

Faktoreiden identifiointi ja rahastotuottojen analysointi Eurooppalaisille Large- Cap osakerahastoille

Mat-2.177 Operaatiotutkimuksen
projektityöseminaari

Kristian Nikinmaa (projektipäällikkö)
Markus Ehrnrooth
Matti Ollila
Richard Nordström
Ville Niskanen

4/23/2007

Sisällysluettelo

1.	Johdanto	4
2.	Kirjallisuuskatsaus	5
2.1	Kirjallisuudessa todettuja faktoreita	5
2.2	Käytetyt mallit varojen tuoton mallintamiseen	6
3.	Rahastojen ja faktoreiden graafinen tulkinta.....	8
3.1	Osakerahastojen tuotto verrattuna markkinaindeksiin	8
3.2	Rahastojen graafinen tulkinta.....	9
3.3	Faktoreiden graafinen tulkinta.....	10
3.4	Aikasarjojen muokkaaminen.....	10
4.	Rahastojen ja faktoreiden välisten korrelaatioiden tunnistaminen ja analysointi.....	13
4.1	Rahastojen ja faktoreiden väliset korrelaatiot	13
4.2	Rahastotuottojen sekä faktorituottojen ristikorrelaatiot	14
4.3	Rahastotuottojen korrelaatio faktoreiden kanssa aikaviiveellä	15
5.	Regressiomalli.....	17
5.1	Lineaarinen regressio	17
5.2	Dynaaminen lineaarinen regressio	18
5.3	Regressiodiagnostiikka.....	18
5.4	Regressiomallin valinta	19
5.5	Valikoiva askellus	19
5.6	Regressiomallin toteutus	20
6.	Tulokset.....	21
6.1	Mallien estimointi rahastojen arvojen perusteella.....	21
6.2	Mallien estimointi rahastojen tuottojen perusteella.....	22
6.2.1	Mallin estimointi päivittäisellä datalla	22
6.2.2	Mallin estimointi kuukausittaisella datalla.....	23
6.2.3	Mallien estimointi kuukausittaisella datalla ilman yritysten indeksejä.....	24
6.2.4	Mallien estimointi Log-tuottojen perusteella	26
6.2.5	Mallien estimointi käyttäen viivästettyjä selittäjiä	26
6.3	Parhaan mallin valinta ja herkkyysanalyysi	28
6.4	Mallin validointi	29
7.	Lineaarinen malli AR(1) virhetermillä.....	32
7.1	Päivittäinen data	32
7.2	Viikottainen data	33

8.	Pääkomponenttianalyysi.....	35
8.1	Tulokset.....	35
8.2	Regressio pääkomponenttien suhteen	37
9.	Vektoriautoregressiomalli	38
9.1	VAR-prosessi	38
9.2	VARX prosessi.....	39
9.3	VAR(p) prosessin estimointi	39
9.4	Rahastojen tuotot VAR-prosessina	41
9.5	Päivätuottojen ennustaminen VAR(1) - mallilla	41
9.6	VAR malli viikkotuotoille.....	46
10.	Yhteenveto ja johtopäätökset	51
11.	Lähteet.....	52
12.	Liitteet	54

1. Johdanto

Rahastosijoittaminen on viime vuosina yleistynyt huomattavasti ja siitä on tämän myötä tullut varteenotettava vaihtoehto suorille sijoituksille. Sijoittajan kannalta on mielenkiintoista tietää, mitkä markkinaindikaattorit vaikuttavat rahaston tuoton kehitykseen ja miten herkkiä rahastot ovat tietyille indikaattoreille. Lisäksi olisi tärkeätä tietää, mikäli rahasto on erityisen haavoittuvainen jollekin tunnistamattomalle faktorille.

Tutkimus yhtenä päätavoitteena on identifioida faktoreita, jotka vaikuttavat rahastojen tuottoihin. Tarkoituksena olisi löytää tilastollinen malli joka pystyy selittämään rahastojen tuotot tiettyjen faktorien avulla.

Tutkimuksen raportissa perehdytään ensiksi tutkimusaiheeseen kuuluvaan kirjallisuuteen, jota seuraa aineiston, eli tutkimukseen valittujen rahastojen ja faktorien, kuvaaminen ja analysointi. Aineiston yleisestä kuvaamisesta siirrytään aineiston riippuvuussuhteiden, tarkemmin sanottuna korrelaatioiden, tutkimiseen. Tämä antaa hyvät lähtöasetelmat lineaarisen mallin kehittämiseen. Tutkimuksessa pyritään luomaan sekä staattisia että dynaamisia lineaarisia regressiomalleja. Mallien diagnostisia ominaisuuksia ja selityskykyä pyrimme parantamaan muun muassa aggregoimalla aikasarjoja eri ajanjaksojen yli. Kokeilemme myös lineaarisia malleja, joissa virhetermi seuraa AR(1)-prosessia normaaleissa lineaarisissa malleissa esiintyvien residuaalien autokorrelaatioiden poistamiseksi.

Yhtenä tutkimuksen päällimmäisistä tavoitteista on identifioida rahastotuottoihin vaikuttavat piilevät faktorit. Näitä piileviä faktoreita pyritään löytää pääkomponenttianalyysin avulla. Tutkimusraportin viimeisessä osassa esitetään vektoriautoregressiomallin tuottamat tulokset. Vektoriautoregressiomallit osoittautuivat yllättäen joidenkin rahastojen kohdalla todella tehokkaiksi ja niiden avulla pystyttiin ennustamaan tarkahkosti kyseisten rahastojen tulevia tuottoja.

Tutkimuskysymyksiksi saatiin seuraavat:

1. Miten rahastoja on aiemmin mallinnettu portfolioteorian keinoin?
2. Miten rahastojen tuottoa ja riskiä sekä näiden kehitystä voidaan kuvata matemaattisin menetelmin?
3. Mitkä faktorit vaikuttavat rahaston tuoton kehitykseen?

2. Kirjallisuuskatsaus

Sijoitus- ja osakemarkkinat ovat perinteisesti erittäin laajasti tutkittu aihe, johon ei ole olemassa yhtä keskeistä teoriaa. Rahastojen tuoton ja riskin mallintaminen joutaa juurensa Jensenin tutkimuksista 1960-luvulla, missä Jensen vertaa amerikkalaisten osakerahastoja markkinaportfolioon yksinkertaisen CAPM-mallin avulla (Jensen, 1968). Tutkimusalueen ristiriitaisia näkemyksiä kuvaa hyvin nykyisten tutkimustenkin ristiriitaiset käsitykset siitä, tuottavatko osakerahastot keskimäärin markkinaindeksiä paremmin. Esimerkiksi on käynyt ilmi, että ”survivorship bias” huomioiden, amerikkalaiset osakerahastot, salkunhoitmaksut vähennettynä, pärjäävät huonommin kuin markkinaindeksi (Elton;Gruber;& Blake, 1996).

Tutkimukset osakerahastojen käyttäytymisestä ovat 2000-lukuun saakka keskittyneet USA:n markkinoihin. Vuosituhannen vaihteesta lähtien on ilmestynyt tutkimuksia eurooppalaisista rahastoista, joissa on todettu, että jo maiden sisällä sijoittavat rahastot päihittävät markkinat keskimäärin. Yksi annettu selitys tälle on, että eurooppalaiset osake- ja rahastomarkkinat ovat kehityksessä amerikkalaisista sen verran jäljessä, että niissä on edelleen joitain epätehokkuuksia (Otten & Bams, 2002).

Tyypillisiä tutkimuskohteita osakerahastoja tutkittaessa ovat olleet erilaiset tyylifaktorit, ajoitusfaktorit, tammikuu- ja maanantaieffektit, sekä markkinaindikaattorien ja piilevien faktoreiden vaikutus rahastojen tuottoon. Tehtävänantoon liittyen mielenkiintoisia tutkimuskohteita ovat lähinnä kirjallisuudessa identifioidut faktorit, sekä ne menetelmät joita on käytetty tuoton ja volatiliteetin ennustamiseen ja faktoreiden löytämiseen. Viimeksi mainitusta johtuen tulemme tarkastelemaan tässä myös menetelmiä, joita on käytetty puhtaasti osaketuottojen ennustamiseen (uskomme, että samanlaiset oletukset soveltuvat myös osakerahastoihin – faktoreiden vaikutukset saattavat kuitenkin diversifionnista johtuen olla heikompia).

Yleisenä havaintona voidaan todeta, että tuoton ennustaminen on osoittautunut erittäin haastavaksi tutkimusalueeksi. Tästä johtuen, suuri osa kirjallisuudesta esitetyistä menetelmistä soveltuu lähinnä volatiliteetin mallintamiseen. Tässä paperissa pääpaino tehtävänannon mukaisesti on tuoton ennustamisessa ja faktoreiden identifioimisessa ja joten volatiliteetin mallintaminen jätetään vähemmälle huomiolle.

2.1 Kirjallisuudessa todettuja faktoreita

Tyypillisesti osakkeiden ja osakerahastojen tuottoon vaikuttavat faktorit on jaettu viiteen ryhmään: makroekonomiset, markkinaindeksi, yrityspohjaiset (fundamental), tekniset, sekä tilastolliset faktorit (Chan;Karceski;& Lakonishok, 1998). Joskus tähän jaotteluun on vielä lisätty sektori/toimialakohtaiset faktorit (Straumann & Garidi, 2007) (tästä lähtien SG).

Makroekonomiset faktorit ovat esimerkiksi korkoja, raaka-ainehintoja, BKT:n kasvu, ja kuluttajahintaindeksit. Tyypillisesti makrofaktorit ovat suosittuja tutkimuksen kohteita, mutta niillä on, muutamaa poikkeusta lukuunottamatta, huonot selitysasteet. Poikkeuksia ovat esimerkiksi konkurssipremio (default premium), joka voidaan laskea high-yield bondin ja pitkän maturiteetin government bondin kuukausittaisen tuoton erotuksena, ja aika premio, joka voidaan laskea pitkän maturiteetin government bondin ja lyhyen maturiteetin joukkovelkakirjan tuoton erotuksena.

Yrityspohjaiset faktorit viittaavat osakkeen takaa löytyvän yrityksen tunnuslukuihin, esimerkiksi koko-faktoreihin (tyypillisesti pienten ja isojen yritysten osakkeiden tuoton erotus), arvo- tai

kasvuosakkeisiin (perinteisesti Fama-French HML, High book-to-market ration minus low book-to-market ratio), ja tuotto-odotukseen (market capitalization rate). Näiden on todettu selittävän osakkeiden käyttäytymistä erittäin hyvin. Kaksi ensimmäistä faktoria, markkinafaktorin kanssa, ovat vakiintuneet suosituksen Fama-French APT mallin selittäjiksi (Fama & French, 1993). Tähän tutkimukseen liittyen tämäntyyppisten faktoreiden soveltaminen osakerahastoihin on hieman kyseenalaista, koska faktorit liittyvät yritystä kuvaaviin tunnuslukuihin, joita rahastoille ei pystytä määrittämään.

Tekniset faktorit konstruoidaan käyttäen historiallista osake/rahastokohtaista tietoa. Esimerkiksi osaketuottojen kuvaamiseen on käytetty likviditeettimittoja ja osakkeiden historiallista tuottodataa. Vaikka teoria tehokkaista markkinoista väittää ettei historiallisen datan avulla voida voittaa markkinoita, on lukuisia tutkimuksia, jotka ovat osoittaneet, että erilaisilla momenttistrategioilla ("momentum strategies") olisi voitu saada markkinoita parempia tuottoja lyhyellä ja keskipitkällä aikavälillä (Chan; Jegadeesh; & Lakonishok, 1996). Momenttistrategioilla viitataan sijoitusstrategioihin joissa valitaan sijoituskohte aikaisemman performanssi perusteella, esimerkiksi evaluoimalla sijoituskohteiden tuotot kuuden viimeisen kuukauden ajalta kuukausittain ja sijoittamalla siihen, joka on pärjännyt parhaiten. Eakins ja Stansell (2004) antavat kattavan listan tämäntyyppisistä tutkimuksista ja esittävät, että tämäntyyppiset strategiat toimivat (Eakins & Stansell, 2004). Vastaavasti osakerahastoihin liittyvissä tutkimuksissa on todettu, että uudet rahastot pärjäävät keskimäärin vakiintuneita paremmin.

Tilastolliset faktorit perustuvat oletukseen, että jotkin piilevät faktorit generoivat havaittavissa olevat tuotot. Näiden faktoreiden estimoimiseen on perinteisesti käytetty pääkomponentti- ja/tai faktorianalyysiä. Sektori/toimialakohtaiset faktorit ovat usein erilaisia indeksejä. Näistä on hyötyä lähinnä eri toimialojen osakkeiden/rahastojen vertailussa.

2.2 Käytetyt mallit varojen tuoton mallintamiseen

Kirjallisuudessa ylivoimaisesti suosituin faktorimalli on APT, eli Rossin (1976) kehittämä Arbitrage Pricing Theory, ja sen muunnelmat. Periaatteessa tämä on CAPM:n (Capital Asset Pricing Model) yleistys. Molempien mallien lähtökohtana on se, että sijoituskohteen tuoton ja markkinafaktoreiden välillä on lineaarinen riippuvuus, jossa kuhunkin faktoriin liitetään riskipremio. Yleistetty APT-malli on siis seuraavaa muotoa:

$$r_i(t) = \alpha_i(t) + \sum_{j=1}^p \beta_{ij}(t) f_j(t) + \varepsilon_i(t), \quad i=1, \dots, K,$$

missä alfaat, betat ja faktorit voivat olla joko ajasta riippuvia tai stationaarisia. CAPM-malli on kyseisen APT:n erikoismuoto, jossa on alphan tilalla riskitön tuotto ja faktorina markkinaportfolion tuotto. Faktorit voivat olla joko ennalta määriteltyjä tai tilastollisin menetelmin estimoituja. SG esittää faktorimallin estimoinnin suoraan aikasarjaregressiota käyttäen, yleistetyn pienimmän neliosumman menetelmän avulla kahdessa askeleessa tai käyttäen aikariippuvaisia faktorilatauksia (Hamilton, 1994). Lisäksi he esittävät hyvän yhteenvedon perinteisistä APT:n faktoreiden

estimointimenetelmistä: faktoreiksi tekeytyvä portfolio, ”cross-sectional” regressio¹ ja pääkomponenttianalyysi, joista keskimäinen ei ole kovinkaan suosittu.

Faktoriksi tekeytyvillä portfolioilla viitataan menetelmään, jossa osakkeista luodaan portfolioita, joita käytetään faktoreina. Esimerkkinä voidaan mainita Fama-French malli SML (Small Minus Large), joka konstruoidaan rankkaamalla kaikki osakkeet tuottovaatimuksen (market capitalization) perusteella ja jakamalla ne kahteen ryhmään. SML faktori on täten portfolioiden tuottojen erotus $f^{SML}(t) = r^A(t) - r^B(t)$. Pääkomponenttianalyysi menetelmänä esitellään myöhemmin.

Alkuperäistä APT-mallia on modifioitu ja parannettu. Siitä on esimerkiksi esitetty dynaamisia versioita ja markkinoille ehdollistettuja malleja. Esimerkiksi Ferson ja Warther (Ferson & Warther, 1996) esittävät markkinoille ehdollistetun APT-mallin ja toteavat tämän avulla, että ehdollistettu malli pärjää keskimäärin paremmin kuin ehdollistamaton malli. Vastaavasti on yleisesti hyväksyttyä että tuotot ovat heteroskedastisia, eli niiden volatilitteetti muuttuu ajassa esim. (Campbell; Lettau; Malkiel; & Xu, 2001), jotka viittaavat lisäksi lukuisiin muihin tutkimuksiin. Nardari ja Scruggs (Nardari & Scruggs, 2003) (NS tästä lähtien) toteavat kuitenkin, että tutkimukset monifaktorimalleista, joihin yhdistetään ajassa muuttuva kovarianssimatriisi, ovat harvinaisia – lähinnä sen vuoksi, että niiden estimointi ja testaus on erittäin hankalaa. NS estimoi monifaktori-APT:n, jonka ylisuurien tuottojen kovarianssimatriisi noudatta stokastista volatilitteettiprosessia monella muuttujalla (MSV, multivariate stochastic volatility process). Tämän estimointi on mahdollista Bayesiläisten Markov Ketju Monte Carlo metodien avulla esim. (Tsay, 2002). Esimerkiksi Engle, Ng ja Rotschild (Ng; Rothschild; & Engle, 1990), sekä King, Sentana ja Wadwhani (King; Sentana; & Wadwhani, 1994) käyttävät MSV:n sijaan GARCH-malleja (Bollerslev, 1986). ARCH-mallin kehittänyt Engle (Engle, 2001) antaa hyvän yleiskatsauksen ja esimerkin GARCH-mallin estimoinnista.

Osakerahastoihin liittyvien tutkimuksien lisäksi tehtiin nopea katsaus menetelmiin, joita on käytetty ekonometriassa yleisesti. Näistä mielenkiintoisilta vaikuttivat erityisesti osaketuottojen ennustamisessa käytetyt vektoriautoregressiomallit ja neuroverkot. Ensinmainittu esitellään myöhemmin, kun sen toimivuutta testataan fundeilla. Neuroverkot vaikuttavat myös ne lupaavilta: Connor ja Madden (O'Connor & Madden, 2005) rakentavat neuroverkkomallin Dow Jones indeksin ennustamiselle ja havaitsivat että käyttämällä kyseistä mallia sijoituspäätöksen teossa, he olisivat saaneet 23% tuoton ajalla jolloin DJIA kasvoi vain 13%. Lisäksi he esittävät yhteenvedon kirjallisuudesta jossa neuroverkkoja sovelletaan talousindikaattoreihin.

Lopulta voisi mainita, että yleinen käsitys on, että matemaattisten mallien avulla on mahdollista päästä hyviin selitystasoihin. Hieman tähän liittyen esimerkiksi Kothari ja Warner (Kothari & Warner, 2001) osoittavat, että perinteiset mallit eivät kykene selittämään huomattavan suuria voittoja – erityisesti kun analysoidaan portfoliot poikkeavat huomattavasti markkinaportfoliosta. Samassa tutkimuksessa he osoittavat, että normaalit tapahtumapohjaiset analyysityökalut parantavat selitysvoimaa huomattavasti.

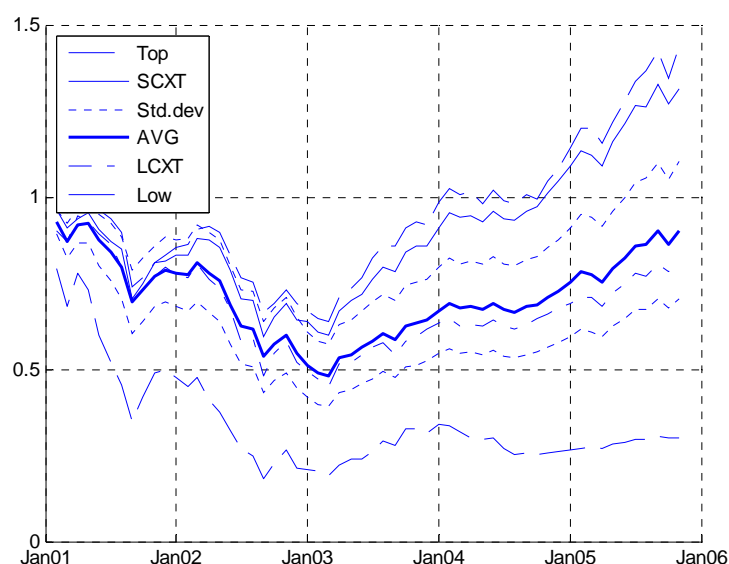
¹ Oletetaan edelleen mallin olevan muotoa $r_i(t) = \alpha_i(t) + \sum_{j=1}^p \beta_{ij}(t)f_j(t) + \varepsilon_i(t)$, mutta normaalista regressiosta poiketen lataukset $\beta_{ij}(t)$ ovat tunnettuja (yritystä kuvaavia lukuarvoja, teknisiä indikaattoreita yms.). Näiden avulla toteutuneet faktoriarvot $f_i(t)$ voidaan estimoida.

3. Rahastojen ja faktoreiden graafinen tulkinta

Tutkittavat rahastot ovat niin kutsuttuja European Large Cap – rahastoja eli ne sisältävät osakkeita suurista eurooppalaisista pörssiyhtiöistä. Suuri osa rahastoista on perustettu vasta vuoden 1998 jälkeen, mutta kaikki niistä olivat olemassa tarkastelujakson viimeisenä päivänä. Lista käytetyistä faktoreista ja fundeista löytyy liitteiden taulukoista 28 ja 29.

3.1 Osakerahastojen tuotto verrattuna markkinaindeksiin

Kirjallisuuden innoittamana lähdettiin perusdata-analyysin ohella tutkimaan, miten osakerahastot ovat pärjänneet markkinoihin verrattuna. Lähtökohtana analyysille oli, että 1.1.2001 sijoitettiin euron verran kuhunkin rahastoon, sekä Large-Cap (LCXT) ja Small-Cap (SCXT) indekseihin ja tutkittiin miten sijoitus tuotti aina 31.12.2005 asti. Kuva 1 havainnollistaa tilannetta.



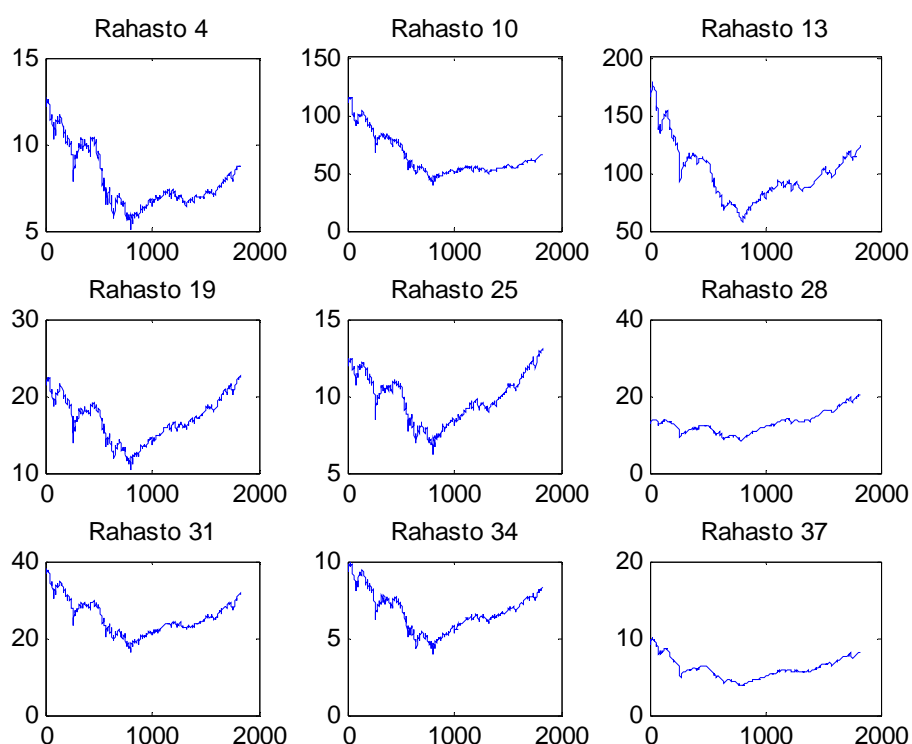
Kuva 1 Osakerahastojen kumulatiivisen tuoton kehitys ajan funktiona. Kuvassa keskiarvon ja standardipoikkeaman lisäksi parhaiten ja huonoiten pärjänneet rahastot, sekä Large-Cap ja Small-Cap indeksit.

