiedetä, joten on suositeltavaa käyttää suhteellisen montaa viivetermiä. (Verbeek 2008, 287.)

Toinen taustaoletuksista, joka täytyy testata, on virhetermin homoskedastisuus. Virhetermi varianssi on heteroskedastinen, jos se vaihtelee ajassa, ja se on homoskedastinen, jos se ei muutu ajassa. Tässä työssä virhetermin heteroskedastisuutta testataan Whiten testillä, kun kyseessä ei ole ARMA-perheen malli. Whiten testissä määritetään alla esitetty malli. (Gujarati & Porter 2009, 64–65 & 365 & 386–388.)

$$\hat{\varepsilon}_t^2 = \gamma_0 + \sum_{i=1}^{K-1} \delta_i x_{i,t} + \sum_{i=1}^{K-1} \zeta_i x_{i,t}^2 + \sum_{i=1}^{K-2} \sum_{j=i+1}^{K-1} \eta_{i,j} x_{i,t} x_{j,t} + \nu_t \quad (44)$$

$$\nu_t \sim WN \quad (45)$$

Yllä esitetystä mallista määritetään selitysaste. Sitten kyseistä selitysastetta testaamalla tehdään päätelmät virhetermin heteroskedastisuudesta. Kyseisen testin testisuure, nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla (Gujarati & Porter 2009, 386–388).

$$H_0: Virhetermi\ on\ homoskedastinen\ (\delta_1 = \cdots = \eta_{K-2,K-1} = 0)$$

$$H_1: Virhetermi\ on\ heteroskedastinen$$

$$NR_W^2 \sim \chi_{\kappa-1}^2, \quad (46)$$

jossa R_W^2 on Whiten testimallin selitysaste ja κ on Whiten testimallin parametrien lukumäärä. Kun taas käytetään ARMA-perheen mallia, niin Whiten testin sijaan käytetään ARCH(q)-testiä. Kyseisessä testissä testataan siis sitä, että onko virhetermin varianssi ARCH-prosessi. Lyhenne ARCH tulee sanoista AutoRegressive Conditional Heteroskedasticity ja kyseisen mallin mukaan virhetermin varianssi riippuu virhetermin edeltävien periodien arvojen neliöistä. ARCH-mallissa lähdetään liikkeelle alla esitetyistä yhtälöistä. (Verbeek 2008, 312.)

$$\forall t: E(\varepsilon_t) = 0 \quad (47)$$

$$\sigma_\varepsilon^2 = E\left[\left(\varepsilon_t - E(\varepsilon_t)\right)^2\right] = E[\varepsilon_t^2] \quad (48)$$

$$\sigma_t^2 = E[\varepsilon_t^2 | \mathrm{I}_{t-1}], \quad (49)$$

jossa σ_ε^2 on virhetermin varianssi, σ_t^2 on virhetermin ehdollinen varianssi ajankohtana t ja I_{t-1} on informaatiojoukko ajankohtana t-1. Yllä esitetyistä yhtälöistä käy ilmi, että virhetermin varianssi on yhtä suuri kuin virhetermin neliön odotusarvo. Täten jos virhetermin neliön odotusarvo on ajankohdasta riippuva, niin vääjäämättä virhetermin varianssi on myös ajankohdasta riippuva. ARCH-testillä testataan juuri tätä. Kyseinen testi lähtee liikkeelle siitä, että määritetään seuraavalla sivulla esitetty malli (Mills & Markellos 2008, 189–191).

$$\hat{\varepsilon}_t^2 = \lambda_0 + \sum_{i=1}^{q} \lambda_i \hat{\varepsilon}_{t-i}^2 + \nu_t \tag{50}$$

$$\nu_t \sim WN \tag{51}$$

Kyseisestä mallista määritetään sitten selitysaste. Selitysasteesta sitten määritetään alla esitetty testisuure, jonka nollahypoteesi ja vaihtoehtoinen hypoteesi on myös esitetty alla (Mills & Markellos 2008, 189–191).

$$H_0: Virhetermi\ on\ homoskedastinen\ (\lambda_1 = \cdots = \lambda_q = 0)$$

$$H_1: Virhetermi\ sisältää\ ARCH - ominaisuuden$$

$$R_A^2 \sim F_{q,N-q-1}, \tag{52}$$

jossa R_A^2 on ARCH testimallin selitysaste ja q on ARCH testimallin viivetermien lukumäärä. Jos käy ilmi, että virhetermi sisältää heteroskedastisuutta, niin parametrien testaamisessa on käytettävä heteroskedastisuuden huomioon ottavia keskivirheitä eli Whiten keskivirheitä tavallisten keskivirheiden sijaan. Syynä tähän on se, että tavalliset keskivirheet muuttuvat harhaisiksi, kun virhetermi sisältää heteroskedastisuutta. Whiten keskivirheet taas ottavat tämän vaikutukset huomioon ja tekevät tällöin parametrien keskivirheet harhattomiksi. Koska jopa silloin, kun virhetermi on homoskedastinen, Whiten keskivirheet ovat harhattomat, niin Whiten keskivirheitä voidaan käyttää myös tapauksissa, joissa ei olla varmoja siitä, että onko virhetermi homoskedastinen vai heteroskedastinen. Whiten keskivirheet voidaan määrittää alla esitetyillä kaavoilla. (Verbeek 2008, 93–94.)

$$\hat{V}_{HC} = \left(\sum_{t=1}^{N} \vec{x}_t \vec{x}_t^T\right)^{-1} \left(\sum_{t=1}^{N} \hat{\varepsilon}_t^2 \vec{x}_t \vec{x}_t^T\right) \left(\sum_{t=1}^{N} \vec{x}_t \vec{x}_t^T\right)^{-1} \tag{53}$$

$$\hat{V}_{HC} = \begin{bmatrix} \sigma_{0,0}^{2,HC} & \cdots & \sigma_{K-1,0}^{2,HC} \\ \vdots & \ddots & \vdots \\ \sigma_{0,K-1}^{2,HC} & \cdots & \sigma_{K-1,K-1}^{2,HC} \end{bmatrix} \tag{54}$$

$$hcse(\widehat{\beta_\iota}) = \sqrt{\sigma_{i,i}^{2,HC}}, \tag{55}$$

jossa $\hat{V}_{HC}$ on parametrien estimoitu heteroskedastisuuden huomioon ottava kovarianssi-varianssi matriisi, $\sigma_{i,i}^{2,HC}$ on β_i:n estimaatin estimoitu heteroskedastisuuden huomioon ottava varianssi ja $hcse(\widehat{\beta_\iota})$ on β_i:n estimaatin Whiten keskivirhe.

Kolmas regressiomallin taustaoletus on virhetermin autokorreloimattomuus. Virhetermi sisältää autokorrelaatiota, jos se korreloi aikaisempien arvojensa kanssa. Virhetermin autokorrelaatiota testataan tässä työssä käyttämällä Ljung-Box Q(M)-testisuuretta. Kyseisen testisuureen kaava, nollahypoteesi sekä vaihtoehtoinen hypoteesi on esitetty alla. (Gujarati & Porter 2009, 66–68 & 413 & 753–754.)

H_0: *Virhetermi ei sisällä autokorrelaatiota* $(\rho_0 = \cdots = \rho_M = 0)$

H_1: *Virhetermi sisältää autokorrelaatiota*

$$LB = N(N+2)\sum_{l=1}^{M}\left(\frac{\hat{\rho}_l^2}{N-l}\right) \sim \chi_M^2, \tag{56}$$

jossa $\hat{\rho}_l$ on otosautokorrelaatiokerroin viiveellä l ja M on käytettyjen viiveiden lukumäärä. Jos käy ilmi, että virhetermi sisältää autokorrelaatiota, heteroskedastisuutta tai molempia, on parametrien testaamisessa käytettävä Newey-Westin keskivirheitä. Newey-Westin keskivirheet ottavat siis huomioon sekä autokorrelaation että heteroskedastisuuden vaikutuksen, ja täten niiden avulla voidaan testata luotettavasti parametreja, vaikka virhetermi olisikin autokorreloitunut, heteroskedastinen tai molempia. Kyseiset keskivirheet voidaan määrittää käyttämällä alla esitettyjä kaavoja. (Verbeek 2008, 118–119.)

$$w_j = 1 - \frac{j}{H} \tag{57}$$

$$S^* = \frac{1}{N}\sum_{t=1}^{N}\hat{\varepsilon}_t^2\vec{x}_t\vec{x}_t^T + \frac{1}{N}\sum_{j=1}^{H-1}w_j\sum_{s=j+1}^{N}\hat{\varepsilon}_s\hat{\varepsilon}_{s-j}\left(\vec{x}_s\vec{x}_{s-j}^T + \vec{x}_{s-j}\vec{x}_s^T\right) \tag{58}$$

$$\hat{V}_{HAC} = \left(\sum_{t=1}^{N}\vec{x}_t\vec{x}_t^T\right)^{-1}NS^*\left(\sum_{t=1}^{N}\vec{x}_t\vec{x}_t^T\right)^{-1} \tag{59}$$

$$\hat{V}_{HAC} = \begin{bmatrix} \sigma_{0,0}^{2,HAC} & \cdots & \sigma_{\mathrm{K}-1,0}^{2,HAC} \\ \vdots & \ddots & \vdots \\ \sigma_{0,K-1}^{2,HAC} & \cdots & \sigma_{K-1,K-1}^{2,HAC} \end{bmatrix} \tag{60}$$

$$hacse(\hat{\beta}_\iota) = \sqrt{\sigma_{i,i}^{2,HAC}}, \tag{61}$$

jossa w_j on käytetty painokerroin, $\hat{V}_{HAC}$ on parametrien estimoitu autokorrelaation ja heteroskedastisuuden huomioon ottava kovarianssi-varianssi matriisi, $\sigma_{i,i}^{2,HAC}$ on β_i:n estimaatin estimoitu autokorrelaation ja heteroskedastisuuden huomioon ottava varianssi ja $hacse(\hat{\beta}_\iota)$ on β_i:n estimaatin Newey-Westin keskivirhe.

Neljäs virhetermin taustaoletus on, että virhetermi on normaalistijakautunut. Virhetermien normaalijakautuneisuutta tässä työssä testataan Jarque-Bera testillä. Kyseisen testin testisuure sekä nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla (Gujarati & Porter 2009, 131–132).

$H_0: Virhetermi\ on\ normaalistijakautunut$

$H_1: Virhetermi\ ei\ ole\ normaalistijakautunut$

$$JB = N\left[\frac{Vinous^2}{6} + \frac{(Huipukkuus-3)^2}{24}\right] \sim \chi_2^2 \qquad (62)$$

Toisaalta jos kyseessä on suurotos, niin virhetermin normaalijakautuneisuuteen liittyvän oletuksen paikkansapitämättömyys ei aiheuta ongelmia. Syynä tähän on se, että estimaattoreiden asymptoottinen jakauma on normaalijakauma (Verbeek 2008, 35). Täten mitä suurempi otos on, sitä enemmän estimaattoreiden jakauma lähestyy normaalijakaumaa.

Multikollineaarisuus, eli se, että selittävät muuttujat korreloivat keskenään, ei aiheuta ongelmia, jos kyseinen korrelaatio ei ole kooltaan liian merkittävää (Verbeek 2008, 43).

3.5 Ilmiön tutkimiseen käytetyt mallit

Yksi tämän työn tutkimuskysymyksistä on, että onko aikasarja-momentum -ilmiötä havaittavissa Suomen osakemarkkinoilla. Toinen tämän työn tutkimuskysymyksistä on, että jos kyseinen ilmiö on havaittavissa, niin syntyykö se rationaalisten vai behavioraalisten tekijöiden seurauksena. Näiden tutkimuskysymyksien selvittämiseksi on tarpeellista käyttää itse ilmiön tutkimiseen käytettäviä malleja. Lisäksi ilmiötä tutkimalla saadaan selville, että mikä on optimaalinen takaisinkatsomisaika koskien luvussa 2.3 esitettyjä sijoitussääntöjä.

