

**RATIONAALISIA VAI ADAPTIIVISIA
ENNUSTEITA?**

**YLIOPISTO-OPISKELIJOIDEN JA PANKIN
SIJOITUSASIAANTUNTIJOIDEN
KÄSITYKSET RISKEISTÄ JA TUOTOISTA**

Jyväskylän yliopisto
Kauppakorkeakoulu
Taloustiede
Pro gradu -tutkielma
Huhtikuu 2014
Tekijä: Anni Kinttu
Ohjaaja: Kari Heimonen

JYVÄSKYLÄN YLIOPISTON KAUPPAKORKEAKOULU

Tekijä: Anni Kinttu	
Työn nimi: Rationaalisia vai adaptiivisia ennusteita? Yliopisto-opiskelijoiden ja pankin sijoitusasiantuntijoiden käsitykset riskeistä ja tuotoista	
Oppiaine: Taloustiede	Työn laji: Pro gradu -tutkielma
Aika: Huhtikuu 2014	Sivumäärä: 109 + 35 liitesivua
<p>Tiivistelmä:</p> <p>Opinnäytetyössä tutkittiin kyselytutkimusten avulla Jyväskylän yliopiston kauppakorkeakoulun tarjoamille kursseille osallistuneiden opiskelijoiden ja Osuuspankin sijoitusasiantuntijoiden käsityksiä tulevista osaketuotoista ja -riskeistä. Kyselyssä vastaajia pyydettiin laatimaan seitsemän ja 13 kuukauden hintaennusteet sekä niille 80 %:n luottamusvälit näytettyjen kuuden osakekuvaajan perusteella. Osakekuvaajissa käytettyä indeksiä tai ajankohtia ei paljastettu vastaajille.</p> <p>Ennusteiden perusteella tutkittiin, oliko vastaajilla taipumusta ekstrapoloida mennyttä hintakehitystä, eli olivatko he trendin seuraajia, ja olivatko heidän ennustamansa subjektiiviset todennäköisyysjakaumat vinoja. Lisäksi selvitettiin taustatekijöiden, kuten iän, sukupuolen, käytyjen rahoituksen kurssien sekä työ- ja sijoituskokemuksen vaikutusta vastaajien ennusteisiin.</p> <p>Keskeisinä tuloksina havaittiin, että vastaajat ennustivat menneen trendin jatkuvan, eli he olivat optimistisempia nousu- kuin laskukausien jälkeen. Keskimääräinen odotettu hintamuutos oli nousukausien jälkeen suurempaa kuin laskukausien jälkeen. Vastaajat myös suojasivat ennusteitaan, mikä ilmeni siten, että luottamusvälien keskimääräinen vinous nousukausien jälkeen oli pienempi kuin laskukausien jälkeen. Näytettyjen osakekuvaajan keskimääräinen hintataso veti ylä- ja alaluottamusvälejä alaspäin nousukausien jälkeen ja laskukausien jälkeen taas ylöspäin.</p> <p>Tutkimuksen tuloksia voidaan peilata perinteisen rahoitusteorian oletuksia vasten. Tulokset osoittivat, etteivät yksilöt aina käyttäydy rationaalisesti riskipitoisten valintojen edessä.</p>	
Asiasanat: Rahoituksen käyttäytymistiede, kyselytutkimus, riski- ja tuottoennusteet, trendinseuraaminen, suojaus	
Säilytyspaikka: Jyväskylän yliopiston kauppakorkeakoulu	

KUVIOT

KUVIO 1 Hypoteettinen arvofunktiio (Kahneman & Tversky 1979, 279).....	17
KUVIO 2 Hypoteettinen painofunktiio (Kahneman & Tversky 1979, 282).....	18
KUVIO 3 Mediaaniestimaattoreihin perustuvat painofunktiot voitoille (w^+) ja tappioille (w^-) (Kahneman & Tversky 1992, 313).....	19
KUVIO 4 Opiskelijoiden jakautuminen pääaineittain	37
KUVIO 5 Suoritetut rahoituksen kurssit sukupuolittain.....	39
KUVIO 6 Sijoituskohteet sukupuolittain.	41
KUVIO 7 Pankin vastaajien jakautuminen ikäryhmiin	43
KUVIO 8 Rahoitusalan työkokemus (vuosina)	44
KUVIO 9 Työskentelyvuodet nykyisessä tehtävässä	45
KUVIO 10 Puuttuvan tiedon kokonaistarkastelu	47

TAULUKOT

TAULUKKO 1 Yksilön suhde riskiin	19
TAULUKKO 2 Opiskelijoiden jakautuminen ikäryhmiin.....	37
TAULUKKO 3 Opiskelijoiden jakautuminen tiedekunnittain	38
TAULUKKO 4 Opiskelijoiden pohjakoulutus	38
TAULUKKO 5 Opiskelijoiden jakautuminen luokkiin suoritettujen opintopisteiden perusteella.....	38
TAULUKKO 6 Opiskelijoiden jakautuminen ryhmiin suorittujen rahoituksen kurssien perusteella.....	39
TAULUKKO 7 Opiskelijoiden ilmoittama opintomenestys rahoituksen kursseissa	40
TAULUKKO 8 Rahoituksen kurssien opintomenestys sukupuolittain	40
TAULUKKO 9 Äidin sosioekonominen asema.....	42
TAULUKKO 10 Isän sosioekonominen asema	42
TAULUKKO 11 Pankin vastaajien jakautuminen pohjakoulutuksen mukaan..	43
TAULUKKO 12 Jakautuminen nykyisten tehtävämikkeiden mukaan.....	44
TAULUKKO 13 Ennustettu osaketuotto osakekuvaajan keskiarvoon verrattuna	52
TAULUKKO 14 Toisinajattelijastrategian testaus, yhdistetty perusaineisto.....	54
TAULUKKO 15 Odotettujen hintamuutosten vertailu ja testitulokset, opiskelijoiden perusaineisto	54
TAULUKKO 16 Odotettujen hintamuutosten vertailu ja testitulokset, pankin perusaineisto	55
TAULUKKO 17 Odotettujen hintamuutosten vertailu ja testitulokset, yhdistetty perusaineisto	55
TAULUKKO 18 Luottamusvälien vinouden vertailu ja testitulokset, opiskelijoiden perusaineisto	56

TAULUKKO 19 Luottamusvälien vinouden vertailu, pankin perusaineisto.....	57
TAULUKKO 20 Luottamusvälien vinouden vertailu, yhdistetty perusaineisto	58
TAULUKKO 21 Regressioanalyysien tulokset, opiskelijoiden perusaineisto	60
TAULUKKO 22 Regressioanalyysin tulokset, pankin perusaineisto	62
TAULUKKO 23 Regressioanalyysien tulokset, yhdistetty perusaineisto	62
TAULUKKO 24 Luottamusvälien kapeuden testaus, opiskelijoiden perusaineisto	65
TAULUKKO 25 Osake-ennusteiden varovaisuuden testaus, opiskelijoiden perusaineisto	66
TAULUKKO 26 Luottamusvälien kapeuden testaus, yhdistetty perusaineisto	67
TAULUKKO 27 Osake-ennusteiden varovaisuuden testaus, yhdistetty perusaineisto	68
TAULUKKO 28 Kaksisuuntaisen varianssianalyysin tulokset.....	74
TAULUKKO 29 Logistisen regressioanalyysin tulokset, opiskelijoiden perusaineisto, malli 1.	75
TAULUKKO 30 Logistisen regressioanalyysin tulokset, opiskelijoiden perusaineisto, malli 2.	76

SISÄLLYS

TIIVISTELMÄ
KUVIOT JA TAULUKOT
SISÄLLYS

1	JOHDANTO.....	9
2	RAHOITUSTEORIAM SIOITTAJIN KÄYTTÄYTYMISEN KUVAAJINA 12	
3	TUTKIMUKSIA SIOITTAJIN EPÄRATIONAALISUUDESTA JA HETEROGEENISUUDESTA.....	21
	3.1 Ankkurointiefekti	21
	3.2 Positive feedback traders, trendinseuraajat.....	23
	3.3 Edustavuusharha	25
	3.4 Yliluottamus	26
	3.5 Koulutus, sijoituskokemus ja lähipiiri	26
	3.6 Sukupuoli ja ikä	29
	3.7 Heterogeeniset sijoittajat.....	30
4	MENETELMÄ JA AINEISTO.....	32
	4.1 Tutkimusmenetelmän valinta.....	32
	4.2 Aineiston hankinta	32
	4.3 Opiskelija-aineiston muokkaus	34
	4.4 Pankin sijoitusasiantuntijoiden aineiston muokkaus.....	35
	4.5 Opiskelija-aineiston kuvailu	36
	4.6 Osuuspankin aineiston kuvailu.....	43
	4.7 Puuttuvan tiedon analyysi, opiskelijoiden osake-ennusteet.....	46
	4.8 Aineistojen normalisuuden tarkistukset ja normalisointi.....	48
5	HAVAINTOJA ENNUSTEISTA	50
	5.1 Perusaineistot	51
	5.1.1 Ankkuroituivatko osake-ennusteet keskiarvotuottoa kohti? .	51
	5.1.2 Erot odotetussa hintamuutoksessa.....	54
	5.1.3 Luottamusvälien vinous	56
	5.1.4 Regressioanalyysit luottamusvälien vinoudelle.....	58
	5.1.5 Yliluottamus.....	64
	5.1.6 Tekivätkö koulutetut henkilöt täsmällisempiä ennusteita?....	68
	5.1.7 Sijoituskokemus ja virheet	71
	5.1.8 Lähipiiri ja virheet.....	72
	5.1.9 Ikä ja virheet.....	77
	5.1.10 Opiskelijatko häirikkösijoittajia?.....	78
	5.2 Herkkyysanalyysit: normalisoidut perusaineistot.....	80

5.2.1	Erot odotetussa hintamuutoksessa	80
5.2.2	Luottamusvälien vinous	81
5.2.3	Regressioanalyysit luottamusvälien vinoudelle.....	82
5.3	Herkkyysanalyysit: puuttuvat tiedot poistettu opiskelija-aineistosta	85
5.3.1	Ankkuroituivatko osake-ennusteet keskiarvotuottoa kohti? .	85
5.3.2	Erot odotetussa hintamuutoksessa.....	86
5.3.3	Luottamusvälien vinous	87
5.3.4	Regressioanalyysit luottamusvälien vinoudelle.....	87
5.3.5	Yliluottamus.....	88
5.3.6	Tekivätkö koulutetut henkilöt täsmällisempiä ennusteita?....	90
5.3.7	Sijoituskokemus ja virheet	91
5.3.8	Lähipiiri ja ennustevirheet	92
5.3.9	Ikä ja ennustevirheet.....	94
5.3.10	Opiskelijatko häirikkösijoittajia?.....	94
5.4	Herkkyysanalyysit: puuttuvat tiedot poistettu opiskelija-aineistosta ja normalisoinnit tehty	95
5.4.1	Erot odotetussa hintamuutoksessa	95
5.4.2	Luottamusvälien vinous	96
5.4.3	Regressioanalyysi luottamusvälien vinoudelle	96
6	TUTKIMUKSEN LUOTETTAVUUDESTA	98
7	JOHTOPÄÄTÖKSET	100
	LÄHTEET	105
	Kirjallisuus	105
	Internet – lähteet.....	109
	LIITTEET	110
	Liite 1 Kyselylomake opiskelijoille	110
	Liite 2 Kyselylomake pankin sijoitusasiantuntijoille	121
	Liite 3 Osake-ennusteiden perustunnusluvut.....	132
	Liite 4 Pankin vastaajien suhtautuminen tehokkaiden markkinoiden hypoteesin liittyviin väitteisiin	134
	Liite 5 Osake-ennusteiden jakaumien normaalisuuden testaus ja symmetrisyyden taskastelu.....	136
	Liite 6 Tutkimuksen hypoteesit.....	138
	Liite 7 Trendinseuraaja- ja toisinajattelijaluokan määritysmuuttujan luonti	139
	Liite 8 Muuttujaluettelo.....	140

1 JOHDANTO

Modernin rahoitusteorian yksi keskeisimmistä tutkimuskohteista liittyy kysymykseen, kuinka sijoittajat muodostavat odotuksia tulevista osaketuotoista. Odotusten ymmärtäminen on tärkeää, sillä sijoittajien käyttäytyminen heiluttaa osakemarkkinoita ja sitä kautta talouden mikro- ja makrorakenteita. Perinteisen arvopapereiden hinnoittelumallin (Capital Asset Pricing Model, CAP-malli) mukaan osakkeen odotettu tuotto liittyy positiivisesti systemaattiseen markkinarisktiin. Mallin taustalta löytyvien oletusten mukaan sijoittajat ovat rationaalisia ja riskiä karttavia hyödyn maksimoijia. Sijoittajan käyttäytymistä riskiä sisältävissä päätöksentekotilanteissa onkin perinteisesti mallinnettu juuri odotetun hyödyn teorian mukaisesti. Teorian olettaa, että rationaalinen sijoittaja pystyy hyödyntämään saatavilla olevaa tietoa tehokkaasti ja päivittämään uskomuksiaan sitä mukaa kun uutta tietoa tulee markkinoille. Markkinoiden tehokkuus on seurausta odotetusta hyödystä ja rationaalisten odotusten paradigmasta.