Kuvasta nähdään, että keskimäärin päädyttiin tappiolle ajanjakson yli. Euroopan suuryrityksiä kuvaava LCXT käyttäytyi samalla tavalla. Nähdään kuitenkin, että rahastot pärjäsivät keskimäärin indeksiä paremmin, mikä viittaa siihen että hyvä salkunhoitaja voi saada aikaiseksi huomattavasti markkinoita parempia tuottoja – tätä tulosta tukee parhaiten pärjänneen, AXA Rosenberg PAN EUR S/C-A rahaston tulos. Toisaalta nähdään, että pienyrityksiin sidottu indeksibondi tuottaisi miltei saman verran kuin paras suurten yritysten rahasto. Huonoiten pärjännyt rahasto, UBS LUX EQTY-European GTH-B, onkin tuottanut huomattavasti huonommin kuin markkinat – kehitystä ei olla saatu nousuun edes markkinoiden noustessa.

3.2 Rahastojen graafinen tulkinta

Alkuperäinen data sisälsi 136:n rahaston aikasarjat aikaväliltä 31.12.1998–13.3.2007, mikä kokonaisuudessaan käsittää 2995 havaintoa. Saatua dataa muokattiin sopivaksi tilastollista analyysia varten. Näin ollen päädyttiin käyttämään 71 rahaston aikasarjadataa aikaväliltä 1.1.2001–31.12.2005. Näille rahastoille on olemassa päivittäinen, aukoton data.

Alla esitetyissä kuvissa kaikki rahastot alkavat indeksistä nolla, mikä tarkoittaa, että aikasarjojen kuvaajia ei ole yhteismitallistettu. Rahastoja on niin monta, ettei niiden kaikkien piirtäminen ole tarkoituksenmukaista. Tutustutaan kuitenkin rahastojen arvon kehitykseen piirtämällä niistä muutamia ja tarkastelemalla esiintykö aikasarjoissa huomattavia eroja.



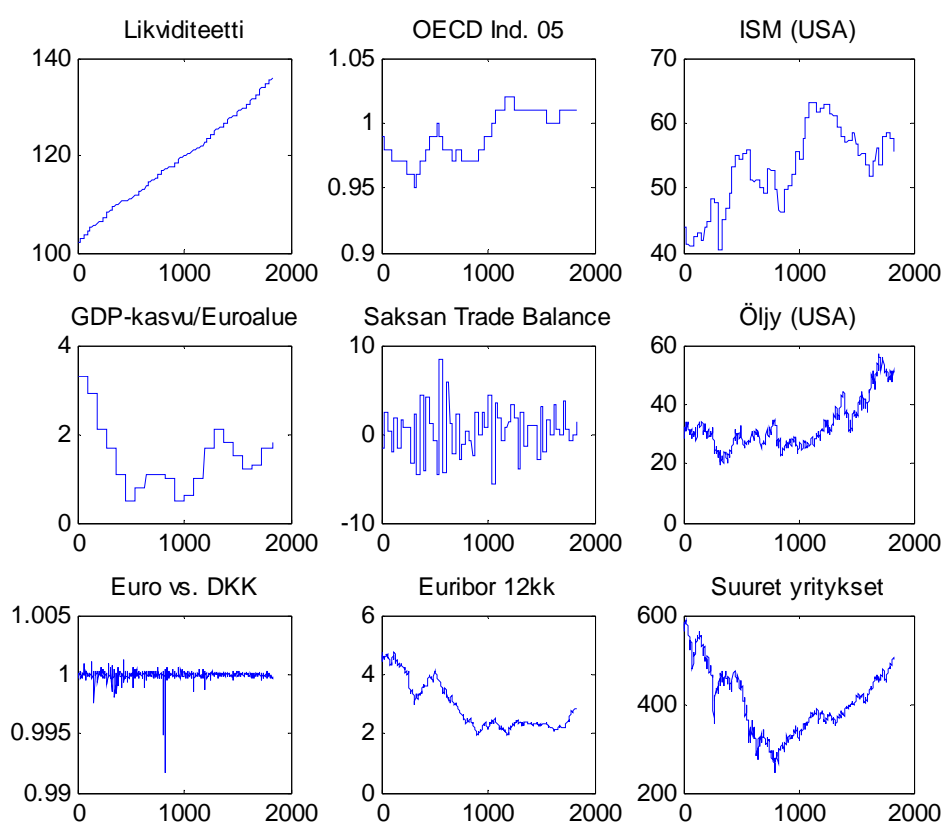
Kuva 2 Rahastojen arvojen kehittyminen 2001-2005.

Kuvassa 1 on esitetty 9 rahaston arvon kehittyminen vuosien 1.1.2001–31.12.2005 välillä. Nopea silmäys paljastaa, että rahastojen arvon kehityksessä on havaittavissa jokseenkin samanlaista trendiä. Hyvänä vertailukohtana voidaan pitää kuvassa 2 esitettävää suurten yritysten indeksiä. On selvästi havaittavissa, että rahastoissa on aikajakson ensimmäisinä vuosina tapahtunut lasku, kuten indeksissäkin, jonka jälkeen rahaston arvo on lähtenyt kasvamaan. Kuvasta nähdään myös, että rahastojen arvojen kehittymisen nopeudessa on kuitenkin huomattavia eroja.

3.3 Faktoreiden graafinen tulkinta

Toimeksiantaja on toimittanut tiedot yli 300 mahdollisesta faktorista. Näistä 123 faktorin aikasarjat osoittautuvat käyttökelpoisiksi etsittäessä selittäjiä rahastojen tuottojen kehitykselle. Pois jätettyjen faktoreiden data on joko ajalta ennen tarkasteltavien rahastojen aikasarjoja, niiden datan laatu on heikkoa tai asialogiset perustelut ovat epärelevantteja työmme kannalta. Faktoreiden aikasarjat ovat samalta ajanjaksolta kuin rahastojenkin. 37:n faktorin datat on saatu päivätasolla ja 86:n faktorin data on kuukausittaista.

Tutustutaan faktoreiden aikasarjoihin piirtämällä kuvat niiden arvoista ajan funktiona. Kuvassa 2 on esitetty 9 faktorin aikasarjat tarkastelujaksolta.

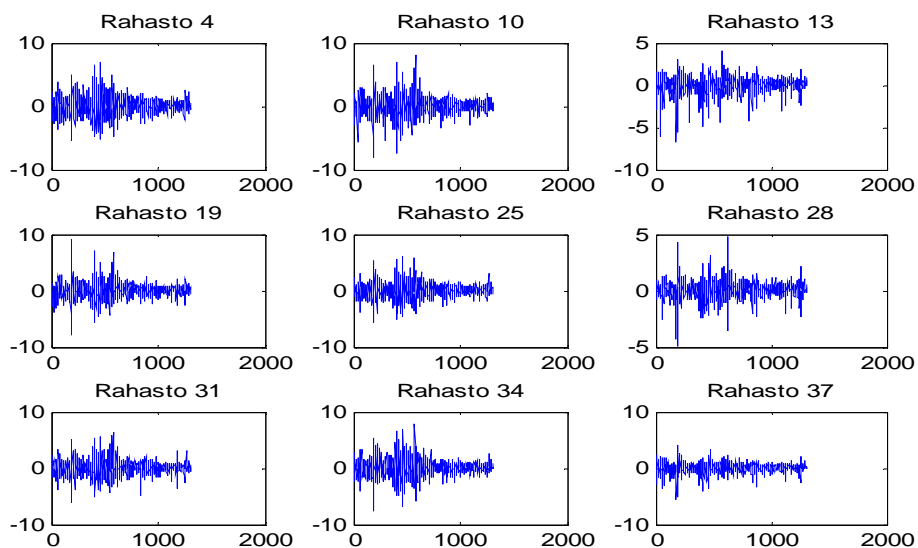


Kuva 3 Yhdeksän faktorin aikasarjat

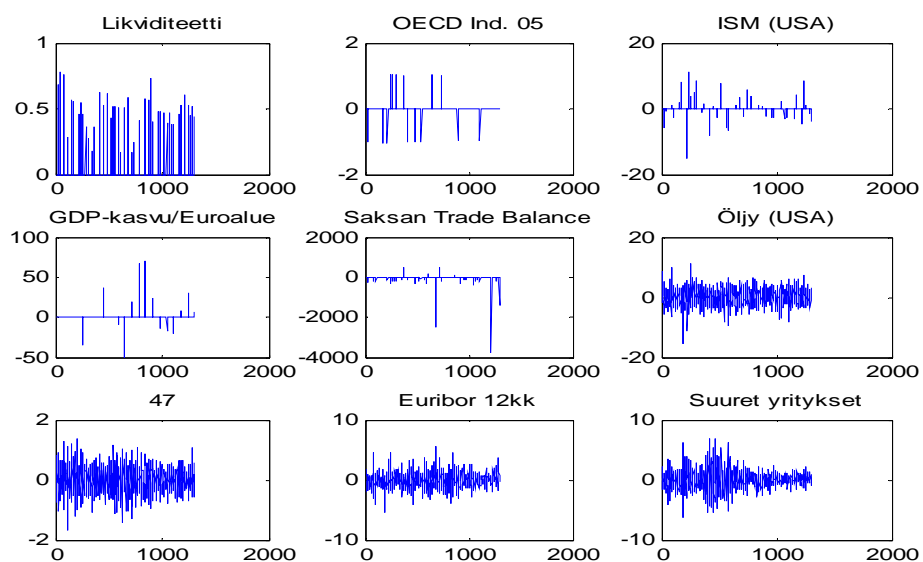
3.4 Aikasarjojen muokkaaminen

Aikasarjojen yhteismitallistamiseksi ja autokorrelaation vähentämiseksi muunnetaan rahaston arvot päivittäisten tuottojen aikasarjoiksi ja faktoreiden arvot faktorien arvojen päivittäisten muutosten aikasarjoiksi. Rahastojen aikasarjojen muuntaminen tuotoiksi vastaa aikasarjan ensimmäisen kertaluvun differensoimista. Differensoimalla voidaan saattaa aikasarja stationaariseksi, jolloin aikasarjamallien soveltaminen on teoreettisilta ominaisuuksiltaan mahdollista.

Kuvaan 3 on piirretty kuvaa yksi vastaavien rahastojen aikasarjojen ensimmäisen kertaluvun differensit, jotka vastaavat rahaston tuottoa. Faktorit on differensoitu kuten rahastot. Kuvassa 4 on esitetty kuvan 2 faktoreista differensoimalla muodostetut aikasarjat.



Kuva 4 Rahastojen tuotot



Kuva 5 Faktoreiden 1. kertaluvun differenssit

Käsittelyssä aikasarjoista poistetaan myös viikonloput. Lisäksi aikasarjadata aggregoidaan kuukausittaiseksi. Aikaisemmissa tutkimuksissa rahastojen tuottoja on usein mallinnettu käyttäen logaritmisia tuottoja (esimerkiksi Tsay, 2002), joten laskemme myös päivittäiset- ja kuukausittaiset logaritmiset tuotot. Logaritmisoinnin tarkoituksena on varianssin vähentäminen.

4. Rahastojen ja faktoreiden välisten korrelaatioiden tunnistaminen ja analysointi

Korrelaatioiden tunnistaminen antaa tietoa siitä, mistä faktoreista rahastojen tuotot riippuvat. Toisaalta rahaston korreloimattomuus tietyn faktorin suhteen voi tarkoittaa sitä, ettei kyseinen rahasto sisällä tähän faktoriin liittyvää riskiä, esimerkiksi öljyn hintaan liittyvää riskiä. Korrelaatioita tarkastellessa käytettiin 36 faktorin ja 71 rahaston aikasarjoja ajalta 1.1.2001-31.12.2005. Faktoreiden ja rahastojen välisiä korrelaatioita tutkittiin Pearsonin korrelaatiokertoimien avulla.

4.1 Rahastojen ja faktoreiden väliset korrelaatiot

Ensiksi tutkittiin rahastojen *arvojen* sekä faktoreiden *arvojen* korrelaatiota. Liitteen kuva 16 esittää saadut tulokset korrelaatiomatriisina. Korrelaatiokertoimet on luokiteltu ryhmiin siten, että tummempi ruutu vastaa voimakkaampaa korrelaatiota. Punainen teksti viittaa negatiiviseen korrelaatioon siinä missä musta tai valkoinen teksti viittaa positiiviseen korrelaatioon.

Kuvan 16 perusteella voimme todeta, että markkina-indeksit LCXT-, MCXT ja SCXT korreloivat voimakkaasti rahastojen kanssa. Lisäksi voimme nähdä, että mitä suuremmista yrityksistä indeksi koostuu, sitä voimakkaammin indeksi korreloi rahastojen arvojen kanssa (*LargeCapXT* vastaan *SmallCapXT*). Selityksenä tälle lienee se, että rahastot koostuvat suurimassa osin suurista yrityksistä ja täten rahastot seuraavat paremmin suurten yritysten indeksejä.

Rahastojen arvot vaikuttaisivat myös korreloivan jonkin verran tiettyjen raaka-aineiden hintojen kanssa. Näistä raaka-aineista mainittakoon ainakin sinkki ja alumiini. Valuuttakurssit ja saksan vaihtotase eivät ainakaan lineaarisessa muodossa ole korreloineet rahastojen arvojen kanssa, mutta Euribor sen sijaan osoittautuu korreloivan rahastojen arvojen kanssa.

Rahastojen arvojen ja faktorien arvojen väliset korrelaatiot voivat kuitenkin antaa jokseenkin väärän kuvan oikeasta tilanteesta. Rahastoarvojen sekä faktoriarvojen aikasarjat voivat kummatkin kasvaa ajan myötä, ja näin ollen näiden *arvojen* välillä on havaittavissa korrelaatiota. Saadaksemme paremman arvion rahastojen todellisista riippuvuuksista faktorien kanssa, on tutkittava sekä rahastojen että faktorien arvojen *tuottoja*, eli *vaihteluita*.

Tutkiaksemme rahastojen sekä faktoreiden *tuottojen* korrelaatioita, muunnamme ensiksi datan niin, että rahastojen sekä faktoreiden *arvot* muutetaan prosentuaaliseksi tuotoksi päivästä T-1 päivään T. Esimerkiksi jos rahaston arvo on 10,0 päivänä T-1 ja päivänä T 11,0, on tuotto 10 (%).

Tehtyämme samat laskut tuotto-aineistolle kuten aikaisemmalle aineistolle, saatiin seuraavat tulokset.

Kuvasta 17 voidaan nähdä, että tutkittaessa arvojen sijasta tuottoja, rahastojen ja faktorien väliset lineaariset riippuvuudet ovat paljon heikommat. Voimme kuitenkin todeta, että rahastot riippuvat vielä voimakkaasti markkina-indekseistä. Lisäksi nähdään, että mitä suuremmista yrityksistä markkina-indeksi koostuu, sitä voimakkaammin se korreloi rahastojen tuottojen kanssa.

Kiinnostavana huomiona voidaan nähdä, että on olemassa joitain rahastoja jotka eivät riipu niin voimakkaasti markkina-indeksistä tai mistään muustakaan. Tiedyt raaka-aineet ja Euribor 12

kuukautta säilyttävät vielä jonkin asteisen korrelaation rahastojen kanssa, kun tarkastelua muutetaan arvoista tuottoihin.

4.2 Rahastotuottojen sekä faktorituottojen ristikorrelaatiot

Kiinnostuksen kohteena on myös tutkia miten rahastotuotot ja faktorituotot korreloivat keskenään. Rahastotuottojen keskenäinen korrelaatio kertoo miten hyvin eri rahastojen tuotot seuraavat toisiaan. Korreloimattomuus sen sijaan kertoo, miten hyvin eri rahastoihin sijoittaen voidaan diversifioida.

Faktoreiden väliset ristikorrelaatiot ovat myös hyvin mielenkiintoisia. Jos joidenkin faktoreiden välillä on huomattavissa korrelaatiota, tulee käyttää toista faktoria kummankin faktorin sijasta. Lineaarista mallia luotaessa on tärkeätä havaita nämä voimakkaasti keskenään korreloivat faktorit, jotta tekijöiden välillä ei olisi multikollinearisuutta. Poistamalla multikollinearisuus selittäjistä saadaan tilastollisesti merkitsevempiä.

Kuvassa 18 on matriisi faktorien ristikorrelaatioista. Matriisista nähdään, että esimerkiksi öljyn hinta USA:ssa korreloi melkein täydellisesti eurooppalaisen öljyn hinnan kanssa. Lisäksi nähdään, että kuparin hinta korreloi voimakkaasti öljyn, lyijyn, hopean ja pienten sekä keskisuurten yritysten tuottojen kanssa.

Pienten yritysten tuotot korreloivat melkein täydellisesti keskisuurten yritysten kanssa. OECD indeksi korreloi miltei täydellisesti likviditeetin kanssa. Kiinnostavana, mutta kenties odotettuna seikkana voidaan nähdä että Euribor 3 kuukautta korreloi *negatiivisesti* melkein täydellisesti likviditeetin kanssa.

Kuvasta 18 nähdään myös, että tietyt faktorit, jotka korreloivat voimakkaasti keskenään, selittävät toistensa vaihtelut. Yhtenä esimerkkinä voidaan nostaa esiin kuparin hinta, joka korreloi muiden metallien hintojen kanssa hyvin voimakkaasti. Tämä tarkoittaa käytännössä sitä, että lineaarista mallia luodessa, yhden metallin hinta tulee selittämään muiden metallien hintavaihtelut. Raportin myöhemmässä vaiheessa, jossa luodaan lineaarinen malli selittämään rahastojen tuottoja, tulemme näkemään, että kupari esiintyy hyvin usein näissä malleissa. Samalla näemme, että muita metalleja ei sisällytetä malliin selittäjiksi juuri ollenkaan. Syy tähän juurtaa kuparin ja muiden metallien korrelaatiosta.

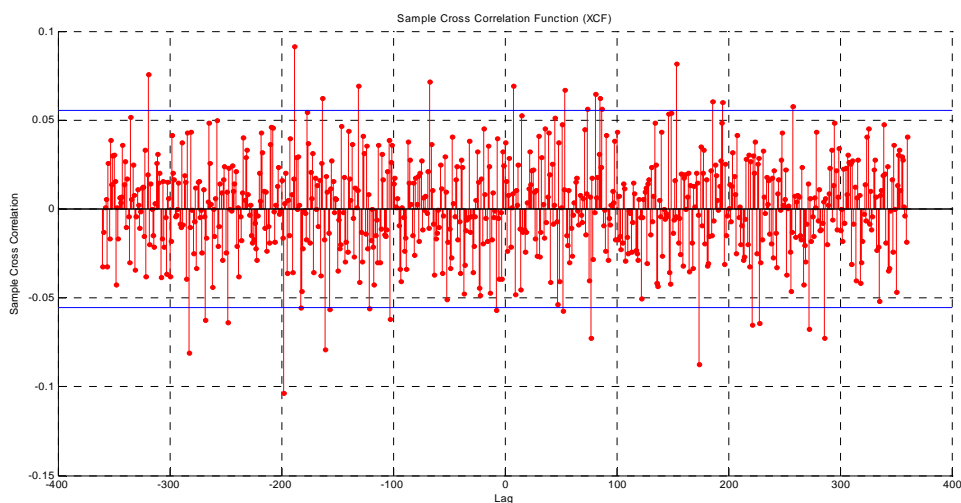
Vastaavalla tavalla emme näe likviditeettiä lineaarisissa malleissa johtuen siitä, että OECD:n indeksit ovat hyviä selittämään rahastojen tuottoja ja koska OECD:n indeksit korreloivat melkein täydellisesti likviditeetin kanssa.

Kuvasta 19 nähdään eri rahastojen tuottojen väliset korrelaatiot. Kuvasta voidaan nähdä, että suurin osa rahastojen tuotoista korreloi keskenään melko voimakkaasti. Tästä päätellen voidaan todeta, että rahastot ovat hyvin samantapaisia. On kuitenkin nähtävissä tiettyjä rahastoja, jotka eivät korreloi juuri yhtään muiden kanssa. Luotaessa rahastoja rahastoista, nämä rahastot voisivat toimia hyvinä työkaluina diversifioitaessa.

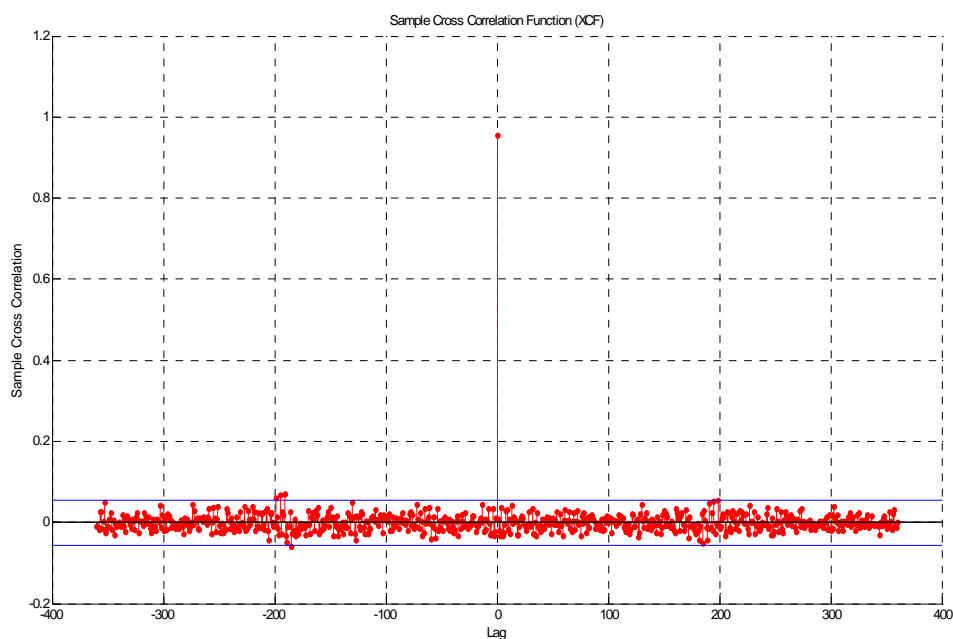
Muiden rahastojen tuottojen kanssa heikosti korreloivat rahastot ovat: LU0119128261 Equity CAF Europe Equities Classic Acc, LU0119106291 Equity HSBC Smaller European Companies Index, LU0093576881 Equity Alterum EurQuality, FI0008802749 Equity SEB Gyllenberg European Index.