Kyseisen ilmiön tutkimiseen käytetään tässä työssä neljää eri mallia. Ensimmäinen näistä malleista on ARMA(p, q)-malli. ARMA-mallin ideana on tutkia, että kuinka jokin stokastinen prosessi käyttäytyy. Tässä mallissa menneitä tuottoja käsitellään siis joidenkin satunnaismuuttujien realisaatioina. Lyhenne ARMA tulee sanoista AutoRegressive Moving Average. Kyseinen malli koostuu siis kahdesta osasta. Ensimmäinen osa on autoregressiivinen eli selitettävän muuttujan omien viiveiden vaikutus selitettävän muuttujan nykyiseen ja tuleviin arvoihin. Toinen osa on liukuva keskiarvo, joka tässä kontekstissa viittaa virhetermin viiveiden vaikutukseen kohdistuen selitettävään muuttujaan. Tämän mallin määrittämiseksi täytyy valita siinä käytettävien viivepituuksien lukumäärä. Tässä työssä viivepituudet on valittu siten, että mallia koskeva AIC-tunnusluku minimoituu. Kyseisen mallin estimointiin käytetään tässä työssä suurimman uskottavuuden menetelmää. Kyseinen malli on esitetty seuraavalla sivulla. (Verbeek 2008, 269–278.)

$$R_t = \alpha_0 + \sum_{j=1}^{p} \varphi_j R_{t-j} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{63a}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2) \tag{64}$$

Yllä esitetystä ARMA-mallista on vaikea tehdä päätelmiä koskien prosessin käyttäytymistä pelkkiä mallin parametrien arvoja tarkastelemalla. Tästä syystä siitä on hyvä määrittää impulssivastefunktio. Impulssivastefunktio kuvaa sitä, että miten selitettävää muuttujaa koskevat shokit vaikuttavat sen nykyiseen ja tuleviin arvoihin. Stationaarisen ARMA-mallin tapauksessa impulssivastefunktio voidaan määrittää käyttämällä alla esitettyjä kaavoja. (Mills & Markellos 2008, 124–125.)

$$\psi(L) = \phi(L)^{-1}\theta(L) \tag{65}$$

$$\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i \tag{66}$$

$$\theta(L) = 1 + \sum_{j=1}^{q} \theta_j L^j \tag{67}$$

$$\psi(L) = \sum_{j=0}^{\infty} \psi_j L^j \tag{68}$$

$$\psi_j = \frac{\partial R_{t+j}}{\partial \varepsilon_t}, \tag{69}$$

jossa ψ_j on impulssivastefunktion arvo koskien ajankohtaa t+j. Toinen käytetyistä malleista on AR(p)-GARCH(v, r)-in-Mean -malli. Rahoitusaikasarjoissa havaitaan useasti volatiliteetin klusteroitumista. Yleensä ottaen siis suuria shokkeja seuraavat suuret shokit ja pieniä shokkeja pienet shokit. Tästä syystä osakemarkkinoilla on sekä korkean volatiliteetin että alhaisen volatiliteetin aikoja. Sijoituskohteiden riskisyys voi siis muuttua ajassa. Tämän takia on luontaista mallintaa virhetermin varianssia siten, että se muuttuu ajassa. Yksi tapa mallintaa ajassa muuttuvaa varianssia on käyttää GARCH-mallia. GARCH-malli (Bollerslev 1986) on ARCH-mallin (Engle 1982) laajennus. Lyhenne GARCH tulee sanoista Generalized AutoRegressive Conditional Heteroskedasticity. GARCH-prosessissa virhetermin ehdollinen varianssi määräytyy virhetermin viiveiden neliöiden ja virhetermin ehdollisen varianssin omien viiveiden perusteella. (Verbeek 2008, 311–313.)

GARCH-in-Mean mallin (Bollerslev, Engle & Wooldridge 1988) ideana on taas ottaa huomioon sijoituskohteen tuottojen mahdollisesti ajassa muuttuvan varianssin vaikutus sijoituskohteen tuottojen odotusarvoon. Rahoitusteorian mukaan tehokkailla markkinoilla tuotot vastaavat riskiä. Täten, jos riski muuttuu ajassa, niin pitäisi myös tuottojen odotusarvon muuttua vastaavalla tavalla pitäen tuotto-riski -suhde vakiona. Tässä tapauksessa käytettyyn

GARCH-in-Mean malliin on lisätty vielä AR(p)-komponentti, koska halutaan saada selville, että onko tuottojen autokovarianssirakenteen takana muutokset sijoituskohteen riskisyydessä vai onko sen sijaan kyseessä jokin behavioraalinen tekijä. AR-komponentin viivepituus on tässä työssä valittu siten, että AIC-tunnusluku on verratuista vaihtoehdoista pienin. GARCH-komponentin viivepituudet on taas valittu siten, että mallin vakiotermi on tilastollisesti merkitsevä ja positiivinen sekä mallin muiden parametrien summa on pienempi kuin yksi. Malleista, jotka täyttävät nämä kriteerit, pyritään valitsemaan sopivin. Mallin sopivuus arvioidaan käyttämällä AIC-tunnuslukua ja parametrien tilastollista merkitsevyyttä. (Verbeek 2008, 269–275 & 312–317.) Tässä työssä kyseisen mallin parametrit estimoidaan käyttämällä suurimman uskottavuuden menetelmää. Kyseinen malli on esitetty alla (Bollerslev, Engle & Wooldridge 1988).

$$R_t = \alpha_0 + \sum_{j=1}^{p} \varphi_j R_{t-j} + \alpha_1 h_t + u_t \qquad (70a)$$

$$u_t = \epsilon_t \sqrt{h_t} \qquad (71)$$

$$\epsilon_t \sim NID(0,1) \qquad (72)$$

$$h_t = \beta_0 + \sum_{i=1}^{v} \omega_i h_{t-i} + \sum_{j=1}^{r} \lambda_j u_{t-j}^2, \qquad (73a)$$

jossa h_t on ehdollinen varianssi ajankohtana t. Yllä esitetystä mallista saatuja parametreja hyödyntämällä on mahdollista saada lisäksi selville tuottojen pitkän aikavälin ehdoton varianssi. Ehdoton pitkän aikavälin varianssi ja keskihajonta on tällöin alla esitettyä muotoa (Verbeek 2008, 313).

$$\sigma^2 = \frac{\beta_0}{1 - \sum_{i=1}^{v} \omega_i - \sum_{j=1}^{r} \lambda_j} \qquad (74)$$

$$\sigma = \sqrt{\sigma^2}, \qquad (75)$$

jossa σ^2 on pitkän aikavälin ehdoton varianssi ja σ on pitkän aikavälin ehdoton keskihajonta. Varianssi on stationaarinen, jos $\sum_{i=1}^{v} \omega_i + \sum_{j=1}^{r} \lambda_j < 1$. Jos kyseinen summa on lähellä ykköstä, niin silloin volatiliteetin pysyvyys on korkea. Jos taas kyseinen summa on tasan yksi, niin volatiliteetti shokit ovat vaikutukseltaan pysyviä ja täten volatiliteetti on epästationaarinen. (Verbeek 2008, 313.)

Kolmas malleista on probit-malli. Probit-mallissa selitettävä muuttuja on binäärinen diskreetti muuttuja. Selittävät muuttujat ovat tässä mallissa taas jatkuva-arvoisia muuttujia, joiden arvoilla yritetään selittää selitettävän muuttujan todennäköisyyttä saada arvokseen nolla tai yksi. Kun estimoidaan selittävien muuttujien arvojen vaikutusta kyseiseen todennäköisyyteen, niin on tarpeellista määrittää jokin todennäköisyysjakauma. Probit-mallin tapauksessa todennäköisyysjakaumana käytetään normaalijakaumaa. Tässä tapauksessa käytetty diskreetti muuttuja saa arvon nolla, kun saatu tuotto on pienempi tai yhtä suuri kuin nolla, ja arvon yksi, kun saatu tuotto on suurempi kuin nolla. Kyseessä on siis latentti muuttuja malli. Ideana on siis saada selville, että onko menneillä tuotoilla vaikutusta todennäköisyyteen saada voittoa tai tappiota tällä periodilla. Mallin estimointiin käytetään suurimman uskottavuuden menetelmää. Käytetty malli on alla esitettyä muotoa. (Verbeek 2008, 199–203.)

$$y_t = \begin{cases} 1, & jos\ R_t > 0 \\ 0, & jos\ R_t \leq 0 \end{cases} \tag{76}$$

$$P(y_t = 1|W_t) = \Phi(W_t) \tag{77}$$

$$W = \beta_0 + \sum_{j=1}^{p} \varphi_j R_{t-j} + \varepsilon_t \tag{78a}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{79}$$

jossa Φ on standardi normaalijakauman kertymäfunktio. Kyseisestä mallista se, että onko yksittäisen selittävän muuttujan vaikutus selitettävän muuttujaan positiivinen vai negatiivinen, saadaan selville kyseisen muuttujan parametria tarkastelemalla, mutta sen vaikutuksen suuruutta taas ei. Syynä tähän on se, että kaikkien selittävien muuttujien suuruudella on vaikutusta yksittäisen selittävän muuttujan rajavaikutukseen. Kyseinen rajavaikutus on alla esitettyä muotoa. (Verbeek 2008, 200–202.)

$$\frac{\partial y_t}{\partial R_{t-j}} = \phi(W_t)\varphi_j, \tag{80}$$

jossa ϕ on standardi normaalijakauman tiheysfunktio. Neljäs ilmiön tutkimiseen käytetyistä malleista on kahden tilan regiimin muutosmalli. Regiimin muutosmallia voidaan pitää ARMA-prosessin epälineaarisena laajennuksena. Kahden tilan regiimin muutosmallissa sijoituskohde voi olla kahdessa eri tilassa. Eri tilassa ollessaan sijoituskohteen tuotoilla on eri odotusarvo ja keskihajonta. Tässä työssä tiloiksi määritetään laskumarkkina ja nousumarkkina. Sijoituskohdetta koskeva tila taas voi vaihtua yhdestä toiseen jollakin todennäköisyydellä. Tässä työssä sijoituskohteen tilaa kuvaava binäärinen diskreetti muuttuja saa arvon nolla, kun kyseessä on nousumarkkina ja arvon yksi, kun kyseessä on laskumarkkinat. Kyseinen tilamuuttuja ja sen arvon vaihtumiseen liittyvät todennäköisyydet on esitetty seuraavalla sivulla. (Mills & Markellos 2008, 216–222.)

$$S_t = \begin{cases} 1, kun\ kyseessä\ on\ laskumarrkina \\ 0, kun\ kyseessä\ on\ nousumarkkina \end{cases} \tag{81}$$

$$P(S_t = 1|S_{t-1} = 1) = p \tag{82}$$

$$P(S_t = 0|S_{t-1} = 1) = 1 - p \tag{83}$$

$$P(S_t = 1|S_{t-1} = 0) = 1 - q \tag{84}$$

$$P(S_t = 0|S_{t-1} = 0) = q, \tag{85}$$

jossa S_t on sijoituskohdetta koskeva tila ajankohtana t. Regiimin muutosmallin johtaminen lähtee liikkeelle siitä, että epästationaarinen tasomuuttuja, tässä tapauksessa sijoituskohteen logaritmoidut hinnat, voidaan esittää alla esitettynä prosessina (Mills & Markellos 2008, 216).

$$logP_t = z_t + u_t \tag{86}$$

$$z_t = \propto_0 + \propto_1 S_t + z_{t-1} \tag{87}$$

$$u_t \sim WN \tag{88}$$

Yllä esitetyistä yhtälöistä voidaan taas johtaa logaritmoituja tuottoja koskeva prosessi, joka on esitetty alla.

$$\Delta logP_t = \Delta z_t + \Delta u_t \tag{89}$$

$$\Delta logP_t = R_t \tag{90}$$

$$\Delta z_t = \alpha_0 + \alpha_1 S_t \tag{91}$$

$$\Delta u_t = \varepsilon_t \tag{92}$$

Yhdistämällä yllä esitetyt tulokset saadaan stationaarisia logaritmoituja tuottoja kuvaava malli. Mallin parametrien estimointiin käytetään tässä työssä suurimman uskottavuuden menetelmää. Kyseinen malli on esitetty alla (Mills & Markellos 2008, 216–222).

$$R_t = \propto_0 + \propto_1 S_t + \varepsilon_t \tag{93}$$

$$\varepsilon_t \sim WN \tag{94}$$

Täten nousumarkkinoiden aikana tuottojen odotusarvo on $\propto_0$ ja laskumarkkinoiden aikana $\propto_0 + \propto_1$. Regiimin muutosmallilla saadaan siis selville, että mikä on sijoituskohteen tuottojen odotusarvo ja keskihajonta missäkin tilassa. Lisäksi saadaan selville, että millä todennäköisyydellä tuottojen odotusarvo muuttuu positiivisesta negatiiviseksi tai päinvastoin.

Täten voidaan päätellä, että jos jollain ajanjaksolla on saatu keskimäärin voittoa tai tappiota, niin kuinka todennäköistä on, että sitä seuraavalla periodilla saadaan myös voittoa tai tappiota. Tässä on toki hyvä pohtia, että kuinka pitkällä aikavälillä otoskeskiarvo konvergoituu lähelle odotusarvoa. Toisin sanoen, millä otoskoolla otoskeskiarvo on hyvä estimaattori odotusarvosta. Toisaalta mitä suurempaa otoskokoa käytetään, sitä suurempi on todennäköisyys sille, että tuottojen odotusarvo on jo ehtinyt vaihtua.

3.6 Sijoitussääntöjen kannattavuuden tutkimiseen käytetyt mallit

Tämän työn kolmas tutkimuskysymys on, että voiko sijoittaja hyödyntää kyseistä ilmiötä sijoitustoiminnassaan. Täten on tarpeellista verrata eri sijoitussääntöjen tuottojen odotusarvojen, riskisyyden ja tuotto-riski -suhteiden eroavaisuuksia erilaisilla malleilla.