Psykologiaan liittyvät tutkimukset ovat kuitenkin osoittaneet, että yksilöt eivät aina käyttydy odotetun hyödyn teorian oletusten mukaisesti. On osoitettu, että rationaalisuuden oletuksia sekä Bayesin sääntöä rikotaan yleisesti päätöksenteossa. Viimeksi kuluneen kolmenkymmenen vuoden aikana on lisääntyvässä määrin julkaistu tutkimuksia, jotka kuvaavat sijoittajien suhdetta riskiin toisella tavalla kuin perinteinen rahoitusteoria. Monet näistä tutkimuksista perustuvat prospektiteoriaan, joka on rahoituksen käyttäytymistieteen keskeinen saavutus.

Rahoituksen käyttäytymistieteen havainnot osoittavat, että sijoittajien tekemät virheet voivat johtaa harhaisiin odotuksiin, mikä aiheuttaa tehottomuutta markkinoille. Markkinahinnat eivät aina heijasta täydellisesti kaikkea informaatiota, eivätkä markkinat siten ole täydellisen tehokkaat. Sijoittajat eivät välttämättä ole täysin rationaalisia ja voivat päätöksenteossaan hyödyntää esimerkiksi ekstrapolointia eli menneiden trendien seuraamista.

Tässä opinnäytetyössä tutkittiin kyselytutkimuksen avulla Jyväskylän yliopiston kauppakorkeakoulun tarjoamille kursseille osallistuneiden opiskelijoiden ja Osuuspankin sijoitusasiantuntijoiden odotuksia tulevista osaketuotoista. Aihe liittyy rahoituksen käyttäytymistieteen alaan ja se pohjautuu Werner De Bondtin tekemiin tutkimuksiin (De Bondt 1993). De Bondt teki useita tutkimuksia liittyen muun muassa yliopisto-opiskelijoiden ja ammattisijoittajien ennusteisiin tulevista osaketuotoista. Eräässä kokeessa vastaajille esitettiin kuusi kuviota, jotka esittivät S&P 500 indeksin hintakehitystä 48 kuukauden ajalta (neljä vuotta). Kuvioden pystyakselin hinnat oli skaalattu, jotta niiden perusteella ei pystynyt identifioimaan käytettyä indeksiä, eikä sitä kerrottu vastaajille etukäteen. Vastaajien tehtävä oli ennustaa näiden indeksien arvo seitsemän ja 13 kuukautta eteenpäin viimeisestä kuviossa esiintyvistä arvosta (piste-ennusteet).

Tämän lisäksi vastaajien tuli laatia 80 %:n luottamusväliennusteet kyseisille ajankohdille. De Bondt oli valinnut S&P 500 indeksistä ajankohdat, joista kolme esitti nousukautta (bull market) ja kolme taas laskukautta (bear market). Vastusten perusteella De Bondt tutki, seurasivatko koehenkilöiden ennusteet mennyttä trendiä ja millaisia muotoja heidän ennustamansa todennäköisyysjakaumat saivat.

Tämän työn keskeinen idea on samanlainen kuin De Bondtin tutkimuksessa. Tutkimus toteutettiin kyselytutkimuksena, ja käytetyt kyselylomakkeet on esitetty liitteenä (liitteet 1 ja 2). Lomakkeet oli jaettu kahteen osaan: ensimmäinen, perustietoja kartoittava osa oli erilainen opiskelijoille ja pankin vastaajille. Opiskelijoilta kysyttiin muun muassa sijoitusharrastusta, suoritettuja rahoituksen kursseja ja niihin liittyvää opintomenestystä. Pankin vastaajilta taas kysyttiin esimerkiksi työkokemusta. Lomakkeen toinen osa oli molemmille vastaajaryhmille samanlainen ja se noudatteli De Bondtin tutkimuksen mallia. Osakekuvaajat olivat samoja S&P 500 indeksin kuvaajia kuin De Bondtilla ja ne oli skaalattu samalla tavalla¹. Kyselyyn osallistuneet vastaajat eivät tienneet, mistä indeksistä tai vuosista oli kyse, sillä vaaka-akselilla näkyivät vain kuu-kaudet. Kuvaajista yksi, neljä ja kuusi (A, D ja F) kuvasivat nousukautta ja kaksi, kolme ja viisi (B, C ja E) laskukautta.

Pelkästään De Bondtin tutkimuksen toisinto tämä opinnäytetyö ei kuitenkaan ole, kuten tutkimuskysymyksistä käy ilmi. Tutkimuskysymykset ovat seuraavat:

- i) onko opiskelijoilla taipumusta seurata mennyttä hintakehitystä, eli ovatko he optimistisia nousukausien jälkeen ja pessimistisiä laskukausien jälkeen,
- ii) ovatko subjektiiviset todennäköisyysjakaumat vinoja vasemmalle nousukausien jälkeen ja oikealle laskukausien jälkeen,
- iii) miten Osuuspankin sijoitusasiantuntijat käyttäytyvät suhteessa opiskelijoihin,
- iv) vaikuttavatko ikä, sukupuoli, käytyjen rahoituskurssien määrä, opintopisteet, opintomenestys, sijoitus- tai työkokemus ennusteisiin ja käyttäytymiseen,

¹ Osakekuvaaja numero yksi esitti S&P 500 indeksiä vuosilta 1964–1967, osakekuvaaja kaksi vuosia 1967–1970, osakekuvaaja kolme vuosia 1971–1974, osakekuvaaja neljä vuosia 1977–1980, osakekuvaaja viisi vuosia 1979–1982 ja osakekuvaaja kuusi vuosia 1983–1986. Kuvaajissa olevat osakehinnat oli skaalattu seuraavasti: kuvaajissa kolme ja neljä (C ja D) hinnat oli jaettu kolmella, kuvaajissa kaksi ja viisi (B ja E) viidellä ja kuvaajassa kuusi (F) kymmellä. Kuvaajassa yksi (A) hinnat olivat alkuperäiset.

v) vaikuttavatko vanhempien sosioekonominen tausta tai vanhempien ja lähimpien ystävien sijoitusharrastus opiskelijoiden ennusteisiin ja käyttäytymiseen.

Tutkimuksen tarkoituksena on siis pääasiallisesti *kuvailla* vastaajien käyttäytymiseen liittyviä ilmiöitä, *selittää* tämän käyttäytymisen syitä taustamuuttujilla (esimerkiksi ikä, sijoituskokemus tai työkokemus) ja *verrata* kahden vastaajaryhmän vastauksia toisiinsa. Opiskelijoiden ja pankin sijoitusasiantuntijoiden odotuksia ei ole aikaisemmin Suomessa tutkittu vastaavalla tavalla, ja tutkimuksen tavoitteena onkin saada lisäymmärrystä siitä, miten yksilöt muodostavat odotuksia tulevista osaketuotoista. Saatuja tuloksia voidaan peilata sijoittajien käyttäytymisestä saatuja aikaisempia tutkimustuloksia vasten. Tutkimuksen kautta voidaan hahmottaa, ovatko vastaajat rationaalisia vai alttiita virheille ja miten mahdolliset virheet ilmenevät. Tulokset antavat oman panoksensa tehokkaiden markkinoiden ja siihen liittyvän rationaalisuuden sekä rahoituksen käyttäytymistieteen välillä käytävään debattiin.

Tutkimuksen rakenne on seuraavanlainen. Luku kaksi esittelee keskeisiä rahoitusteorioita: tehokkaiden markkinoiden hypoteesin ja siihen liittyvän odotetun hyödyn teorian sekä näille vastakkaisen rahoituksen käyttäytymistieteen lähestymistavan, jossa keskeisenä osana on prospektiteoria. Luvussa kolme esitellään tutkimuksen kannalta tärkeimmät käsitteet ja näkökulmat sekä niihin liittyvä aikaisempi tutkimus. Samassa luvussa muodostetaan tutkimuksen hypoteesit. Luvussa neljä selostetaan tutkimusmenetelmän valintaa ja aineiston hankinnan periaatteita. Lisäksi siinä tuodaan esiin, miten aineistoa muokattiin ennen analyysien tekoa ja kuvaillaan aineistoa. Luku viisi on omistettu tulosten analysoinnille ja luku kuusi tutkimuksen luotettavuuden arvioimiselle. Seitsemäs ja viimeinen luku kokoaa tutkimuksen johtopäätökset yhteen.

2 RAHOITUSTEORIAT SIOITTAJIEN KÄYTTÄYTYMISEN KUVAAJINA

Yksi rahoitusmarkkinoiden keskeisimmistä teorioista on Eugene Faman muotoilema tehokkaiden markkinoiden hypoteesi (1970, 1991). Teorian mukaan arvopapereiden hinnat heijastavat täydellisesti kaiken saatavilla olevan informaation. Kun uusi tieto saapuu markkinoille, se leviää nopeasti, ja arvopapereiden hinnat sopeutuvat viiveettä uuden tiedon mukaisiksi. Tehokkaiden markkinoiden hypoteesi liittyy läheisesti teoriaan osakekurssien satunnaiskulusta. Satunnaiskulkuteorian mukaan osakekurssija ei voi ennustaa, sillä huomisen päivän hintamuutokset heijastavat vain huomista tietoa ja ne ovat riippumattomia tämän päivän hintamuutoksista (Kendall & Bradford Hill 1953).

Fama muotoili tehokkaiden markkinoiden hypoteesin ehdot heikkoihin, keskivahvoihin ja vahvoihin ehtoihin. Heikot tehokkuusehdot tarkoittavat, että kaikki historiallinen tieto heijastuu rahoitusinstrumenttien hintoihin siten, ettei tähän tietoon perustuva investointistrategia voi tuottaa ylimääräistä tuottoa. Keskivahvat ehdot tarkoittavat, ettei minkään julkisesti saatavan tiedon avulla ole mahdollista saavuttaa ylimääräistä tuottoa. Vahvat ehdot taas tarkoittavat, ettei millään tiedolla voi muodostaa sijoitusstrategiaa, jonka avulla olisi mahdollista saavuttaa ylimääräistä tuottoa. (Fama 1970, 383.)

Hypoteesin taustalla olevien oletusten mukaan markkinoilla ei ole kaupankäynnistä aiheutuvia kustannuksia, sijoittajat voivat hyödyntää kaikkea markkinatietoa ilmaiseksi ja markkinaosapuolet hyväksyvät kaiken saatavilla olevan tiedon heijastumisen nykyisiin hintoihin (Fama 1970, 387). Sijoittamisen kannalta kysymys markkinoiden tehokkuudesta on hyvin tärkeä. Jos markkinat ovat tehokkaat, se tarkoittaa, etteivät hintojen muutokset ole ennustettavissa. Tällöin sijoittajan kannattaa valita passiivinen sijoitusstrategia ja sijoittaa perinteisen arvopapereiden hinnoitteluteorian (CAP-malli) mukaisesti markkinasalkkuun ja riskittömään kohteeseen.

Tehokkaiden markkinoiden hypoteesi implikoi, että sijoittajat käyttäytyvät rationaalisesti. Rationaalisuus voidaan ymmärtää Bayesin säännön hengessä siten, että sijoittajat päivittävät uskomuksiaan aina sitä mukaa, kun uutta tietoa saapuu markkinoille. Bayesin sääntö on seuraava:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)},$$

jossa $p(A)$ on A:n priori-todennäköisyys, joka ei riipu B:stä, $p(A|B)$ on A:n todennäköisyys ehdolla B (posteriori-todennäköisyys), $p(B|A)$ on B:n todennäköisyys ehdolla A ja $p(B)$ on B:n priori-todennäköisyys. Lisäksi rationaalisuus voi tarkoittaa sitä, että sijoittajien valinnat ovat normatiivisesti hyväksytyjä, koska

ne ovat yhteensopivia odotetun hyödyn teorian kanssa. (Barberis & Thaler 2002, 2.)

Odotetun hyödyn teoria muodostaa normatiivisen lähtökohdan tarkastella riskiä sisältäviä päätöksentekotilanteita. Teorian mukaisesti sijoittajat pyrkivät maksimoimaan hyötyä, eivätkä heidän ennustamansa hinnat sisällä virheitä. Odotetun hyödyn keskeiset käsitteet ja hyötyfunktio konstruointiin John von Neumannin ja Oskar Morgensternin toimesta vuonna 1947. Hyötyfunktion muodostamista varten he tekivät oletuksia tai aksiomia yksilöiden preferensseistä. Preferenssit määriteltiin riskiä sisältävien prospektien kautta. Prospektilla he tarkoittivat lopputuloksia, eräänlaisia onnenpelejä, joihin liittyi todennäköisyys. Aksiomat, jotka asettavat ehdot yksilön käyttäytymiselle, ovat seuraavat:

1. Täydellisyys: oletetaan, että yksilöllä on selvä preferenssi niin, että jokin tulos on mieluisampi kuin jokin toinen. Lisäksi yksilö kykenee aina valitsemaan kahden vaihtoehdon väliltä. Esimerkiksi, jos x on mieluisampi tulos kuin y , niin ($x > y$), tai jos y on mieluisampi tulos kuin x , niin ($y > x$). Yksilö voi olla myös indifferentti x :n ja y :n välillä ($x \sim y$).

2. Transitiivisuus: oletetaan, että yksilön valitessa täydellisyys aksioman mukaisesti, hän valitsee aina johdonmukaisesti. Jos yksilö pitää tulosta x parempana kuin tulosta y ja tulosta y parempana kuin z , on x :n oltava mieluisampi kuin z :n. Jos yksilö on indifferentti x :n ja y :n sekä y :n ja z :n välillä, hänen on oltava indifferentti myös x :n ja z :n välillä. Toisin sanoen, jos $x > y$ ja $y > z$, niin tulee olla $x > z$ sekä, jos $x \sim y$ ja $y \sim z$, niin tulee olla $x \sim z$.