4.3 Rahastotuottojen korrelaatio faktoreiden kanssa aikaviiveellä

Tutkimme myös riippuuko jonkin rahaston tuotto jostain faktorista tietyllä viiveellä (lag). Voi olla että löytyisi jokin rahasto, joka riippuu jostain faktorista esimerkiksi puolen vuoden viiveellä.



Kuva 6 korrelaatio eri viiveillä rahaston FI0008803614 Equity ja öljyn hinnan (USA) välillä.



Kuva 7 Korrelaatiot eri viiveillä rahaston LU0094557526 Equity ja LCXT Index (Suuret yritykset) kanssa.

Tutkittiin eri faktoreiden ja rahastojen välisiä korrelaatioita eri viiveillä. Tulokseksi saadaan, että faktorit ja rahastot korreloivat useimmiten voimakkaimmin keskenään, kun emme käytä mitään viivettä. Pystytään myös toteamaan, että korrelaatiot viiveillä tarkasteltujen faktorien ja rahastojen kanssa oli miltei aina hyvin heikkoa. Kuvissa 7 ja 8 on esimerkkejä joidenkin rahastojen ja faktoreiden korrelaatioista, kun ollaan huomioitu mahdollisia viiveitä ja niiden vaikutukset.

5. Regressiomalli

5.1 Lineaarinen regressio

Tarkastellaan lineaarisen regressiomallin sopivuutta mallintaa rahastojen aikasarjojen ja tuottojen kehitystä. Usean selittäjän lineaarisessa mallissa oletetaan, että selitettävän muuttujan y havaittujen arvojen vaihtelu halutaan selittää selittävien muuttujien x_1, x_2, \dots, x_k arvojen vaihtelun avulla. Mallissa on tarkoituksena käyttää selittävinä tekijöinä saatuja faktoreita ja selitettävänä tekijöinä rahastojen aikasarjoja. Regressiota voidaan pitää staattisena sillä selitettävän muuttujan arvot eivät muutu ellei selittävien muuttujien arvot muutu ja toisaalta selitettävän muuttujan arvot reagoivat viiveettä, kun selittävien muuttujien arvot muuttuvat.

Lineaarinen regressio etsii relaatiota selittävän muuttujan x_1, x_2, \dots, x_k ja selitettävän muuttujan y välillä, mutta ei ota kantaa relaation kausaalisuuteen. Tutkimuksissa on, esimerkiksi Liljeblom ja Stenius (Liljeblom & Stenius, 1993), saatu tuloksia joiden mukaan pörssikurssien volatiliiteetti selittää makroekonomista volatiliiteettiä. Selitämme tässä kuitenkin makroekonomisilla faktoreilla ja taloudellisilla indekseillä rahastojen arvojen ja tuottojen kehitystä.

Yleinen lineaarinen malli voidaan esittää muodossa:

$$y_t = \beta_0 + \beta_1 x_{t1} + \beta_2 x_{t2} + \dots + \beta_k x_{tk} + \varepsilon_t, \quad t = 1, 2, \dots, n,$$

missä y_t on selitettävän muuttujan havaittu arvo, x_t selittävän muuttujan havaittu arvo, β_0 vakioselittäjän regressiokerroin, β_j on selittäjän x_j $j = 1, 2, \dots, k$ regressiokerroin ja ε_t on jäännöstermi.

Yleistä lineaarista mallia koskee kuusi standardioletusta, joiden voimassaolo takaa sopivien estimointi- ja testausmenetelmien käytettävyyden mallille. Nämä standardioletukset ovat:

- (i) Selittävät muuttujat ovat kiinteitä eli ei satunnaisia muuttujia.
- (ii) Selittäjien välillä ei ole lineaarisia riippuvuuksia.
- (iii) $E(\varepsilon_t) = 0, \quad t = 1, 2, \dots, n$
- (iv) $Var(\varepsilon_t) = \sigma^2, \quad t = 1, 2, \dots, n$
- (v) $Cor(\varepsilon_t, \varepsilon_l) = 0, \quad t \neq l$
- (vi) $\varepsilon_t \sim N(0, \sigma^2), t = 1, 2, \dots, n$

Tarkastellaan standardioletusten tulkintaa ja vaikutuksia. Standardioletuksia käytetään testattaessa mallin hyvyttä. Lineaaristen mallien perusteoria nojaa voimakkaasti oletukseen (i). Koska kyseessä on lineaarisen regression soveltaminen aikasarjoihin, oletuksesta seuraa, että regressiokertoimet eivät muutu ajassa. Jos kertoimet muuttuvat ajassa, mallissa tapahtuu rakenteen muutos. Toinen standardioletus on tärkeä, sillä jos selittäjä x_j riippuu vahvasti toisesta selittäjästä, voidaan selittäjää x_j pitää turhana ja se voidaan poistaa mallista. Kolmas oletus sanoo, että kaikilla jäännöstermeillä tulee olla sama odotusarvo, 0. Oletus takaa, ettei mallissa ole systemaattista virhettä. Oletuksen (iv) mukaan kaikilla jäännöstermeillä tulee olla sama varianssi ja tätä oletusta

kutsutaan homoskedastisuusoletukseksi. Jos oletus ei ole päde, ovat jäännöstermit heteroskedastisia, mikä tekee estimoiduista estimaattoreista tehottomia. Oletusta (v) kutsutaan korreloimattomuusoletukseksi. Jos oletus ei ole voimassa ovat regressiokertoimen estimaattorit tehottomia ja jopa harhaisia. Oletuksen (vi) mukaan jäännöstermien tulee olla normaalijakautuneita.

5.2 Dynaaminen lineaarinen regressio

Voisi kuvitella, että taloudessa tapahtuvat muutokset eivät tapahdu viiveettä. Toisin sanoen, kestää jonkun aikaa, ennen kuin makroekonomisten muutosten vaikutukset näkyvät täysipainoisina osakekurseissa. Näin ollen olisi järkevää käyttää malleja, joissa selitettävän muuttujan ehdollinen odotusarvo reagoi vähitellen tai asteittain selittävien arvojen muutoksiin. Dynaaminen regressiomalli vastaa tähän haasteeseen. Yhden selittäjälle muotoa:

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \dots + \beta_k x_{t-k} + \varepsilon_t, \quad t = k+1, k+2, \dots, n$$

Tällaista mallia kutsutaan jakautuneen viipymän malliksi ja sen jäännöstermin oletetaan toteuttavan seuraavat modifioidut standardioletukset:

- (i) $E(\varepsilon_t | x_t, x_{t-1}, \dots, x_{t-k}) = 0, \quad t = k+1, k+2, \dots, n$
- (ii) $Var(\varepsilon_t | x_t, x_{t-1}, \dots, x_{t-k}) = \sigma^2, \quad t = k+1, k+2, \dots, n$
- (iii) $Cor(\varepsilon_t, \varepsilon_s | x_t, x_{t-1}, \dots, x_{t-k}, x_s, x_{s-1}, \dots, x_{s-k}) = 0, \quad t \neq s$

Verrattuna staattiseen tilanteeseen, jakautuneen viipymän mallin määrittelemän systeemin sopeutuminen uuteen tasapainoon kestää k ajanhetkeä. Regressiokertoimen β_0 voidaan siis tulkita kuvaavan muuttujan x välitöntä vaikutusta ja regressiokertoimien $\beta = \beta_0 + \beta_1 + \dots + \beta_k$ kuvaavan muuttujan x saaman yhden yksikön lisäyksen pitkän ajan vaikutusta.

5.3 Regressiodiagnostiikka

Regressiomallia koskevien oletusten tarkistamista kutsutaan regressiodiagnostiikaksi. Saatua mallia pidetään tilastollisesti oikeana, jos mallista saadut estimointitulokset ovat sopuinnussa mallia koskevien oletusten kanssa.

Malleissa joudutaan tekemään oletus (oletuksen (i) mukaisesti), että parametrit ovat vakioita ajassa. Aikasarjassa tapahtuvaa rakenteen muutosta voidaan testata esimerkiksi jakamalla tarkasteltava aikaväli kahtia, suorittamalla regressio molemmille periodeille ja vertaamalla selittäjiä keskenään. Täysin eri selittäjät eri periodeille voisivat viitata jonkin asteiseen rakenteen muutokseen.

Mallin estimoinnin jälkeen tarkistetaan, selittäjien lineaarinen riippuvuus. Multikollineaarisuuden ilmetessä on syytä yhdistää vahvasti korreloivat selittäjät keskenään ja suorittaa uusi regressio.

Oletuksia (iii) – (vi) voidaan tutkia tilastollisten testien ja regressiodiagrammien avulla.

(iv) Homoskedastisuusoletusta voidaan tutkia apuregression avulla, joka on muotoa:

$$e_i^2 = \alpha_0 + \alpha_1 \hat{y}_i + \delta_i \quad i = 1, 2, \dots, n,$$

Missä ε_i on mallin jäännöstermi ja \hat{y}_i mallin tuottama sovite. Homoskedastisuusoletuksen pätiessä

$nR^2 \sim \chi^2(1)$, missä n on havaintojen lukumäärä ja R^2 mallin selitysaste.

(v) Autokorrelaatiota voidaan testata Durbin ja Watsonin testisuureella, joka on muotoa:

$$DW = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2}, \quad i = 1, 2, \dots, n \quad \text{Tulkinta:}$$

$$(i) \quad 0 \leq DW \leq 4$$

$$(ii) \quad DW \approx 0 \leftrightarrow \rho \approx +1$$

$$(iii) \quad DW \approx 2 \leftrightarrow \rho \approx 0$$

$$(iv) \quad DW \approx 4 \leftrightarrow \rho \approx -1$$

Siis, jos $DW \approx 2$ nollahypoteesi pätee ja residuaalit eivät ole autokorreloituneita. Residuaalien tarkasteluissa oletetaan, että autokorrelaatio ei ole ongelma jos $1,5 \leq DW \leq 2,5$.

(vi) Normaalisuustestejä on useita, joista tässä käytetään Bowmanin ja Shentonin testiä. Se on muotoa:

$$\chi^2 = \frac{n}{6}c_1^2 + \frac{n}{24}c_2^2, \quad \text{missä } c_1 \text{ ja } c_2 \text{ ovat estimoidun mallin residuaalien vinous ja huipukkuus.}$$

Testisuure on $\chi^2(2)$ jakautunut. Suurilla testisuureen arvoilla nollahypoteesi normaalisuudesta hylätään.

5.4 Regressiomallin valinta

Hyvän regressiomallin tulisi sekä selittää hyvin havaitut tapahtumat että ennustaa tarkasti tulevia tapahtumia. Käytettäessä useita selittäviä muuttujia saadaan tutkittava data yleensä hyvin kuvattua, mutta malli voi olla tällöin monimutkainen ja tehoton (regressiokertoimien varianssit suuria). Malliin tulisi täten valita sopiva määrä selittäjiä, jotka kaikki ovat tilastollisesti merkitseviä ja joilla saadaan selitysvoimainen, mutta yksinkertainen malli. Paras malli löytyisi kokeilemalla kaikki selittäjäehdokkaiden lineaarikombinaatiot. Kaikkien kombinaatioiden kokeileminen on kuitenkin hyvin suuritöistä kun selittäjiä on paljon ja tämän vuoksi joudutaan käyttämään tehokkaampia mallinvalintatapoja, joilla voidaan löytää hyvä malli. Rahastojen tuottoja kuvaavan regressiomallin valinta on toteutettu valikoivalla askelluksella.

5.5 Valikoiva askellus

Valikoivalla askelluksella löydetään tilastollisesti merkitsevä regressiomalli, jossa kaikki selittävät muuttujat ovat tilastollisesti merkitseviä. Mallia määritettäessä valitaan aluksi merkitsevyystaso kuvaamaan selittäjien tilastollista merkitsevyyttä. Menetelmä koostuu seuraavista vaiheista:

1. Lähtöasetelmassa mallissa ei ole muuttujia, eli malli on tyhjä.
2. Ensimmäiseksi selittäjäksi valitaan se, jonka korrelaatio selitettävän muuttujan kanssa on suurin (selittäjä, jonka p-arvo mallissa on pienin).
3. Valitaan malliin uusi selittäjä, lisäämällä malliin yksi kerrallaan jokainen malliin kuulumattomista selittäjistä ja valitsemalla uudeksi selittäjäksi se, joka nostaa mallin selityksastetta eniten (p-arvo pienin).

4. Uuden selittäjän lisäyksen jälkeen tutkitaan, onko jonkin malliin aiemmin lisätyn selittäjän p-arvo noussut yli valitun merkitsevyystason. Mikäli näin on, poistetaan kyseinen selittäjä.
5. Toistetaan vaiheita 2 ja 3 kunnes malliin ei ole lisättävissä selittäjiä, joiden p-arvo mallissa on alle valitun tason.

Yllä esitellyllä menetelmällä saadaan muodostettua hyvä regressiomalli selitettävälle muuttujalle. Mallin ja sen selittäjien tilastollisen merkitsevyys ei kuitenkaan vielä takaa mallin käyttökelpoisuutta. Käyttökelpoisuuden tutkimiseksi on suoritettava regressiodiagnostiikkaa.

5.6 Regressiomallin toteutus

Regressiomallin valinnat rahastoille on toteutettu Matlabin valikoivaa askellusta hyödyntävällä Stepwisefit-metodilla. Stepwisefit-metodilla on jokaisella rahastolle valittu selittävät faktorit. Regress-metodilla saadaan tämän jälkeen laskettua regressiodiagnostiikkaa varten residuaalit, virherajat.

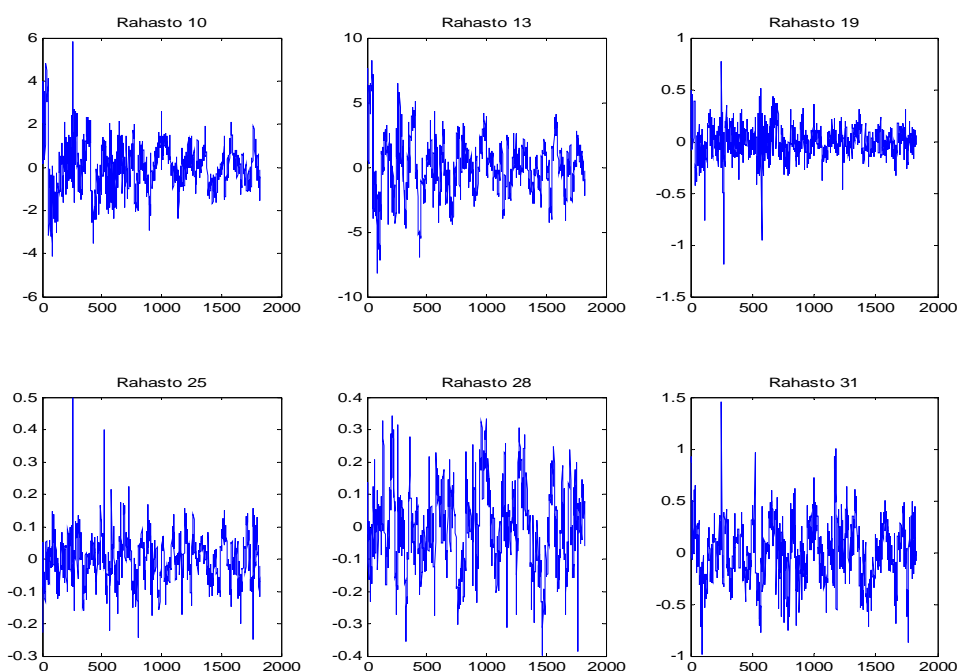
6. Tulokset

6.1 Mallien estimointi rahastojen arvojen perusteella

Selitetään faktorien aikasarjoilla rahastojen aikasarjoja. Aikasarjadataalle tehdyt ajot osoittavat, pelätyn ja kirjallisuudesta ilmi tulleen tosiasian residuaalien autokorrelaatiosta. Taulukossa 1 on esitetty tuloksia 10 rahastolle. Durbinin ja Watsonin testisuureen arvosta nähdään, että residuaaleissa on vahvaa autokorrelaatiota. Muiden 64 rahaston residuaalit käyttäytyvät samaan tapaan. Piirretään vielä kuvaan 8 muutaman estimoidun mallin residuaalit kuudelle rahastolle.

Taulukko 1 Testit residuaaleille.

Rahaston indeksi	Mallin R ²	Bowm.-Sh. (0 Ok, 1 Ei)	Apuregression χ^2	Testi (1 Ok, 0 Ei)	Durbin Watson
4	0.9956	0	0.0411	1	0.4056
10	0.9956	0	0.0441	1	0.1479
13	0.9919	1	0.1513	1	0.1085
19	0.9976	0	0.03	1	0.4809
25	0.998	0	0.0263	1	0.3533
28	0.9979	0	0.0022	1	0.2857
31	0.9962	0	0.014	1	0.1742
34	0.9946	0	0.3308	1	1.0934
37	0.9911	0	0.0228	1	0.2259
40	0.9947	0	0.0018	1	0.3048



Kuva 8 Residuaalien kuvaajat.

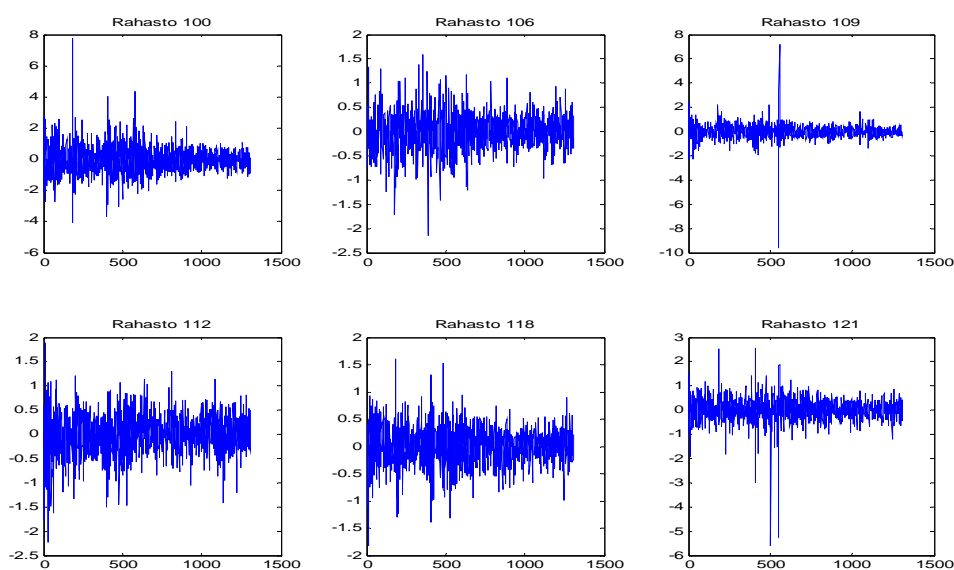
6.2 Mallien estimointi rahastojen tuottojen perusteella

6.2.1 Mallin estimointi päivittäisellä datalla

Edellä todettiin, että faktoreiden aikasarjojen avulla estimoidut mallit eivät läpäise tilastollisia testejä. Selitetään nyt faktoreiden muutoksilla rahastojen tuottoja. Estimoitujen mallien tuloksia on esitetty taulukossa 2. Tuloksista nähdään, että autokorrelaatio on pääosin poistunut mallista, mutta mallin residuaaleissa on vahvaa heteroskedastisuutta. Kuva 9 vahvistaa tuloksemme varianssin muutoksesta residuaaleille. Esimerkiksi rahaston 100 residuaalien varianssi vaimenee huomattavasti aikavälin loppupuolella.

Taulukko 2 Testit residuaaleille päivittäisellä datalla.

Rahaston indeksi	Mallin R ²	Bowm.-Sh. (0 Ok, 1 Ei)	Apuregression χ^2	Testi (1 Ok, 0 Ei)	Durbin Watson
100	0.6113	1	157.6627	0	2.6153
106	0.7764	1	291.1484	0	2.0363
109	0.7943	1	17.1675	0	2.1124
112	0.7779	1	326.9426	0	1.91
118	0.8584	1	324.7747	0	2.0382
121	0.7676	1	51.337	0	2.1183
124	0.9335	1	287.8521	0	2.2568
130	0.1706	1	311.2809	0	2.1181
133	0.8426	1	46.0314	0	2.3567
145	0.5092	1	106.4275	0	2.8075



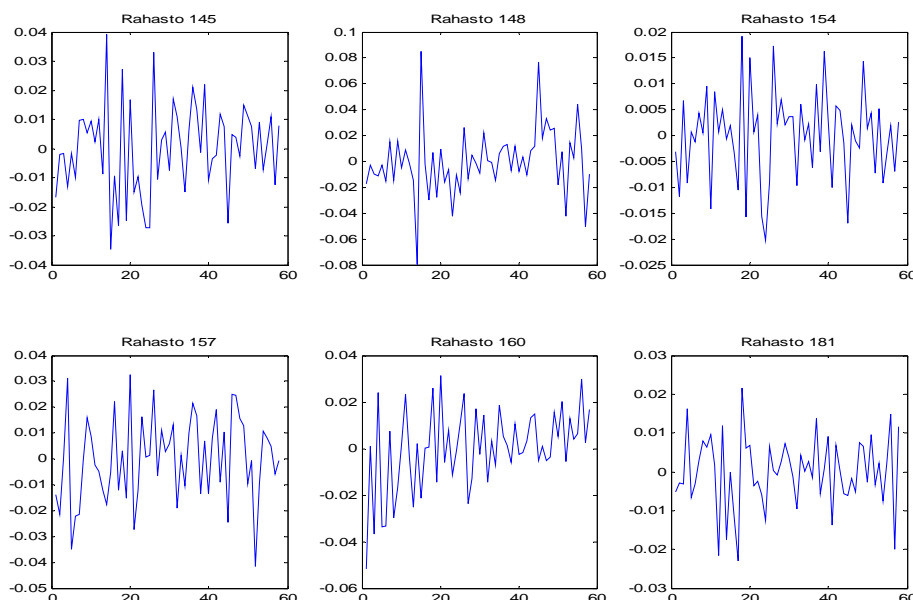
Kuva 9 Residuaalien kuvaajat päivittäisellä datalla.

6.2.2 Mallin estimointi kuukausittaisella datalla

Kun aikaväliä kasvatetaan, voidaan autokorrelaation ja varianssin vaihtelun olettaa vähenevän. Kuukausittaisesta aikaväliä käyttämällä malliin saadaan lisää mahdollisia selittäjiä. Nyt selitetään faktoreiden kuukausittaisilla muutoksilla rahastojen kuukausittaisia tuottoja. Valikoiva askellus antaa tuloksena 10 valitun rahaston residuaaleille taulukon 3 mukaiset arvot. Tulokset näyttävät varsin hyviltä. Mallin residuaalit ovat homoskedastisia ja 7 esitetystä rahastosta läpäisee autokorrelaomattomuustestin. Lisäksi 6 rahastoista läpäisee residuaalien normaalisuustestin. Piirretään vielä kuvaan 10 mallien residuaaleja. Kuvista nähdään, että residuaalit käyttäytyvät varsin hyvin ja koska residuaalit läpäisevät tilastolliset testit useiden rahastojen kohdalla, olemme löytäneet hyvän mallin.