Ensimmäiseksi testataan, että tuottavatko aikasarja-momentum -sijoitussääntö tai liukuva keskiarvon sijoitussääntö odotusarvoisesti korkeampaa tuottoa kuin osta-ja-pidä -sijoitussääntö. Tämän testaamiseksi käytetään riippumattomien otosten t-testiä. Tässä työssä riippumattomien otosten t-testissä oletetaan, että eri sijoitussäännöillä on eri varianssit. Riippumattomien otosten t-testin nollahypoteesi hylätään, jos havaittu t-arvo on suurempi kuin taulukko arvo. Tässä tapauksessa havaittua arvoa ei siis verrata itseisarvoon. Kyseisen testin kaava sekä sen nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla (Newbold, Carlson & Thorne 2010, 426–427).

$H_0: E(R_s) \leq E(R_o)$

$H_1: E(R_s) > E(R_o)$

$$t_{havaittu} = \frac{\bar{R}_S - \bar{R}_O}{\sqrt{\frac{s_S^2}{N} + \frac{s_O^2}{M}}} \sim t_v \tag{95}$$

$$v = \frac{\left[\left(\frac{s_S^2}{N}\right) + \left(\frac{s_O^2}{M}\right)\right]^2}{\frac{\left(\frac{s_S^2}{N}\right)^2}{N-1} + \frac{\left(\frac{s_O^2}{M}\right)^2}{M-1}}, \tag{96}$$

jossa $E(R_j)$ on sijoitustyyli j:n tuottojen odotusarvo, $\bar{R}_s$ on vaihtoehtoisen sijoitussäännön tuottojen keskiarvo, $\bar{R}_o$ on osta-ja-pidä -tuottojen keskiarvo, s_s^2 on vaihtoehtoisen sijoitussäännön tuottojen varianssi, s_o^2 on osta-ja-pidä -tuottojen varianssi, N on vaihtoehtoisen sijoitussäännön tuottojen lukumäärä ja M on osta-ja-pidä -tuottojen lukumäärä. Sitten mitataan eri sijoitussääntöjen riskisyyttä. Sijoitussääntöjen riskisyys mitataan tässä työssä GARCH(v, r)-mallilla. GARCH-malli on avattu tarkemmin luvussa 3.5. Käytetty malli on esitetty alla (Bollerslev 1986).

$$R_t = \alpha_0 + u_t \tag{97}$$

$$u_t = \epsilon_t \sqrt{h_t} \tag{71}$$

$$\epsilon_t \sim NID(0,1) \tag{72}$$

$$h_t = \beta_0 + \sum_{i=1}^{v} \omega_i h_{t-i} + \sum_{j=1}^{r} \lambda_j u_{t-j}^2, \tag{98}$$

Yllä esitetyn mallin viivepituudet on valittu samalla tavalla kuin 3.5 luvussa käsitellyn AR-GARCH-in-Mean mallin tapauksessa. Kyseisestä mallista saaduista parametreista voidaan sitten määrittää pitkän aikavälin keskihajonta. Pitkän aikavälin keskihajonnan kaava on esitetty luvussa 3.5. Eri sijoitussääntöjen riskisyyden vertailu tapahtuu sitten vertailemalla niiden pitkän aikavälin keskihajontoja. Mitä pienempi se on, sitä pienempi on siihen sisältyvä riskisyys.

Viimeiseksi mitataan eri sijoitussääntöjen tuotto-riski -suhteet. Tuotto-riski -suhteiden määrittäminen tapahtuu tässä työssä kahdella eri keinolla. Ensimmäiseksi määritetään yksinkertainen tuotto-riski -suhdetta kuvaava tunnusluku. Tässä työssä sellainen on Sharpen luku. Sharpen luku on yleisesti käytetty tuotto-riski -suhteen mitta, jossa suhteutetaan sijoituskohteen tuottama riskipreemio siihen liittyvään volatiliteettiin. Mitä suurempi on Sharpen luku, sitä parempi on sijoituskohteen tuottama tuotto-riski -suhde. (Bodie, Kane & Marcus 2014, 134.) Kyseinen kaava on esitetty alla (Brealey, Myers & Allen 2014, 196).

$$SR = \frac{\overline{RP}}{\sigma_{RP}}, \tag{99}$$

jossa $\overline{RP}$ on logaritmoidun riskipreemion keskiarvo ja σ_{RP} on logaritmoidun riskipreemion otoskeskihajonta. Riskipreemioiden ollessa autokorreloituneita, Sharpen luku vääristyy. Jos riskipreemioiden autokorrelaatio on positiivista, niin Sharpen luku yliarvioi sijoitussäännön tuotto-riski -suhdetta. Jos taas autokorrelaatio on negatiivista, niin Sharpen luku aliarvioi sijoitussäännön tuotto-riski -suhdetta. Täten, jos havaitaan tilastollisesti merkitsevää autokorrelaatiota, on tarpeellista ottaa huomioon mahdollinen autokorrelaatio Sharpen luvun määrittämisessä. Ensimmäisen viiveen autokorrelaation huomioon ottavan Sharpen luvun kaava on esitetty seuraavalla sivulla. (Lo 2002, 36–41.)

$$SR_{AC} = SR\left[1 + \frac{2\rho_1}{1-\rho_1}\left(1 - \frac{1-\rho_1^N}{N(1-\rho_1)}\right)\right]^{-\frac{1}{2}}, \tag{100}$$

jossa SR_{AC} on autokorrelaation huomioon ottava Sharpen luku ja ρ_1 on ensimmäisen viiveen autokorrelaatiokerroin. Sharpen luku perustuu lisäksi implisiittisesti normaalijakauma-oletukseen. Täten, jos kyseisen sijoitussäännön tuottojakauma on normaalijakaumasta poikkeava, niin täytyy tämä ottaa huomioon Sharpen lukua määrittäessä. Sharpen luku voidaan tällöin määrittää käyttäen alla esitettyä kaavaa. (Pézier & White 2008, 42.)

$$SR_{DC} = SR\left[1 + \left(\frac{Vinous}{6}\right)SR - \left(\frac{[Huipukkuus-3]}{24}\right)SR^2\right], \tag{101}$$

jossa SR_{DC} on jakauman muodon huomioon ottava Sharpen luku. Sharpen luvun tilastollinen merkitsevyys voidaan taas testata t-testisuureella. Kyseinen testisuure ja siihen liittyvä nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla. (Sharpe 1994, 51.)

$$H_0: SR_i = 0$$

$$H_1: SR_i \neq 0$$

$$t_{havaittu} = \sqrt{N} * SR_i \sim t_N, \tag{102}$$

jossa SR_i on Sharpen luku tyyppiä i. Sen jälkeen mitataan sijoitussääntöjen tuotto-riski -suhdetta kahdella eri riskimallilla. Molempien riskimallien kohdalla määritetään Jensenin alfa, joka on epänormaalien tuottojen mitta. Jos Jensenin alfa on positiivinen ja tilastollisesti merkitsevä, niin kyseinen sijoitussääntö tuottaa systemaattisesti ylituottoa. Jos se on negatiivinen ja tilastollisesti merkitsevä, niin kyseinen sijoitussääntö tuottaa systemaattisesti alituottoa. Jos taas se ei ole tilastollisesti merkitsevä, niin kyseinen sijoitussääntö ei tuota epänormaalia tuottoa. Tehokkailla markkinoilla Jensenin alfa olisi kaikilla sijoitussäännöillä nolla. (Bodie, Kane & Marcus 2014, 840.) Ensimmäinen tässä työssä käytetyistä riskimalleista on CAPM. Lyhenne CAPM tulee sanoista Capital Asset Pricing Model. Kyseinen malli on alla esitettyä muotoa. Kyseinen malli estimoidaan tässä työssä pienimmän neliösumman menetelmällä. (Bodie, Kane & Marcus 2014, 302.)

$$RP_t = \alpha + \beta_M RP_{M,t} + \varepsilon_t \tag{103}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{104}$$

jossa RP_t on sijoituskohteen riskipreemio ajankohtana t, α on sijoituskohteen Jensenin alfa ja $RP_{M,t}$ on markkinaportfolion m riskipreemio ajankohtana t. Ongelmana CAPM:ssä on se, että systemaattista riskiä mitataan kyseisessä mallissa vain yhdellä tekijällä eli kyseessä on yksitekijämalli. Todellisuudessa systemaattinen riski koostuu monesta eri tekijästä ja täten käytetyn mallin pitäisi reflektoida tätä. Parempi riskimalli on siis monitekijämalli, eikä yksitekijämalli. Monitekijämallista käy lisäksi ilmi tarkemmin, että juuri mitkä riskit ovat kyseisen sijoituskohteen kohdalla kuinka suuret. (Bodie, Kane & Marcus 2014, 326.)

Malli, minkä uskotaan kuvaavan hyvin systemaattisen riskin eri komponentteja, on Fama-French kolmen tekijän malli, joka on CAPM-mallin laajennus. Fama & French (1996.) argumentoivat, että mallissa käytetyt muuttujat eivät ole itsessään riskitekijöitä, vaan ne vain edustavat joitain tällä hetkellä tuntemattomia riskitekijöitä. Kyseinen malli on esitetty alla.

$$RP_t = \alpha + \beta_M RP_{M,t} + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \varepsilon_t \tag{105}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{106}$$

jossa SMB_t on pienen markkina-arvon yritysten tuotto miinus suuren markkina-arvon yritysten tuotto ajankohtana t, HML_t on korkean BM-luvun omaavien yritysten tuotto miinus alhaisen BM-luvun omaavien yritysten tuotto ajankohtana t. Kyseisen mallin muuttujat lasketaan seuraavalla tavalla. Ensin jaetaan joidenkin markkinoiden yritykset markkina-arvon mukaisesti niin, että jakajana toimii mediaani yrityksen markkina-arvo. Sitten luodaan kuvitteellinen portfolio, jossa ostetaan pienen markkina-arvon yrityksiä eli yrityksiä, joiden markkina-arvo on pienempi kuin mediaani yrityksen, ja lyhyeksi myydään suuren markkina-arvon yrityksiä eli yrityksiä, joiden markkina arvo on suurempi kuin mediaani yrityksen. Kyseisen kuvitteellisten portfolion tuotto on SMB. HML lasketaan muuten täysin samalla tavalla kuin SMB, paitsi että jakavana tekijänä käytetään BM-lukua. Monet empiiriset tutkimukset ovat osoittaneet, että Fama-French kolmen tekijän malli pystyy ennustamaan tuottoja paremmin kuin CAPM. (Bodie, Kane & Marcus 2014, 426–433.) Tässä työssä ei käytetä Fama-French kolmen tekijän mallia, koska kyseiseen malliin tarvittavaa aineistoa ei löytynyt euroina vaan pelkästään Yhdysvaltojen dollareina.

Toisena riskimallina sen sijaan käytetään tässä työssä APT-mallia. Lyhenne APT tulee sanoista Arbitrage Pricing Theory. APT-mallissa mallin riskitekijät voidaan määrittää tapauskohtaisesti. APT-malleissa riskitekijöitä voi esimerkiksi olla inflaatio ja bruttokansantuote. (Ross 1976.) Tässä työssä määritetään riskitekijöiksi maailman osakemarkkinoiden tuottoindeksin log-tuotto ja euron ja dollarin valuuttakurssin log-muutos. Syynä siihen, että miksi tässä työssä käytetään euron valuuttakurssia riskitekijänä, on se, että euron valuuttakurssi kuvaa hyvin euroalueen yleistä riskisyyttä. Tässä työssä käytetty kahden tekijän APT-malli on esitetty alla. Kyseinen malli estimoidaan tässä työssä pienimmän neliösumman menetelmällä.

$$RP_t = \alpha + \beta_M RP_{M,t} + \beta_{EURUSD} R_{EURUSD,t} + \varepsilon_t \tag{107}$$

$$\varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{108}$$

jossa $R_{EURUSD,t}$ = EUR/USD valuuttakurssin log-muutos ajankohtana t.

4 Tulokset

4.1 Tulokset koskien ilmiötä

4.1.1 Ilmiö-malleissa käytetty aineisto

Ilmiö-mallien tapauksessa aineistona käytetään Suomen osakemarkkinoiden hintaindeksiä. Niiden parametrit estimoidaan käyttämällä pelkästään in-sample ajanjaksoa eli 01/1987-12/2008. Päätelmät koskien sijoitussääntöjen optimaalisia takaisinkatsomisaikoja tehdään siis pelkästään kyseisen ajanjakson perusteella.

Alla esitetyssä taulukossa on esitetty tämän työn ADF-testien p-arvot sekä mallien optimaaliset viivepituudet. ADF-testi on toteutettu vakiolla ja kuudella viiveellä ja mallien sopivuus on mitattu käyttäen AIC-tunnuslukua. Mallin sopivuustestin rajoitteeksi on määritetty, että malli voi sisältää enintään kaksikymmentä termiä.