3. Jatkuvuus: oletetaan, että on kolme valintaa, jossa x on mieluisampi kuin tulos y , mutta y on mieluisampi kuin tulos z , niin voidaan löytää todennäköisyys α siten, että yksilö on indifferentti tuloksen y ja pelin suhteen, missä saadaan x todennäköisyydellä α ja z todennäköisyydellä $(1-\alpha)$. Toisin sanoen, jos $x > y \succcurlyeq z$ tai $x \succcurlyeq y > z$, niin voidaan löytää α siten, että $y \sim G(x, z; \alpha)$.

4. Itsenäisyys: oletetaan, että kahdesta pelistä toinen on mielekkäämpi ja vaikka molempiin peleihin lisättäisiin kolmas peli, säilyy preferointijärjestys siitä huolimatta. Jos x on mieluisampi valinta kuin y , α on valinnan todennäköisyys ja z on kolmas peli, niin $\alpha x + (1 - \alpha)z > \alpha y + (1 - \alpha)z$.

Yllä mainituista aksiomista voidaan muodostaa mitattava hyöty (hyötyfunktio). Hyötyfunktio osoittaa järjestyksen: jos vaihtoehdosta x saatava hyöty on suurempi kuin y :stä saatu, $U(x) > U(y)$, niin x on mieluisampi kuin y ($x > y$). Odotetun hyödyn avulla voidaan asettaa riskipitoiset valinnat paremmuusjärjestykseen:

$$U[G(x, y; \alpha)] = \alpha U(x) + (1 - \alpha)U(y).$$

Yleisemmin odotetun hyödyn malli voidaan esittää:

$$E[U(W)] = \sum_i p_i U(W_i),$$

jossa $U(W_i)$ on päätöksentekijän saama hyöty varallisuuden tasosta W_i , ja p_i on todennäköisyys tietylle tulokselle. (Von Neumann & Morgenstern 1967, 26–27; Copeland, Weston & Shastri 2005, 46–50.)

Reaalimaailmassa yksilöt eivät kuitenkaan aina käyttäydy rationaalisesti. Monissa tutkimuksissa on raportoitu, että ihmiset systemaattisesti rikkovat odotetun hyödyn teoriaa ja edellä kuvattuja aksioomia. Kenties kuuluisin näistä on Maurice Allais'n (Allais 1953) tutkimus.

Allais osoitti, että ihmiset rutiininomaisesti rikkoivat itsenäisyyden aksioomaa. Allais tutki todennäköisyysteorian tuntevien ihmisten käyttäytymistä hypoteettisessa pelissä, jonka tunnetuin asetelma oli seuraavanlainen. Osallistujia pyydettiin valitsemaan oheisista peleistä mieluisin.

Peli A: 100 %:n mahdollisuus saada 1 miljoonaa FF

Peli B: 10 %:n mahdollisuus saada 5 miljoonaa FF, 89 %:n mahdollisuus saada 1 miljoonaa FF ja 1 %:n mahdollisuus, ettei saa mitään.

Tämän jälkeen osallistujille esitettiin toiset kaksi peliä, joista pyydettiin valitsemaan mieluisin.

Peli C: 11 %:n mahdollisuus saada 1 miljoonaa FF ja 89 %:n mahdollisuus, ettei saa mitään

Peli D: 10 %:n mahdollisuus saada 5 miljoonaa FF ja 90 %:n mahdollisuus, ettei saa mitään.

Yksilön, joka maksimoi odotettua hyötyä, tulisi valita $A > B$ ja $C > D$, mutta Allais'n tutkimuksessa koehenkilöt valitsivatkin säännönmukaisesti vaihtoehdot $A > B$ ja $C < D$. Odotusarvo pelissä A oli 1 miljoonaa FF ja pelissä B 1,39 miljoonaa FF, joten valitsemalla A:n B:n sijaan koehenkilöt maksimoivat odotettua hyötyä, eivätkä odotusarvoa. Koska useimmat valitsivat $A > B$, voidaan odotetun hyödyn mukaisesti kirjoittaa:

$$u(1) > 0,1 * u(5) + 0,89 * u(1) + 0,01 * u(0), \text{ eli} \\ 0,11 * u(1) > 0,1 * u(5) + 0,01 * u(0).$$

Lisätään $0,89 * u(0)$ molemmille puolille, ja saamme

$$0,11 * u(1) + 0,89 * u(0) > 0,1 * u(5) + 0,9 * u(0),$$

mikä osoittaa, että odotettua hyötyä maksimoivan koehenkilön tulisi valita $C > D$. Koska koehenkilöt kuitenkin valitsivat $C < D$, he maksimoivat odotusarvoa, sillä pelin C odotusarvo oli 110 000 FF ja pelin D 500 000 FF. Tämä oli kuitenkin

ristiriidassa pelien A ja B osalta tehtyjen valintojen kanssa. (Allais 1953, 527–528.)

Monet tutkimukset liittyen sijoittajien käyttäytymiseen ovat omalta osaltaan vahvistaneet poikkeamia odotetun hyödyn teoriasta. Rahoituksen käyttäytymistiede on tieteenala, joka tutkii miten psykologiset ilmiöt vaikuttavat sijoittajien käyttäytymiseen. Sen mukaan jotkut rahoituksen ilmiöt ovat paremmin ymmärrettävissä, kun vähennetään rationaalisuuden vaatimuksia. Rahoituksen käyttäytymistiede ottaa huomioon ne irrationaalisuudet ja virheet, joille sijoittajat altistuvat päätöksentekotilanteissa. (Shefrin 2008, 1; Barberis & Thaler 2002, 2; Baker & Nofsinger 2002, 98.)

Rahoituksen käyttäytymistieteen puolestapuhujien mielestä huonosti informoidut ja ei-sofistikoituneet sijoittajat voivat käyttäytymisellään aiheuttaa markkinoiden tehottomuuden. Tieteenalana kyseinen tiede, joka käyttää tutkimusmetodeinaan usein induktiivisia menetelmiä, kokeita, kyselyitä ja kenttä-tutkimuksia, on melko nuori: ensimmäiset teoreettiset tutkimukset alalta julkaistiin 1980-luvun alussa. Tutkijat kiinnostuivat anomaliaista, jotka eivät selittyneet tehokkaiden markkinoiden teorian avulla. Anomalialla tarkoitetaan rahoituksessa ilmiötä, johon ei löydy rationaalista selitystä. (De Bondt, doĝlu, Shefrin & Staikouras 2008, 7-8; Shiller 2003, 84.)

Rahoituksen käyttäytymistieteen kaksi keskeistä tulosta liittyvät arbitraasin rajallisuuteen ja markkinapsykologiaan. Arbitraattoreilla viitataan rahoituksessa hyvin informoituihin sijoittajiin, jotka hallitsevat fundamentteihin perustuvan tiedon. Arbitraattorit muodostavat rationaalisia odotuksia ja etsivät riskipitoisia sijoituskohteita. Tehokkaisiin markkinoihin sisältyy ajatus, että sijoittaja saa poistettua salkustaan markkinariskin, jos arvopapereille on olemassa täydelliset substituuutit. Arbitraasi on riskitön, koska jos esimerkiksi arvopaperin A hinta laskee alle arvopaperin B hinnan, ja B on täydellinen A:n substituuuti, arbitraattori myy arvopaperin B ja ostaa lisää A:ta, kunnes hinnat ovat tasoittuneet. Tämä ei kuitenkaan ole aina realistista. Jos esimerkiksi osakkeet ja joukkovelkakirjat ovat jostain syystä väärin hinnoiteltuja, eivät ne ole täysin riskittömiä, koska kyseisille kohteille ei ole olemassa täydellisiä substituuutteja. Näin ollen sijoittaja ei saa poistettua salkustaan kokonaan markkinariskiä eli fundamenttiriskiä. Arbitraasin rajallisuudella viitataan lisäksi siihen, että näiden arbitraattoreiden voi olla vaikeaa korjata markkinoilla olevia väärin hinnoitteluja, jotka ei-rationaaliset sijoittajat ovat aiheuttaneet. Esimerkiksi, jos sijoittaja löytää alihinnoitellun kohteen ja ostaa kohdetta, hän altistuu silti sille riskille, että valtaosa sijoittajista valitsee trendinseuraamisstrategian ja jatkaa kohteen myyntiä, jolloin hinta laskee edelleen. Markkinapsykologialla taas viitataan heuristiikkaan ja löydettyihin virheisiin, joita etenkin ei-sofistikoituneet sijoittajat, noise traders, tekevät. Kun virheitä tehdään systemaattisesti ja laajassa mittakaavassa, ne aiheuttavat hintakuplia markkinoille. (Barberis & Thaler 2002, 2-8; Shleifer & Summers 1990, 20–21; De Bondt ym. 2008, 9.) Sijoitusinstrumenttien hinnat siis heijastavat sijoittajien tekemiä virheitä eli virheellisiä päätelmiä tulevista tuotoista ja riskeistä. Nämä tulokset saavat tärkeän merkityksen, kun niitä tarkastellaan viimevuosien talouskehityksen yhteydessä. Esimerkiksi sijoittajan yli-

luottamus voi aiheuttaa liiallista investointihalukkuutta ja piittaamattomuutta markkinoiden selvistä, varoittavista signaaleista.

Rahoituksen käyttäytymistieteen heikkous on se, että siitä puuttuu yksi yhtenäinen ydinteoria, joka nivoisi yhteen keskeiset tulokset (Barberis & Thaler 2002, 15; De Bondt ym. 2008, 15). Prospektiteoriaa pidetään kuitenkin yhtenä tärkeimmistä tämän tieteenalan teorioista. Sen ovat kehittäneet psykologit Daniel Kahneman ja Amos Tversky (Kahneman & Tversky 1979). Teoria osoittaa, että ihmiset systemaattisesti rikkovat odotetun hyödyn teorian oletuksia. Tämän opinnäytetyön teoreettinen viitekehys on juuri prospektiteoria, ja seuraavaksi teoria esitellään tarkemmin.

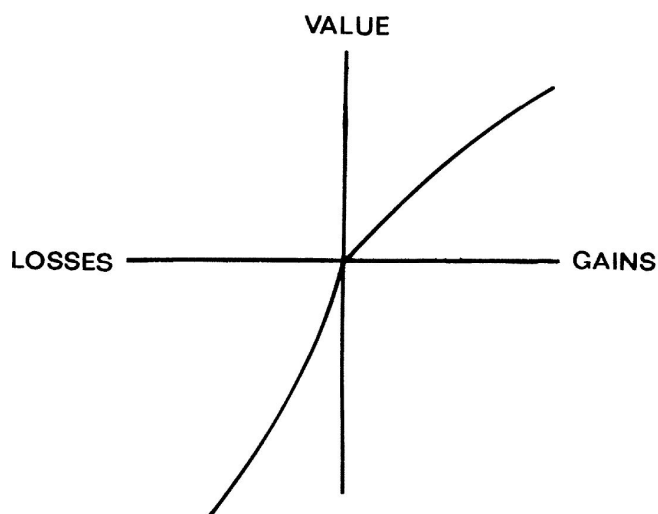
Prospektiteoria pyrkii selittämään sijoittajien käyttäytymistä kognitiivisen psykologian alaan liittyvien kokeiden perusteella. Teorian mukaan sijoittajalle arvoa tuovat voitot ja tappiot suhteessa johonkin referenssipisteeseen (usein henkilön nykyinen vauraus), eikä lopullinen vauraus, kuten perinteisessä odotetun hyödyn teoriassa. Teoria on eräänlainen deskriptiivinen runko, eikä niinkään normatiivinen, ja siinä pyritään selittämään riskipitoisten vaihtoehtojen valintaa. Prospektiteorian mukaan sijoittajan riskinsieto on erilaista riippuen siitä, kokeeko hän olevansa tietyllä hetkellä voitolla vai tappiolla. Yksilöt alipainottavat lopputuloksia, jotka ovat pelkästään todennäköisiä suhteessa lopputuloksiin, jotka ovat saavutettavissa varmuudella. Tämä johtaa riskin karttamiseen valintatilanteissa, joihin liittyy varma voitto ja riskin etsimiseen niissä tilanteissa, joihin liittyy varma tappio. Teoriassa todennäköisyydet on korvattu päätöksenteon painoilla. (Shefrin 2008, 391–392; Kahneman & Tversky 1979, 263.)

Prospektiteoria erottaa valintatilanteesta kaksi eri vaihetta: ensimmäinen vaihe on editointivaihe ja toinen on evaluointivaihe. Editointivaiheessa yksilö muodostaa muistin ja kokemuksen perusteella kuvauksen päätöksentekotilanteesta. Evaluointivaiheessa arvofunktiio ja päätöksentekokertoimet liitetään kuhunkin vaihtoehtoon, ja korkeimman arvon tuova strategia valitaan. Tästä prosessista syntyy prospekti, eli kunkin vaihtoehdon subjektiivinen arvostus painotettuna päätöksentekokertoimella. (Kahneman & Tversky 1979, 274–275.)

Teoriassa on kaksi keskeistä osaa: arvofunktiio eli valintafunktiio (value function) ja painofunktiio eli päätöskerroinfunktiio (weighting function). Olkoot x ja y kaksi mahdollista lopputulosta jollekin sijoitusstrategialle (voitto ja tappio), $v(\cdot)$ funktiio, joka mukaan sijoittaja arvostaa tilaansa, $\pi(p)$ sijoittajan todennäköisyysarviota kuvaava funktiio, p ja q tilojen x ja y todennäköisyydet. Prospektiteorian mukaan sijoittaja valitsee sen strategian, joka tuottaa korkeimman arvon arvofunktiolle:

$$\pi(p)v(x) + \pi(q)v(y),$$

jossa $v(0) = 0$, $\pi(0) = 0$ ja $\pi(1) = 1$.