Taulukko 3 Testit lineaarisen mallin residuaaleille kuukausittaisella datalla.

Rahaston indeksi	Mallin R ²	Bowm.-Sh. (0 Ok, 1 Ei)	Apuregression χ^2	Testi (1 Ok, 0 Ei)	Durbin Watson
145	0.91	0	0.0016	1	2.4431
148	0.7727	1	0.0009	1	2.2077
154	0.9714	0	0.019	1	2.5165
157	0.9272	0	0.0083	1	2.073
160	0.9092	0	0.0032	1	2.1014
181	0.9686	0	0.0015	1	2.6144
187	0.892	1	0.0004	1	1.765
190	0.877	1	0.0021	1	1.4812
193	0.8964	0	0.0022	1	2.2667
196	0.9174	1	0.1146	1	1.5903



Kuva 10 Residuaalien kuvaajat kuukausittaisella datalla.

Tarkastellaan, mitkä selittäjät ovat selittäneet rahastojen kehitystä. Valitaan tarkasteluun malleista vain sellaiset, jotka läpäisevät kaikki residuaaleja koskevat mallin hyvyystestit. Tällaisia rahastoja 71:stä on 48 kappaletta. Taulukossa 4 on 14 faktoria, jotka esiintyvät yleisimmin merkitsevinä selittävinä tekijöinä malleissa. Yhteensä selittäviä faktoreita rahastojen tuotoille löytyy 38. Yleisimpänä selittäjänä on keski suurten yritysten arvonkehitystä kuvaava indeksi ja toiseksi yleisempänä suurten yritysten indeksi. Nämä selittäjät ovat varsin loogisia, mutta eivät anna arvokasta tai yllättävää tietoa siitä, mikä vaikuttaa rahastojen tuottoon. Näitä yritysten indeksejä voidaan jopa pitää endogeenisinä selittäjinä, sillä rahastot ja selittäjät sisältävät samojen yritysten osakkeita. Jotta, löytäisimme ulkoisia selittäjiä, poistamme yritysten arvonkehitystä kuvaavat faktorit: suuret yritykset, keski suuret yritykset ja pienet yritykset.

Taulukko 4 Todetut faktorit kuukausittaisella datalla. Yritysindeksit sisälletty.

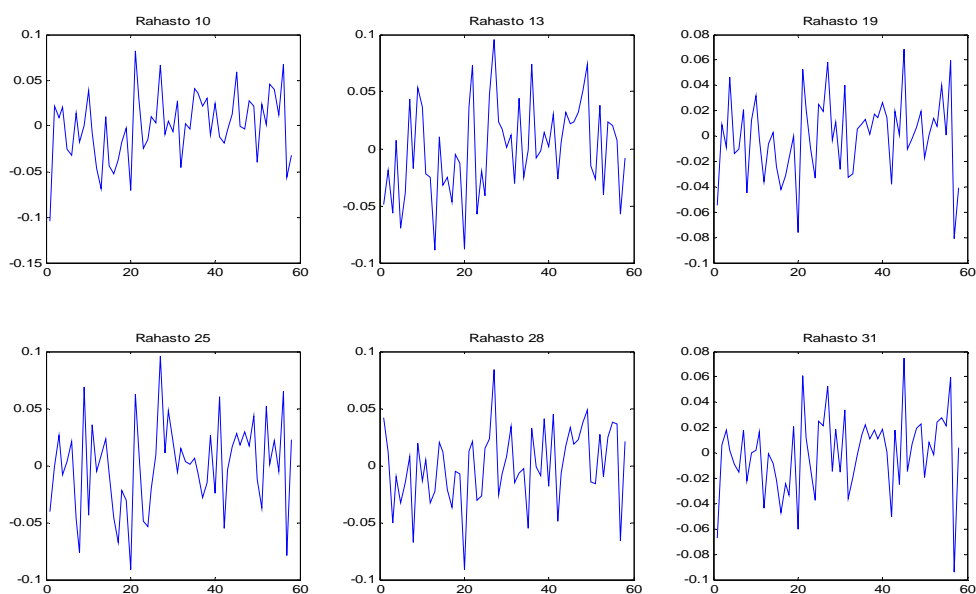
Lkm rahastoissa	Faktori-indeksi	Faktori
35	71	Keski suuret yritykset
28	73	Suuret yritykset
10	59	Euro vs. NZD
7	5	OECD Ind. 05
6	1084	Current Account, Services, Net, Total, SA, EUR
6	1071	Total reserves Incl gold, EUR
5	69	Pienet yritykset
3	1055	PMI, Manufacturing Sector, Stocks of Finished Goods, SA
3	45	Euro vs. USD
2	1079	Current Account, Income, Net, Total, EUR
2	1074	Financial Account, Direct Investment, Balance, Total, EUR
2	1026	Credit to oth euro area residents, Chg YY, SA
2	1005	CPI, Education, Total, Index
2	65	Euribor 3kk

6.2.3 Mallien estimointi kuukausittaisella datalla ilman yritysten indeksejä

Taulukossa 5 on esitetty estimoinnin tulokset. On huomattava, että mallien selitysasteet ovat laskeneet huomattavasti, verrattuna edelliseen kohtaan, mutta normaalisuustestit ja autokorreloituneisuustestit läpäistään useammin. Kuvaan 11 piirretyt kuvat residuaaleista varmistavat käsityksen, että estimoidut mallit ovat hyviä. Verrattaessa residuaalien kuvaajia kohdan 6.2.2 kuvaan, todetaan, että residuaalit käyttäytyvät jopa paremmin.

Taulukko 5 Testit residuaaleille kun malli on estimoitu ilman yritysindeksijä.

Rahaston indeksi	Mallin R ²	Bowm.-Sh. (0 Ok, 1 Ei)	Apuregression χ^2	Testi (1 Ok, 0 Ei)	Durbin Watson
4	0.4985	0	0.0438	1	1.9893
10	0.5294	0	0.0006	1	1.8084
13	0.5322	0	0.0041	1	1.7943
19	0.6467	0	0.0006	1	2.1877
25	0.3705	0	0.0225	1	2.2085
28	0.5773	0	0.001	1	2.1238
31	0.6171	0	0.0002	1	2.3273
34	0.6273	0	0.0119	1	1.7953
37	0.6151	0	0.0063	1	2.2017
40	0.8232	1	0.1174	1	2.5064



Kuva 11 Residuaalien kuvaajat.

Alla olevaan taulukkoon on kerätty lista faktoreista, jotka ovat useimmin olleet selittävinä tekijöinä mallissa. Huomataan, että faktorit eroavat selvästi kohdan 6.2.2 faktoreista. Nyt selittäviä faktoreita on myös useampia, yhteensä 48 erilaista. Yleisimmäksi selittäjäksi nousee yllättäen kupari. On vaikeaa löytää selvää asialoogista perustelua sille, miksi juuri kupari on hyvä selittämään rahastojen tuottojen kehitystä. OECD:n johtava indikaattori 06 on mielenkiintoinen selittäjä.

Taulukko 6 Todetut faktorit kun datasta on poistettu yritysindeksit.

Lkm rahastoissa	Faktori-indeksi	Faktori
40	25	Kupari
32	7	OECD Ind. 06
26	1038	Net external assets, SA, EUR
22	1047	New Orders, Aggregates, Consumer non-durables, Index
15	9	ISM(USA)
14	1062	PMI, Whole Economy, Input prices Whole Economy, SA
13	1028	Credit to oth euro area residents, SA, EUR
12	67	Euribor 12kk
7	27	Lyijy
7	1074	Financial Account, Direct Investment, Balance, Total, EUR
6	23	Alumiini
5	1078	Current Account, Goods, Net, Total, SA, EUR
4	33	Sinkki
4	1056	PMI, Manufacturing Sector, Input prices, SA
3	5	OECD Ind. 05
3	1026	Credit to oth euro area residents, Chg YY, SA
2	37	Sähkö
2	1009	CPI, Housing, Water, Electricity, Gas & Other Fuels,Tot,Idx
2	1041	New Orders, Aggregates, Consumer goods, Index
2	1068	PMI, Manufacturing Sector, Supplier's delivery times, SA

6.2.4 Mallien estimointi Log-tuottojen perusteella

Estimoidaan mallit logaritmoiduille tuotoille käyttäen päivittäistä dataa. Estimointituloksista voidaan todeta, etteme saa tilastollisesti merkitseviä malleja päivittäiselle logaritmoidulle datalle. Ongelmiksi nousevat residuaalien vahva heteroskedastisuus sekä residuaalien autokorrelaatio.

Myöskään kuukausittaisilla logaritmisilla tuotoilla estimoidut regressiomallit eivät ole merkitseviä. Malleissa on ongelmina normaalisuus, heteroskedastisuus ja autokorrelaatio.

6.2.5 Mallien estimointi käyttäen viivästettyjä selittäjiä

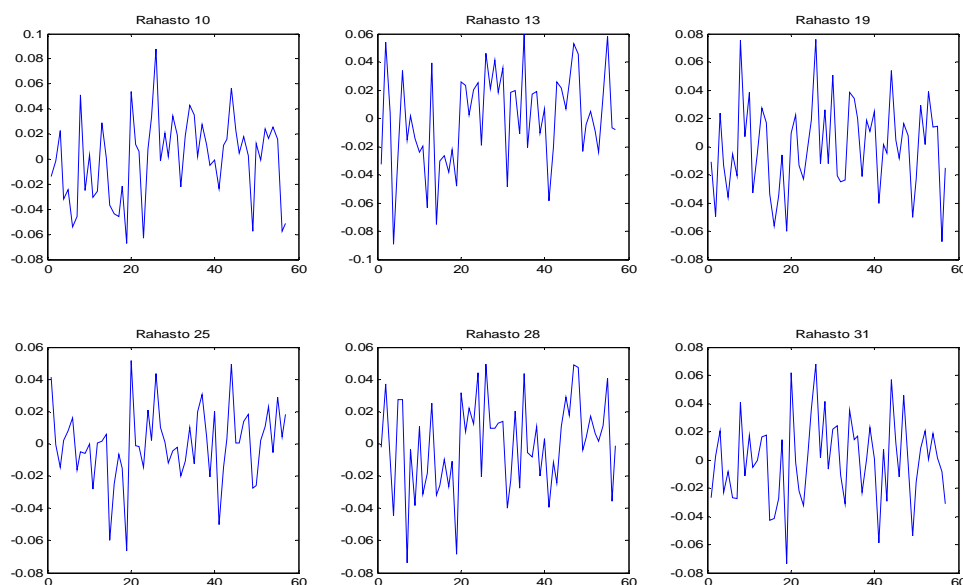
Estimoidaan kohdan 5.2 mukainen jakautuneen viivästyksen malli viivästyksen faktoreita yhdellä yksiköllä. Malli on tällöin siis muotoa:

$$y_t = \beta_0 x_t + \beta_1 x_{t-1} + \varepsilon_t$$

Tulokset valituille rahastoille on esitetty taulukossa 7. Tuloksista selviää, että estimoidut mallit läpäisevät kaikki residuaalien testit. Kuvaan 12 on piirretty kuvia residuaaleista ja kuvista varmistuu residuaalinen hyvä käyttäytyminen.

Taulukko 7 Testit residuaaleille viivästetyillä selittäjillä.

Rahaston indeksi	Mallin R ²	Bowm.-Sh. (0 Ok, 1 Ei)	Apuregression χ^2	Testi (1 Ok, 0 Ei)	Durbin Watson
100	0.8731	0	0.0031	1	2.0576
106	0.8537	0	0.0017	1	1.9771
109	0.891	0	0.0014	1	1.5157
112	0.6719	0	0.0013	1	1.5637
118	0.7421	0	0.007	1	2.2845
121	0.9887	0	0.0007	1	2.018
124	0.9943	0	0.0182	1	2.0467
130	0.9343	0	0.0144	1	1.7965
133	0.5551	0	0.1665	1	1.5615
145	0.4924	0	13.5466	0	1.7889



Kuva 12 Residuaalien kuvaajat.

Taulukkoon 8 on tehty lista faktoreista, jotka esiintyvät useimmin selittävinä tekijöinä estimoiduissa yhden viivymän mallissa viipeellä 0. Lista on samansuuntainen kohdan 6.2.3 listan kanssa. Kupari on jälleen selittävistä tekijöistä yleisin. Järkevänä selittäjänä voidaan pitää kolmanneksi yleisintä 12kk Euriboria, joka esiintyy 24:ssä rahastossa. Taulukossa 9 on esitetty yleisimmät faktorit yhden aikayksikön viipeellä. Käytännön tulkinta on, että kyseisen indikaattorin edellisen kuukauden arvo selittää tämän kuukauden rahaston tuottoa. Yleisin faktori on ISM (USA), joka on selittäjänä 15 rahastossa.

Taulukko 8 Nollan aikayksikön viipeellä esiintyvät faktorit dynaamisessa regressiomallissa.

Lkm	Indeksi	Faktori	Keskimäär. Reg. Kerroin
46	25	Kupari	0.3465
31	7	OECD Ind. 06	8.1317
27	67	Euribor 12kk	0.3652
24	1038	Net external assets, SA, EUR	0.0239
19	9	ISM(USA)	-0.4210
18	1062	PMI, Whole Economy, Input prices Whole Economy	-0.6459
17	1047	New Orders, Aggregates, Consumer non-durables	-0.2105
14	1074	Financial Account, Direct Investment, Balance	0.0032
8	1056	PMI, Manufacturing Sector, Input prices	-0.2844
7	33	Sinkki	0.3340
6	1028	Credit to oth euro area residents	4.9980
5	1053	PMI, Services Sector, Employment	0.8165
5	1078	Current Account, Goods, Net, Total	0.0130
4	1083	Current Account, Services, Net, Total	-0.0051
3	5	OECD Ind. 05	2.2666
3	27	Lyijy	0.1733
3	1045	New Orders, Manuf, Overall, Total, Index	-0.1081

Taulukko 9 Yhdellä viipeellä esiintyvät faktorit.

Indeksi Lag	Lkm	Faktori	Keskimäär. Reg. Kerroin
10009	17	ISM(USA)	-0.5071
11078	12	Current Account, Goods, Net, Total, SA, EUR	-0.0296
11035	6	Assets, loans to euro area residents Incl Eurosystem, EUR	5.9462
11055	4	PMI, Manufacturing Sector, Stocks of Finished Goods, SA	-0.8280
10003	2	Likviditeetti	9.4170
11005	4	CPI, Education, Total, Index	4.7445
11012	3	CPI, Restaurants and Hotels, Total, Index	-2.0788
11076	3	Financial Account, Other Investment, Balance, Total, EUR	0.0011
10025	3	Kupari	0.2214
10003	2	Likviditeetti	9.4170

Yhden kuukauden viipeen lisäksi kokeiltiin selittäjiä kolmen ja kuuden kuukauden viipeillä. Estimoidut mallit eivät kuitenkaan olleet hyviä. Myös päivittäiseen dataan sovitettiin jakautuneen viipymän malli, mutta residuaalien heteroskedastisuus säilyi vahvana eikä hyvää mallia saatu estimoitua.

6.3 Parhaan mallin valinta ja herkkyysanalyysi

Tarkoituksena on ollut mallintaa rahastojen arvojen ja tuottojen käyttäytymistä lineaarisella regressiomallilla. Tuloksista voidaan todeta, ettei päivittäisdatalle ole onnistuttu estimoimaan hyvää mallia, eli sellaista, joka läpäisee standardioletusten vaatimat testit. Ongelmaksi nousee residuaalien heteroskedastisuus, mikä tarkoittaa residuaalien varianssin muuttumista tarkasteluajanjaksolla. Tämä ongelma oli odotettavissa kirjallisuuskatsauksen perusteella.

Kuukausittaisella datalla tehdyt estimoinnit antavat merkitsevän mallin. Ensimmäinen estimointi osoitti, että yritysten indeksit ovat hallitsevia selittäjiä mallissa. Ulkoisten selittäjien löytämiseksi kyseiset indeksit poistettiin ja saatiin uusia merkitseviä selittäjiä. Lopulta testattiin jakautuneen viipymän malleja. Tilastollisesti paras malli löydettiin selittämällä rahastojen kuukausittaisia tuottoja

faktoreiden kuukausittaisilla muutoksilla samassa aikayksikössä ja yhdellä viipeellä. Taulukossa 10 on esitetty esimerkkinä korkeimman selityksasteen omaava malli, sen selittäjät ja niiden regressiokertoimien arvo. Malli kuvaa rahaston 124 tuoton selittämistä mainituilla faktoreilla. Regressiokertoimien tulkinta on se, että yhden yksikön muutos x :n arvossa saa regressiokertoimen kokoisen muutoksen aikaan selitettävässä tekijässä y . Koska selittäjien x arvot ovat faktoreiden muutoksia, ovat regressiokertoimet yhteismitallisia. Näin voidaan suoraan tehdä herkkyyksanalyysiä eri faktoreiden vaikutuksesta. Rahaston 124 tuottoon voidaan siis sanoa vaikuttavan positiivisesti OECD:n johtava indikaattori 06, 12Kk Euribor ja Kuparin hinta. ISM(USA)- ja PMI, Whole Economy, Input Prices –faktoreilla on negatiivinen kerroin, mikä tarkoittaa, että niiden noustessa rahaston tuoton tulisi laskea.

Taulukko 10 Jakautuneen viipymän mallin faktorit rahastolle 124.

Rahasto 124	
Faktori	Regressiokerroin
Euribor 12Kk	0.4073
Kupari	0.4309
ISM (USA)	-0.3975
OECD Ind. 06	5.6961
New Orders, Aggr, Consumer non-durables	-0.1809
PMI, Whole Economy, Input pr. Whole Econ.	-0.6407
Current Account, Goods, Net, Total, SA, EUR	0.0173
Viipeellä	
Retail Sales, Turnover, Total	-0.0043
Financial Account, Other Investment, Balance	0.0009
Current Account, Goods, Net, Total, SA, EUR	-0.0254
Term premium	-0.0007
Vakiotermi	-4.3111

Rahastojen herkkyys faktoreille näkyy regressiokertoimien suuruuksina, tulkittaessa faktoreiden tuottojen muutokset yhtä suuriksi. Taulukoihin 8 ja 9 on kerätty lista faktoreista ja niiden keskiarvoisista regressiokertoimista. Regressiokertoimista nähdään, minkä suuntainen vaikutus kullakin faktorilla on rahaston tuottoon. Positiivinen relaatio rahaston tuoton kanssa on kuparin hinnalla, OECD:n johtavalla indikaattorilla 06, 12Kk Euriborilla ja Net External Assets:illa. Negatiivinen relaatio rahaston tuoton kanssa on ISM (USA):lla ja PMI Whole Economy:lla.

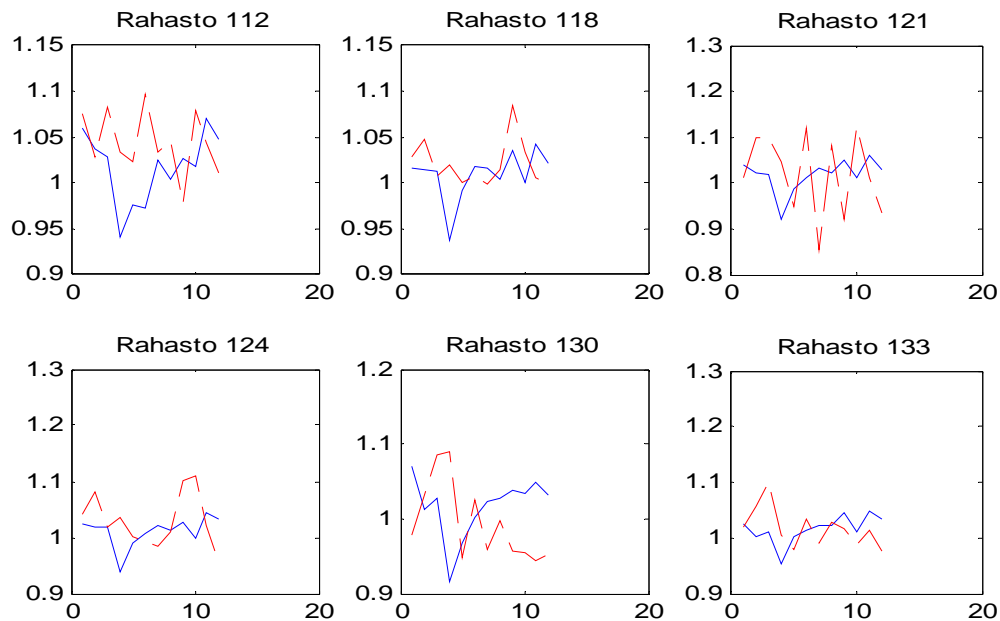
6.4 Mallin validointi

Testataan nyt estimoituja malleja vuoden 2006 toteutuneilla arvoilla. Käytetään jälleen kohdan 5.4 mallia. Validoinnissa siis sijoitetaan faktoreiden vuonna 2006 toteutuneet arvot estimoituun regressiomalliin, jolloin saada sovite ja soviteen ympärille piirretään 95% luottamusvälit. Saatua sovitea verrataan vuoden 2006 toteutuneisiin arvoihin, jolloin nähdään, miten hyvin estimoitu malli on vastannut todellisuutta.

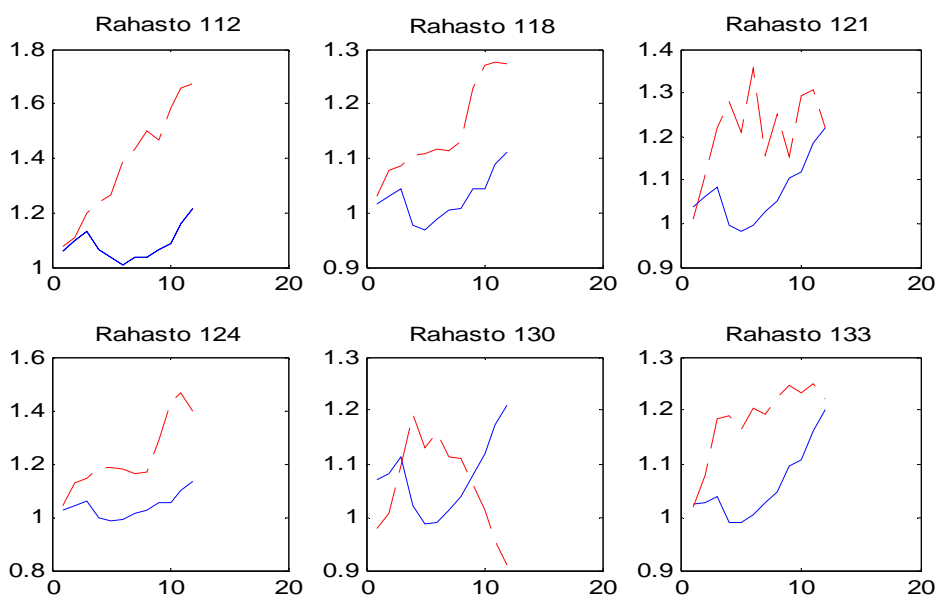
Kuvaan 13 on piirretty kuuden rahaston estimoidut sovitteet punaisella ja toteutuneet arvot sinisellä. Kuvat sisältävät arvot kahdeltatoista kuukaudelta aikaväliltä 1.1.2006–1.2.2007. Kuviin ei ole piirretty 95% luottamusvälejä, koska luottamusvälit ovat niin laajat, että soviteen ja toteutuneiden arvojen käyttäytymisestä ei saa selvää. Tämän perusteella voidaan siis todeta, että estimoidut mallit eivät ole kovinkaan luotettavia. Kuvasta huomataan, että estimoidut sovitteet eivät seuraa kovinkaan eksaktisti

toteutuneita arvoja. Etenkään korkeimman selityksasteen omaava malli, malli rahastolle 124, tuntuu täsmäävän heikosti toteutuneiden arvojen kanssa.