Taulukko 1: Suomen osakemarkkinoiden hintaindeksi, stationaarisuus ja optimaalinen viivepituus (AIC)

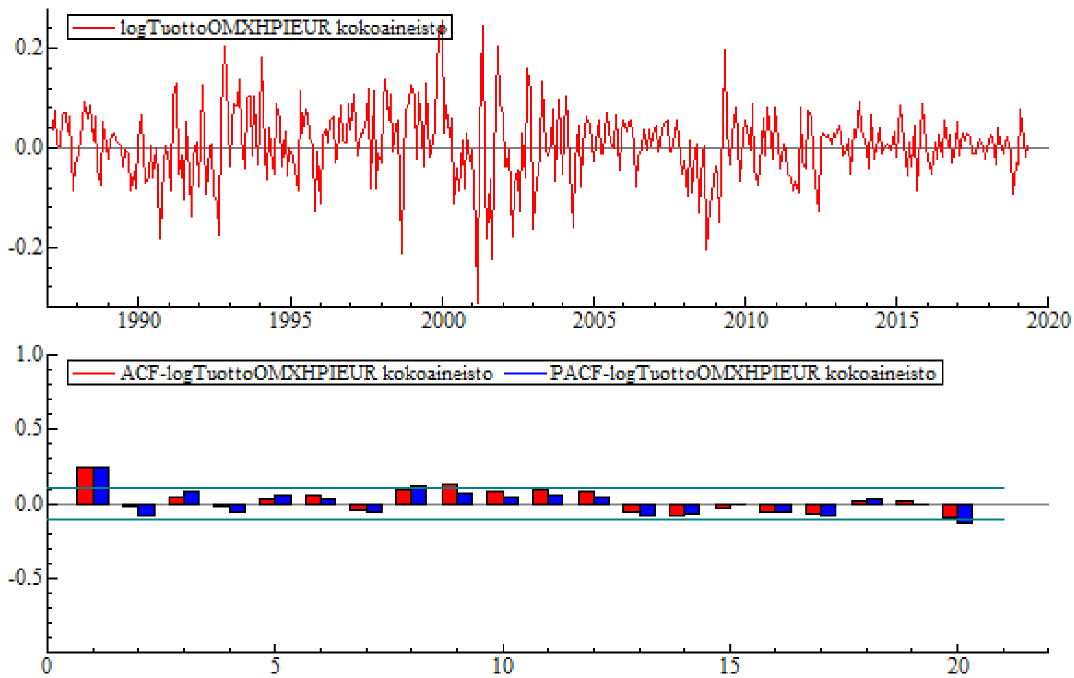
	01/1987-04/2019	01/1987-12/2008	01/2009-04/2019
ADF-testi vakiolla ja 6 viiveellä (p-arvo)	-7.109 (<0.01)**	-5.326 (<0.01)**	-5.036 (<0.01)**
Sopivin AR(p)	9	1	2
Sopivin ARMA(p, q)	8, 4	8, 4	2, 0

Yllä esitetyistä tuloksista voidaan päätellä, että Suomen osakemarkkinoiden hintaindeksin log-tuotot ovat stationaariset. Seuraavalla sivulla olevassa taulukossa on esitetty kyseinen hintaindeksin tunnusluvut sekä kokoaineiston kohdalla että jaettuna kahteen osaan eli 01/1987-12/2008 ja 01/2009-04/2019.

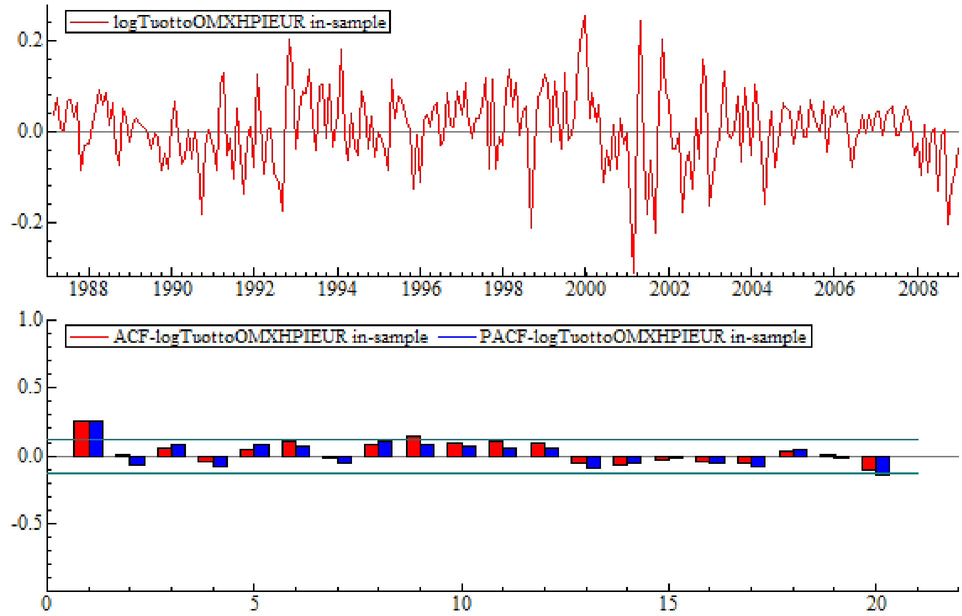
Taulukko 2: Suomen osakemarkkinoiden hintaindeksi, perustunnusluvut log-kuukausituotto

	01/1987-04/2019	01/1987-12/2008	01/2009-04/2019
Havaintojen lukumäärä	387	263	124
Keskiarvo	0.006	0.006	0.005
Keskihajonta	0.072	0.081	0.049
Vinous	-0.214	-0.243	0.081
Huipukkuus	4.855	4.232	4.602
95% Value-at-Risk (normaalijakauma)	-0.113	-0.127	-0.076
95% Expected Shorfall (normaalijakauma)	-0.135	-0.151	-0.093
95% Value-at-Risk (otosjakauma)	-0.112	-0.131	-0.083
95% Expected Shorfall (otosjakauma)	-0.167	-0.186	-0.104

Sitten esitetään kyseisen indeksin aikasarja- ja autokorrelaatiokuviot. Ensimmäinen kuvio kuvaa kyseistä indeksiä kokoaineiston kohdalla ja toinen kuvio kuvaa sitä in-sample ajankohtana. Ensimmäinen kuvio on esitetty alla ja toinen seuraavalla sivulla.



Kuvio 1: Suomen osakemarkkinoiden hintaindeksin aikasarja- ja autokorrelaatiokuvio, 01/1987-04/2019.



Kuvio 2: Suomen osakemarkkinoiden hintaindeksin aikasarja- ja autokorrelaatiokuvio, 01/1987-12/2008.

4.1.2 ARMA-mallin tulokset

In-sample ajanjaksoa koskevien tuloksien perusteella ARMA-mallin viiveiksi valikoituu 8 ja 4. Käytetty malli on siis alla esitettyä muotoa.

$$R_t = \alpha_0 + \sum_{j=1}^{8} \varphi_j R_{t-j} + \sum_{j=1}^{4} \theta_j \varepsilon_{t-j} + \varepsilon_t \tag{63b}$$

Esitetään ensin kyseisen mallin diagnostisten testien tulokset. Ne löytyvät alla olevasta taulukosta.

Taulukko 3: Diagnostiset testit, ARMA(8, 4), Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

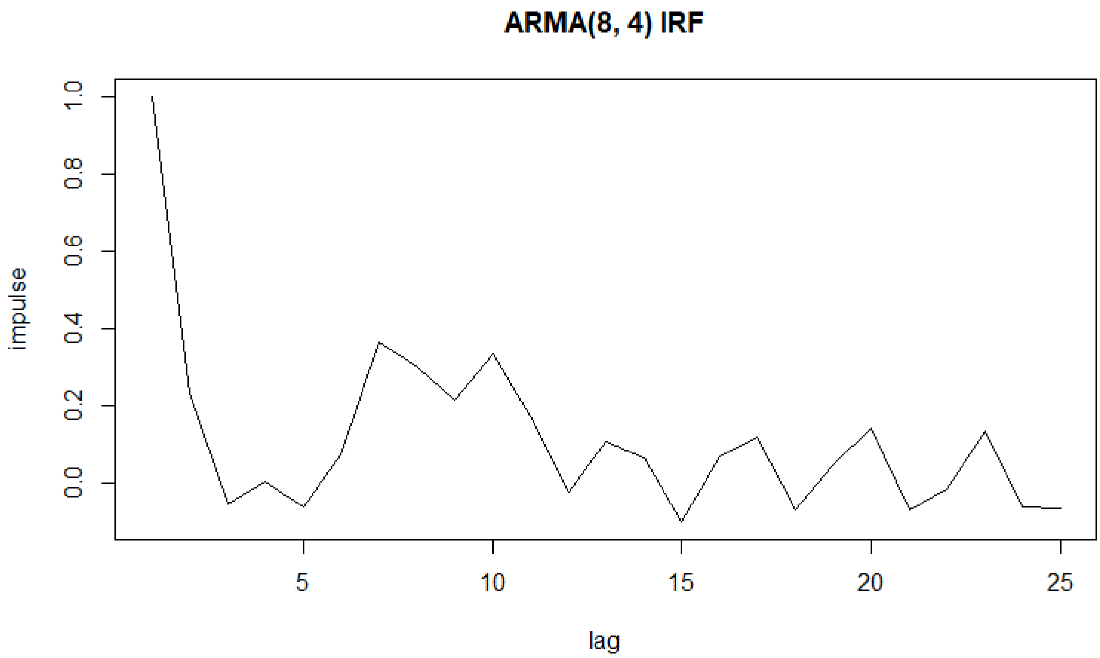
	Testisuure (p-arvo)
Ljung Box Q(20) testi	4.8085 (0.7778)
Jarque-Bera testi	8.3880 (0.0151)*
ARCH(1) testi	2.1379 (0.1450)

Yllä esitetystä taulukosta havaitaan, että virhetermi ei ole normaalistijakautunut. Täten vedotaan suurotosominaisuuteen sitä koskien. Esitetään seuraavaksi kyseisen mallin estimoidut parametrit. Ne on esitetty seuraavalla sivulla olevassa taulukossa.

Taulukko 4: Tulokset log-kuukausituotto, Suomen osakemarkkinoiden hintaindeksi, ARMA(8, 4), 01/1987-12/2008

	Parametrin arvo (p-arvo)
α_0	0.00604616 (0.442)
ϕ_1	1.33299 (0.000)**
ϕ_2	-0.998751 (0.000)**
ϕ_3	1.30875 (0.000)**
ϕ_4	-1.33795 (0.000)**
ϕ_5	0.54623 (0.000)**
ϕ_6	-0.233368 (0.074)
ϕ_7	0.113872 (0.305)
ϕ_8	0.0156764 (0.816)
θ_1	-1.09685 (0.000)**
θ_2	0.629560 (0.000)**
θ_3	-0.996585 (0.000)**
θ_4	0.909337 (0.000)**

Yllä olevasta taulukosta havaitaan, että kaikki muut parametrit ovat tilastollisesti merkitseviä jopa 1 prosentin riskitasolla, paitsi vakio ja viiveet 6, 7 ja 8. Seuraavaksi esitetään edellä esitettyjä tilastollisesti merkitseviä parametrien arvoja käyttäen määritetty impulssivastefunktio. Se on esitetty alla.



Kuvio 3: Impulssivastefunktio, ARMA(8, 4), Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008.

Edellisellä sivulla esitetystä kuviosta käy ilmi, että shokin vaikutus lähestyy lähellä nollaa parissa kuukaudessa. Shokin vaikutus näyttää nousevan taas suuremmaksi viisi kuukautta myöhemmin, mutta lähestyy sen jälkeen takaisin nollaa kohti. Kyseisen mallin perusteella selkein shokin vaikutus siis on yhden kuukauden päästä shokin syntymisestä. Lisäksi todennäköisyys sille, että syntyy uusi shokki, on suurempi viiden kuukauden päästä aiemmasta shokista kuin yhden kuukauden päästä. Täten kyseisien tuloksien perusteella yhden kuukauden takaisinkatsomisaikaa voidaan pitää parhaimpana vaihtoehtona.

4.1.3 AR-GARCH-in-Mean-mallin tulokset

Optimaaliseksi viivepituudeksi koskien tavallista AR-mallia saadaan yksi viive. Täten estimoidaan kyseinen malli niin, että siihen on lisätty GARCH-in-Mean komponentti. GARCH-komponentin käytetyiksi viivepituuksiksi taas saadaan kaksi ja yksi. Saatu malli on alla esitettyä muotoa.

$$R_t = \alpha_0 + \varphi_1 R_{t-1} + \alpha_1 h_t + u_t \quad (70b)$$

$$u_t = \epsilon_t \sqrt{h_t} \quad (71)$$

$$\epsilon_t \sim NID(0,1) \quad (72)$$

$$h_t = \beta_0 + \omega_1 h_{t-1} + \omega_2 h_{t-2} + \lambda_1 u_{t-1}^2, \quad (73b)$$

Kyseisen mallin diagnostisten testien tulokset on esitetty alla.