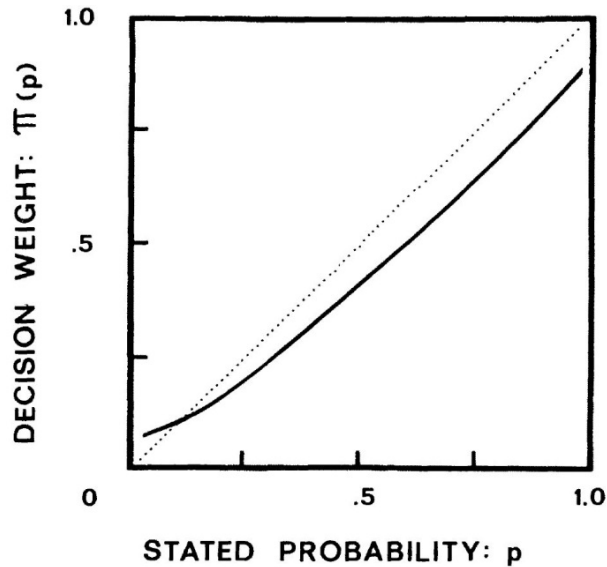
KUVIO 1 Hypoteettinen arvofunktiio (Kahneman & Tversky 1979, 279)

Arvofunktiio on konkaavi voittojen alueella, $v''(x) < 0$, kun $x > 0$, mikä kuvastaa riskin karttamista, ja konveksi tappioiden alueella, $v''(x) > 0$, kun $x < 0$, mikä kuvastaa riskin hakemista. Funktiion muodon perusteella hyvinvoinnin herkkyys on tappion alueella suurempi, toisin sanoen hyvinvointi laskee tappion alueella voimakkaammin. Yksilön kokema harmitus jonkun tietyn summan menettämisestä on suurempi kuin saman summan saamisesta aiheutuva mielihyvä. Nollapisteen kohdalla, eli referenssipisteessä, funktiiossa on kulma ja funktiio kaartuu tappioiden alueella jyrkemmin, $v'(x) < v'(-x)$, kun $x \geq 0$. Referenssipiste edustaa status quo -tilaa tai yksilön nykyisiä varoja. Funktiion yleinen muoto on:

$$v(x) = \begin{cases} x^\alpha, & \text{jos } x \geq 0 \\ -\lambda(-x)^\beta, & \text{jos } x < 0, \end{cases}$$

jossa x on voitto tai tappio suhteessa referenssipisteeseen, λ on tappion karttamisen kerroin, α on riskinkarttamisen kerroin ja β on riskin etsimisen kerroin. (Kahneman & Tversky 1979, 278–279; Shefrin 2008, 404; Kaustia 2003, 2-3.)

Prospektiteoriassa jokaisen lopputuloksen arvo kerrotaan päätöksenteon painoilla. Painot pystytään päättelemään yksilön tekemistä valinnoista ja ne mittaavat tapahtuman vaikutusta tietyn valinnan houkuttelevuuteen. Painofunktion tai päätökskerroinfunktion, π , kulmakerroin välillä $[0,1]$ kuvaa preferenssien herkkyyttä todennäköisyyksien muutoksille. (Kahneman & Tversky 1979, 280 ja 282.)



KUVIO 2 Hypoteettinen painofunktio (Kahneman & Tversky 1979, 282)

Painofunktio esitetään konkaavina kehikossa, jossa vaaka-akselilla on todennäköisyys välillä [0,1] ja pystyakselilla päätöksenteon painot, $\pi(p)$, välillä [0-1]. Yksilöt eivät ota huomioon suoraan objektiivisia todennäköisyyksiä, vaan painoja. Painofunktio on monotonisesti kasvava p :n suhteen ja epäjatkuva kohdissa $p = 0$ ja $p = 1$. Funktion kuvaajan perusteella ilmenee, että ihmiset ylipainottavat pieniä todennäköisyyksiä ja alipainottavat suuria todennäköisyyksiä. (Shefrin 2008, 401-404 ja 407-408.)

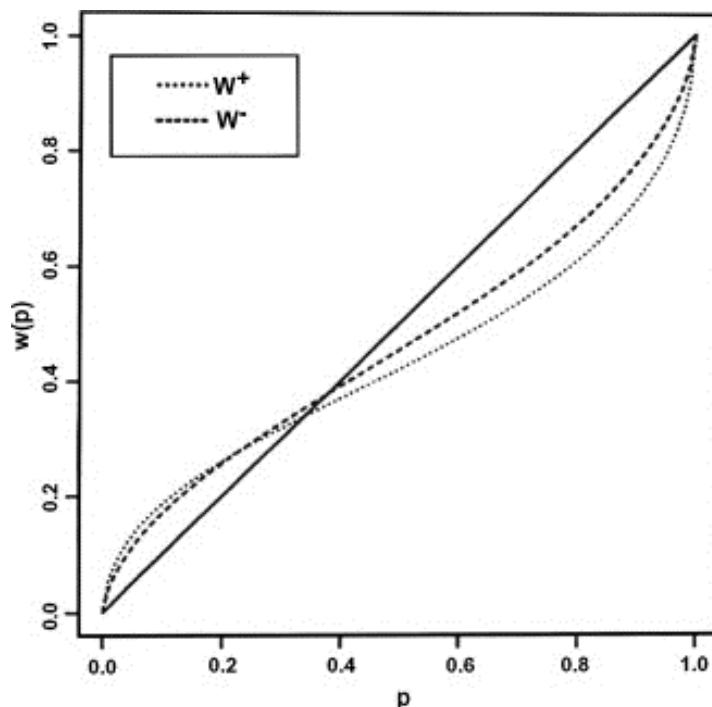
Kahneman ja Tversky modifioivat teoriaansa vuonna 1992, jotta se vastaisi paremmin perinteistä odotetun hyödyn teoriaa. Modifioinnin tuloksena syntyi kumulatiivinen prospektiteoria, jonka mukaan funktion arvo voittojen alueella on:

$$w^+(p) = \frac{p^\gamma}{\sqrt[\gamma]{p^\gamma + (1-p)^\gamma}}$$

ja tappioiden alueella:

$$w^-(p) = \frac{p^\delta}{\sqrt[\delta]{p^\delta + (1-p)^\delta}}$$

Estimointien tuloksena tutkijakaksikko sai seuraavat mediaaniestimaattorit: $\lambda = 2,25$, $\gamma = 0,61$ ja $\delta = 0,69$.



KUVIO 3 Mediaaniestimaattoreihin perustuvat painofunktiot voitoille (w^+) ja tappioille (w^-) (Kahneman & Tversky 1992, 313)

Kuviosta 3 voidaan nähdä, että yksilöt ylipainottavat pieniä todennäköisyyksiä ja alipainottavat kohtalaisia sekä suuria todennäköisyyksiä. Painofunktiot voitoille ja tappioille ovat myös melko lähellä tosiaan, mutta ensiksi mainittu kaa-reutuu hiukan enemmän. Riskin karttaminen voittojen alueella on selvempää kuin riskin etsiminen tappioiden alueella sekä kohtalaisissa että suurissa to-dennäköisyyksissä. Pienien todennäköisyyksien ylipainottaminen selittää esi-merkiksi arpajaisten ja vakuutusten suosiota. (Kahneman & Tversky 1992, 312 ja 316.)

Kahnemanin ja Tverskyn tutkimusten perusteella voidaan muodostaa ne-likenttä, joka kuvaa yksilön suhdetta riskiin (Kahneman & Tversky 1992, 297 ja 306):

TAULUKKO 1 Yksilön suhde riskiin

	Merkittävä voitto	Merkittävä tappio
Suuri todennäköisyys	Riskin karttaminen	Riskin etsiminen
Pieni todennäköisyys	Riskin etsiminen	Riskin karttaminen

Olennaista mille tahansa rahoituksen alaan liittyvälle mallille on yksilön odo-tusten mallintaminen. Rahoituksen käyttäytymistieteen yksi peruspilareista on nojautuminen psykologiaan ja kokeellisten tutkimusten antamiin vastauksiin siitä, miten yksilöt muodostavat uskomuksia. Kahneman ja Tversky (1974) osoittivat, että ihmiset luottavat heuristisiin periaatteisiin, kun heidän tehtävä on esimerkiksi arvioida epävarman tapahtuman todennäköisyyttä. Heuristiikka liittyy intuition ja siihen liittyvien periaatteiden käyttö yksinkertaistaa päätök-

sentekoa, mutta johtaa samalla systemaattisiin virheisiin. Kahneman ja Tversky havaitsivat ja nimesivät psykologisten testien perusteella lukuisia eri virheitä, jotka kuvaavat yksilön käyttäytymistä epävarmoissa, riskiä sisältävissä valintatehtävissä. Tämän tutkimuksen kannalta keskeisiä, rahoituksen käyttäytymistieteen alaan kuuluvia virheitä ovat ankkurointi, trendin seuraaminen, ekstrapolointi, edustavuus ja ylikuottamus. Nämä virheikäyttäytymisen muodot osoittavat omalta osaltaan tehokkaiden markkinoiden hypoteesin puutteelliseksi: hinnat poikkeavat todellisesta arvostaan, koska sijoittajat eivät kaikissa päätöksentekotilanteissa toimi rationaalisesti.

3 TUTKIMUKSIA SIOITTAJIEN EPÄRATIONAALISUUDESTA JA HETEROGEENISUUDESTA

Tässä luvussa käydään läpi tämän tutkimuksen kannalta tärkeitä käsitteitä, jotka rajaavat työn näkökulmaa ja joiden avulla perehdytään aikaisempaan tutkimukseen. Ankkurointiefektiin, trendinseuraamiseen, edustavuusharhaan ja ylikuottamukseen liittyvät tutkimukset haastavat markkinoiden tehokkuuden ja siihen liittyvän odotetun hyödyn maksimoimisen. Ne myös osoittavat, että sijoittajajoukko on heterogeeninen. Tässä luvussa tarkastellaan lisäksi tutkimuksia, jotka kuvaavat sukupuolen, iän, sijoituskokemuksen sekä lähipiiriin vaikutusta yksilön käyttäytymiseen rikispitoisissa valintatilanteissa.

3.1 Ankkurointiefekti

Psykologit Kahneman ja Tversky (1974) tutkivat ankkurointiefektiä kokeella, jossa pyöritettiin ensin pyörää ja sen jälkeen kysyttiin koehenkilöiltä, oliko heidän mielestään YK:ssa olevien afrikkalaisten jäsenvaltioiden määrä prosentteina suurempi vai pienempi kuin pyörän osoittama luku (1-100). Tutkijat huomasivat, että pyörän antama satunnainen luku vaikutti koehenkilöiden vastauksiin. Esimerkiksi, kun pyörä antoi luvun 10, vastaajien keskimääräinen arvio afrikkalaisista jäsenmaista oli 25 %, ja kun pyörä antoi luvun 60, oli vastaajien arvio keskimäärin 45 %. Pyörän antama satunnainen luku ankkuroi vastaajien arvioita ja veti niitä juuri nähtyä lukua kohti.

Rahoituksen käyttäytymistieteessä ankkurointiefektillä tarkoitetaan sijoittajan tekemää virhettä, jossa sijoittaja hyödyntää päätöksenteossaan ja ennusteissaan jotakin irrelevanttia asiaa tai arvoa, joka on ennestään tuttu, vaikka tällä ei olisi mitään tekemistä varsinaisen sijoituspäätöksen kanssa. Sijoittaja voi esimerkiksi ankkuroitua hintaan, jolla osti osakkeen.

Rahoituksessa ankkurointiefektin yhteyttä osake-ennusteisiin on tutkittu verrattain paljon. Japanilaisia osakemarkkinoita ovat tutkineet esimerkiksi Nakazono (2012) ja Fujiwara, Ichiue, Nakazono & Shigemi (2013). Nakazono sai selville, että suuri osa Japanin osakemarkkinoilla toimivista sijoittajista painotti merkittävästi menneitä ennustearvoja. Hänen mukaansa ankkurointiefekti oli riippuvainen siitä, mihin yhtiöön sijoittaja kuului. Fujiwara ym. puolestaan tutkivat ammattisijoittajien ennustekäyttäytymistä ja osoittivat, että ennustekäyttäytyminen oli erilaista osake- ja velkakirjamarkkinoilla ja että markkinoille osallistuneet olivat taipuvaisia ankkurointiefektiin.

Suomalaiset Kaustia, Alho ja Puttonen (2008) tutkivat 300 rahoitusalan ammattilaisten ja yli 200 yliopisto-opiskelijoiden ennusteita osaketuotoista. He havaitsivat merkittävän ankkurointiefektin opiskelijoiden keskuudessa. Sijoi-

tusasantuntijoiden osalta ankkurointiefekti oli pienempi, mutta tilastollisesti merkitsevä. Ankkurointiefekti ilmeni siten, että lähtöarvot vaikuttivat tehtyihin ennusteisiin. Kaustia ym. raportoivat myös, että sijoitusasiantuntijat eivät olleet tietoisia historiallisten tuottojen vaikutuksista heidän ennusteisiinsa.