Kuvassa 14 on esitetty rahastojen tuoton kehitys tilanteelle, jossa vuoden 2006 alussa on sijoitettu yksi yksikkö. Kuvasta nähdään, että estimoidut mallit antavat samansuuntaisen tuoton toteutuneeseen verrattuna, mutta tuotot ovat toteutuneita paljon suurempia. Tämän validoinnin perusteella joudutaan mallin ennustuskyky ja toimivuus kyseenalaistamaan.



Kuva 13 Ennuste vuodelle 2006 plotattu yhtenäisellä viivalla, toteutuneet arvot katkoviivalla.



Kuva 14 Ennuste vuodelle 2006 plotattu yhtenäisellä viivalla, toteutuneet arvot katkoviivalla.



7. Lineaarinen malli AR(1) virhetermillä

Aiemmin todettiin että päivittäiseen dataan sovitetun lineaarisen regressiomallin residuaaleissa ilmenee huomattavasti autokorrelaatiota. Tämän perusteella päätettiin kokeilla, mikäli lineaarinen malli aikasarjavirheellä antaisi parempia tuloksia. Päädyttiin kokeilemaan lineaarista mallia AR(1) virhetermillä (esim. Hamilton 1994, s. 223-4). Malli on muotoa:

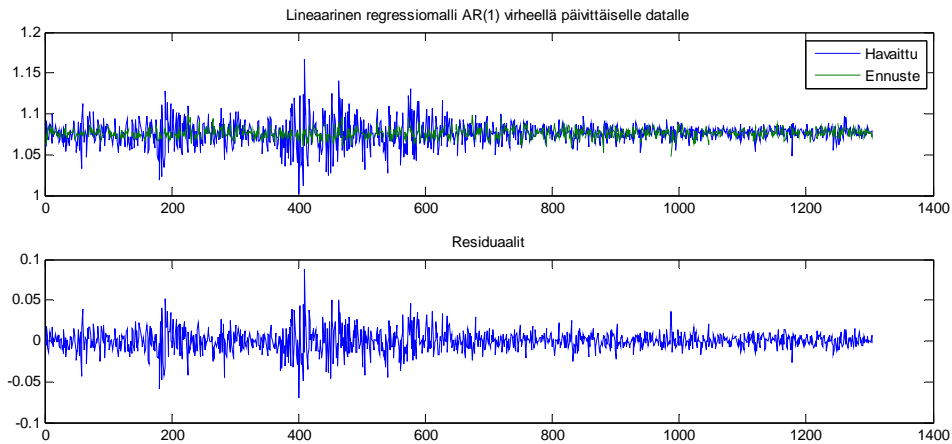
$$y_t = \alpha + \beta x_t + u_t, u_t = \rho u_{t-1} + \varepsilon_t$$

Missä virhetermi u_t siis on ensimmäisen asteen autoregressiivinen malli. Kirjoittamalla tämä matriisimuotoon $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$ voidaan malli estimoida maksimoimalla log likelihoodfunktiota $-\left[\frac{T-1}{2}\right] \log(2\pi) - \left[\frac{T-1}{2}\right] \log(\sigma^2) - \left[\frac{1}{2\sigma^2}\right] \sum_{t=2}^T [(y_t - \mathbf{x}_t^T \boldsymbol{\beta}) - \rho(y_{t-1} - \mathbf{x}_{t-1}^T \boldsymbol{\beta})]^2$. Tämä voidaan ratkaista iteratiivisesti käyttämällä ns. Cochrane-Orcutt menetelmää: havaitaan että jos ρ tunnetaisiin (vakiot eivät tietenkään vaikuta maksimointiin), voitaisiin $\boldsymbol{\beta}$ ratkaista regressoimalla $(y_t - \rho y_{t-1})$ pienimmän neliösumman menetelmällä $(\mathbf{x}_t - \rho \mathbf{x}_{t-1})$ suhteen kun $t = 2, 3, \dots, T$. Vastaavasti, jos $\boldsymbol{\beta}$ tunnetaisiin, voitaisiin ratkaista ρ regressoimalla $(y_t - \mathbf{x}_t^T \boldsymbol{\beta})$ pienimmän neliösumman menetelmällä $(y_{t-1} - \mathbf{x}_{t-1}^T \boldsymbol{\beta})$ suhteen kun $t = 2, 3, \dots, T$. Estimointi voidaan siis suorittaa valitsemalla alkuarvaukseksi $\rho = 0$ ja ratkaista $\boldsymbol{\beta}$ normaalisti pienimmän neliösumman menetelmällä, jonka jälkeen ρ :n arvo voidaan päivittää jälkimmäisen regression avulla. Näin voidaan jatkaa kunnes haluttu tarkkuus on saavutettu.

Käytännössä estimointi suoritettiin soveltamalla aiemmin esiteltyä askeltavaa regressiomenetelmää faktoreiden valitsemiseen ja $\boldsymbol{\beta}$:n alkuarvauksen saamiseksi. Tämä syötettiin Matlabin Econometrics Toolboxin (LeSage, 2005) olsc-funktioon, joka suoritti Cochrane-Orcutt iteraation. Analyysi suoritettiin ensin päivittäiselle datalle 71 fundille ajalla 1.1.2001-31.12.2005, jonka jälkeen analyysi uusittiin käyttämällä viikottaisia aggregaatteja.

7.1 Päivittäinen data

Tuloksista kävi ilmi, että AR(1) virhetermin lisääminen malliin ei juuri parantanut mallia päivätiedon suhteen. Parhaimmillaan selitysasteeksi saatiin 0.3299, huonoimmillaan selitysaste oli vain 0.0343 ja keskimäärin 0.1462 (ilman AR(1) virhetermiä 0.2921, 0.0279 ja 0.1455). Residuaalien autokorrelaatio koitui edelleenkin ongelmaksi: kahta fundia lukuunottamatta Ljung-Box Q-testi autokorrelaatiolle johti nollahypoteesin hylkäämiseen suurimmalle osalle viiveistä. Huomattavaa on lisäksi että päivittäiselle datalle askeltava regressioalgoritmi sisältää faktorit malliin huomattavasti herkemmin kuin kuukausittaiselle datalle – samalla selitysasteet jäävät paljon alhaisemmiksi, joten tämä ei tuo lisäarvoa faktoreiden etsinnälle.

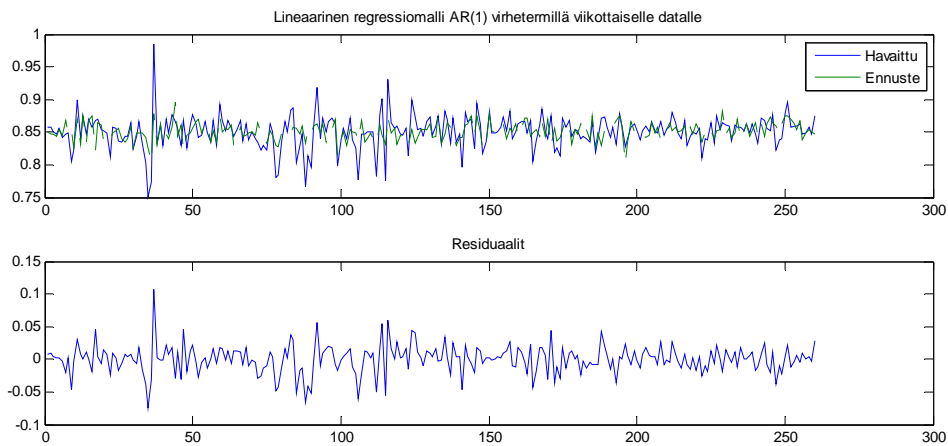


Kuva 15 esimerkki lineaarisen regressiomallin AR(1) virheellä soveltamisesta päivätataan. Malli ei nähtävästi pärjää kovinkaan hyvin (selitysaste 0.14).

7.2 Viikottainen data

Kuten mainittu, päivittäisellä datalla residuaalien autokorrelaatio AR(1) termi huomioiden oli edelleen liiallista. Tästä syystä päätettiin kokeilla mallia viikottaisella aggregoinnilla. Tässä tapauksessa voitiin todeta, että normaalille lineaariselle regressiolle residuaalien autokorrelaatio oli edelleen liiallista, mutta että AR(1)-virhetermin lisääminen vähensi sitä huomattavasti: ilman AR(1)-termiä yli puolella fundien residuaaleista oli autokorrelaatiota siinä määrin että Ljung-Box Q-testi hylkäsi nollahypoteesin korreloimattomuudesta, kun taas AR(1) termi huomioiden autokorrelaatio katosi lähes täysin.

Vaikka residuaalit käyttäytyivät hyvin, ei malli kuitenkaan antanut suuria selitysasteita: selitysasteet olivat keskiarvoltaan 0.2216, maksimi 0.2808 ja minimi 0.0836 (ilman AR(1) termiä vastaavasti 0.2181, 0.2774, 0.0837). Tulokset viittaavat siis siihen, että viikottaisella datalla satunnainen vaihtelu on edelleen niin suuri, että sitä ei yksinkertaisella lineaarisella mallilla saada kuvattua. Vastaavasti tämä malli ei juuri anna lisäarvoa faktoreiden etsintää ajatellen: selitysasteet jäävät alhaisiksi ja ilmenevät faktorit ovat lähinnä samoja kuten aiemmin: kupari-indeksi (66 kertaa), OECD Total Money Aggregates (63 kertaa) mittari ja 3 kuukauden Euribor korko (35 kertaa).



Kuva 16 esimerkki lineaarisen regressiomallin ennusteesta AR(1) virhetermillä. Nähdään, että malli pärjää huonosti (selityksaste 0.2802).

Yhteenvetona voidaan todeta, että tässä tapauksessa AR(1) virhetermin lisääminen lineaariseen malliin ei juuri tuo lisäarvoa, joten tällaisen mallin implikaatiot jätetään vähemmälle huomiolle.

8. Pääkomponenttianalyysi

Pääkomponenttianalyysin avulla pyritään löytämään jonkin datajoukon takana piilevät faktorit. Kyseessä on menetelmä dimensioiden redusoinniksi, joka perustuu siihen että havaittavissa olevien muuttujien välillä on huomattavasti korrelaatiota. Toisin sanoen, oletetaan että havaittavissa olevat muuttujat (osakerahastojen arvo/tuotto) riippuvat pienestä määrästä riippumattomia faktoreita, jotka generoivat havaittavan datan, ja joita lähdetään etsimään maksimointitehtävän avulla. Kyseistä menetelmää ollaan käytetty m.m. fundien tyylifaktoreiden löytämiseen (Fung & Hsieh, 1998), tai esimerkiksi osakkeiden toimialakohtaisten faktoreiden löytämiseen (Tsay, 2002). Analyysi itsessään ei ota kantaa siihen, mitä löydettyt faktorit ovat, vaan tämä jää analyysin suorittajan pohdittavaksi. Menetelmä muistuttaa paljon faktorianalyysiä ja antaa useimmissa tapauksissa samoja tuloksia, mutta voidaan suorittaa yleisemmin (ei vaadi esimerkiksi kovarianssimatriisin kääntymistä tai aikasarjan heikkoa stationarisuutta kuten faktorianalyysi).

Merkitään fundien muodostamaa satunnaisvektoria $\mathbf{r} = (r_1, r_2, \dots, r_k)^T$ - jokainen havainto on siis fundia vastaavan tuottovektorin r_i jokin alkio. Pääkomponenttianalyysin tarkoituksena on siis löytää havaittujen alkioiden lineaarikombinaatio $\boldsymbol{\beta}^T \mathbf{r}$ jonka varianssi on suurin mahdollinen. Pääkomponentit ovat tällöin vektoreita $\mathbf{y}_i = \boldsymbol{\beta}_i^T \mathbf{r}$, jotka voidaan ratkaista maksimointitehtävästä

$$\max \text{Var}(\boldsymbol{\beta}^T \mathbf{r}), \quad \text{s. t. } \boldsymbol{\beta}^T \boldsymbol{\beta} = 1$$

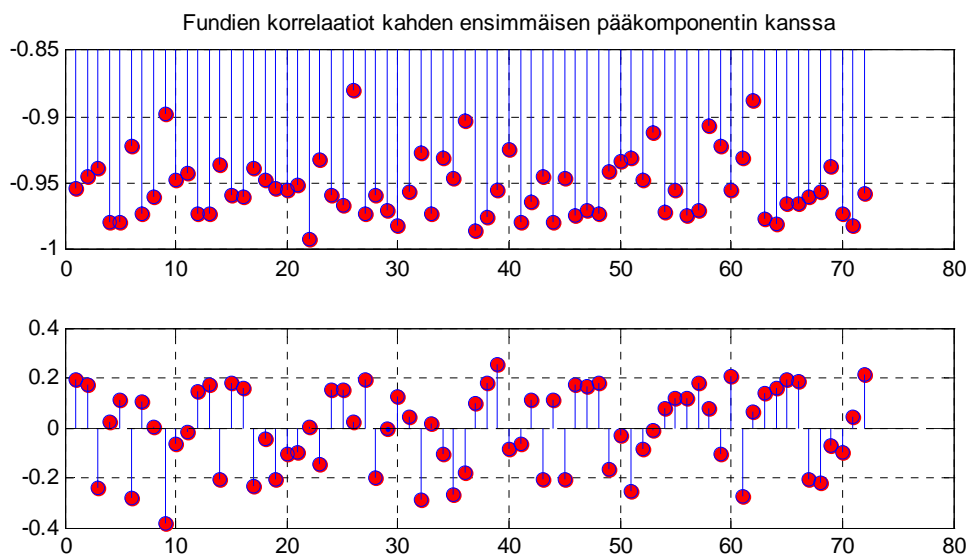
missä ehto normeeraa painovektorit $\boldsymbol{\beta}$ ykkösen mittaisiksi. Lisäksi käytetään ehtoa että pääkomponenttien kovarianssit toistensa kanssa ovat nolla.

Osoittautuu, että varianssin maksimoiva painovektori saadaan kovarianssimatriisin $\boldsymbol{\Sigma}$ pääakselihajotelmasta $\boldsymbol{\Sigma} = \mathbf{B}^T \boldsymbol{\Lambda} \mathbf{B}$, missä \mathbf{B} on kovarianssimatriisin ominaisvektoreista koostuva matriisi ja $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_k)$ vastaavat ominaisarvot. Järjestetään matriisi \mathbf{B} uudestaan: $\mathbf{B} = [\boldsymbol{\beta}_1 : \boldsymbol{\beta}_2 : \dots : \boldsymbol{\beta}_k]$, missä $\boldsymbol{\beta}_1$ on suurinta ominaisarvoa vastaava ominaisvektori, $\boldsymbol{\beta}_2$ on toiseksi suurinta ominaisarvoa vastaava ominaisvektori jne. Täten haetut pääkomponentit saadaan siten, että ensimmäinen pääkomponentti, joka selittää suurimman osan varianssista, $\mathbf{y}_1 = \boldsymbol{\beta}_1^T \mathbf{r}$, toinen $\mathbf{y}_2 = \boldsymbol{\beta}_2^T \mathbf{r}$ jne (Mellin, 2007).

Käytännössä pääkomponenttianalyysi kannattaa suorittaa joko jollain tilasto-ohjelmistolla tai Matlabilla. Näiden tärkein ulostulema on niin sanotut faktorilataukset, eli alkuperäisten muuttujien korrelaatio pääkomponenttien kanssa. Lisäksi voidaan kovarianssimatriisin ominaisarvojen avulla päätellä kuinka suuren osan havaintoaineiston vaihtelusta kukin pääkomponentti selittää.

8.1 Tulokset

Pääkomponenttianalyysi suoritettiin 71 fundille kuukausittaisella tuottodatalla ajalta 1/2001-12/2005. Saatujen ominaisarvojen perusteella ensimmäinen pääkomponentti selittää noin 90.5% havaintojen varianssista, toinen noin 2.7% ja kolmas noin 1.5%. Faktorilataukset (alkuperäisten muuttujien korrelaatio pääkomponenttien suhteen) ensimmäisen pääkomponentin suhteen olivat keskiarvoltaan -0.95 standardipoikkeamalla 0.024. Suurimmalla osalla rahastoista korrelaatio oli suunnilleen tätä luokkaa, muutamaa poikkeusta lukuunottamatta: Fidelity FDS-FID SELE joka korreloi paljon muita voimakkaimmin (yli 0.99), sekä Fidelity FDS-European AGGR-A, Credit Suisse EQ-ST INV EU-B ja UBS LUX EQTY-European GTH-B jotka korreloivat huomattavasti vähemmän ensimmäisen pääkomponentin kanssa.



Kuva 17 fundien korrelaatio kahden ensimmäisen pääkomponentin kanssa. Ylempi kuva on korrelaatio ensimmäisen pääkomponentin suhteen, alempi on korrelaatio toisen pääkomponentin suhteen. Eri osakerahastot näkyvät x-akselilla, kun taas y-akseli vastaa näiden korrelaatiota.

Niin kuin yleensä pääkomponenttianalyysissä ja faktorianalyysissä, ensimmäinen pääkomponentti vaikuttaisi olevan jonkinlainen yleisfaktori, eli tässä tapauksessa markkinafaktori. Kyseinen pääkomponentti selittää fundien kehityksestä ylivoimaisesti suurimman osan, mikä ei sinällään ole kovin yllättävää kun kyseessä on erittäin samantapaisia rahastoja. Tulokselle voidaan löytää muutama tulkinta: ensinnäkin voidaan juuri todeta että fundit ovat erittäin samantapaisia (tämä on myös selvää korrelaatiomatriisista, jossa pienin fundien välinen korrelaatio on 0.74), toisaalta voidaan myös todeta että ne ovat onnistuneet diversifioinnissa koska vaihtelu (ja riski) koostuu lähinnä markkinoista. Tämä ei kuitenkaan tuo faktoreiden etsinnälle lisäarvoa. Poikkeavat havainnot eivät nekään valaise tilannetta; UBS LUX rahastoa lukuunottamatta, joka on kaikista huonoiten tuottanut rahasto, ne ovat pärjänneet keskinkertaisesti.

Toinen pääkomponentti vaikuttaa vuorostaan jakavan havainnot kahtia ja sille saattaakin löytyä käypä ja järkevä tulinta. Kannattaa kuitenkin huomata, että vaikka esitämme faktorille tulkinnan, tulkinta on erittäin subjektiivinen. Lisäksi, koska komponentti vastaa vain 2.7% kokonaisvaihtelusta, kyseessä saattaa tulkinnan kannalta olla vain sattumaa. Havaitaan kuitenkin, että rahastoista osa korreloi positiiviset ja osa suunnilleen yhtä paljon negatiivisesti tämän pääkomponentin kanssa. Jos fundit ryhmitellään siten, että ne joiden korrelaatio toisen pääkomponentin kanssa on yli 0.1 kuuluvat yhteen ryhmään ja ne joiden korrelaatio on alle -0.1 toiseen, havaitaan että jälkimmäiseen ryhmään kuuluvat tuottivat keskimäärin noin 16 % paremmin tarkasteluajanjaksona – ero oli tilastollisesti merkittävä vielä 99% luottamustasolla². Lisäksi

² Suoritettiin Milton ja Arnold (Milton & Arnold, 2003) s. 347-8 mukainen T-testi kahden keskiarvon samankaltaisuudelle testisuurella $\frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$, missä $\bar{X}_1 = 0.833, S_1 = 0.098, n_1 = 27$ ja $\bar{X}_2 = 0.97, S_2 =$

tarkastelemalla mitä rahastoja kyseessä oli, todettiin että paremmin tuottavaan ryhmään kuului lähinnä pieniin ja keskisuuriin yrityksiin sijoittavat (datasta löytyi näitä, vaikka kyseessä periaatteessa on Large-Cap rahastoja) ja kasvu-rahastot. Vaikuttaisi siis siltä, että kyseinen faktori jakaa rahastot kasvu ja arvoraheistoihin – tosin poikkeuksia löytyi, joten tulkinta ei välttämättä ole sataprosenttisesti vedenpitävä.

8.2 Regressio pääkomponenttien suhteen

Pääkomponenttianalyysin tuloksena saatiin uudet muuttujat \mathbf{y}_i , jotka periaatteessa sisältävät suurimman mahdollisen määrän informaatiota datajoukosta. Mielenkiintoista on kuitenkin riippuko kaikki alkuperäinen data suunnilleen yhtä paljon näistä pääkomponenteista. Tämän selvittämiseksi, suoritimme regression alkuperäiselle datalle kolmen ensimmäisen pääkomponentin suhteen. Tällainen lähtökohta tuottaa tietenkin suuria selitysasteita, mutta kyseessä onkin lähinnä kätevä menetelmä löytää poikkeavia havaintoja. Tarkastelemme lähinnä rahastojen selitysasteita, selvittääksemme mitkä fundeista käyttäytyvä huomattavasti eri tavalla kuin muut. Malli on tällöin muotoa $\mathbf{X} = \boldsymbol{\mu} + \boldsymbol{\Gamma}\mathbf{Y}$, missä \mathbf{Y} on pääkomponenttien muodostama matriisi $\boldsymbol{\Gamma}$ kerroinmatriisi näiden suhteen.

Voimme todeta, että R^2 -arvot ovat keskimäärin luokkaa 0.944 standardipoikkeamalla 0.04. Tämä on sama kuin kolmen ensimmäisen pääkomponentin kumulatiivinen osuus havaintojen vaihtelusta kuten pitääkin. Joukosta löytyy kuitenkin huomattavasti poikkeavia havaintoja: pienimmät selitysasteet on Gartmore-PAN European FD-A (0.77) ja SEB Europe 2 FD (0.83) rahastoilla. Nämä eivät kuitenkaan osoita mitään poikkeavaa performanssia, vaan ne ovat pärjänneet keskimuotoisesti.

Yhteenvetona voi sanoa, ettei pääkomponenttianalyysi anna mitään hyödyllisiä tai yllättäviä tuloksia – menetelmä vaatisi jonkin asialoogisen syyn miksi havaintojen tulisi poiketa toisistaan, mitä ei oikein Large-Cap osakerahastojen tapauksessa toteudu.