Taulukko 5: Diagnostiset testit, AR(1)-GARCH(2, 1)-in-Mean,
Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

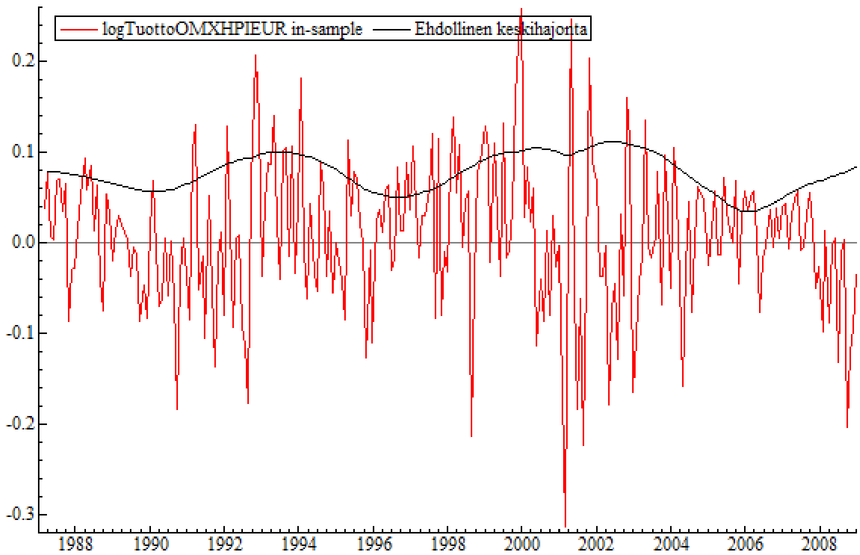
	Testisuure (p-arvo)
Ljung Box Q(20)-testi	18.621 (0.4814)
Jarque-Bera testi	6.0019 (0.0497)*
ARCH(1) testi	0.067524 (0.7952)

Diagnostisista testeistä havaitaan, että virhetermi ei ole normaalistijakautunut. Täten suurotosominaisuuteen vedotaan myös tämän mallin kohdalla. Esitetään sitten kyseisen mallin estimoidut parametrit p-arvoineen. Saadut tulokset on esitetty seuraavalla sivulla.

Taulukko 6: Tulokset log-kuukausituotto, AR(1)-GARCH(2, 1)-in-Mean, Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

	Parametrin arvo (p-arvo)
α_0	0.0119358 (0.104)
ϕ_1	0.222991 (0.000)**
α_1	-1.02333 (0.454)
β_0	0.0000284464 (0.000)**
ω_1	1.9654 (0.000)**
ω_2	-0.974074 (0.000)**
λ_1	0.00437987 (0.044)**

Yllä esitetyistä tuloksista käy ilmi, että ehdollisella varianssialla ei ole vaikutusta Suomen osakemarkkinoilla saatujen tuottojen keskiarvoihin ainakaan indeksitasolla. Toisin sanoen sillä, että onko korkean tai alhaisen riskin ajat, ei ole vaikutusta tuottojen odotusarvoihin. Lisäksi havaitaan, että tuottojen ensimmäisen viiveen parametri on selkeästi tilastollisesti merkitsevä ja positiivinen. Tulokset siis indikoivat, että tuottojen autokovarianssirakenteen takana ovat jotkin behavioraaliset tekijät. Lisäksi tuloksista havaitaan, että aikasarja-momentum -sijoitussäännössä yhden viiveen takaisinkatsomisaika on hyvä vaihtoehto. Esitetään seuraavaksi kyseisen mallin ehdollisen keskihajonnan ja log-tuottojen aikasarja-kuvio. Se on esitetty alla.



Kuvio 4: Ehdollinen keskihajonta ja log-kuukausituotto, AR(1)-GARCH(2, 1)-in-Mean, Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008.

Edellisellä sivulla esitetystä kuviosta havaitaan, etteivät ehdollinen keskihajonta ja saadut log-tuotot seuraa toisiaan. Toisin sanoen voidaan pitää epätodennäköisenä sitä, että aikasarja-momentumin takana olisi ajassa muuttuva keskihajonta.

4.1.4 Probit-mallin tulokset

Probit-mallissa käytetään myös yhtä viivettä, koska yhden viiveen AR-malli oli sopivin AR-malli in-sample ajanjaksona. Kyseinen malli on siis alla esitettyä muotoa.

$$y_t = \begin{cases} 1, & jos\ R_t > 0 \\ 0, & jos\ R_t \leq 0 \end{cases} \tag{76}$$

$$P(y_t = 1|W_t) = \Phi(W_t) \tag{77}$$

$$W = \beta_0 + \varphi_1 R_{t-1} + \varepsilon_t \tag{78b}$$

Kyseisestä probit mallista saadaan alla esitetyt tulokset.

Taulukko 7: Tulokset log-kuukausituotto, probit-AR(1), Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

	Parametrin arvo (p-arvo)
β_0	0.10822 (0.170)
ϕ_1	2.9787 (0.003)**

Yllä esitetystä taulukosta havaitaan, että ensimmäisellä viiveellä on tilastollisesti merkitsevästi positiivinen vaikutus todennäköisyyteen saada voittoa seuraavalla periodilla. Tämä edelleen indikoi sitä, että yhden viiveen takaisinkatsomisaika koskien aikasarja-momentum -sijoitussääntöä on paras vaihtoehto.

4.1.5 Regiimin muutosmallin tulokset

Tässä työssä käytetty regiimin muutosmalli on alla ja seuraavalla sivulla esitettyä muotoa.

$$R_t = \propto_0 + \propto_1 S_t + \varepsilon_t \tag{93}$$

$$S_t = \begin{cases} 1, kun\ kyseessä\ on\ laskumarrkina \\ 0, kun\ kyseessä\ on\ nousumarkkina \end{cases} \tag{81}$$

$$P(S_t = 1|S_{t-1} = 1) = p \tag{82}$$

$$P(S_t = 0|S_{t-1} = 1) = 1 - p \tag{83}$$

$$P(S_t = 1|S_{t-1} = 0) = 1 - q \tag{84}$$

$$P(S_t = 0|S_{t-1} = 0) = q \tag{85}$$

Esitetään ensimmäiseksi kyseisen mallin diagnostiset testit. Ne löytyvät alla olevasta taulukosta.

Taulukko 8: Diagnostiset testit, Regiimin muutosmalli, Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

	Testisuure (p-arvo)
Ljung Box Q(20)-testi	32.222 (0.0410)*
Jarque-Bera testi	34.547 (0.0000)**
ARCH(1) testi	2.6832 (0.1026)

Yllä esitetystä taulukosta havaitaan, että tässä mallissa virhetermi sisältää autokorrelaatiota ja ei ole normaalistijakautunut. Parametrien testisuureisiin on siis suhtauduttava varauksellisesti. Esitetään sitten kyseisen mallin estimoidut parametrit. Mallista saadut tulokset löytyvät alla olevasta taulukosta.

Taulukko 9: Tulokset log-kuukausituotto, Regiimin muutosmalli, Suomen osakemarkkinoiden hintaindeksi, 01/1987-12/2008

	Parametrin arvo (p-arvo)
α_0	0.0323168 (0.000)**
α_1	-0.0598709 (0.000)**
p	0.7687
1-p	0.2313
q	0.90657
1-q	0.093426

Yllä esitetyistä tuloksista päätellään, että jos tämän kuukauden tuotoilla on positiivisen odotusarvo, niin seuraavan kuukauden tuotoilla on noin 90,7 prosentin todennäköisyydellä positiivinen odotusarvo. Määritetään tämän perusteella liukuvan keskiarvon sijoitussäännön takaisinkatsomisaika niin, että seuraavan kuukauden tuotoilla on yli 60 prosentin todennäköisyys olla edelleen odotusarvoltaan positiiviset. Indikaationa siitä, että onko nyt positiivisen vai negatiivisen odotusarvon kuukausi, käytetään siis aiempien ja nykyisen kuukauden indeksiarvojen liukuvaa keskiarvoa. Tätä indikaattori sitten verrataan tämän kuukauden indeksiarvoon. Tämä indikaattori ei ole täydellinen, koska kesken kyseisen laskenta-ajanjakson odotusarvo on saattanut jo kääntyä. Täten todennäköisyys sille, että seuraavan kuukauden tuoton odotusarvo on positiivinen, voi sen sijaan olla noin 23 prosenttia.

Optimaalinen takaisinkatsomisaika koskien liukuvaa keskiarvoa on määritetty alla esitetyllä yhtälöllä.

$$l = \ln(0.6)\,/\ln(q), \tag{109}$$

jossa l = liukuvan keskiarvon sijoitussäännön takaisinkatsomisaika. Yllä olevasta yhtälöstä saatu tulos pyöristetään aina alaspäin. Ideana on siis selvittää, että montako kuukautta yhteen menoon yli 60 prosentin todennäköisyydellä on positiivisen odotusarvon kuukausia. Tässä tapauksessa optimaalinen takaisinkatsomisaika on viisi kuukautta.

4.2 Tulokset koskien sijoitussääntöjen kannattavuutta

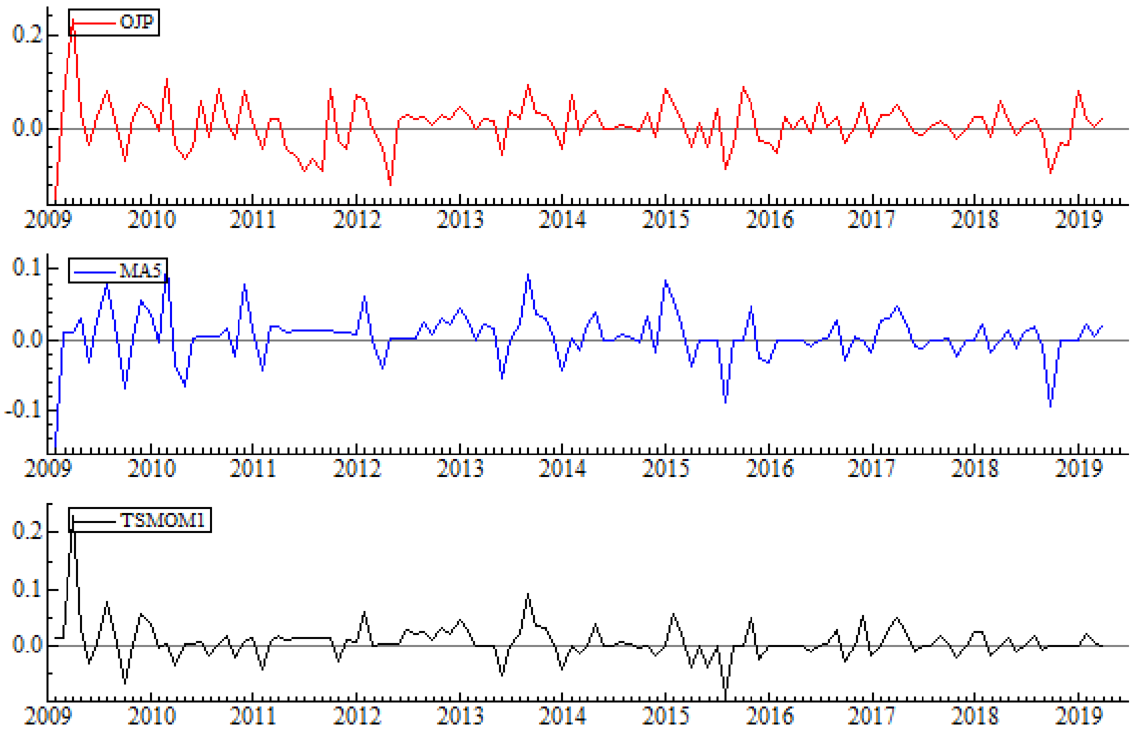
4.2.1 Kannattavuus-malleissa käytetty aineisto

Sijoitussääntöjen kannattavuus on mitattu käyttämällä Suomen osakemarkkinoiden tuottoindeksiä. Ajanjaksona tässä käytetään pelkästään out-of-sample ajanjaksoa eli 01/2009-04/2019. Alla esitetystä taulukosta löytyvät eri sijoitussääntöjen perustunnusluvut kyseisenä ajanjaksona. Aikasarja-momentumin takaisinkatsomisajaksi on määritetty yksi kuukausi ja liukuvan keskiarvon viisi kuukautta.