Yhdysvaltalaiset tutkijat Campbell ja Sharpe (2009) puolestaan tutkivat ammattilaisten talousennusteita käyttäen hyväksi Money Market Servicen (MMS) kokoamaa aineistoa. He raportoivat, että ammattilaisten tekemissä kuukausittaisissa talousjulkaisuissa olevat arvot olivat ankkuroituneita edellisten kuukausien arvioihin. Esimerkiksi vähittäiskaupan myynnin arvoissa 25 %:n osuus oli ennustettavissa edellisten kuukausien arvojen perusteella.

Edellisten lisäksi Cen, Hilary ja Wei (2013) havaitsivat amerikkalaisen aineiston perusteella ankkurointiefektin analyytikkojen osinkotuottoennusteissa. Analyytikkojen osake-ennusteet olivat optimistisempia, kun yhtiön ennustettu osinkotuotto (EPS = Earnings per share) oli teollisuuden keskiarvoon nähden matala, ja päinvastoin. Ennusteet olivat siis ankkuroituneet kohti teollisuuden keskiarvo-osaketuottoa.

Tämän tutkimuksen kannalta tärkeimmät ankkurointiefektiin liittyvät tutkimukset ovat De Bondtin tutkimukset (De Bondt 1993). De Bondt suoritti useita tutkimuksia, joista osa tehtiin yliopisto-opiskelijoille ja osa ammattisijoittajille. Tutkimuksissa vastaajia pyydettiin ennustamaan sekä piste-ennusteet että luottamusväliennusteet. Kuuluisaan Livingstonin aineistoon perustuen De Bondt huomasi, että ammattisijoittajat olivat taipuvaisia ennustamaan markkinoiden täyskäännöstä, jolloin ennusteet seurasivat keskiarvotuottoa, kun taas muissa tutkimuksissa ammatilliset ja opiskelijat näyttivät käyttäytyvän trendinseuraajan tavoin. (Shefrin 2008, 79–80.)

Kahnemanin ja Tverskyn (1974) mukaan ankkurointiefekti selittää suurelta osin sen, miksi luottamusvälit ovat usein liian kapeita. De Bondt oletti, että kokeelliseen tutkimukseen osallistuneiden vastauksia ankkuroi kaksi tekijää. Ensimmäinen ankkuri muodostui odotetusta, menneiden hintamuutosten jatkumosta (osakekuvaajan jyrkkyys). Toinen ankkuri muodostui keskiarvohinnasta, joka oli enemmän tai vähemmän irrallinen viimeisimmistä hintamuutoksista. Esimerkiksi, jos oletetaan jonkun yrityksen osakkeen hinnan nousevan kuudessa viikossa 23 dollarista 41 dollariin, ja nousua oletetaan olevan kolme dollaria viikossa, yksilön paras ennuste seitsemännen viikon hinnaksi olisi luultavasti 44 dollaria. De Bondt oletti kuitenkin, että keskiarvo 32 dollaria toimisi ikään kuin menneiden hintojen edustajana ja vetäisi luottamusvälejä kapeammaksi. (De Bondt 1993, 357.)

Ankkurointiefektiin liittyvien aikaisempien tutkimusten perusteella voidaan muotoilla tähän tutkimukseen liittyvä ensimmäinen hypoteesi:

H₁: Osake-ennusteet ankkuroituvat osakkeen hintakehityksen keskiarvoa kohti.

3.2 Positive feedback traders, trendinseuraajat

Monet sijoittajat ovat taipuvaisia etsimään trendiä menneestä hintakehityksestä ja odottavat trendin jatkuvan myös tulevaisuudessa. Näitä sijoittajia kutsutaan trendin seuraajiksi, positive feedback traders (De Long, Shleifer, Summers & Waldmann 1990). Trendin seuraajat ostavat hintojen noustessa ja myyvät hintojen laskiessa. Trendin seuraamiseen liittyy käsite ekstrapolointi, mikä tarkoittaa, että esimerkiksi jonkin funktion kuvaajan muotoa arvioidaan niin, että voidaan havaita arvoja, joihin alkuperäinen kuvaaja ei yllä.

Trendin seuraamiseen liittyvistä aikaisemmista tutkimuksista tämän työn kannalta tärkeimpiä ovat De Bondtin (1993) tutkimukset. Niiden mukaan suurin osa opiskelijoista ennusti osakkeiden hintojen nousun tai laskun jatkuvan, eli he ekstrapoloivat kuviossa näkyviä trendejä. De Bondt jakoi opiskelijat ennusteiden perusteella vielä vahvoihin ja heikkoihin trendin seuraajiin. Vahvat trendin seuraajat ennustivat nousu- tai laskutrendin jatkuvan sekä seitsemän että 13 kuukautta, kun taas heikot trendin seuraajat ennustivat saman trendin jatkuvan seitsemän, mutta ei enää 13:a kuukautta. (Shefrin 2008, 50–55.) Ammattisijoittajatkin olivat taipuvaisia seuraamaan trendiä, mutta osa heistä ennusti myös markkinoiden täyskäännöstä. Ammattisijoittajien ennustamat osakkehinnat olivat korkeampia kuin opiskelijoiden ennustamat hinnat, ja heidän ennusteissaan oli enemmän vaihtelua laskukausien jälkeen. (Shefrin 2008, 56–59.)

Seuraavien hypoteesien muodostamiseksi oletetaan De Bondtin esimerkin mukaisesti, että P_0 on viimeksi tiedetty indeksin hintataso ja F on yksilön piste-estimaatti. Siten odotettu hintamuutos EPC , expected price change, on yhtä kuin $F - P_0$. Jos vastaajat keskimäärin odottavat menneen trendin jatkuvan, voidaan hypoteesipariksi muotoilla:

$$H_{20}: \overline{EPC}_{\text{Bull}} = \overline{EPC}_{\text{Bear}}$$

$$H_{21}: \overline{EPC}_{\text{Bull}} > \overline{EPC}_{\text{Bear}},$$

eli keskimääräisen hintamuutoksen nousukausien jälkeen tulisi olla suurempi kuin keskimääräisen hintamuutoksen laskukausien jälkeen. (De Bondt 1993, 357.)

Hypoteesipari liittyy De Bondtin luomaan teoriaan luottamusvälien suojauksesta, hedging theory of confidence intervals. Toinen osa teoriaa liittyy luottamusvälien tutkimiseen. Oletetaan, että H olisi yksilön ennustaman todennäköisyysjakauman yläarvo, L ala-arvo ja F piste-estimaatti. Siten ylempi luottamusväli UCI (upper confidence interval) = $H - F$ ja alempi luottamusväli LCI (lower confidence interval) = $F - L$. Näin ollen vinous, Δ , on UCI :n ja LCI :n erotus eli $\Delta = H + L - 2F$. Jos yksilö arvioi tulevaisuuden hintojen jakauman olevan symmetrinen hänen ennustamansa piste-estimaatin ympärillä, niin $\Delta = 0$, eli tällöin $F = (H + L) / 2$. Jos yksilö taas odottaa hinnan olevan lähempänä H :ta kuin L :ää, niin $\Delta < 0$ ja edelleen, jos yksilö odottaa tulevan hinnan olevan lä-

hempänä L:ää, niin $\Delta > 0$. Jos useimmat vastaajat ekstrapoloivat mennyttä trendiä, niin:

$$H_{30}: \bar{\Delta}_{\text{Bull}} = \bar{\Delta}_{\text{Bear}}$$

$$H_{31}: \bar{\Delta}_{\text{Bull}} < \bar{\Delta}_{\text{Bear}},$$

eli keskimääräisen vinouden tulisi olla pienempi nousukausien jälkeen kuin laskukausien jälkeen. (De Bondt 1993, 357–358.)

De Bondt huomasi teoriaansa liittyen, että koehenkilöt suojasivat ennustettaan: tulevien hintojen subjektiiviset todennäköisyysjakaumat olivat vasemmalle vinoja, jos suurta hinnannousua ennustettiin. Näin ikään kuin implisiittisesti otettiin huomioon mahdollinen hintojen romahdus. De Bondt selitti tutkimuksissaan vinouden parametreja vastaavilla odotetuilla hintamuutoksilla ja trendinseurantaluokituksella ja havaitsi, että vinouden parametrien etumerkki oli aina vastakkainen odotetun hintamuutoksen parametreihin nähden. Tämän perusteella voidaan muodostaa seuraava hypoteesi (De Bondt 1993, 360 ja 368–369):

H₄: Luottamusvälit kasvavat enemmän oikealle vinoiksi, kun suurempaa hinnannousua ennustetaan ja toisinpäin.

Muradoğlu (2002) hyödynsi tutkimuksessaan De Bondtin alkuperäisen tutkimuksen mallia ja suoritti kokeellisen tutkimuksen turkkilaisten yliopistop opiskelijoiden ja sijoitusalan ammattilaisten keskuudessa. Hän pyysi tutkimukseen osallistuvilta ennusteet piste-estimaatteina ja luottamusväliennusteina, aivan kuten De Bondt. Muradoğlun tutkimus erosi De Bondtin tutkimuksesta siten, että tutkimukseen osallistuville annettiin lisätehtäväksi tehdä reaali maailman ennusteet todellisista pörssiosakkeista. Lisäksi ennusteet tuli laatia lyhyelle, keskipitkälle ja pitkälle aikavälille. Muradoğlu raportoi, että ammattisijoittajat ennustivat hintamuutosta pitkällä aikavälillä. Lyhyellä aikavälillä ammattisijoittajien käyttäytyminen oli samanlaista kuin De Bondtin tutkimuksessa opiskelijoilla, eli he olivat trendin seuraajia. Muradoğlun mukaan luottamusvälit olivat yleisesti ottaen suurempia pitemmän aikavälin ennusteissa. Yksilöt spekuloiivat ennustaessaan nousukauden jatkumista ja arvioivat korkeammat ylävaihteluvälit. Toisaalta he suojautuivat ennustaessaan hintamuutosta laskukausien jälkeen arvioimalla pienempiä luottamusvälejä. Muradoğlu raportoi lisäksi, että indeksiennusteissa noviisit, eli opiskelijat, olivat optimistisia kaikissa tapauksissa ja kaikilla ennustehorisonteilla. He ennustivat nousukauden jatkuvan ja hintojen lähtevän nousuun laskukausien jälkeen. Pörssiosakkeiden ennusteiden kohdalla myös ammattilaiset olivat optimistisia. Ammattilaisilla oli lisäksi taipumus osoittaa suurempaa optimismia nousukausilla ja pitemmillä aikaväleillä. Yksilöt odottivat nousukausien jälkeen suurempaa hinnannousua ja spekuloiivat asettamalla korkeammat ylemmät luottamusvälit.

Trendinseuraamisilmiötä on edellä mainittujen lisäksi havainnut tutkimuksissaan muun muassa Bange (2000). Hän tutki piensijoittajien osakesalkku-

jen muutoksia American Association of Individual Investorsin keräämän aineiston perusteella. Hän havaitsi, että osakesalkkujen omistuksissa tapahtuneet muutokset heijastivat trendin seuraamista. Sijoittajat ostivat ja myivät osakkeita, jos niiden hintojen liikkeet pystyttiin ekstrapoloimaan tulevaisuuteen.

3.3 Edustavuusharha

Trendin seuraaminen ja ekstrapolointi aiheutuvat edustavuuden harhasta. Kahneman ja Tversky (1974) kysyivät koehenkilöiltä, mikä on todennäköisyys, että asia A kuuluu luokkaan B, tai mikä on todennäköisyys, että prosessi B tuottaa tapahtuman A. Koehenkilöiden vastaukset osoittivat, että ihmisten arvioihin vaikutti se, missä määrin A heijasteli B:n ominaisuuksia, tai missä suhteessa A muistutti B:tä. Kahneman ja Tversky huomasivat myös, että käyttäytyessään edustavuusharhan mukaisesti koehenkilöt luottivat stereotypioihin tehdessään arvioita todennäköisyyksistä ja rikkoivat Bayesin sääntöä jättämällä huomiomatta aikaisemmat todennäköisyydet, eli kiinnittävät huomion vain säännön jaettavan ensimmäiseen termiin. Edustavuus ilmeni myös silloin, kun henkilöiden oli arvioitava jonkun ominaisuuden todennäköisyyttä perusjoukosta poimitussa otoksessa. He arvioivat otoksen ominaisuuksia perusjoukon ominaisuuksien perusteella, eikä otoksen koolla ollut tällöin merkitystä. Edustavuutta kuvasti myös virheellinen ajattelutapa ruletinpeluussa: jos punainen oli osunut kohdalle monta kertaa peräkkäin, koehenkilöt pitivät seuraavaa pysähtymiskohtaa todennäköisemmin mustana kuin punaisena. Tämä ilmiö tunnetaan yleisesti uhkapelaajan harhana, gambler's fallacy. (Kahneman & Tversky 1974, 1124–1125.)

Kahnemanin ja Tverskyn mukaan osake-ennusteissa edustavuuden harha voi näkyä taustainformaation heijastumisena ennusteisiin. Esimerkiksi jos sijoittajalle annetaan yhtiötä koskevaa, hyvin myönteistä informaatiota, se voi heijastua korkeana tuottoennusteena. Henkilöt voivat luottaa pelkästään annetun tiedon suosiollisuuteen, eivätkä kiinnitä huomiota sen luotettavuuteen tai tarkkuuteen. Edustavuuden harha vaikuttaa taloudellisiin ennusteisiin väärin, koska ennusteet tehdään etsimällä samankaltaisuuksia menneestä kehityksestä. (Kahneman & Tversky 1974, 1126.)