0.21, $n_2 = 21$. Nollahypoteesin pätiessä tämän tulisi noudattaa T-jakaumaa vapausasteilla $\gamma \doteq \frac{\left[\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right]^2}{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$

(Smith-Satterthwaite vapausasteet). Testisuureksi saatiin 2.72 ja P-arvoksi yksisuuntaiselle testisuurelle ($H_0: \mu_0 = \mu_1$ ja $H_1: \mu_0 > \mu_1$) 0.005.

9. Vektoriautoregressiomalli

Osakkeiden volatiliiteetin ja osittain myös tuottojen mallintamiseen on käytetty erilaisia vektoriautoregressio-menetelmiä (Tsay 2002). Mallien toimivuuden perusteena on se, että markkinoilla toimivat agentit sijoittavat lähes poikkeuksetta useampaan kohteeseen, joten yksittäisen sijoituskohteen autoregression looginen jatke on yleistäminen autoregressio kaikille sijoituskohteille samanaikaisesti. Erityisesti sijoituskohteiden volatiliiteetin on todettu leviävän markkinoilla sijoituskohteiden välillä (Hamilton, Stockmarket volatility and the business cycle, 1996). Samoin voisi kuvitella, että tuotot eri sijoituskohteiden välillä saattaisivat ”tartuttaa” toisiaan. Tämän innoittamana kokeilemme vektoriautoregressiota rahastojen tuottojen selittämiseen.

9.1 VAR-prosessi

Tutkimme seuraavaksi aikasarjamenetelmiä vektoriarvoisille aikasarjoille. Vektori autoregressio (VAR) on aikasarjamenetelmä jota käytetään systeemien mallintamiseen kun systeemin muuttujat riippuvat systeemin muiden muuttujien viipeistä omiensa viipeidensä lisäksi. Yksinkertainen esimerkki on seuraava kahden muuttujan systeemi. Ensimmäinen muuttuja riippuu ensimmäisestä viipeestään sekä toisen muuttujan ensimmäisestä viipeestä, ja toinen muuttuja seuraa AR(2) prosessia:

$$\begin{aligned}x_{1t} &= \phi_{11}x_{1,t-1} + \phi_{12}x_{2,t-1} + \varepsilon_{1t} \\x_{2t} &= \phi_{21}x_{2,t-1} + \phi_{22}x_{2,t-2} + \varepsilon_{2t}\end{aligned}$$

missä $E(\varepsilon_{1t}\varepsilon_{2s}) = \sigma_{12}$ kun $t = s$ ja $E(\varepsilon_{1t}\varepsilon_{2s}) = 0$ kun $t \neq s$. Systeemi voidaan kirjoittaa muotoon

$$\begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ 0 & \phi_{21} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & \phi_{22} \end{pmatrix} \begin{pmatrix} x_{1,t-2} \\ x_{2,t-2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix},$$

eli matriisimuodossa

$$\mathbf{x}_t = \mathbf{\Phi}_1 \mathbf{x}_{t-1} + \mathbf{\Phi}_2 \mathbf{x}_{t-2} + \boldsymbol{\varepsilon}_t$$

missä

$$E(\boldsymbol{\varepsilon}_t) = \mathbf{0}, \quad E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_s^T) = \mathbf{0} \quad \text{kun } s \neq t \quad (1) \text{ ja}$$

$$E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t^T) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}.$$

Yleinen VAR(p) prosessi n -vektorille \mathbf{x}_t , jonka residuaalit ovat valkoista kohinaa, voidaan täten kirjoittaa muotoon

$$\begin{aligned}\mathbf{x}_t &= \mathbf{c} + \Phi_1 \mathbf{x}_{t-1} + \Phi_2 \mathbf{x}_{t-2} + \dots + \Phi_p \mathbf{x}_{t-p} + \boldsymbol{\varepsilon}_t \\ &= \sum_{i=1}^p \Phi_i \mathbf{x}_{t-i} + \boldsymbol{\varepsilon}_t\end{aligned}\quad (2)$$

missä $\mathbf{c} \in \mathbb{R}^{n \times 1}$ on vakiovektori ja $\Phi_i \in \mathbb{R}^{n \times n}$.

9.2 VARX prosessi

Lisäämällä VAR prosessiin riippuvuus ulkoisista muuttujista saadaan VARX malli (Vector Auto Regression with eXternal variable):

$$\begin{aligned}\mathbf{x}_t &= \mathbf{c} + \Phi_1 \mathbf{x}_{t-1} + \Phi_2 \mathbf{x}_{t-2} + \dots + \Phi_p \mathbf{x}_{t-p} + \Psi \mathbf{z} + \boldsymbol{\varepsilon}_t \\ &= \sum_{i=1}^p \Phi_i \mathbf{x}_{t-i} + \Psi \mathbf{z} + \boldsymbol{\varepsilon}_t\end{aligned}$$

missä vektori $\mathbf{z} \in \mathbb{R}^{k \times 1}$ sisältää k ulkoista muuttujaa ja $\Psi \in \mathbb{R}^{n \times k}$.

9.3 VAR(p) prosessin estimointi

Olkoot meillä $(T+p)$ havaintoa tutkimastamme vektori-prosessista $\mathbf{x}_t \in \mathbb{R}^{n \times 1}$ (\mathbf{x}_t siis vektori, joka sisältää havainnot n :stä tutkimastamme muuttujasta hetkellä t). Oletamme, että \mathbf{x}_t seuraa VAR(p) prosessia (2), jonka residuaalivektori on $\boldsymbol{\varepsilon}_t \sim i.i.d \mathbf{N}(\mathbf{0}, \Omega)$.

Ehdollistamalla T tuoreinta havaintoa (merkitään $\mathbf{x}_1, \mathbf{x}_{T-1}, \dots, \mathbf{x}_T$) p ensimmäiselle havainnolle (merkitään $\mathbf{x}_{-p+1}, \mathbf{x}_{-p+2}, \dots, \mathbf{x}_0$) voidaan muodostaa ehdollinen likelihood-funktio

$$f_{\mathbf{x}_T, \mathbf{x}_{T-1}, \dots, \mathbf{x}_1 | \mathbf{x}_0, \mathbf{x}_{-1}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta}}$$

missä $\boldsymbol{\theta}$ on parametrivektori, joka sisältää VAR prosessimme parametrit $\mathbf{c}, \Phi_1, \Phi_2, \dots, \Phi_p$ sekä kovarianssimatriisin Ω . Maksimoimalla ehdollista likelihood-funktiota parametrivektorin $\boldsymbol{\theta}$ suhteen saadaan täten VAR prosessin parametrit.

Kirjoitetaan VAR(p) prosessimme

$$\mathbf{x}_t = \Phi_1 \mathbf{x}_{t-1} + \Phi_2 \mathbf{x}_{t-2} + \dots + \Phi_p \mathbf{x}_{t-p} + \boldsymbol{\varepsilon}_t$$

hieman kompaktimpaan muotoon

$$\mathbf{x}_t = \Pi \mathbf{y}_t + \boldsymbol{\varepsilon}_t$$

missä $\Pi = (\mathbf{c}^T \quad \Phi_1^T \quad \Phi_2^T \quad \dots \quad \Phi_p^T)$ ja $\mathbf{y}_t = (1 \quad \mathbf{x}_{t-1} \quad \mathbf{x}_{t-2} \quad \dots \quad \mathbf{x}_{t-p})^T$.

Koska virhetermimme $\boldsymbol{\varepsilon}_t \sim i.i.d \mathbf{N}(\mathbf{0}, \Omega)$, on ehdollinen odotusarvo

$$E(\mathbf{x}_t | \mathbf{x}_{t-1}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta}) = \mathbf{\Pi} \mathbf{y}_t,$$

joten

$$\mathbf{x}_t | \mathbf{x}_{t-1}, \dots, \mathbf{x}_{-p+1} \sim N(\mathbf{\Pi} \mathbf{y}_t, \boldsymbol{\Omega}).$$

Täten ehdollinen tiheysjakauma t :n nelle havainnolle on

$$f_{\mathbf{x}_t | \mathbf{x}_{t-1}, \dots, \mathbf{x}_{-p+1}}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta}) = (2\pi)^{-n/2} |\boldsymbol{\Omega}^{-1}|^{1/2} \exp\left[-(1/2)(\mathbf{x}_t - \mathbf{\Pi} \mathbf{y}_t)^T (\boldsymbol{\Omega}^{-1})(\mathbf{x}_t - \mathbf{\Pi} \mathbf{y}_t)\right]$$

Aikaisempiin havaintoihin ehdollistettu yhteisjakauma havainnoille ajanhetkestä 1 ajanhetkeen t toteuttaa

$$f_{\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_0, \mathbf{x}_{-1}, \dots, \mathbf{x}_{-p+1}}(\mathbf{x}_t, \mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_{-1}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta}) = f_{\mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_{-1}, \dots, \mathbf{x}_{-p+1}}(\mathbf{x}_{t-1}, \dots, \mathbf{x}_1 | \mathbf{x}_{-1}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta}) * f_{\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{-p+1}}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta})$$

Soveltamalla tätä rekursiivisesti koko otokseen $\mathbf{x}_T, \mathbf{x}_{T-1}, \dots, \mathbf{x}_1$ ja logaritmoimalla saadaan log-likelihood-funktio

$$L(\boldsymbol{\theta}) = \sum_{t=1}^T \log[f_{\mathbf{x}_t | \mathbf{x}_{t-1}, \dots, \mathbf{x}_{-p+1}}(\mathbf{x}_t | \mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \dots, \mathbf{x}_{-p+1}; \boldsymbol{\theta})] = (Tn/2) \log(2\pi) + (T/2) \log |\boldsymbol{\Omega}^{-1}| - (1/2) \sum_{t=1}^T [(\mathbf{x}_t - \mathbf{\Pi} \mathbf{y}_t)^T (\boldsymbol{\Omega}^{-1})(\mathbf{x}_t - \mathbf{\Pi} \mathbf{y}_t)]$$

Osoittautuu, että maximum likelihood estimaattori (MLE) parametrivektorille $\mathbf{\Pi}$ saadaan muodossa

$$\hat{\mathbf{\Pi}} = \left[\sum_{t=1}^T \mathbf{x}_t \mathbf{y}_t^T \right] \left[\sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^T \right]^{-1}.$$

Täten k :nnes rivi on muotoa

$$\hat{\boldsymbol{\pi}}_k = \left[\sum_{t=1}^T x_{kt} \mathbf{y}_t^T \right] \left[\sum_{t=1}^T \mathbf{y}_t \mathbf{y}_t^T \right]^{-1}$$

mikä on yksinkertaisesti PNS regression kertoimet selitettäessä x_{kt} :tä \mathbf{y}_t :llä. Täten maximum likelihood estimaatit k :n nelle yhtälölle VAR prosessissa saadaan yksinkertaisesti PNS regressiolla jossa yksittäistä selitettävää muuttujaa x_{kt} selitetään vakiolla ja p :llä viipeellä kaikista systeemin muuttujista. Voidaan osoittaa, että tulos pätee myös, vaikkei oikeat residuaalivektorit $\boldsymbol{\varepsilon}_t$ ole normaalijakautuneita, kuten edellä oletettiin. Ehdon (1) tulee kuitenkin päteä.

9.4 Rahastojen tuotot VAR-prosessina

Tutkitaan seuraavaksi rahastojen tuottojen keskeistä dynamiikkaa sovittamalla niihin VAR malleja. Ottamalla n :n tuottoaikasarjan otos kaikista tuottoaikasarjoista ja sovittamalla siihen VAR malli voimme siis tutkia valittujen n :n rahastojen välistä dynaamista hinnanmuodostusta.

VAR mallien sovittamiseksi käytämme Econometrics toolbox:in Matlab scripteistä kehitettyjä työkaluja. VAR työkalumme sovittaa PNS sovituksen avulla jokaisen valitun rahaston tuoton omien tuottojensa viipeillä p :nteen viipeeseen asti sekä kaikkien muiden tuottovektoriin valittujen rahastojen tuottojen viipeillä:

$$\begin{pmatrix} r_{1,t} \\ \vdots \\ r_{n,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + \begin{pmatrix} \Phi_{1,11} & \cdots & \Phi_{1,1n} \\ \vdots & \ddots & \vdots \\ \Phi_{1,n1} & \cdots & \Phi_{1,nn} \end{pmatrix} \begin{pmatrix} r_{1,t-1} \\ \vdots \\ r_{n,t-1} \end{pmatrix} + \dots + \begin{pmatrix} \Phi_{p,11} & \cdots & \Phi_{p,1n} \\ \vdots & \ddots & \vdots \\ \Phi_{p,n1} & \cdots & \Phi_{p,nn} \end{pmatrix} \begin{pmatrix} r_{1,t-p} \\ \vdots \\ r_{n,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (3)$$

Tuloksena on täten jokaiselle rahaston tuotolle lineaarinen regressiomalli vakiotermeillä ja pn :llä selittäjällä, jotka ovat tuottovektorin rahastojen viivästettyjä tuottoja. Mallien laajuudesta selvinnee, että suuri osa selittäjistä ei tule olemaan tilastollisesti merkitseviä. Tämän takia työkalumme voi myös soveltaa ”askeltavaa regressiota”, estimoimalla ensin täysilajajuinen malli (3), jonka jälkeen yksi kerrallaan tilastollisesti merkityksettömät muuttujat tiputetaan mallista pois ja uusi malli estimoidaan. Täten lopputuloksena on malli, jossa tilastollisesti merkityksettömät parametrit ovat nolliä. Työkalua voidaan täten käyttää itse mallin estimoimisen lisäksi myös tietyn rahastojoukon tuotonmuutokseen tilastollisesti merkitsevien muuttujien (historiallisten tuottojen) tunnistamiseen.

Työkalumme venyy puhtaan VAR mallin estimoimista pidemmälle mahdollistaen ulkoisten muuttujien lisäämisen, mahdollisesti myös viivästettyinä. Tuloksena on täten VARX – malli

$$\begin{pmatrix} r_{1,t} \\ \vdots \\ r_{n,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} + \begin{pmatrix} \Phi_{1,11} & \cdots & \Phi_{1,1n} \\ \vdots & \ddots & \vdots \\ \Phi_{1,n1} & \cdots & \Phi_{1,nn} \end{pmatrix} \begin{pmatrix} r_{1,t-1} \\ \vdots \\ r_{n,t-1} \end{pmatrix} + \dots + \begin{pmatrix} \Phi_{p,11} & \cdots & \Phi_{p,1n} \\ \vdots & \ddots & \vdots \\ \Phi_{p,n1} & \cdots & \Phi_{p,nn} \end{pmatrix} \begin{pmatrix} r_{1,t-p} \\ \vdots \\ r_{n,t-p} \end{pmatrix} + \\ + \begin{pmatrix} \mathbf{X}_{0,11} & \cdots & \mathbf{X}_{0,1n} \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{0,k1} & \cdots & \mathbf{X}_{0,kn} \end{pmatrix} \begin{pmatrix} x_{1,t} \\ \vdots \\ x_{k,t} \end{pmatrix} + \begin{pmatrix} \mathbf{X}_{1,11} & \cdots & \mathbf{X}_{1,1n} \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{1,k1} & \cdots & \mathbf{X}_{1,kn} \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ \vdots \\ x_{k,t-1} \end{pmatrix} + \dots + \begin{pmatrix} \mathbf{X}_{l,11} & \cdots & \mathbf{X}_{l,1n} \\ \vdots & \ddots & \vdots \\ \mathbf{X}_{l,k1} & \cdots & \mathbf{X}_{l,kn} \end{pmatrix} \begin{pmatrix} x_{1,t-l} \\ \vdots \\ x_{k,t-l} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$$

missä on VAR malliin lisätty k ulkoista muuttujaa, sekä l näiden viivettä. Tämäkin malli voidaan estimoida PNS regressiolla yhtälöittäin, ja halutessa työkalumme tekee edellä mainitun ”askeltavan regression” myös VARX mallille.

9.5 Päivätuottojen ennustaminen VAR(1) -mallilla

Käytämme seuraavaksi VAR työkaluamme tutkiaksemme päivätuottojen sisäistä dynamiikkaa eri rahastoilla. Aloitamme estimoimalla VAR mallin työkalullamme, jossa selitettäväksi tuottovektoriksi otetaan kaikki 72 rahastoa, ja mallin viipeeksi valitaan 1. Estimoimme siis VAR(1) mallin kaikkien rahastojen tuottojen muodostamalle vektoriaikasarjalle. Askeltavan regression

kynnysarvona muuttujan merkitsevyyden t-testissä käytetään 1% jotta valittujen selittäjien määrä olisi pieni.

Askeltava regressio on työkaluna erittäin hyvä laajojen mallien estimoimisessa. Se tuottaa kuitenkin tiettyjä ongelmia tulosten tulkitsemisen kohdalla, sillä rahastojen tuottojen vahvojen korrelaatioiden takia lopullisesti malliin valitut selittäjät muuttuu riippuen siitä, missä järjestyksessä ei-merkitsevät muuttujat tiputetaan pois mallista. Pitäen tämä mielessä estimoidaan seuraavaksi malli, jossa rahastojen tuotot syötetään kasvavassa järjestyksessä indeksiansä mukaan. Työkalumme on ohjelmoitu siten, että tällöin ”pieni-indeksiset” rahastot tulevat herkemmin valituksi malliimme kun ”suuri-indeksiset”.

Suurimmalle osalle rahastoista saatujen mallien selitysaste on 0.01 ja 0.3 välillä, ja kaikkien mallien vakiotermi on 99% luottamustasolla nolla. Jokaiselle rahastolle löytyy kuitenkin vähintään yksi oman tai toisen rahaston historiallinen tuotto, joka 99% luottamustasolla on tilastollisesti merkitsevä selittäjä nykyiselle tuotolle³. Yllätyksiä kuitenkin löytyy. 72 rahastosta kuudelle löydetään mallit, joiden selitysasteet ylittävät 50%. Näistä kolmen selitysasteet ovat yli 0.8, ja yhden yli 0.9.

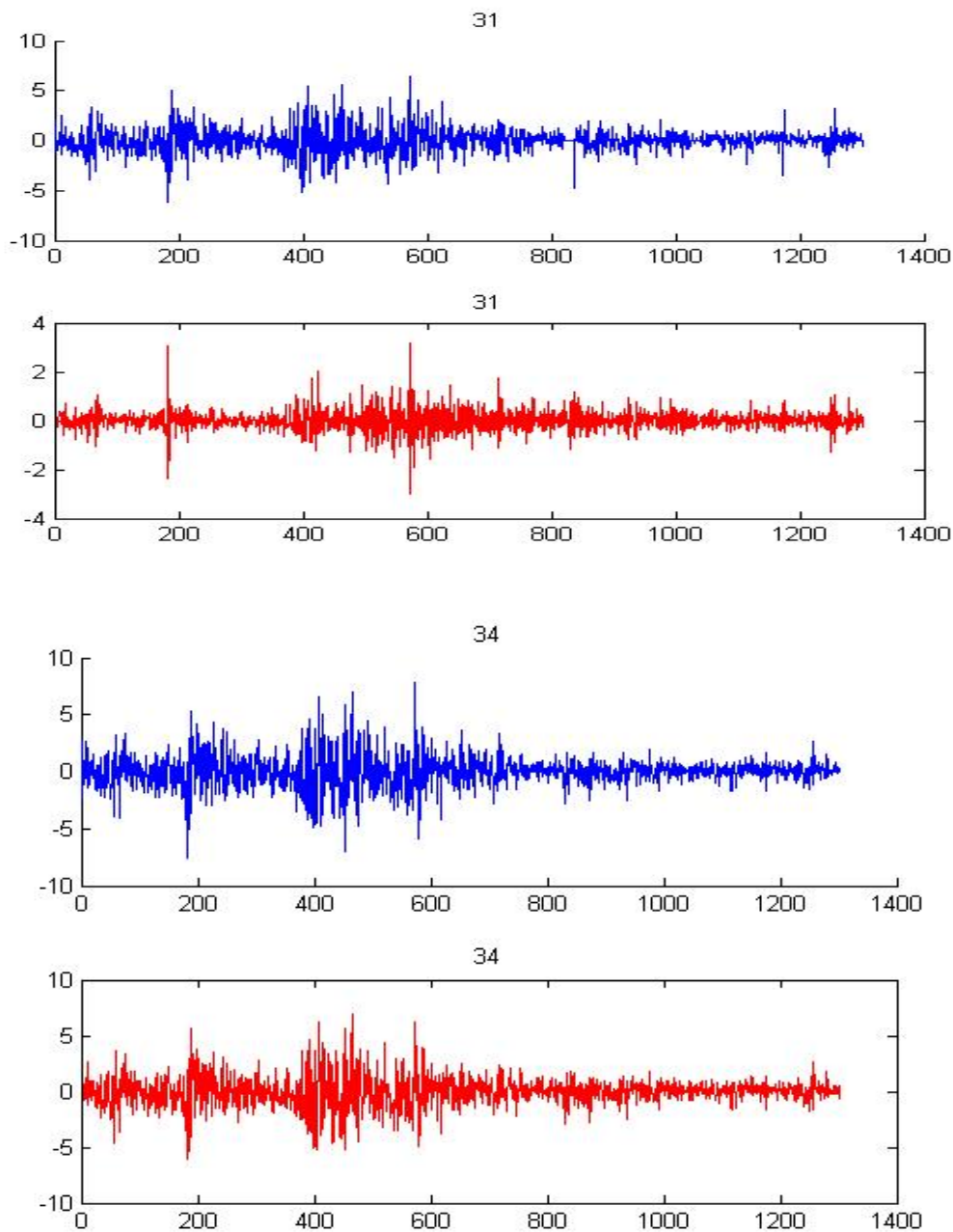
Kuvassa Kuva 18 on esitetty kahden rahaston toteutuneet päivätuotot periodilta 1.1.2001 – 31.12.2005, sekä VAR mallimme antamat sovitteet.

Tutkimme seuraavaksi mallimme residuaaleja. Ensinnäkin kuvasta Kuva 18 näemme, että tuotot ovat selvästi heteroskedastisia, joten residuaalimme tulevat todennäköisesti olemaan heteroskedastisia. Sovellamme Englen hypoteesitestiä residuaaleillemme Matlabin archtest-funktion avulla. Kuten odotettu nollahypoteesi residuaalien homoskedastisuudesta tulee hylättyä jokaiselle mallille.

Tutkimme myös residuaalien autokorreloituneisuutta. Sovellamme jälleen Matlabissa valmiiksi implementoitua Ljung-Box Q-testiä residuaalien autokorrelaivattomuudelle. Kaikissa malleissa nollahypoteesi residuaalien autokorrelaivattomuudesta tulee hylättyä viimeistään viipeellä 6.

Yllä esitettyjen regressiodiagnostisten testien valossa voimme siis todeta, että ensimmäinen VAR – mallimme ei selitä aikasarjan kaikkea rakennetta. Saadut estimaattorit ovat myös tehottomia, joten edellä esitetyt estimaattorien luottamusrajat ovat turhan leveitä. Saadut mallit ovat siis yhteenvetona diagnostisten testien perusteella huonoja.