Taulukko 10: Suomen osakemarkkinoiden tuottoindeksi, perustunnusluvut log-kuukausituotto, 01/2009-04/2019

	Osta-ja-pidä	TSMOM1	MA5
ADF-testi vakiolla ja 6 viiveellä (p-arvo)	-5.109 (<0.01)**	-4.545 (<0.01)**	-5.216 (<0.01)**
Havaintojen lukumäärä	123	123	123
Keskiarvo	0.009	0.008	0.006
Keskihajonta	0.050	0.033	0.034
Vinous	0.269	2.524	-0.681
Huipukkuus	5.928	20.161	7.213
95% Value-at-Risk (normaalijakauma)	-0.073	-0.046	-0.050
95% Expected Shorfall (normaalijakauma)	-0.090	-0.058	-0.062
95% Value-at-Risk (otosjakauma)	-0.083	-0.039	-0.052
95% Expected Shorfall (otosjakauma)	-0.103	-0.055	-0.084

Yllä esityistä tuloksista havaitaan, että parhaiten tuottoa on tuottanut osta-ja-pidä -sijoitussääntö. Toiseksi parhaiten on tuottanut yhden viiveen aikasarja-momentum ja huonoiten viiden viiven liukuva keskiarvo. Keskihajonnan kohdalla havaitaan selkeää eroavaisuutta osta-ja-pidä -sijoitussäännön ja vaihtoehtoisten sijoitussääntöjen välillä. Molemmat vaihtoehtoiset sijoitussäännöt omaavat selkeästi alhaisemman keskihajonnan kuin osta-ja-pidä. Täten on havaittavissa indikaatiota siitä, että vaihtoehtoiset sijoitussäännöt omaavat pienemmän riskin kuin pelkkä osta-ja-pidä. Yhden viiveen aikasarja-momentum -sijoitussäännön tuottojakauma omaa lisäksi selkeästi suuremman vinouden kuin osta-ja-pidä. Tämä taas indikoi, että aikasarja-momentumin keskihajonta yliarvioi siihen liittyvän riskin. Toisaalta kyseisen sijoitussäännön kohdalla myös huipukkuus on selkeästi suurempi kuin osta-ja-pidä -sijoitussäännöllä. Korkean huipukkuuden tapauksessa keskihajonta aliarvioi riskiä, mutta kun tämän yhdistää positiiviseen vinouteen, niin kokonaisvaikutus voi silti olla riskiä yliarvioiva. Täten yhden viiveen aikasarja-momentumin kohdalla löytyy indikaatiota pienemmästä riskistä. Liukuvan keskiarvon kohdalla taas vinous on negatiivinen ja huipukkuus suurempi kuin osta-ja-pidä -sijoitussäännön kohdalla. Tämä indikoi sitä, että kyseisen sijoitussäännön kohdalla keskihajonta aliarvioi riskin suuruutta. Täten ei voida varmuudella todeta, että omaako liukuva keskiarvo todellisuudessa pienemmän riskin kuin osta-ja-pidä vai ei. Sekä normaalijakauma-oletusta käyttämällä että otosjakaumasta johdettuna molemmat vaihtoehtoiset sijoitussäännöt näyttävät omaavan pienemmän häntäriskin kuin osta-ja-pidä. Esitetään seuraavaksi sijoitussääntöjen tuottojen aikasarja-kuvio. Kyseinen kuvio on esitetty alla.



Kuvio 5: Log-kuukausituotto, Suomen osakemarkkinoiden tuottoindeksi, sijoitussäännöt, 01/2009-04/2019.

Esitetään sitten tässä luvussa käytettyjen riskitekijöiden perustunnusluvut. Ne on esitetty seuraavalla sivulla.

Taulukko 11: Riskimalleissa käytetyt riskitekijät, 01/2009-04/2019

	LogRiskipreemioMaailma	LogTuottoEURUSD
ADF-testi vakiolla ja 6 viiveellä (p-arvo)	-4.416 (<0.01)**	-4.092 (<0.01)**
Havaintojen lukumäärä	123	123
Keskiarvo	0.009	0.001
Keskihajonta	0.034	0.027

Yllä esitetystä taulukosta havaitaan, että molemmat tässä työssä käytetyt riskitekijät ovat stationaarisia.

4.2.2 Riippumattomien otosten t-testin tulokset

Riippumattomien otosten t-testillä nollahypoteesi ja vaihtoehtoinen hypoteesi on esitetty alla.

$H_0: E(R_s) \leq E(R_o)$

$H_1: E(R_s) > E(R_o)$

Kyseisen testin tulokset löytyvät alla olevasta taulukosta.

Taulukko 12: log-kuukausituotto riippumattomien otosten t-testi, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

	Havaittu t-arvo
MA5 vs. Osta-ja-pidä	-0.5534 (H_0 hyväksytään)
TSMOM1 vs. Osta-ja-pidä	-0.2563 (H_0 hyväksytään)

Yllä esitetyistä tuloksista havaitaan, ettei kumpikaan vaihtoehtoisista sijoitussäännöistä tuota odotusarvoisesti suurempaa tuottoa kuin osta-ja-pidä -sijoitussääntö.

4.2.3 GARCH-mallin tulokset

GARCH-mallilla ei saatu estimoitua hyväksyttäviä tuloksia liukuvan keskiarvon tapauksessa, joten sitä päätettiin käyttää vain osta-ja-pidä -sijoitussäännön ja aikasarja-momentum -sijoitussäännön tapauksessa. Käytetyt viivepituudet GARCH-mallissa vaihtelevat molempien sijoitussääntöjen kohdilla. Osta-ja-pidä -sijoitussäännön kohdalla viivepituudet ovat kolme ja kaksi ja aikasarja-momentum -sijoitussäännön kohdalla taas yksi ja yksi. Käytetty GARCH-malli on seuraavalla sivulla esitettyä muotoa.

$$R_t = \alpha_0 + u_t \tag{97}$$

$$u_t = \epsilon_t \sqrt{h_t} \tag{71}$$

$$\epsilon_t \sim NID(0,1) \tag{72}$$

$$h_t = \beta_0 + \sum_{i=1}^{v} \omega_i h_{t-i} + \sum_{j=1}^{r} \lambda_j u_{t-j}^2, \tag{98}$$

Määritetään ensimmäiseksi kyseisen mallin diagnostiset testit. Saadut tulokset on esitetty alla.

Taulukko 13: Diagnostiset testit GARCH(v, r),
01/2009-04/2019

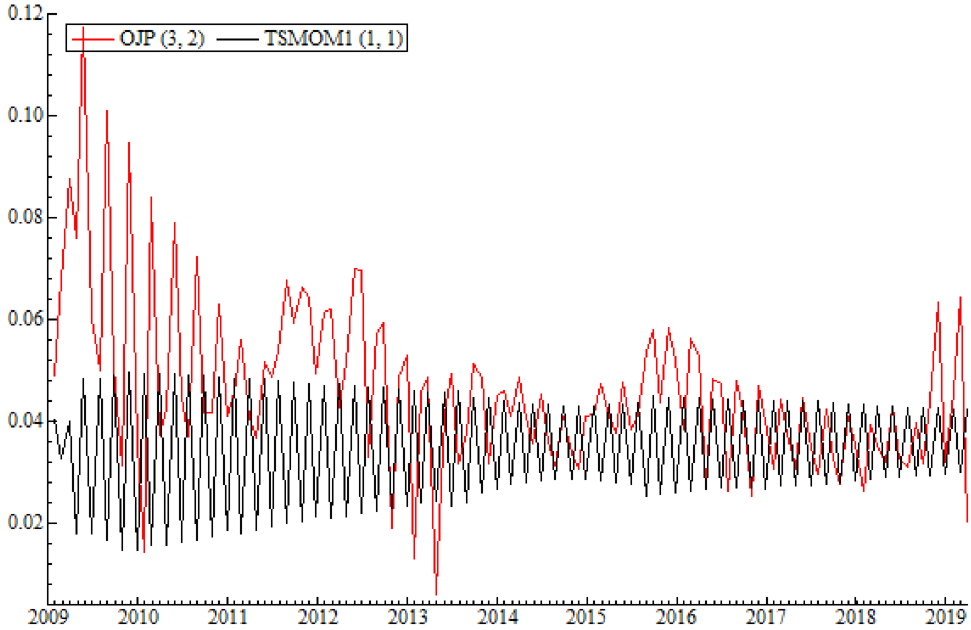
	Osta-ja-pidä (3, 2)	TSMOM1 (1, 1)
Ljung Box Q(10)-testi	6.8480 (0.7397)	18.171 (0.0521)
Jarque-Bera testi	6.6442 (0.0361)*	731.84 (0.0000)**
ARCH(1)-testi	0.14161 (0.7074)	0.016566 (0.8978)

Kyseisistä diagnostisista testeistä havaitaan, että sekä aikasarja-momentum -sijoitussäännön että osta-ja-pidä -sijoitussäännön kohdilla virhetermi ei ole normaalistijakautunut. Täten vedotaan suurotosominaisuuteen. Määritetään sitten mallin parametrit. Saadut tulokset on esitetty alla.

Taulukko 14: Tulokset OMXHTREUR log-kuukausituotto GARCH(v, r), 01/2009-04/2019

	Osta-ja-pidä (3, 2)	TSMOM1 (1, 1)
α_0	0.0109812 (0.000)**	0.00918143 (0.000)**
β_0	0.000533673 (0.024)**	0.00267625 (0.000)**
ω_1	-0.1373 (0.001)**	-0.988536 (0.000)**
ω_2	-0.161042 (0.000)**	-
ω_3	0.732709 (0.000)**	-
λ_1	0.0940418 (0.005)**	-0.0152588 (0.000)**
λ_2	0.241767 (0.000)**	-
λ_3	-	-

Seuraavalla sivulla olevassa kuviossa on esitetty molempien sijoitussääntöjen ehdolliset keskihajonnat aikasarja-kuviona.



Kuvio 6: Ehdollinen keskihajonta, GARCH(v, r), log-kuukausituotto, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019.

Määritetään sitten pitkän aikavälin varianssi ja keskihajonta käyttäen alla esitettyjä kaavoja.

$$\sigma^2 = \frac{\beta_0}{1-\sum_{i=1}^{v}\omega_i-\sum_{j=1}^{r}\lambda_j} \tag{74}$$

$$\sigma = \sqrt{\sigma^2} \tag{75}$$

Kyseisien sijoitussääntöjen kohdilla käyttäen niiden tilastollisesti merkitseviä parametreja saadaan alla olevassa taulukossa esitetyt tulokset.

Taulukko 15: Tulokset OMXHTREUR log-kuukausituotto GARCH(v, r), pitkän aikavälin keskihajonta, 01/2009-04/2019

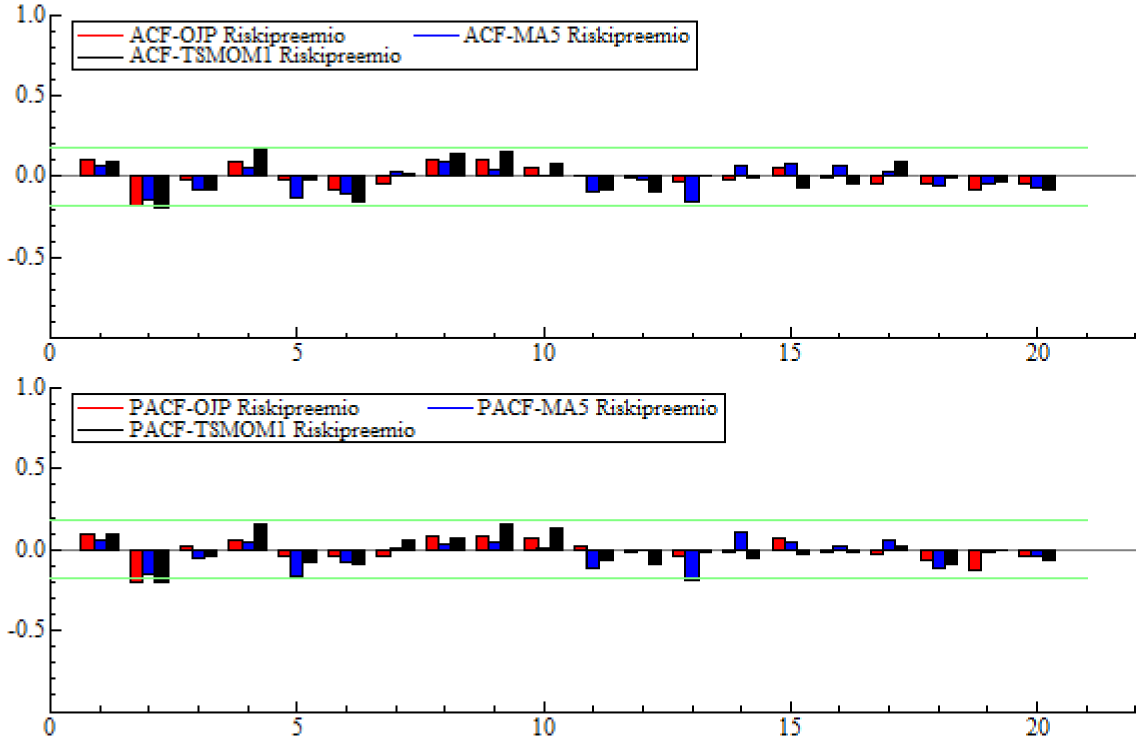
	Osta-ja-pidä (3, 2)	TSMOM1 (1, 1)
β_0	0.000533673	0.00267625
$\Sigma\, \omega_i + \Sigma\, \lambda_j$	0.7701758	-1.0037948
σ^2	0.002322092	0.001335591
σ	0.048188093	0.036545736

Yllä esitystä taulukosta havaitaan, että osta-ja-pidä -sijoitussäännön log-tuotot sisältävät suuremman pitkän aikavälin keskihajonnan kuin aikasarja-momentum -sijoitussääntö.

Saaduista tuloksista voidaan siis päätellä, että yhden viiveen aikasarja-momentum -sijoitussääntö sisältää alhaisemman riskin kuin osta-ja-pidä -sijoitussääntö.