Edustavuusharhaa ovat tutkineet esimerkiksi Nguyen ja Schüßler (2012); Chen, Kim, Nofsinger ja Rui (2007); sekä Dhar ja Kumar (2001). Nguyen ja Schüßler tutkivat saksalaisten yksityisijoittajien käyttäytymiseen liittyviä virheitä. He saivat selville, että sijoittajat arvioivat todennäköisyyksiä väärin edustavuuden harhan takia ja että korkeampi koulutus vähensi tätä taipumusta. Chen ym. puolestaan tutkivat kiinalaisen aineiston pohjalta yksityisten ja institutionaalisten sijoittajien altistumista rahoituksen käyttäytymistieteen paljastamille virheille. He saivat selville, että kiinalaiset sijoittajat olettivat menneiden tuottojen olevan indikaatio tulevasta tuotoista, mikä kertoi edustavuuden harhasta. Tämä ilmeni siten, että kiinalaiset sijoittajat olivat taipuvaisia ostamaan viimeaikoina menestyneitä osakkeita. Chen ym. havaitsivat myös, ettei koke-

mus merkittävästi vähentänyt sijoittajien tekemiä virheitä. Dhar ja Kumar sen sijaan tutkivat yhdysvaltalaisien kotitalouksien kaupankäyntipäätöksiä ja saivat selville, että sijoittajien osto- ja myyntipäätöksiin vaikutti lyhyen ajan menneet hintatrendit. Mennyt hintatrendi näytti aiheuttavan edustavuuden harhan tuleviin hintoihin.

Kaiken kaikkiaan edustavuuden harha näyttää niin läheisesti liittyvän trendin seuraamiseen, ettei siitä muodosteta erillistä hypoteesia.

3.4 Yliluottamus

De Bondt (1993) huomasi tutkimuksissaan, että vastaajat olivat tyypillisesti yli luottavaisia. Suhtautuminen riskiin käy ilmi luottamusvälin leveydestä. Yliluottavaiset yksilöt asettavat luottamusvälit, jotka ovat liian kapeita.

Sijoittajien yliluottamusta ovat De Bondtin lisäksi tutkineet muun muassa Odean (1998); Barber ja Odean (1999, 2001); Van de Verter ja Michayluk (2008) sekä Glaser, Langer ja Weber (2005). Odeanin (1998) mukaan yliluottamus johtaa riskin aliarvioimiseen. Barber ja Odean (1999) päättelivät, että yliluottamuksen vuoksi sijoittajat olivat liian vakuuttuneita omista arvioistaan, eivätkä siksi huomioineet muiden mielipiteitä. Sijoittajat myös näkivät omien päätöstensä sisältävän vähemmän riskiä. Van de Verter ja Michayluk raportoivat, että australialaisten rahoitusalan suunnittelijoiden laatimiin tulevaisuuden ennusteisiin liittyi virheitä yliluottamuksen vuoksi. Glaser, Langer ja Weber puolestaan tutkivat yliluottamusta saksalaisten rahoitusmarkkinoiden ammattilaisten (esimerkiksi traiderit ja investointipankkiirit) ja opiskelijoiden keskuudessa. Heidän tutkimuksensa johtopäätös oli, että ammattilaisten luottamusväliennusteet olivat kapeampia kuin opiskelijoiden.

Kaustia, Lehtoranta ja Puttonen (2013) puolestaan tutkivat rahoitusneuvojien testituloksia ja saivat selville, että sofistikoituneemmat sijoitusneuvonantajat olivat taipuvaisia tekemään alhaisempia tuottoennusteita. Heidän mukaansa sijoitusneuvojien työkokemus käyttäytyi ennusteissa oppimiskäyrän mukaisesti: ensimmäinen työkokemusvuosi näytti opettavan enemmän kuin toinen ja niin edelleen. Myös Vissing-Jørgensen (2003) havaitsi kokemuksen yhteyden alhaisempiin tuottoennusteisiin.

Edellä olevien tutkimusten perusteella voidaan olettaa, että yliluottamusta esiintyy myös tämän tutkimuksen vastaajajoukossa. Muodostetaan hypoteesiksi:

H₅: Pankin sijoitusasiantuntijat ovat yliluottavaisempia kuin opiskelijat.

3.5 Koulutus, sijoituskokemus ja lähipiiri

Rahoituksen näkökulmasta koulutus ja sijoituskokemus määritellään usein sillä, kuinka sofistikoitunut sijoittaja on. Mitä sofistikoituneempi sijoittaja on, sitä

enemmän hänellä on kokemusta ja taitoa arvioida sijoitusinstrumentteihin liittyviä riskejä. Tutkimukset ovat osoittaneet, että rahoituksen alaan liittyvä yleisivistys on yhteydessä sijoittajan käyttäytymiseen ja osallistumiseen. Riley ja Chow (1992) raportoivat riskin karttamisen ja koulutuksen välisestä käänteisestä riippuvuudesta. Rahoituksen alaan liittyvän koulutuksen ja riskipitoisiin arvopapereihin sijoittamishalukkuuden välillä näyttää yleisesti olevan positiivinen yhteys (esimerkiksi Hibbert, Lawrence & Prakash 2013; Joo & Grable 2005; Christiansen, Joensen & Rangvid 2008 sekä Bernheim, Skinner & Weinberg 2001). Hibbert ym. (2013) tutkivat kahden aineiston perusteella korkeasti koulutettujen naisten ja miesten eroja suhtautumisessa riskiin. Ensimmäisen tutkimuksen tulosten mukaan (Federal Reserve Board's Survey of Consumer Finances:n aineisto) koulutus ei näyttänyt yleisesti ottaen vähentävän naisten ja miesten eroja suhtautumisessa riskiin. Toisen tutkimuksen mukaan (yhdysovaltalaisen yliopistojen professoreihin liittyvä aineisto) suhtautumisessa riskiin ei ollut eroja samantasoisien koulutuksen hankkineiden naisten ja miesten välillä. Jälkimmäisen aineiston perusteella korkeampi koulutus näytti lisäävän mieltymystä riskinottoon ja lieventävän sukupuolten välisiä eroja suhteessa riskin karttamiseen. Christiansen ym. puolestaan osoittivat tanskalaisen aineiston perusteella, että ekonomistin koulutuksen saaneet henkilöt sijoittivat muun koulutuksen saaneita henkilöitä enemmän osakkeisiin.

Van Rooijn, Lusardin ja Alessien (2011) tutkimuksen mukaan ne henkilöt, joilla oli vähäinen rahoitukseen alaan liittyvä sivistys, olivat haluttomampia osallistumaan osakesijoittamiseen. Calvet, Campbell ja Sodini (2009) puolestaan havaitsivat ruotsalaisen aineiston perusteella, että rikkaat, paremmin koulutetut kotitaloudet olivat vähemmän taipuvaisia tekemään sijoittamiseen liittyviä virheitä kuin vähemmän koulutetut, köyhemmät ja maahanmuuttajakotitaloudet. Sofistikaation puute siis näytti liittyvän sijoitusvirheisiin. Alexander ja McElreath taas (1999) tutkivat, vaikuttaako koulutus siihen, kuinka hyvin opiskelijat ennustavat markkinoiden käyttäytymistä. He osoittivat, että ne opiskelijat, jotka saivat keskimääräistä parempia arvosanoja sijoittamisen peruskurssilta, laativat tarkempia ennusteita. Heikommin suoriutuneet opiskelijat olivat pessimistisempiä ennusteiden laatimisessa. Alexanderin ja McElreathin tulokset osoittavat, että koulutus voi parantaa opiskelijoiden ennustekykä. Myös lukio-opintojen suorittamisen ja rahoituksen alaan liittyvän sivistyksen välille on löydetty positiivinen yhteys (Lusardi, Mitchell & Curto 2010). Yllä olevien tutkimusten perusteella voidaan johtaa viides hypoteesi:

H₆: Paremmin koulutetut henkilöt tekevät vähemmän virheitä kuin heikommin koulutetut.

Sijoituskokemus näyttää koulutuksen ohella vähentävän sijoittajan tekemiä virheitä (Feng & Seasholes 2005; Shapira & Venezia 2001). Feng ja Seasholes saivat selville, että sijoittajien sofistikoituneisuus ja kaupankäyntikokemus eliminoivat sijoittajien halukkuuden realisoida tappioita. Ilmiö tunnetaan dispositioefekti-

nä². Heidän mukaansa sofistikoituneet sijoittajat olivat 67 % vähemmän taipuvaisia dispositioefektiin kuin keskivertosijoittajat, ja kaupankäyntikokemus lievensi 72 % tätä taipumusta. Shapira ja Venezia osoittivat, että ammattisijoittajien koulutus ja kokemus voivat vaikuttaa dispositioefektiä vähentävästi. Kaustia ym. (2008) puolestaan huomasivat, ettei ankkurointiefekti hävinnyt ammattilaisten ennusteista, vaikka se olikin heikompa kuin opiskelijoilla. Muodostetaan seuraavaksi hypoteesiksi:

H7: Sijoituskokemus vähentää sijoittajien tekemiä virheitä.

Tiedetään lisäksi, että vanhempien ja muun lähipiirin, esimerkiksi ystävien, yhteys henkilön rahoitusalan sivistykseen voi olla suuri ja sillä on mahdollisesti vaikutusta henkilön käyttäytymiseen. Lusardi, Mitchell ja Curto (2010) havaitsivat, että perheen ja etenkin äidin sofistikoituneisuudella oli suuri vaikutus siihen, miten laaja rahoitusalaan liittyvä sivistys nuorella oli. Mandell (2008) raportoi, että niillä lukiolaisilla, joiden vanhemmatkin olivat käyneet lukion, oli korkeampi rahoituksen alaan liittyvä sivistys. On myös osoitettu, että osa rahoitusalaan liittyvästä sivistyksestä voi siirtyä suoraan vanhemmilta lapsille (esimerkiksi Chiteji & Stafford 1999 sekä Li 2009). Lusardi ja Mitchell (2006) sekä van Rooij, Lusardi ja Alessie (2011) saivat selville, että rahoitukseen liittyvissä päätöksissä ystävien ja kollegoiden merkitys oli suuri. Tiedon jakamista ja oppimista voi tapahtua merkittävästi sosiaalisen vuorovaikutuksen kautta. Hongin, Kubikin ja Steinin (2004) sekä Brownin, Ivkovicin, Smithin ja Weisbennerin (2008) saamat tulokset osoittavat lisäksi, että lähipiiri vaikuttaa vahvasti esimerkiksi salkun valintaan.

Vanhempien sosioekonominen asema vaikuttaa lasten koulutusvalintoihin inhimillisen pääoman akkumulaatioteorian mukaisesti. Sosiologian alalla perinteinen näkemys on ollut pysyvä eriarvoisuus, mikä tarkoittaa sitä, että vanhempien matala sosioekonominen asema on yhteydessä lasten alhaiseen koulutustasoon (esimerkiksi Shavit & Blossfield 1993). Breen, Luijkx, Müller ja Pollak (2005) tutkivat kahdeksan eurooppalaisen valtion aineistolla sosiaaliluokasta johtuvaa eriarvoisuutta 1970-luvulta 2000-luvun alkuun ja raportoivat eriarvoisuuden laskeneen selvästi Ruotsissa, Alankomaissa, Saksassa ja Ranskassa. Suomi ei ollut tuossa tutkimuksessa mukana, mutta myös Suomessa erilaisista kotitaustoista lähtöisin olevien koulutukselliset erot ovat vähentyneet. Esimerkiksi yliopisto-opintoihin osallistumisella mitattuna koulutuksellinen tasa-arvo on lisääntynyt. Kivinen, Hedman ja Kaipainen (2012) raportoivat, että vuonna 1970 odds ratio, joka tarkoittaa ikäluokkien suhteellisia eroja mahdollisuuksissa osallistua opintoihin, oli 19,1, kun taas vuonna 2010 se oli 6,8. Tämä tarkoittaa, että akateemisesti koulutettujen vanhempien lapset päätyivät vuonna 2010 yliopistoon opiskelemaan 6,8 kertaa useammin kuin ei-akateemisten vanhempien.

Voidaan siis olettaa, että suurin osa tämän tutkimuksen kyselyyn vastaavista opiskelijoista on kotoisin perheistä, joissa vanhempien sosioekonominen

² Dispositioefektillä tarkoitetaan rahoituksen käyttäytymistieteessä epärationaalisuutta, jossa sijoittajilla on taipumus realisoida voittoja mieluummin kuin tappioita.

status on korkea. Edellä olevien tutkimusten perusteella voidaan myös loogisesti päätellä, että korkea sosioekonominen asema on yhteydessä hyvään rahoituksen alaan liittyvään sivistykseen ja sitä kautta sijoittamiseen (esimerkiksi Christiansen ym. 2008). Voidaan siis päätellä, että ne henkilöt, joiden vanhempien sosioekonominen asema on korkea ja joiden vanhemmat ja lähipiiri harrastavat sijoittamista, ovat muita vastaajia sofistikoituneempia ja tekevät vähemmän sijoittamiseen liittyviä virheitä.

Hs: Ne vastaajat, joiden vanhemmillä on korkea sosioekonominen asema ja joiden vanhemmat ja lähipiiri harrastavat sijoittamista tekevät vähemmän sijoittamiseen liittyviä virheitä.