³ Huomaa, että tässä vaiheessa mallille ei ole tehty minkäänlaista regressiodiagnostiikkaa, joten saadut tulokset eivät välttämättä perustu korrektiin tilastolliseen päättelyyn!



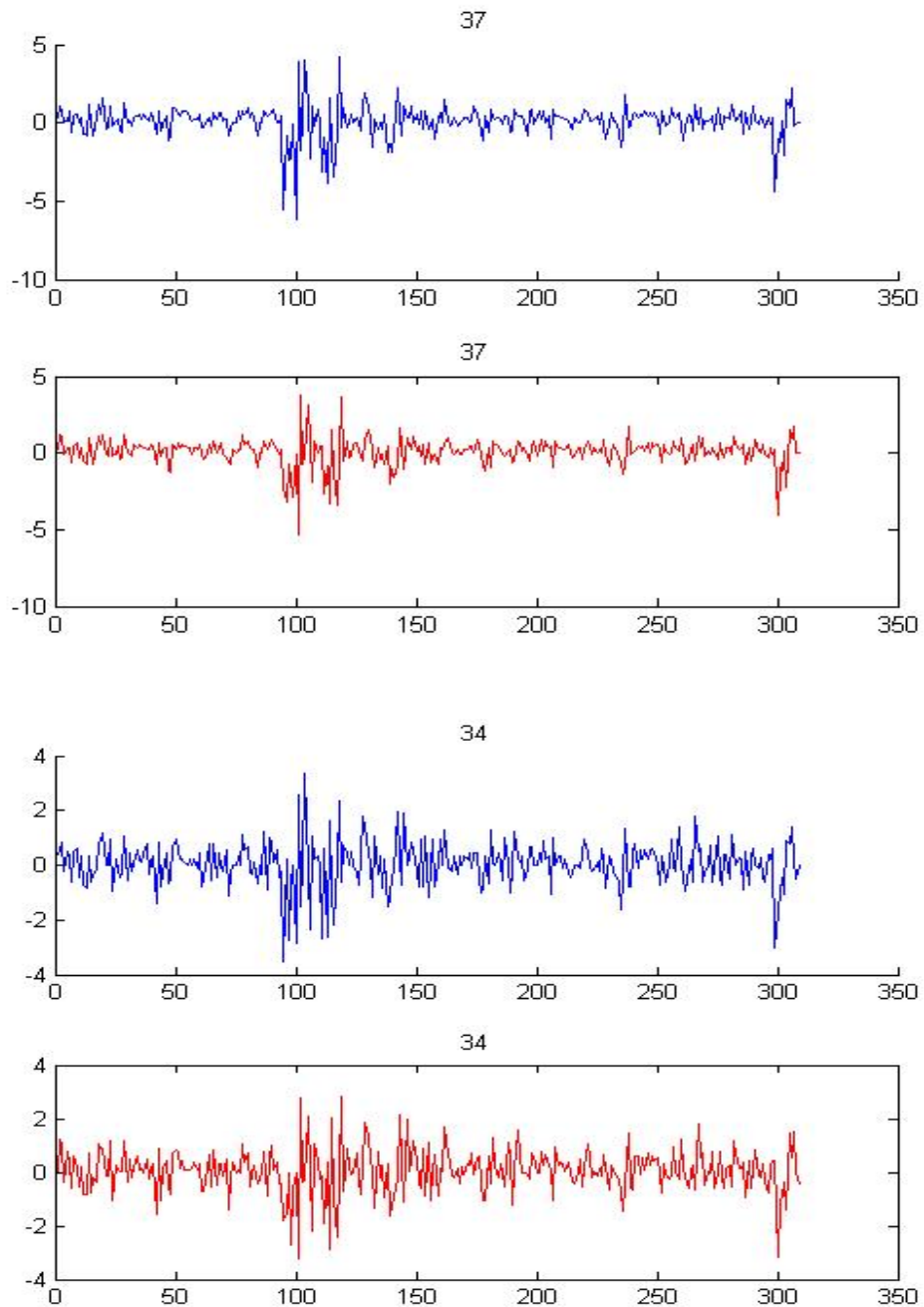
Kuva 18 Rahastojen 31 ja 34 toteutuneet tuotot ylempänä, sekä VAR(1) mallimme tuottamat sovitteet alempana. Vaakaakselilla on juokseva numerointi arkipäiville lähtien 1.1.2001. Vaakaksieli kuvaa prosentuaalista päivätuottoa. Rahastolle 34 mallimme selitysaste on 0.92 ja rahastolle 31 se on 0.12.

Yllä esitetyistä puutteista huolimatta estimoitujen mallien estimaattorit eivät ole harhaisia, joten mallien tulisi olla ”suuntaa antavia” vaikeivät olekaan riittäviä siinä mielessä, että kaikkea

havaittua vaihtelua ei tule niiden avulla selitettyä. On kuitenkin huomattava, että estimoitujen muuttujien tilastollisten testien tulkitseminen testien normaalissa mielessä ei käy. Tutkimme kiinnostuksen vuoksi miten estimoidut mallit pärjäävät vuoden 2006 datalla, jota ei siis ole käytetty mallien estimoisissa.

Kuvassa Kuva 19 on esitetty samojen rahastojen kun kuvassa Kuva 18 tuottoa kuvaavien mallien sovitteet vuoden 2006 datalle, sekä toteutuneet tuotot vuonna 2006. Näemme, että selitysasteensa perusteella paras vuosien 01-05 dataan sovitettu malli ennustaa hyvin tarkasti kyseisen rahaston tuotot vuoden 2006 ensimmäiselle noin 125 päivälle. Tämän tienoilla aikasarjassa on havaittavissa selvä varianssin kasvu, aikasarjan rytmin muutos. Tämän jälkeiseltä periodilta mallimme ennusteet vaikuttavat olevan ”myöhässä”. Mallimme siis ennustaa hyvin niin kauan kun aikasarjan luonne ei muutu. Malli rahaston 31 tuotolle, jonka selitysaste oli vain 0.12, ei tietenkään pärjää ennustamisessakaan.

Liitteessä **Error! Reference source not found.** on tutkittu kaksi vuosien 01-05 dataan sovittamisen perusteella toiseksi parasta mallin pärjäämistä vuoden 2006 tuottojen ennustamisessa. Rahastojoiden tuottojen ennustaminen sujuu hyvin, ovat 130 ja 37. Teknillisistä syistä myös kaksi muuta huonommin pärjäävää mallia on otettu mukaan kuviin. Voimme todeta, että toinen malli (rahaston 37 tuottoa kuvaava) ennustaa hyvin edellä mainittuun aikasarjan luonteenmuutokseen asti, jonka jälkeen ennusteet vaikuttavat olevan todellisuutta edellä. Toinen (rahaston 130 tuottoa kuvaava) malli taas pärjää ennustamisessa kohtuullisen hyvin myös varianssin kasvun jälkeen.



Kuva 19 VAR(1) mallimme sovitteet vuodelle 2006 (ylempänä) sekä toteutuneet tuotot (alempänä), rahastoille 31 ja 34. Vaaka-akselilla on juokseva indeksointi arkipäiville 1.12006 lähtien ja pysty-akselilla rahastojen prosentuaalinen päivätuotto.

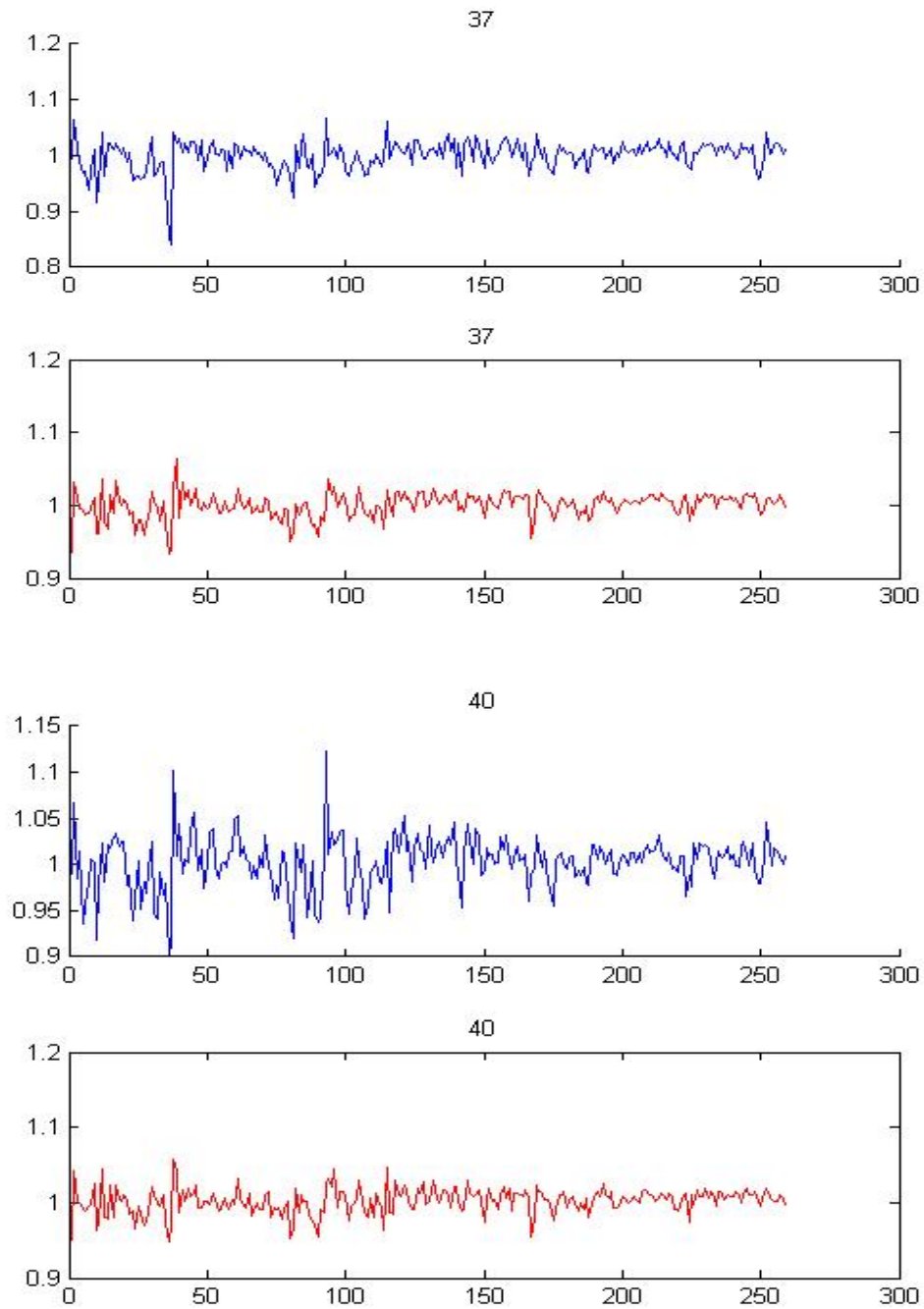
Mallit, joille saatiin korkeat selitysasteet niiden estimoinnin yhteydessä vaikuttavat puutteistaan huolimatta toimivan hyvin vuoden 2006 dataan. On kuitenkin huomioitava, että suurimmalle osalle rahastoista VAR(1) mallimme ennusteet toimivat varsin huonosti, eikä yksikään malli ole tilastollisessa mielessä riittävä.

Kokeilemme seuraavaksi VARX(1) mallia. Otamme malliimme mukaan kaikki faktorit, sekä viivästyttäen että yhdellä viipeellä. Askeltavan regression kynnsarvona muuttujan merkitsevyyden t-testissä käytetään 1% jotta valittujen selittäjien määrä olisi edelleen pieni. Malli on jo varsin laaja ja tulokset melko vaikeasti käsiteltäviä. Todettakoon tässä tämän vuoksi vaan, että huomattavaa parannusta ei nähdä mallin selitysasteissa, ja että vain harvaan malliin tuli lisättyä ulkoinen muuttuja. Ulkoisista muuttujista useimmiten tuli lisättyä faktori 21, hopea, ilman viivettä. Tämä tuli merkitseväksi tekijäksi 8 malliin 72 mallistamme. Ainoastaan kolmeen malliin tuli valittua ulkoinen viivästetty muuttuja. Nämä olivat kaikki faktoria 21 eli hopeaa. Mallien residuaalien ominaisuudet eivät parantuneet puhtaasta VAR mallin tuloksiin verrattuna. Voimme siis todeta, että rahastojen päivätuottoihin vaikuttavat vain hyvin heikosti tutkimamme faktoreiden päivätuotot (prosentuaaliset muutokset päivästä toiseen), sekä tuskin lainkaan tutkimamme faktoreiden viivästytyt tuotot.

Kokeilimme myös VAR(2) malleja päivädatalle. Tulokset eivät parantuneet merkitsevästi niin selitysasteiltaan, kun myöskään regressiodiagnostisilta ominaisuuksiltaan.

9.6 VAR malli viikkotuotoille

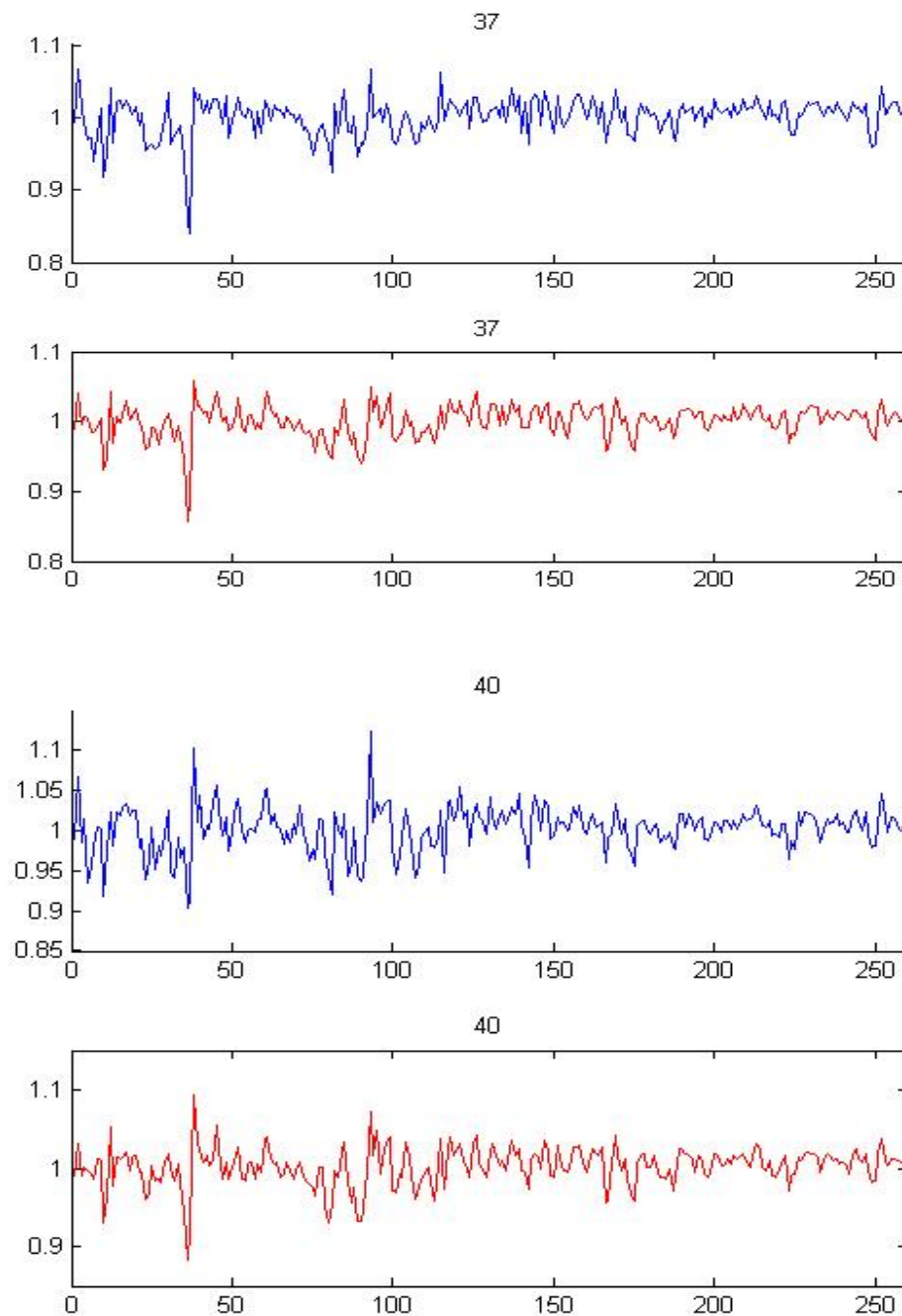
Kokeilimme puhdasta VAR(1) mallia myös viikkotuotoille. Saatujen mallien selitysasteet pysyvät keskimäärin samoina kun päivädatalle tehdyille VAR(1) malleille, mutta yhtäkään hyvää mallia ei saada aikaiseksi. Yhdenkään mallin selitysaste ei ylitä 0.5. Residuaalien ominaisuudet kuitenkin parantuvat huomattavasti, ja kaikkien mallien residuaalit läpäisevät Ljung-Boxin Q testin autokorreloimattomuudelle 50 viipeellä. Suurin osa saatujen mallien residuaaleista läpäisevät lisäksi myös Englen hypoteesitestin residuaalien homoskedastisuudesta 15 viipeellä, ja kaikki läpäisevät testin 10 viipeellä. Tämä tarkoittaa, että saatujen mallien selittäjien tilastollinen testaaminen on hyväksyttävää. Alla kuvassa Kuva 20 on esitetty parhaiten selittävän mallin sovitteet sekä toteutuneet tuotot samassa kuvassa. Myös toinen, huonommin selitettävä rahasto on esitetty.



Kuva 20 Viikkotuottojen perusteella estimoitu VAR(1) mallimme sovitteet rahastojen tuotoille alempana, sekä toteutuneet tuotot ylempänä. Rahastolle 37 saavutettiin korkein selitysaste, 0.46. Alemman rahaston mallin selitysaste on 0.25.

Kokeillaan miten VARX malli ilman viivästettyjä ulkoisia muuttujia toimii viikkotuottojen selittämiseen. Vaikkei saataisikaan kovin hyvää mallia, voimme mikäli saatujen mallien residuaalit ovat autokorreloimattomia ja homoskedastisia tehdä päätelmiä siitä mitkä ulkoiset tekijät ovat tilastollisesti merkitseviä rahaston tuoton muodostumiseen.

Käytämme jälleen VAR-työkaluamme VARX mallin estimoimiseen kaikille rahastojen viikkotuotoille, jossa ulkoisiksi muuttujiksi otetaan kaikki faktorimme ilman viivettä. Askeltavasta regressiosta johtuen lopullisiin malleihin ei kuitenkaan tule selittäjiksi kuin pieni osa ulkoisista muuttujista. Saatujen mallien selitysasteet eivät kasva huomattavasti verrattuna edellä estimoituun VAR mallin selitysasteisiin, vaikka kohoavatkin hieman. Residuaalit säilyvät autokorreloimattomina ja homoskedastisina. Kuvassa Kuva 21 on esitetty samojen kahden rahaston tuotot sekä VARX mallimme sovitteet jotka kuvassa Kuva 20 esitettiin VAR mallimme sovitteita vastaan.



Kuva 21 VARX(1) mallimme sovitteet kahden rahaston viikkotuotoille alempana sekä toteutuneet viikkotuotot ylempänä. Vaakaakselilla viikkojen juokseva numerointi 1.12001 lähtien. Pystyakselilla rahastojen viikkotuotot desimaalimuodossa. Rahaston 37 viikkotuottoa kuvaavaa mallin selitysaste nousee ulkoisen mallin lisäämisen myötä 0.56:een ja rahaston 40 mallin selitysaste kohoaa 0.44 ään.

VARX mallissamme ulkoisista muuttujista selkeästi eniten esiintyvä on faktori 67, 12kk euriborin viikkomuutos, joka esiintyy 69 mallissa 71 mallista. 12kk euriborin muutos on siis 99% luottamustasolla merkitsevä selittäjä rahastojen viikkotuottojen selittäjä melkein kaikille rahastoille. Kaikissa malleissa kerroin on positiivinen ja 0.208 ja 0.372 välillä. Muita malleissa yleisesti esiintyviä ulkoisia selittäjiä ovat faktori 25, kuparin hinta, joka esiintyy 38/71 mallissa ja faktori 3, OECD total monetary aggregates, joka esiintyy 29/71 mallissa. Faktorin 25 kertoimet ovat kaikki positiivisia niissä malleissa jossa se esiintyy. Sama pätee faktorin 3 kertoimille.

10. Yhteenveto ja johtopäätökset

Tutkimuksessa löydettiin lineaaristen mallien avulla eurooppalaisten large CAP rahastojen tuottoihin vaikuttavia tekijöitä.

Suurimmalle osasta rahastoista päivätuottojen selittäminen kokeilemillamme lineaarisilla malleilla onnistui heikosti. Vektoriautoregressiolla löydettiin kuitenkin muutamalle rahastolle yllättävänkin hyviä tuloksia. Päivätuottojen kohdalla ei pystytty identifioimaan luotettavasti yhtäkään ulkoista faktoria, joka vaikuttaisi rahastojen tuottoihin. Ainoana mahdollisesti rahastojen päivittäisten hinnanmuutoksien kanssa tilastollisesti merkitsevän yhteisvaikutuksen omaavana ulkoisena muuttujana todettiin hopean päivittäinen hinnanmuutos.

Rahastojen viikko-tuottojen selittäminen onnistui lineaarisilla malleillamme huonohkosti. Suurimmassa osassa malleistamme ei saatu regressiodiagnostisilta ominaisuuksiltaan hyviä malleja. Samalla mallien selitysasteet jäivät mataliksi. Vektoriautoregressiolla saadut mallit viikkotuotoille olivat regressiodiagnostisilta ominaisuuksiltaan hyviä, mutta mallien selitysasteet jäivät pääasiassa hyvin mataliksi. Viikko-tuottoihin tilastollisesti merkitsevästi vaikuttaviksi ulkoisiksi selittäjiksi pystyttiin identifioimaan muutama tutkimistamme faktoreista. Muun muassa 12kk euribor-koron muutos, kuparin hinta, sekä OECD total monetary aggregates tunnistettiin tilastollisen yhteisvaikutuksen omaaviksi.

Rahastojen kuukausituotoille pystyimme muodostamaan tilastollisilta ominaisuuksiltaan hyviä lineaarisia malleja joiden selittäjinä käytettiin ulkoisia faktoreita sekä niitä viipeillä. Viipeellä vaikuttavina muuttujina todettiin ISM USA indeksi, joka kuvaa teollisuuden ja palvelualan luottamusta. Myös Euroopan vaihtotase todettiin viipeellä vaikuttavaksi tekijäksi. Ilman viivettä vaikuttavina muuttujina todettiin kuparin hinta, OECD total leading indicators Euribor 12kk korko, sekä Net external assets-indeksi.