4.2.4 Sharpen luvun tulokset

Koska Sharpen luku vääristyy, jos riskipreemioihin sisältyy autokorrelaatiota, niin tarkastetaan ensin sijoitussääntöjen autokorrelaatiokuvio. Kyseinen kuvio on esitetty alla.



Kuvio 7: Autokorrelaatio ja osittaisautokorrelaatio, log-kuukausiriskipreemio, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

Yllä esitetyssä kuviossa autokorrelaation tilastollisen merkitsevyyden raja 5 prosentin riskitasolla on kuvattu vaalean vihreinä viivoina. Kyseisestä kuviosta havaitaan, että ensimmäinen viiveen autokorrelaatio ei ole kyseisellä ajanjaksolla milläkään työn sijoitussäännöllä tilastollisesti merkitsevä. Toisen viiveen autokorrelaatio taas on sekä osta-ja-pidä -sijoitussäännöllä että yhden viiveen aikasarja-momentum -sijoitussäännöllä tilastollisesti merkitsevä ja negatiivinen. Voidaan siis päätellä, että näiden kahden sijoitussäännön kohdalla Sharpen luku aliarvioi niiden tuotto-riski -suhteita. Yhden viiveen aikasarja-momentumin kohdalla toisen viiveen autokorrelaatio näyttää olevan vieläkin enemmän negatiivinen kuin osta-ja-pidä -sijoitussäännön kohdalla. Täten voidaan päätellä, että yhden viiveen aikasarja-momentumin tapauksessa Sharpen luku aliarvioi enemmän tuotto-riski -suhdetta kuin osta-ja-pidä -sijoitussäännön tapauksessa.

Sharpen luku on esitelty luvussa 3.6. Alla esitetystä taulukossa on esitelty sekä Sharpen luku (kaava 99) että sen muunnelma, joka ottaa huomioon tuottojakauman poikkeaman normaalijakaumasta (kaava 101). Koska tilastollisesti merkitsevää autokorrelaatiota ei havaita ensimmäisellä viiveellä yhdelläkään tämän työn sijoitussäännöllä, niin mahdollisen ensimmäisen viiveen autokorrelaation huomioon ottavaa Sharpen lukua (kaava 100) ei määritetä tässä työssä ollenkaan.

Taulukko 16: OMXHTREUR Sharpen luvut, 01/2009-04/2019

	SR (p-arvo)	SR_{DC} (p-arvo)
Osta-ja-pidä	0.1402 (0.123)	0.1407 (0.121)
TSMOM1	0.1771 (0.052)	0.1863 (0.041)*
MA5	0.1187 (0.191)	0.1168 (0.198)

Kaikista sijoitussäännöistä suurimman tavallisen Sharpen luvun saa yhden viiveen aikasarja-momentum. Tämä indikoi, että se tarjoaa näistä parhaimman tuotto-riski -suhteen. Viiden viiveen liukuva keskiarvo taas näyttää tuottavan näistä kaikista huonoimman tuotto-riski -suhteen. Toisaalta täytyy ottaa huomioon, että 5 prosentin riskitasolla mikään yllä esityistä Sharpen luvuista ei ole tilastollisesti merkitsevä. Täten niiden perusteella ei voi tehdä täysin luotettavia päätelmiä. Kun normaalijakaumasta poikkeava tuottojakauman muoto otetaan huomioon ero yhden viiven aikasarja-momentum -sijoitussäännön ja osta-ja-pidä -sijoitussäännön välillä kasvaa entisestään. Lisäksi 5 prosentin riskitasolla yhden viiveen aikasarja-momentum -sijoitussäännön Sharpen luku on tilastollisesti merkitsevä, kun otetaan huomioon havaittu tuottojakauman muoto. Täten kyseinen tunnusluku tarjoaa todisteita sen puolesta, että yhden viiveen aikasarja-momentum tarjoaa paremman tuotto-riski -suhteen kuin osta-ja-pidä. Viiden viiveen liukuva keskiarvo taas huononee entisestään, joka edelleen todistaa, että kyseinen sijoitussääntö tarjoaa huonomman tuotto-riski -suhteen kuin osta-ja-pidä.

4.2.5 CAPM-mallin tulokset

CAPM-malli on esitetty luvussa 3.6. Kyseinen kaava on alla esitettyä muotoa.

$$RP_t = \alpha + \beta_M RP_{M,t} + \varepsilon_t \tag{103}$$

Tarkastetaan ensin mallin diagnostiset testit. Ne on esitetty seuraavalla sivulla.

Taulukko 17: Diagnostiset testit CAPM, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

	Osta-ja-pidä	TSMOM1	MA5
Ljung Box Q(10) testi	4.8031 (0.9039)	12.084 (0.2794)	2.9592 (0.9824)
Jarque-Bera testi	0.25077 (0.8822)	505.29 (0.0000)**	8.3668 (0.0152)*
Whiten testi	4.2634 (0.0163)*	23.670 (0.0000)**	34.399 (0.0000)**

Yllä esitetyistä testisuureista saadaan selville, että kaikkien sijoitussääntöjen kohdilla tuottojen virhetermit sisältävät heteroskedastisuutta. Täten niiden kaikkien kohdilla käytetään Whiten keskivirheitä. Lisäksi havaitaan, että molempien vaihtoehtoisten sijoitussääntöjen kohdalla virhetermi ei noudata normaalijakaumaa. CAPM-malli perustuu normaalijakauma-oletukseen, joten kyseisiä sijoitussääntöjä koskeviin tuloksiin täytyy suhtautua varauksellisesti. CAPM-mallista saadut tulokset on esitetty alla.

Taulukko 18: Tulokset log-riskipreemio CAPM, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

	Osta-ja-pidä (HCSE)	TSMOM1 (HCSE)	MA5 (HCSE)
α	-0.00319459 (0.3132)	0.00204143 (0.4149)	-0.000784283 (0.8123)
β_M	1.12735 (0.0000)**	0.400163 (0.0113)*	0.532918 (0.0004)**
R^2	0.581801 (0.000)**	0.180213 (0.000)**	0.281074 (0.000)**
Korjattu R^2	0.578345	0.173438	0.275132

Yllä esitetyistä tuloksista voidaan päätellä, että molemmat vaihtoehtoiset sijoitussäännöt sisältävät selkeästi vähemmän systemaattista riskiä kuin osta-ja-pidä -sijoitussääntö. Kaikkia yllä esitettyjä sijoitussääntöjä koskien alfaa koskeva nollahypoteesi hyväksytään 5 prosentin riskitasolla. Toisin sanoen mikään työssä käsiteltävistä sijoitussäännöistä ei tuota ylituottoa tilastollisesti merkitsevästi CAPM-mallin mukaan.

4.2.6 APT-mallin tulokset

Kahden tekijän APT-malli on alla esitettyä muotoa.

$$RP_t = \alpha + \beta_M RP_{M,t} + \beta_{EURUSD} R_{EURUSD,t} + \varepsilon_t \tag{107}$$

Kyseisen mallin diagnostisten testien tulokset on esitetty seuraavalla sivulla.

Taulukko 19: Diagnostiset testit APT, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

	Osta-ja-pidä	TSMOM1	MA5
Ljung Box Q(10) testi	12.338 (0.2631)	12.678 (0.2422)	1.9190 (0.9969)
Jarque-Bera testi	5.5233 (0.0632)	500.76 (0.0000)**	6.4103 (0.0406)*
Whiten testi	2.9763 (0.0145)*	10.199 (0.0000)**	15.670 (0.0000)**

Yllä esitetyistä tuloksista saadaan selville, että kaikkien sijoitussääntöjen kohdalla jälleen havaitaan heteroskedastisuutta. Täten kaikkien sijoitussääntöjen kohdilla käytetään Whiten keskivirheitä. Molempien vaihtoehtoisen sijoitussäännön kohdalla virhetermi ei näytä noudattavan normaalijakaumaa. Koska myös APT-malli perustuu normaalijakauma-oletukseen, täytyy vaihtoehtoisia sijoitussääntöjä koskeviin tuloksiin suhtautua varauksellisesti. APT-mallista saadut tulokset löytyvät alla esitetystä taulukosta.

Taulukko 20: Tulokset log-riskipreemio, kahden muuttujan APT, Suomen osakemarkkinoiden tuottoindeksi, 01/2009-04/2019

	Osta-ja-pidä (HCSE)	TSMOM1 (HCSE)	MA5 (HCSE)
α	-0.0028562 (0.2971)	0.0020826 (0.4108)	-0.00071583 (0.8273)
β_M	1.15873 (0.0000)**	0.403977 (0.0112)*	0.539264 (0.0004)**
β_{EURUSD}	-0.571914 (0.0000)**	-0.0695229 (0.3737)	-0.115679 (0.2890)
R^2	0.677144 (0.000)**	0.183677 (0.000)**	0.289507 (0.000)**
Korjattu R^2	0.671763	0.170071	0.277665

APT-mallista saatavat tulokset näyttävät antavan suhteellisen saman kuvan kuin CAPM-mallista saatavat tulokset. APT-mallista saadaan lisäksi selville, että euron valuuttakurssilla on selkeä vaikutus osta-ja-pidä -sijoitussääntöä koskien, mutta ei taas tilastollisesti merkitsevää vaikutusta koskien vaihtoehtoisia sijoitussääntöjä. Nämä tulokset siis myös indikoivat, että systemaattisen riskin määrä on pienempi koskien vaihtoehtoisia sijoitussääntöjä.

4.3 Tutkimuskysymyksien vastaukset

Tämän työn tutkimuskysymyksiä oli yhteensä kolme. Ensimmäinen koski ilmiön havaittavuutta. Kaikkien ilmiö-mallien tuloksista voidaan päätellä, että aikasarja-momentum -ilmiö on havaittavissa Suomen osakemarkkinoilla.

Toinen niistä kosketti ilmiön lähdettä. AR-GARCH-in-Mean mallin tuloksista havaittiin indikaatiota siitä, että ilmiön lähde olisi behavioraalinen eikä rationaalinen.

Kolmas tutkimuskysymyksistä oli, että voiko yksittäinen sijoittaja hyödyntää kyseistä ilmiötä omassa sijoitustoiminnassaan. Riippumattomien otosten t-testistä käy ilmi, että kyseisen ilmiön hyödyntäminen ei tuota parempaa tuottoa odotusarvoisesti kuin osta-ja-pidä -sijoitussääntö. Toisaalta yhden viiveen aikasarja-momentum -sijoitussäännön noudattaminen näyttää GARCH-mallin perusteella sisältävän vähemmän kokonaisriskiä kuin osta-ja-pidä -sijoitussääntö. Sharpen luku, sekä tavallisessa muodossa että jakauman muoto huomioon ottavana, oli näistä sijoitussäännöistä suurin yhden viiveen aikasarja-momentum -sijoitussäännöllä. Tavallisessa muodossa kyseinen Sharpen luku ei ollut tilastollisesti merkitsevä, mutta jakauman muoto huomioon ottavana Sharpen luku oli taas tilastollisesti merkitsevä pelkästään yhden viiveen aikasarja-momentumin tapauksessa. Viiden viiveen liukuvan keskiarvon sijoitussääntö taas tuotti näistä sijoitussäännöistä huonoimman Sharpen luvun. Täten Sharpen luvusta saatujen tuloksien perusteella voidaan päätellä, että yhden viiveen aikasarja-momentum -sijoitussääntö tarjoaa paremman tuotto-riski -suhteen kuin osta-ja-pidä -sijoitussääntö. Kun taas viiden viiveen liukuva keskiarvo tarjoaa huonomman.

Sekä CAPM-mallista että kahden muuttujan APT-mallista taas käy ilmi, että ilmiötä hyödyntämällä voidaan pienentää systemaattisen riskin määrää suhteessa osta-ja-pidä -sijoitussääntöön. Alfa ei toisaalta ollut tilastollisesti merkitsevä yhdenkään sijoitussäännön kohdalla, joten käytettyjen riskimallien kohdilla ei voitu tehdä päätelmiä tuotto-riski -suhteiden eroavaisuuksista. Lisäksi molemmissa riskimalleissa kumpaakin vaihtoehtoista sijoitussääntöä koskien residuaali ei noudata normaalijakaumaa, joka on molempien käytettyjen riskimallien taustaoletus. Täten niistä saatuihin tuloksiin täytyy suhtautua varauksellisesti.