3.6 Sukupuoli ja ikä

Sukupuolen yhteyttä sijoittajan käyttäytymiseen on tutkittu paljon. Barber ja Odean (2001) tutkivat osakesijoittamista ja saivat selville, että miehet ovat yli-luottavaisempia kuin naiset. Miehet myös kävivät kauppaa enemmän kuin naiset ja pitivät hallussaan riskisempiä positioita. Powell ja Ansic (1997) huomasivat vakuutuksia koskevassa tutkimuksessaan, että naiset kaihtoivat riskiä enemmän kuin miehet. Naiset ja miehet myös omaksuivat erilaisia strategioita päätöksenteossa. Myös Felton, Gibson ja Sanbonmatsu (2003) raportoivat samankaltaisia tuloksia: miehet tekivät riskipitoisempia sijoituspäätöksiä kuin naiset. Jianakoplos ja Bernasek (1998) tutkivat naisten ja miesten eroja riskinsiedossa ja saivat selville, että naiset pitävät halussaan miehiä pienempää osuutta riskipitoisemmissa arvopapereissa. Schubert, Brown, Gysler ja Brachinger (1999) eivät puolestaan havainneet sijoitus- ja vakuutusvalintoja käsitelleessä tutkimuksessaan, että miehet tekisivät riskipitoisempia rahoituspäätöksiä kuin naiset.

Voidaan olettaa, että ikä on yhteydessä sijoituskokemukseen: mitä vanhempi yksilö on kyseessä, sitä enemmän hänellä on sijoituskokemusta ja sitä sofistikoituneempi hän on ja sitä vähemmän hän tekee virheitä sijoituspäätöksiinsä liittyen. Dorn ja Huberman (2002) havaitsivat kuitenkin, että sijoitusmarkkinoille osallistuvat sijoittajat ovat useimmiten koulutettuja, noin nelikymppisiä miehiä, eli eivät kovin vanhoja. Agarwal, Driscoll, Gabaix ja Laibson (2009) raportoivat, että rahoituksen alaan liittyvät virheet ovat yleisimpiä nuorilla ja vanhemmilla henkilöillä. Näillä ryhmillä oli kaikista heikoin rahoituksen alaan kuuluva yleissivistys. Van Rooij, Lusardi ja Alessie (2011) saivat niin ikään tuloksia, joiden mukaan rahoituksen alaan liittyvä korkeampi sivistys oli parhainta keski-ikäisillä henkilöillä. Kyseisessä tutkimuksessa myös osakeomistus oli suurinta keski-ikäisillä.

Taloustieteilijät, biologit ja antropologit ovat tutkineet elinkaariomuuttujien, kuten iän, vaikutuksia riskiä sisältäviin valintoihin. Näiden tutkimusten perusteella ikääntymisen pitäisi olla yhteydessä vähentyneeseen riskinottohalukkuuteen. (Mata, Josef, Samanez-Larkin & Hertwig 2011, 19.) Mata ym. tulkitsivat

kuitenkin, että ikään liittyvät eroavuudet yksilöiden välillä vaihtelivat tehtävästä riippuen, kun kyseessä oli riskipitoiset valinnat. Tehtävien edellyttäessä kokemuksen perusteella oppimista, vanhemmat ihmiset osoittautuivat enemmän riskiä hakeviksi kuin nuoremmat ihmiset silloin, kun oppiminen johti riskiä karttavaan käytökseen, mutta olivat enemmän riskiä karttavia silloin, kun oppiminen johti riskiä etsivään käyttäytymiseen. Kun tehtävät eivät edellyttäneet kokemusta, ei vanhempien ja nuorempien ihmisten välillä ollut eroja riskikäyttäytymisessä.

Tiedetään myös, että ikääntymiseen liittyvä joidenkin peruskognitiivisten taitojen heikentyminen voi johtaa muutoksiin riskikäyttäytymisessä. Esimerkiksi työmuistin heikentyminen iän myötä voi vaikuttaa strategian valintoihin. Vanhemmat ihmiset näyttävät hyödyntävän helpompia päätöksenteonstrategioita, jotka edellyttävät vähemmän tietojen yhdistelemistä. Vanhemmat ihmiset tekevät lisäksi enemmän virheitä kun useita tiedonpalasia pitää yhdistellä. Motivaatioteoriat ovat myös tuoneet esille, että ikääntyminen voi johtaa kasvaneeseen emotionaaliseen säätelyyn, mikä voi johtaa virheisiin tiedon käsittelyssä. On esimerkiksi osoitettu, että vanhemmat aikuiset osoittivat reagoivansa vähemmän tappioihin kuin nuoremmat aikuiset, mutta myös vastakkaisesta käytöksestä on todisteita. (Mata ym. 2011, 20–21; Hibbert, Lawrence & Prakash 2013, 9.)

Tässä tutkimuksessa ollaan kiinnostuneita siitä, miten sukupuoli ja ikä vaikuttavat sijoittajan käyttäytymiseen. Muotoillaan yllä olevien tutkimusten perusteella seuraavat hypoteesit:

H₉: Miehet ovat ylikuottavaisempia kuin naiset.

H₁₀: Keski-ikäiset tekevät kaikista vähiten virheitä osake-ennusteissa.

3.7 Heterogeeniset sijoittajat

Kaikki edellä esitetyt käsitteet ja aikaisempi tutkimus osoittavat, ettei sijoittajia enää pidetä homogeenisena joukkona. Rahoituksen käyttäytymistieteen tutkimustulosten myötä on yhä enemmän alettu kiinnittää huomiota rahoitusmarkkinoiden mikrorakenteeseen. Se tosiseikka, etteivät kaikki sijoittajat aina käytädy rationaalisesti, on johtanut kasvavaan kiinnostukseen markkinoilla toimivia erilaisia sijoittajia kohtaan. Sijoittajat ovat heterogeenisiä: heillä on erilainen osaamistausta ja he käsittelevät eri tavoilla tietoa. Heillä on myös erilaiset preferenssit ja odotukset esimerkiksi tulevista tuotoista. Tämä on ristiriidassa odotetun hyödyn ja perinteisen arvopapereiden hinnoitteluteorian, CAP-mallin kanssa. Niissä molemmissa oletetaan, että kaikki sijoittajat jakavat samat, homogeeniset odotukset.

1980-luvun lopulta lähtien on muotoiltu runsaasti arvopapereiden hintojen dynamiikkaa kuvaavia behavioralistisia malleja, joissa lähtökohtana ovat heterogeeniset agentit (Heterogeneous Agent Models, HAMs). Monet näistä malleista perustuvat heuristiikkaan ja sijoittajien rajallisen rationaalisuuden peri-

aatteeseen. Yksinkertaiset HAM-mallit voivat selittää esimerkiksi väliaikaisten hintakuplien syntyä, äkillisiä romahduksia tai liikaa volatiilisuutta. (Hommes 2005.)

HAM-malleihin liittyviä ja tämän tutkimuksen kannalta tärkeitä käsitteitä ovat vielä häirikkösijoittaja, noise trader, ja toisinajattelija, contrarian investor. Häirikkösijoittajat ovat ns. tavallisia sijoittajia, jotka ovat trendin seuraajia, ylitai alireagoijia ja huonoja ajoittajia. Nämä sijoittajat näyttävät reagoivan erilaisiin pseudo-singnaaleihin, eivätkä he käyttäydy rationaalisesti. Joskus arbitraattorit saattavat seurata häirikkösijoittajien trendinseuraamista ja tämä voi aiheuttaa markkinoille hintakuplia lyhyellä aikavälillä (Shleifer & Summers 1990, 23 ja 28–29.) Toisinajattelija taas tarkoittaa sijoittajaa, joka tekee sijoituspäätöksiä vastoin markkinatrendejä, eli hän ostaa silloin, kun osakkeiden hinta on alhaalla ja myy silloin, kun hinnat ovat ylhäällä. Joskus toisinajattelijastrategiaa kutsutaan arvostrategiaksi (value strategy), jossa sijoittajat etsivät osakkeita, joiden reaaliarvo alhaisempi kuin kirjapitoarvo. Lakonishok, Shleifer ja Vishny (1994) tutkivat arvostrategian toimivuutta vuosina 1968–1990 ja saivat selville, että kyseinen strategia tuotti merkittävästi parempia vuosituottoja kuin investoiminen niin sanottuihin glamour-osakkeisiin³. Osa De Bondtinkin tutkimuksen (1993) ammattisijoittajista ennusti markkinoiden täyskäännöstä ja olivat siis toisinajattelijoita. Tämän tutkimuksen kannalta voidaan teoriaan liittyen selvittää, osoittautuvatko opiskelijat enemmän häirikkösijoittajiksi kuin pankin sijoitusasiantuntijat. Muotoillaan hypoteesiksi:

H₁₁: Opiskelijat käyttäytyvät häirikkösijoittajien tavoin.

De Bondtin tutkimukseen (1993) perustuen voidaan loogisesti ajatella, että pankin sijoitusasiantuntijat olisivat rationaalisia tai informoituja vastaajia, eli sofistikoituneita. Tässä työssä yhtenä rationaalisuuden ja sofistikoituneisuuden osoituksena voidaan pitää tehokkaiden markkinoiden hypoteesin tuntemista. On mielenkiintoista nähdä, mitkä tekijät pankin aineistossa selittävät sitä, tuntevatko vastaajat tehokkaiden markkinoiden hypoteesin vain eivät. Ei muotoilla tästä kuitenkaan omaa hypoteesia.

³ Glamour-osake tarkoittaa osaketta, johon liittyy voimakas kasvuodotus.

4 MENETELMÄ JA AINEISTO

Ensimmäiseksi tässä luvussa käydään läpi tutkimuksen menetelmälliset ratkaisut ja aineiston hankinnan periaatteet. Sen jälkeen käsitellään aineistojen muokkausta ja tuodaan esille muun muassa vastauksissa ilmenneiden väärinymmärrysten käsittelyperiaatteet. Lisäksi aineistoja kuvaillaan käyttäen hyväksi taulukoita ja kuvioita. Tämän jälkeen tuodaan esille, miten opiskelija-aineistosta puuttuva tieto on käsitelty. Lopuksi paneudutaan osake-ennusteiden normaalisuuden testaukseen ja työssä käytettyyn normalisointiratkaisuun.

4.1 Tutkimusmenetelmän valinta

Tutkimuksella on aina jokin tarkoitus tai tehtävä. Tutkimusstrategialla tarkoitetaan tutkimuksen menetelmällisten ratkaisujen muodostamaa kokonaisuutta. Tutkimuksen tarkoitus ja tehtävät määräävät tutkimusstrategian valinnan. (Hirsjärvi, Remes & Sajavaara 1997, 126 ja 134.) Kuten tämän työn johdannosta ja tutkimuskysymyksistä käy ilmi, on tutkimusote pääasiassa kuvaileva ja selitettävä. Tutkimusstrategiaksi valittiin tämän perusteella kyselytutkimus. Kysely eli survey-tutkimukselle on tyypillistä, että tietyistä ihmisjoukosta poimitaan otos yksilöitä ja otokseen poimituilta kerätään tietoa standardoidussa muodossa eli kaikilta samalla tavalla (Hirsjärvi ym. 1997, 130).

Tämän työn varsinainen tutkimusmenetelmä on kvantitatiivinen, eli kyseessä on määrällinen tutkimus, jossa sovelletaan tilastollisia menetelmiä. Kvantitatiivisessa tutkimuksessa keskeistä on johtopäätösten ja hypoteesien esittäminen aiemman tutkimuksen perusteella, käsitteiden määrittely, aikaisemmat teorit, aineiston keruun suunnitelmat siten, että havaintoaineisto soveltuu määrälliseen mittaamiseen, otantasuunnitelmat, tulosten saattaminen taulukkomuotoon ja päätelmien teko tilastollisia menetelmiä hyväksikäyttäen (Hirsjärvi ym. 1997, 137). Tässä opinnäytetyössä käytettäviä tilastollisia menetelmiä ovat t-testit ja niiden parametrittomat vastineet, varianssianalyysit, korrelaatiotarkastelut sekä regressioanalyysit.

4.2 Aineiston hankinta

Tutkimuksen aineisto muodostuu siis kahdesta survey-menetelmällä kerätystä aineistosta: opiskelija-aineistosta ja Osuuspankin sijoitusasiantuntijoiden aineistosta. Aineistot kerättiin eri tavoilla.

Opiskelija-aineisto kerättiin luentojen aikana paperisella kyselylomakkeella (liite 1), koska vastauksia oli tällöin odotettavissa enemmän kuin jos olisi käytetty...

tetty esimerkiksi satunnaisotantaa. Opiskelija-aineisto ei näin ollen täytä otoksen kriteerejä, joten kyseessä on harkinnanvarainen näyte (Vehkalahti 2008, 46). Etukäteen valittiin kyselyn toteuttamista varten seuraavat Jyväskylän yliopiston kauppakorkeakoulun syyslukukaudella 2013 järjestämät luontokurssit: Johdatus rahoitukseen (KTTP 320), Kustannuslaskenta (YLA 220) ja Advanced Financial Economics (KTTS 330). Nämä kurssit päätettiin valita, koska tutkimukseen haluttiin saada mukaan mahdollisimman paljon sellaisia opiskelijoita, jotka tulisivat jatkossa opiskelemaan rahoituksen kursseja, tai jotka tulisivat tulevaisuudessa työskentelemään rahoitukseen liittyvien tehtävien parissa. Laskentatoimen opiskelijoiden saamista mukaan tutkimukseen pidettiin tärkeänä, koska osake-ennusteet liittyvät keskeisesti myös laskentatoimen tutkimusalueisiin. Useissa laskentatoimen alaan liittyvissä tutkimuksissa ollaan kiinnostuneita tekijöistä, jotka vaikuttavat ennusteiden tarkkuuteen ja edelleen pääomamarkkinoiden odotuksiin tulevista osinkotuotoista (esim. Clement 1999). Lisäksi yksi kysely toteutettiin viikkoa myöhemmin vielä Syventävä mikrotaloustiede II -kurssin harjoituksissa (KTTS 157), jotta enemmän rahoituksen kursseja käyneitä opiskelijoita saataisiin mukaan kyselyyn.