Saatuihin tuloksiin pitää suhtautua varauksella. Ensinnäkin on huomattava, että tutkittu ajanjakso on melko lyhyt. Kyseisen ajanjakson aikana on esiintynyt sekä vahva markkinoiden notkahdus että toipuminen. Tämä on voinut olla osasyynä hyvien mallien löytämisen vaikeudelle. Malli, joka toimii sekä yleisen hintakehityksen ollessa nouseva että sen ollessa laskeva voidaan kuitenkin pitää paremmin havaittujen ilmiöiden taustatekijöitä selittävänä.

Tutkimuksen kohteena olevan ilmiön erittäin monimutkainen luonne on myös vaikuttanut siihen, että työn toteuttamisessa on jouduttu kokeilemaan useaa lähestymistapaa ennen soveltuvien työkalujen löytämistä. Tutkimustamme on tämän takia leimannut määrätietoisen etenemisen puute. Tähän on myös vaikuttanut aiheen vieraus projektiryhmän jäsenille. Jälkikäteen voimme todeta, että edistyneen tason ekonometristen menetelmien tunteminen ennen työn aloittamista olisi selkeyttänyt tutkimusta ja vapauttanut paljon aikaa olennaisten asioiden tutkimiseen.

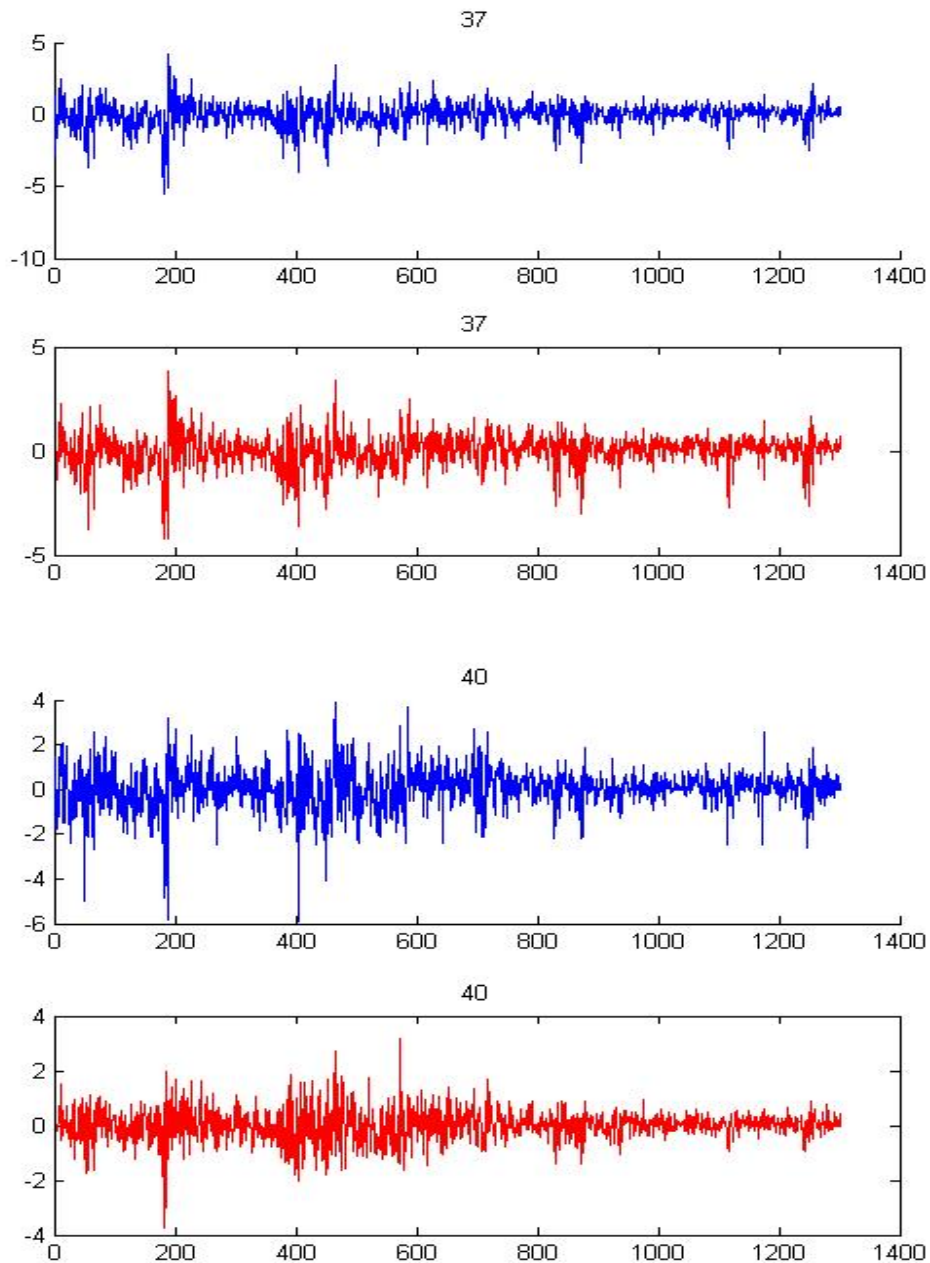
Kiinnostaviksi ja valitettavasti kokeilemattomiksi jääneet menetelmät ovat vektori- MA ja -ARMA-mallit, sekä kaikki epälineaariset mallit. Toinen työn aikana ilmennyt mielenkiintoinen tutkimusaihe on rahastojen volatiliiteetin tutkiminen, joka myös kirjallisuudessa on todettu mielekkäämmäksi ongelmaksi tuottojen tutkimiseen verrattuna.

11. Lähteet

- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* .
- Campbell, J. Y., Lettau, M., Malkiel, B. G., & Xu, Y. (2001). Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk. *The Journal of Finance* .
- Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum Strategies. *The Journal of Finance* .
- Chan, L. K., Karceski, J., & Lakonishok, J. (1998). The Risk and Return from Factors. *Journal of Financial and Quantitative Analysis* .
- Eakins, S. G., & Stansell, S. R. (2004). Do Momentum Strategies Work? *The Journal of Investing* .
- Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Survivorship Bias and Mutual Fund Performance. *The Review of Financial Studies* .
- Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives* .
- Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* .
- Ferson, W. E., & Warther, V. A. (1996). Evaluating Fund Performance in a Dynamic Market. *Financial Analyst Journal* .
- Fung, W., & Hsieh, D. A. (1998). Performance Attribution and Style Analysis: From Mutual Funds to Hedge Funds.
- Hamilton, J. D. (1996). Stockmarket volatility and the business cycle. *Journal of Applied Econometrics* .
- Hamilton, J. D. (1994). *Time Series Analysis*. Princeton.
- Jensen, M. C. (1968). Problems In Selection Of Security Portfolios - The Performance Of Mutual Funds In The Period 1945-1964. *The Journal of Finance* .
- King, M., Sentana, E., & Wadwhani, S. (1994). Volatility and Links Between National Stock Markets. *Econometrica* .
- Kothari, S., & Warner, J. B. (2001). Evaluating Mutual Fund Performance. *The Journal of Finance* .
- LeSage, J. P. (2005). *Econometrics Toolbox*. Retrieved from <http://www.spatial-econometrics.com/>.
- Liljeblom, E., & Stenius, M. (1993). Macroeconomic volatility and stock market volatility; Empirical evidence on Finnish data. *Meddelande från Svenska handelshögskolan* .

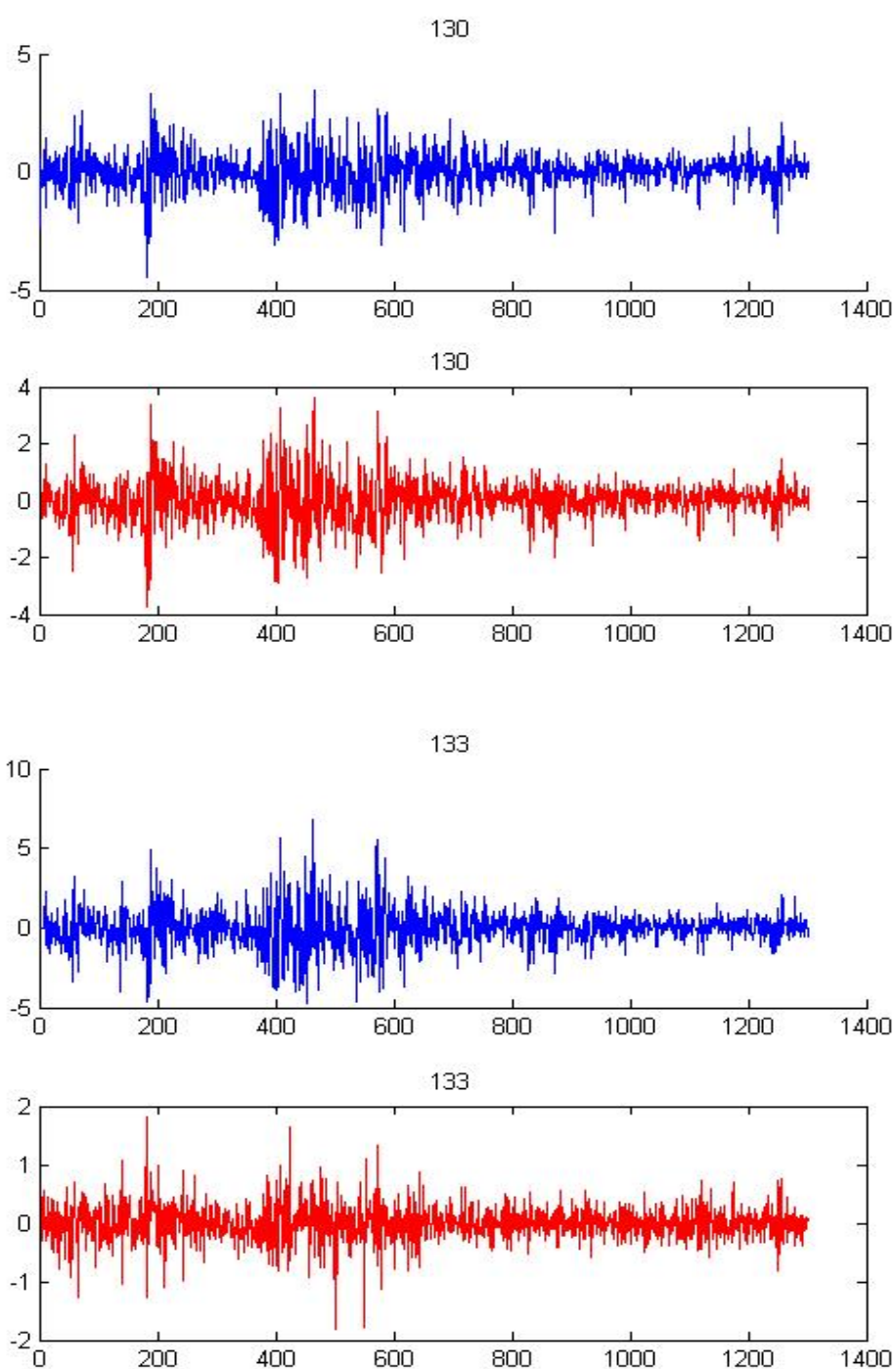
- Mellin, I. (2007). *Pääkomponenttianalyysi*. Retrieved from Systemianalyysin laboratorion www-sivut: <http://www.sal.tkk.fi/Opinnot/Mat-2.112/K07/PCOMP10.pdf>
- Milton, J. S., & Arnold, J. C. (2003). *Introduction to Probability and Statistics, 4th Edition*. McGraw-Hill Higher Education.
- Nardari, F., & Scruggs, J. T. (2003). Analysis of Linear Factor Models with Multivariate Stochastic Volatility for Stock and Bond Returns.
- Ng, V. K., Rothschild, M., & Engle, R. F. (1990). Asset pricing with a Factor-ARCH covariance structure: empirical estimates for Treasury bills. *Journal of Econometrics* .
- O'Connor, N., & Madden, M. G. (2005). A neural network approach to predicting stock exchange movements using external factors. *ScienceDirect* .
- Otten, R., & Bams, D. (2002). European Mutual Fund Performance. *European Financial Management* .
- Straumann, D., & Garidi, T. (2007). Developing an Equity Factor Model for Risk. *RiskeMetrics Journal* .
- Tsay, R. S. (2002). *Analysis of Financial Time Series (1st Edition)*. John Wiley & Sons, Inc.

12. Liitteet



Kuva 22 VAR(1)-mallien ennuste rahastojen päivätuotoille ja toteutuneet arvot vuodelle 2006.

23. huhtikuuta 2007



Kuva 23 VAR(1)-mallien ennuste rahastojen päivätuotoille ja toteutuneet arvot vuodelle 2006.

1	Belgia BCI	1004	CPI, Communication, Total, Index	1045	New Orders, Manuf, Overall, Total, Index
3	Likviditeetti	1005	CPI, Education, Total, Index	1046	New Orders, Manuf, Overall, Total, SA, Index
5	OECD Ind. 05	1006	CPI, Food and Non-alcoholic Beverages, Total, Index	1047	New Orders, Aggregates, Consumer non-durables, Index
7	OECD Ind. 06	1007	CPI, Health, Total, Index	1048	New Orders, Aggregates, Consumer non-durables, SA
9	ISM(USA)	1008	CPI, Furnishing and Household Equipment, Total, Index	1049	Business Surveys, PMI, Manufacturing Sector, Total
11	GDP-kasvu/Euroalue	1009	CPI, Housing, Water, Electricity, Gas & Other Fuels, Total, Index	1050	PMI, Services Sector, Total business activity, SA
13	Saksan trade balance	1010	CPI, Miscellaneous Goods and Services, Total, Index	1051	PMI, Services Sector, Business expectations
15	Öljy (USA)	1011	CPI, Recreation and Culture, Total, Index	1052	PMI, Manufacturing Sector, Employment, SA
17	Öljy	1012	CPI, Restaurants and Hotels, Total, Index	1053	PMI, Services Sector, Employment, SA
19	Kulta	1013	CPI, Transport, Total, Index	1054	PMI, Whole Economy, Employment Whole Economy, SA
21	Hopea	1014	CPI, Total Excl alcohol and tobacco, Index	1055	PMI, Manufacturing Sector, Stocks of Finished Goods
23	Alumiini	1015	CPI, Total Excl energy, Index	1056	PMI, Manufacturing Sector, Input prices, SA
25	Kupari	1016	CPI, Total Excl education, health & social protection, Index	1057	PMI, Services Sector, Input prices, SA
27	Lyijy	1017	CPI, Total Excl energy, food, alcohol and tobacco, Index	1058	PMI, Manufacturing Sector, Output, SA
29	Nikkeli	1018	CPI, Total Excl liquid fuels and lubricants, Index	1059	PMI, Services Sector, New business, SA
31	Tina	1019	CPI, Total Excl energy and seasonal food, Index	1060	PMI, Manufacturing Sector, New orders, SA
33	Sinkki	1020	CPI, Total Excl seasonal food, Index	1061	PMI, Whole Economy, New orders Whole Economy, SA
35	Hilli	1021	CPI, Total Excl energy and unprocessed food, Index	1062	PMI, Whole Economy, Input prices Whole Economy, SA
37	Sähkö	1022	CPI, Total Excl housing, water, electr, gas & other fuels, Index	1063	PMI, Services Sector, Outstanding business, SA
39	Teräs	1023	Credit to Gov, EUR	1064	PMI, Whole Economy, Output Whole Economy, SA
41	Etanoli	1024	Credit to Gov, SA, EUR	1065	PMI, Services Sector, Prices charged, SA
43	Euro vs. JPY	1025	Credit, Loans, Households & NPISH, from MFIs, Total, EUR	1066	PMI, Manufacturing Sector, Stocks of purchases, SA
45	Euro vs. USD	1026	Credit to oth euro area residents, Chg YY, SA	1067	PMI, Manufacturing Sector, Quantity of Purchases, SA
47	Euro vs. GBP	1027	Credit to oth euro area residents, SA, Index	1068	PMI, Manufacturing Sector, Supplier's delivery time
49	Euro vs. SEK	1028	Credit to oth euro area residents, SA, EUR	1069	PMI, Manufacturing Sector, Total, SA
51	Euro vs. AUD	1029	Current Account, Balance, Total, EUR	1070	Business Surveys, PMI, Whole Economy, Composite, SA
53	Euro vs. CHF	1030	Current Account, Balance, Total, SA, EUR	1071	Total reserves Incl gold, EUR
55	Euro vs. CNY	1031	Business Surveys, ZEW, Indicator of Economic Sentiment	1072	Retail Sales, Turnover, Total, IA, Cal Adj, Chg Y
57	Euro vs. CAD	1032	Errors and Omissions, Total, EUR	1073	Business Surveys, PMI, Services Sector, Total buss act
59	Euro vs. NZD	1033	Foreign exchange reserves, EUR	1074	Financial Account, Direct Investment, Balance, Total, EUR
61	Euro vs. NOK	1034	CPI, All Items, Total Excl tobacco, Index	1075	Financial Account, Balance, Total, EUR
63	Euro vs. DKK	1035	Assets, loans to euro area residents Incl Euros, Total, EUR	1076	Financial Account, Other Investment, Balance, Total, EUR
65	Euribor 3kk	1037	Currency in circulation, SA, EUR	1077	Current Account, Goods, Net, Total, EUR
67	Euribor 12kk	1038	Net external assets, SA, EUR	1078	Current Account, Goods, Net, Total, SA, EUR
69	Pienet yritykset	1039	New Orders, Aggregates, Capital goods, Index	1079	Current Account, Income, Net, Total, EUR
71	Keskisuuret yritykset	1040	New Orders, Aggregates, Capital goods, SA, Index	1080	Current Account, Income, Net, Total, SA, EUR
73	Suuret yritykset	1041	New Orders, Aggregates, Consumer goods, Index	1081	Financial Account, Portfolio Investment, Balance, Total, EUR
1001	Business Surveys, ZEW, Indicator of Current Sentiment	1042	New Orders, Aggregates, Consumer goods, SA, Index	1082	Financial Account, Reserve Assets, Total, EUR
1002	CPI, Alcoholic Beverages, Tobacco and Narcotics, Total, Index	1043	New Orders, Aggregates, Intermediate goods, Index	1083	Current Account, Services, Net, Total, EUR
1003	CPI, Clothing and Footwear, Total, Index	1044	New Orders, Aggregates, Intermediate goods, SA, Index	1084	Current Account, Services, Net, Total, SA, EUR
				1085	Term premium
				1086	Default premium

Kuva 28 Lista käytetyistä faktoreista

1	LU0158927490 Equity	136	LU0138075311 Equity	1067	IE00B0S5PT05 Equity
4	LU0095325956 Equity	139	LU0140363002 Equity	1070	FI0008800396 Equity
7	LU0124675678 Equity	142	IE00B04ZPD85 Equity	1073	FI0008802194 Equity
10	FI0008803614 Equity	145	LU0113992829 Equity	1076	FI0008800974 Equity
13	FI0008803630 Equity	148	LU0113993470 Equity	1079	LU0064319337 Equity
16	IE0030412666 Equity	151	LU0122972895 Equity	1082	FI0008806369 Equity
19	FI0008804836 Equity	154	LU0122973430 Equity	1085	FI0008802301 Equity
22	FI0008805049 Equity	157	LU0122972382 Equity	1088	FI0008805122 Equity
25	IE0008365730 Equity	160	FI0008800750 Equity	1091	LU0111518840 Equity
28	IE0008366142 Equity	163	FI0008805742 Equity	1094	LU0066794719 Equity
31	IE0004866772 Equity	166	FI0008808795 Equity	1097	LU0099624685 Equity
34	LU0119128261 Equity	169	LU0165073775 Equity	1100	LU0111491469 Equity
37	LU0119106291 Equity	172	LU0164906959 Equity	1103	LU0177332227 Equity
40	LU0083347228 Equity	175	LU0165306035 Equity	1106	FI0008808449 Equity
43	LU0093943974 Equity	178	LU0199060350 Equity	1109	FI0008802988 Equity
46	FI0008808373 Equity	181	LU0082087510 Equity	1112	LU0161305163 Equity
49	LU0101102464 Equity	184	LU0194779913 Equity	1115	LU0227178638 Equity
52	LU0193725610 Equity	187	LU0028118809 Equity	1118	LU0086811550 Equity
55	LU0193725966 Equity	190	LU0028119013 Equity	1121	LU0030166333 Equity
58	LU0157028266 Equity	193	LU0119750205 Equity	1124	LU0030166507 Equity
61	LU0193726345 Equity	196	LU0114642118 Equity	1127	LU0088160857 Equity
64	LU0193726857 Equity	199	LU0064070138 Equity	1130	FI0008802665 Equity
67	LU0168301199 Equity	202	LU0119063039 Equity	1133	FI0008802749 Equity
70	LU0055729916 Equity	1001	LU0248032871 Equity	1136	FI0008801774 Equity
73	LU0048365026 Equity	1004	LU0117858596 Equity	1139	IE0031332483 Equity
76	LU0168302833 Equity	1007	LU0168352267 Equity	1142	IE0031387719 Equity
79	LU0161912059 Equity	1010	LU0079556006 Equity	1145	GB0008787136 Equity
82	LU0168301868 Equity	1013	LU0070215420 Equity	1148	LU0139292543 Equity
85	LU0123485178 Equity	1016	LU0119092640 Equity	1151	LU0118128569 Equity
88	LU0088125512 Equity	1019	LU0105624737 Equity	1154	LU0006391097 Equity
91	LU0123484957 Equity	1022	FI0008810221 Equity	1157	LU0210075874 Equity
94	LU0112687495 Equity	1025	FI0008806203 Equity	1160	LU0169528261 Equity
97	FI0008808936 Equity	1028	FI0008802962 Equity	1163	LU0130731390 Equity
100	FI0008802046 Equity	1031	LU0185060844 Equity	1166	LU0130731986 Equity
103	FI0008806914 Equity	1034	LU0185060927 Equity	1169	LU0130732109 Equity
106	FI0008801170 Equity	1037	LU0174484120 Equity	1172	LU0144509717 Equity
109	LU0083291335 Equity	1040	LU0094557526 Equity	1175	LU0130732364 Equity
112	LU0061175625 Equity	1043	LU0125946151 Equity	1178	LU0131725367 Equity
115	LU0119124278 Equity	1046	LU0125944966 Equity	1181	LU0143905700 Equity
118	LU0048578792 Equity	1049	LU0125951151 Equity	1184	LU0117904960 Equity
121	LU0119124781 Equity	1052	LU0154234636 Equity	1187	LU0117858752 Equity
124	LU0103194394 Equity	1055	LU0171280604 Equity	1190	LU0215048744 Equity
127	FI0008806252 Equity	1058	LU0011846440 Equity	1193	LU0117858679 Equity
130	LU0093576881 Equity	1061	LU0006061161 Equity	1196	LU0210531801 Equity
133	LU0122612848 Equity	1064	IE00B0S5PS97 Equity	1199	LU0210531983 Equity
				1202	LU0171281750 Equity

Kuva 29 Lista käytetyistä fundeista.