Pelkkiä sijoitussääntöjen tunnuslukuja tarkkailemalla taas havaitaan, että molempien vaihtoehtoisten sijoitussääntöjen tuotot ovat suuruudeltaan pienemmät kuin osta-ja-pidä -sijoitussäännön tuotot. Vaihtoehtoiset sijoitussäännöt taas molemmat omaavat pienemmän otoskeskihajonnan kuin osta-ja-pidä -sijoitussääntö, joten niitä voidaan pitää alhaisemman riskin sisältävinä tämän perusteella. Lisäksi, kun katsotaan eri sijoitussääntöjen tunnuslukuja, havaitaan, että yhden viiveen aikasarja-momentum -sijoitussäännön tuottojakauma on reilusti enemmän positiivisesti vino kuin osta-ja-pidä -sijoitussäännön tuottojakauma. Täten voidaan päätellä, että aikasarja-momentum -sijoitussäännössä keskihajonta liioittelee riskin suuruuttaa. Toisaalta yhden viiveen aikasarja-momentum -sijoitussäännöllä tuottojakauma on myös huipukkaampi kuin osta-ja-pidä -sijoitussäännöllä. Puolestaan liukuvan keskiarvon sijoitussäännön tuottojakauma on negatiivisesti vino. Keskihajonta siis aliarvioi siihen liittyvää riskiä ja täten ei voida tehdä suoraviivaisia päätelmiä siitä, että sisältääkö kyseinen sijoitussääntö vähemmän kokonaisriskiä kuin osta-ja-pidä vai ei. Toki tässä on hyvä ottaa huomioon kaupankäyntikustannusten vaikutus. Tässä työssä kaupankäyntikustannukset asetettiin vain 0,1 prosentin suuruisiksi. Jos siis käytettäisiin erikokoisia kaupankäyntikustannuksia, tulokset voisivat muuttua radikaalisti.

5 Johtopäätökset

Tämä työ lähti liikkeelle finanssimarkkinoiden perusteorioista. Tehokkaiden markkinoiden hypoteesin mukaan kaikkien sijoituskohteiden ja sijoitussääntöjen tuotto-riski -suhteiden kuuluisi olla samat ajankohdasta riippumatta. Tätä näkemystä on sittemmin kritisoitu niin sanotun behavioraalisen rahoituksen muodossa, jonka mukaan markkinatoimijoiden toimintaa ohjaavat myös vahvasti psykologiset vaikuttimet ja näillä on vaikutusta jopa markkinatasapainossa.

Tässä työssä tutkittiin, että onko aikasarja-momentumin kaltaista ilmiötä havaittavissa Suomen osakemarkkinoilla. Kyseinen ilmiö on niin sanottu anomalia eli poikkeama tehokkaista markkinoista. Tämän tutkimiseen käytettiin kuukausikohtaista Suomen osakemarkkinoiden hintaindeksiä ja tuottoindeksiä sekä muita lisäaineistoja. Kyseisen aineiston analysointi tapahtui käyttämällä erilaisia tilastotieteellisiä malleja. Käytetyt mallit voitiin jakaa kahteen kategoriaan eli ilmiötä tutkiviin ja sijoitussääntöjen kannattavuutta tutkiviin. Aineisto jaettiin myös kahteen osaan eli in-sample ajankohta ja out-of-sample ajankohta. In-sample ajankohtaa eli 01/1987-12/2008 käytettiin ilmiötä tutkivissa malleissa, koska näiden avulla määritettiin eri sijoitussääntöjen takaisinkatsomisajat. Out-of-sample ajankohtaa eli 01/2009-04/2019 käytettiin taas sijoitussääntöjen kannattavuuden arviointiin.

Ilmiötä tutkivia malleja käyttämällä havaittiin, että aikasarja-momentum -ilmiö koskettaa Suomen osakemarkkinoita. Lisäksi löydettiin indikaatiota siitä, että kyseisen ilmiön takana on enemmänkin behavioraaliset tekijät kuin rationaaliset tekijät.

Kyseisen ilmiön hyödyntämiseksi voidaan määrittää erilaisia sijoitussääntöjä. Tässä työssä niitä olivat yhden viiveen aikasarja-momentum ja viiden viiveen liukuva keskiarvo, joiden viiveet siis perustuivat ilmiötä tutkivien mallien tuloksiin. Näiden sijoitussääntöjen tuottoja verrattiin osta-ja-pidä -sijoitussäännön tuottoihin. Kannattavuutta tutkivilla malleilla havaittiin, että vaihtoehtoiset sijoitussäännöt eivät kumpikaan tuota odotusarvoisesti suurempia tuottoja kuin osta-ja-pidä -sijoitussääntö. Toisaalta kyseiset sijoitussäännöt näyttivät omaavan pienemmän riskin kuin osta-ja-pidä -sijoitussääntö. Sharpen luku oli työssä käsiteltävistä sijoitussäännöistä suurin yhden viiveen aikasarja-momentum -sijoitussäännöllä. Lisäksi, jos otetaan huomioon käsiteltävien sijoitussääntöjen tuottojakaumien vinoudet ja huipukkuudet, Sharpen luku on vieläkin selkeämmin suurempi yhden viiveen aikasarja-momentum -sijoitussäännöllä kuin tavallisella osta-ja-pidä -sijoitussäännöllä. Viiden viiveen liukuvalla keskiarvon sijoitussäännöllä oli taas selkeästi huonompi Sharpen luku kuin osta-ja-pidä -

sijoitussäännöllä. Sharpen luvun mukaan siis yhden viiveen aikasarja-momentum -sijoitussäännöllä on parempi tuotto-riski -suhde kuin osta-ja-pidä -sijoitussäännöllä. Tämä tarkoittaa sitä, että jos molempien sijoitussääntöjen tuottojen keskihajonnat asetetaan samalla tasolle, niin yhden viiveen aikasarja-momentum saavuttaa suurempaa tuottoa kuin osta-ja-pidä -sijoitussääntö kyseisellä aikavälillä. Tämä taas indikoi jonkin suuruisesta markkinatehottomuudesta. Toki on huomioitava se, että aina kun testataan markkinatehottomuutta on olemassa yhteishypoteesiongelma. Ei voida siis täydellä varmuudella sanoa, että ottaako käsiteltävä tunnusluku tai malli kaikkia sijoituskohteen tai sijoitussäännön tuotto-riski -suhteeseen vaikuttavia tekijöitä huomioon.

Puolestaan tässä työssä käytettyjen riskimallien mukaan mikään työn sijoitussääntö ei tuottanut ylituottoa tilastollisesti merkitsevästi. Täten riskimalleista saatuja tuloksia ei voida käyttää eri sijoitussääntöjen tuotto-riski -suhteiden vertailuun.

Kyseisessä työssä käytetyt kaupankäyntikustannukset olivat 0,1 prosenttia eli suhteellisin alhaiset. Täten, jos markkinatoimija joutuu maksamaan korkeampia kaupankäyntikustannuksia omista kaupoistaan, niin tulokset voisivat muuttua merkitsevästi. Kyseisessä työssä käytettiin lisäksi kuukausikohtaista aineistoa. On mahdollista, että saadut tulokset muuttuisivat, jos sen sijaan käytettäisiin viikko, päivä- tai jopa mikrosekunttikohtaista aineistoa. Täten mahdollinen jatkotutkimuskohde voisi olla tehdä samat testit, mutta kuukausikohtaisen aineiston sijaan käytettäisiin esimerkiksi päiväkohtaista aineistoa. Tällöin saataisiin tarkempi kuva kyseisen ilmiön käyttäytymisestä ja täten voitaisiin paremmin arvioida sen implikaatiota markkinatehokkuuden kannalta.

Käytettynä aineistona voitaisiin lisäksi käyttää yleisindeksin sijaan esimerkiksi markkina-arvojen mukaan jaettuja indeksejä. Tällöin saataisiin selville se, että käyttäytyykö kyseinen ilmiö eri tavalla Suomen osakemarkkinoilla esimerkiksi pienen markkina-arvon yrityksissä kuin suuren markkina-arvon yrityksissä. Indeksien sijaan voitaisiin myös käyttää yksittäisien yrityksien osakkeita ja seurata kuinka kyseisiä sijoitussääntöjä käyttämällä portfolion arvo kehittyy. Tässä työssä ei myöskään pystytty käyttämään Fama-Frenchin kolmen tekijän mallia, koska tässä työssä käytettiin euromääräisiä indeksiarvoja ja Fama-Frenchin riskitekijäaineisto oli määritetty dollareissa. Täten, jos jatkotutkimuksessa käytettäisiin dollarimääräistä aineistoa, niin voitaisiin lisäksi kokeilla, että minkälaiset tulokset Fama-French mallista saadaan. Lisäksi jatkotutkimuksessa voitaisiin määrittää APT-malli, jossa käytetään eri riskitekijöitä kuin tässä työssä on käytetty.

Lähdeluettelo

Bodie, Z., Kane A. & Marcus, A. (2014). *Investments* (10th global edition). Maidenhead, United Kingdom: McGraw-Hill Education.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, vol. 31, no. 3, 307–327.

Bollerslev, T., Engle, R. F. & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, vol. 96, no. 1, 116–131.

Brealey, R., Myers S. & Allen, F. (2014). *Principles of Corporate Finance* (11th global edition). Maidenhead, United Kingdom: McGraw-Hill Education.

Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, vol. 50, no. 4, 987–1007.

Engle, R. F. & Granger, C. W. J. (1987). Co-Integration and Error Correction: Representation, Estimation and Testing. *Econometrica*, vol. 55, no. 2, 251–276.

Eurostat (2019). *Money market interest rates - monthly data (01/2009-04/2019).* https://ec.europa.eu/eurostat/web/products-datasets/-/irt_st_m. 22.05.2019.

Fama, E. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, vol. 25, no. 2, 383–417.

Fama, E. (1991). Efficient Capital Markets II. *The Journal of Finance*, vol. 46, no. 5, 1575–1617.

Fama, E. F. & French K. R. (1996). Multifactor explanations of asset pricing anomalies. *The Journal of Finance*, vol. 51, no. 1, 55–84.

Frazzini, A. (2006). The disposition effect and underreaction to news. *The Journal of Finance*, vol. 61 no. 4, 2017–2046.

Gordon, M. J. & Shapiro, E. (1956). Capital equipment analysis: The required rate of profit. *Management Science*, vol. 3, no. 1, 102–110.

He, X. & Li, K. (2015). Profitability of time series momentum. *Journal of Banking and Finance*, vol. 53, 140–157.

Hong, K. J. & Satchell, S. (2015). Time series momentum trading strategy and autocorrelation amplification. *Quantitative Finance*, vol. 15, no. 9, 1471–1487.

Jegadeesh, N. & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of Finance*, vol. 48, no. 1, 65–91.

Knüpfer, S. & Puttonen V. (2014). *Moderni rahoitus* (8. painos). Helsinki, Finland: Talentum Media Oy.

Lim, B. Y., Wang, J. G. & Yao, Y. (2018). Time-series momentum in nearly 100 years of stock returns. *Journal of Banking and Finance*, vol. 97, 283–296.

Lo, A. W. (2002). The statistics of Sharpe ratios. *Financial Analysts Journal*, vol. 58, no.4, 36–52.

Lütkepohl H., Krätzig M. & Phillips P. (2004). *Applied Time Series Econometrics*. New York, United States of America: Cambridge University Press.

Malkiel, B. G. (2003). The efficient market hypothesis and its critics. *The Journal of Economic Perspectives*, vol. 17, no. 1, 59–82.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance*, vol. 7, no. 1, 77–91.

Marshall, B. R., Nguyen, N. H. & Visaltanachoti, N. (2017). Time series momentum and moving average trading rules. *Quantitative Finance*, vol. 17, no. 3, 405–421.

Mills, T. C. & Markellos, R. N. (2008). *The Econometric Modelling of Financial Time Series* (3rd edition). New York, United States of America: Cambridge University Press.

Moskowitz, T. J., Ooi, Y. H. & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics*, vol. 104, no. 2, 228–250.

Newbold, P., Carlson, W. L. & Thorne, B. M. (2010). *Statistics for Business and Economics* (7th global edition). New Jersey, United States of America: Pearson Education, Inc.

Pézier, J. & White, A. (2008). The relative merits of alternative investments in passive portfolios. *The Journal of Alternative Investments*, vol. 10, no. 4, 37–49.

Gujarati, D. N. & Porter, D. C. (2009). *Basic Econometrics* (5th international edition). Singapore, Singapore: McGraw-Hill Education.

Ross, S. A. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory*, vol. 13, no. 3, 341–360.

Samuelson, P. A. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review*, vol. 6, no. 2, 41–49.

Samuelson, P. A. (1973). Proof that properly discounted present values of assets vibrate randomly. *Bell Journal of Economics*, vol. 4, no. 2, 369–374.

Sharpe, W. (1994). The Sharpe Ratio. *Journal of Portfolio Management*, vol. 21, no. 1, 49–58.

Shiller, R. J., Fischer, S. & Friedman, B. M. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity*, vol. 2, 457–510.

Shiller, R. J. (2014). Speculative asset prices. *The American Economic Review*, vol. 104, no. 6, 1486–1517.

Thomson Reuters (2019). *EURO TO US $ (WMR&DS) - EXCHANGE RATE (01/1987-04/2019)*. 20.05.2019.

Thomson Reuters (2019). *FTSE GLOBAL E :A - TOT RETURN IND (09/2003-04/2019)*. 20.05.2019.

Thomson Reuters (2019). *OMX Helsinki PI (01/1987-04/2019)*. 20.05.2019.

Thomson Reuters (2019). *OMX HELSINKI (OMXH) - TOT RETURN IND (01/1991-04/2019)*. 20.05.2019.

Verbeek, M. (2008). *A Guide To Modern Econometrics* (3rd edition). Chichester, England, United Kingdom: John Wiley & Sons, Ltd.