Opiskelijoille tarkoitettua kyselyä testattiin ennen varsinaisen kyselyn toteuttamista kansantaloustieteen aineopinnoihin kuuluvassa seminaarissa lokakuussa 2013. Testikyselyyn osallistui 22 opiskelijaa, joilla lähes kaikilla oli pääaineena kansantaloustiede. Testikyselystä saadun palautteen perusteella kyselylomakkeen kysymyksiin tehtiin muutamia muutoksia, lähinnä kysymyksenasettelun selkeyttämiseen liittyen.

Varsinaiset opiskelijakyselyt toteutettiin marraskuussa viikoilla 45 ja 46. Kyselyt pidettiin heti luento- ja harjoituksen alussa, eli ne olivat niin sanottuja informoituja kyselyitä (Hirsjärvi, Remes & Sajavaara 1997, 192–193). Tutkimuksen tekijä oli itse paikan päällä ohjeistamassa ja toteuttamassa kyselyn. Näin varmistettiin, että kaikki saivat saman sisältöisen alkuinformaation. Alkuinformaatio annettiin suullisesti ja siinä käytiin läpi kyselyn tarkoitus ja annettiin lyhyet, mutta selkeät ohjeet vastaamiseen. Vastaamiseen annettiin aikaa noin 20 - 25 minuuttia. Vastaajien ei sallittu tehdä kyselyyn vastaamisen aikana yhteistyötä. Ennen kuin kysymyslomakkeita alettiin kerätä pois, varmistettiin vielä kysymyksellä, ettei kukaan ollut vastaaminen kesken. Mikäli näin oli, annettiin vastaamiseen lisää aikaa maksimissaan viisi minuuttia.

Osuuspankin sijoitusasiantuntijoille suunnattu kysely (liite 2) siirrettiin internetissä vastattavaksi kyselyksi Webropol-ohjelmalla. Tutkimukseen saatiin mukaan Etelä-Karjalan, Pohjois-Savon ja Keski-Suomen Osuuspankit. Näissä pankeissa kyselyn kohderyhmäksi valittiin sijoitusneuvojat, -päälliköt ja johtajat, joille pankin sisäinen yhteyshenkilö lähetti linkin, jota kautta kyselyyn pystyi vastaamaan. Kyselylinkkiä pidettiin auki 12.11. - 20.12.2013. Tutkimuksessa mukana olleiden pankkien yhteyshenkilöiden ilmoitusten mukaan kysely oli välitetty eteenpäin 106 henkilölle: Etelä-Karjalan Osuuspankissa 31:lle, Pohjois-Savon Osuuspankissa 21:lle ja Keski-Suomen Osuuspankissa 54:lle. Vastausprosentin kohottamiseksi muistutus- ja kannustusviestejä lähetettiin kyselylinkin

aukioloaikana kolme kappaletta. Myöskään pankin aineisto ei täytä satunnaisotoksen kriteereitä.

Aineistojen hankinnassa käytettiin kannustimena sekä kirjapalkintoa että lahjakorttia, koska niiden uskottiin lisäävän motivaatiota vastata kyselyyn ja vähentävän pilavastausten määrää. Kirjapalkinnot arvottiin erikseen sekä opiskelija- että pankin vastausten kesken. Lahjakortin sai molemmissa ryhmissä vastaaja, jolla ennustevirheen neliösumma oli pienin.

Taloustieteilijöiden keskuudessa yleinen käsitys on, että kokeelliseen tutkimukseen osallistuvat henkilöt eivät tee mitään ilmaiseksi ja että he tarvitsevat kannustimia suoriutuakseen paremmin. Kannustimien käyttämisestä koekes- telmissä on kuitenkin olemassa vaihtelevia tuloksia. Camerer ja Hogarth (1999) kävivät läpi 74 koetta, joissa käytettiin eriasteisia rahallisia kannustimia. He raportoivat, että joissakin tapauksissa kannustimet paransivat koehenkilöiden suoriutumista (esimerkiksi muistiin ja mieleen painamiseen liittyvät tehtävät), kun taas joissakin tapauksissa kannustimilla ei ollut merkitystä keskiver- tosuoriutumisen kannalta (esimerkiksi riskipitoisia valintoja sisältävät tehtävät). De Bondtin (1993) alkuperäisessä tutkimuksessa oli käytetty rahallisia kannus- timia, joten tässäkin työssä niitä päädyttiin käyttämään.

4.3 Opiskelija-aineiston muokkaus

Opiskelijavastauksia saatiin yhteensä 146 kappaletta. Yhteensä kuusi henkilöä jätti tyhjän vastauslomakkeen, joten vastausaktiivisuutta voitaneen pitää erin- omaisena. Luennoilta myöhästyneitä henkilöitä (yhteensä viisi henkilöä) ei otet- tu mukaan lopulliseen vastaajajoukkoon.

Vastauslomakkeissa oli puuttuvia tai epäselviä tietoja seuraavasti. Neljä henkilöä oli merkinnyt pohjakoulutukseksi sekä ylioppilastutkinnon ja ammat- tikoulun. Tämän tulkittiin tarkoittavan kaksoistutkintoa ja se lisättiin muuttu- jiin. Kaksi vastaajaa ei ollut merkinnyt kokonaisopintopistemääräänsä. Näitä puuttuvia opintopistemääriä ei paikattu mitenkään. Samoin kaksi vastaajaa ei ollut ilmoittanut ollenkaan äitinsä sosioekonomista asemaa, ja kuusi vastaajaa ei ollut merkinnyt ollenkaan tai oli merkinnyt epäselvästi isänsä sosioekonomi- sen aseman⁴. Näitä puuttuvia tietoja ei muokattu mitenkään, vaan ne jätettiin tyhjiksi.

Joissakin opiskelijavastauksissa oli ilmeisesti ymmärretty luottamusvä- liennuste väärin, sillä ylä- ja alaluottamusvälit olivat menneet väärinpäin 14 opiskelijalla, toisin sanoen yläluottamusväli merkitty oli pienemmäksi kuin ala- luottamusväli. Nämä väärinymmärrykset näyttivät kasaantuvan suurimmaksi osaksi samoille vastaajille ja ne korjattiin oikeinpäin⁵. Yhdessä opiskelijavasta- uksessa oli ensimmäiselle osakkeelle (A) annettu seitsemän kuukauden piste-

⁴ Epäselvällä tarkoitetaan tässä sitä, että vastaaja oli esimerkiksi merkinnyt kaksi luokkaa isänsä sosioekonomiseksi asemaksi.

⁵ Virheellisiä luottamusväliennusteita oli yhteensä 37 kappaletta. Kaikkiaan luottamusvä- liennusteita oli koko kyselyssä 4×6 osaketta $\times 146 = 3504$ kappaletta.

estimaatiksi 80 ja 13 kuukauden piste-estimaatiksi 105. Luottamusväliennusteiksi oli kuitenkin merkitty 13 ja 18 (seitsemän kuukauden ala- ja yläluottamusvälit) sekä 11 ja 28 (13 kuukauden ala- ja yläluottamusvälit). Tämä epäloogisuus korjattiin siten, että seitsemän kuukauden yläluottamusväliksi merkittiin 98 (80 + 18) ja alaluottamusväliksi 67 (80 - 13). 13 kuukauden luottamusvälit muutettiin samalla periaatteella.

Yksi vastaaja oli antanut neljännelle osakkeelle (D) seitsemän kuukauden piste-estimaatiksi tasan neljä. Vastaaja oli ilmeisesti tarkoittanut arvoksi neljäkymmentä jotain, joten tämä tulkittiin kirjoitusvirheeksi. 13 kuukauden estimaatti oli 40 ja luottamusvälit muutenkin loogisia, joten tuo seitsemän kuukauden epälooginen estimaatti korjattiin neljännen osakkeen seitsemän kuukauden piste-estimaatin koko vastaajajoukon keskiarvolla (46,7).

Opiskelijoiden antamia ennusteita oli yhteensä $146 \times 6 \times 6$ eli 5256 kappaletta. Näistä 40 oli epäloogisia vastauksia, joissa piste-estimaatti ei mahtunut annetun luottamusvälin sisään. Ongelmana näissä tapauksissa oli se, ettei voitu tietää, oliko opiskelija ennustanut piste-estimaatin vai luottamusvälin väärin, joten nämä kohdat jätettiin aineistossa ennalleen.

Ennen varsinaisia analyysyjä muuttujamuunnoksia ja ryhmittelyjä tehtiin melko vähän. Opiskelijat ilmoittivat pääaineekseen 16 eri pääainetta, joista muodostettiin kuusi ryhmää tiedekuntajakoa noudattaen. Ryhmä 1. koostui kauppatieteilijöistä (sisältäen pääaineet KTT, YLA, YJO, YMA, YRI ja ne yrityksen taloustieteiden opiskelijat, jotka eivät olleet vielä valinneet pääainettaan), ryhmä 2. koostui humanisteista (sisältäen pääaineet RU ja HIS), ryhmä 3. koostui yhteiskuntatieteilijöistä (sisältäen pääaineet VO, YHTKPOL, FIL ja SOS), ryhmä 4. koostui informaatioteknoloogeista (pääaineet TJT ja TT), ryhmä 5. koostui matemaattisluonnontieteellisen tiedekunnan opiskelijoista (pääaineet MA ja TILA) ja ryhmä 6. koostui liikuntatieteilijöistä (pääaine VTE).⁶ Aineiston kuvailua varten opiskelijoita ryhmiteltiin esimerkiksi suoritettujen kokonaisopintopisteiden ja suoritettujen rahoituksen kurssien perusteella. Syntymävuoden perusteella muodostettiin uusi IKÄ-muuttuja. Muut muuttujamuunnokset, uudet ryhmittelyt ja uusien muuttujien luonti on kerrottu tarkemmin analyysien yhteydessä.

4.4 Pankin sijoitusasiantuntijoiden aineiston muokkaus

Osuuspankin sijoitusasiantuntijoille suunnatun kyselyn tulokset siirrettiin analysoitavaksi SPSS-ohjelmaan. Vastauksia kertyi yhteensä 37 kappaletta eli vastausprosentti oli noin 35. Kyselyn linkki oli kuitenkin avattu 88 kertaa lähettä-

⁶ Lyhenteet tarkoittavat seuraavaa: KTT = kansantaloustiede, YLA = laskentatoimi, YJO = johtaminen, YMA = markkinointi, YRI = yrittäjäyys, RU = ruotsi, HIS = historia, VO = valtio-oppi, YHTKPOL = yhteiskuntapolitiikka, FIL = filosofia, SOS = sosiologia, TJT = tietojärjestelmätiede, TT = tietotekniikka, MA = matematiikka, TILA = tilastotiede ja VTE = valmennus- ja testausoppi.

mättä kyselyä, joten on mahdollista, että kyselyn pituus vaikutti negatiivisesti vastaamismotivaatioon.

Sijoitusasiantuntijoiden vastuksissa ei ollut puuttuvia vastuksia, koska kysely oli laadittu Webropol-ohjelmalla siten, että kaikki kysymykset olivat pakollisia. Näin ollen kyselyssä eteneminen vaati oikeantyyppisen vastuksen.

Osake-ennusteista löytyi virheellisyyksiä tai väärinymmärryksiä, kuten opiskelijoillakin. Kaiken kaikkiaan sijoitusasiantuntijat antoivat $6 \times 6 \times 37 = 1332$ ennustetta. Kaikkiaan 39 ennustetta oli epäloogista siinä mielessä, että annettu piste-estimaatti ei mahtunut annetun luottamusvälin sisään. Kolmesta ennusteesta luottamusvälit korjattiin oikein päin, eli muutettiin yläluottamusväli suuremmaksi kuin alaluottamusväli. Yksi vastaaja oli systemaattisesti ymmärtänyt kaikkien kuuden osakkeen kohdalla luottamusvälin erona piste-estimaattiin. Nämä korjattiin lisäämällä tai vähentämällä annettu luottamusväli estimaattista.⁷ Korjausten jälkeen aineistoon jäi vielä 24 edellä kuvatuslaista epäloogista ennustetta. Näitä ei korjattu mitenkään, kuten ei opiskelijoidenkaan kohdalla.

Pankin aineistoon tehtiin aineiston kuvailua varten vähäisiä muuttujamuunnoksia. Vastaajia ryhmiteltiin iän, rahoitusalan työkokemuksen ja nykyisessä tehtävässä saadun työkokemuksen perusteella. Analyysien yhteydessä on myöhemmin esitetty muut muuttujamuunnokset, uusien muuttujien luonti ja ryhmittelyt.

4.5 Opiskelija-aineiston kuvailu

Ennen kuin paneudutaan varsinaisesti aineistojen kuvailuun, on paikallaan tehdä muutama huomautus ristiintaulukointien esittämiseen liittyen. Yleinen tapa ristiintaulukoinnin esittämisessä on, että selitettävä eli riippuva muuttuja valitaan rivimuuttujaksi ja selittävä eli riippumaton muuttuja sarakemuuttujaksi. Tässä ja seuraavassa alaluvussa esitettyjen ristiintaulukointien osalta ei jako selittävään ja selitettävään muuttujaan ole yksiselitteistä, joten yllä mainittu sääntö ei välttämättä päde. Kyse on vain aineiston kuvailusta, jolloin muuttujien sijoittelussa ristiintaulukoinnissa ei ole taulukon havainnollistavuuden kannalta niin suurta merkitystä.

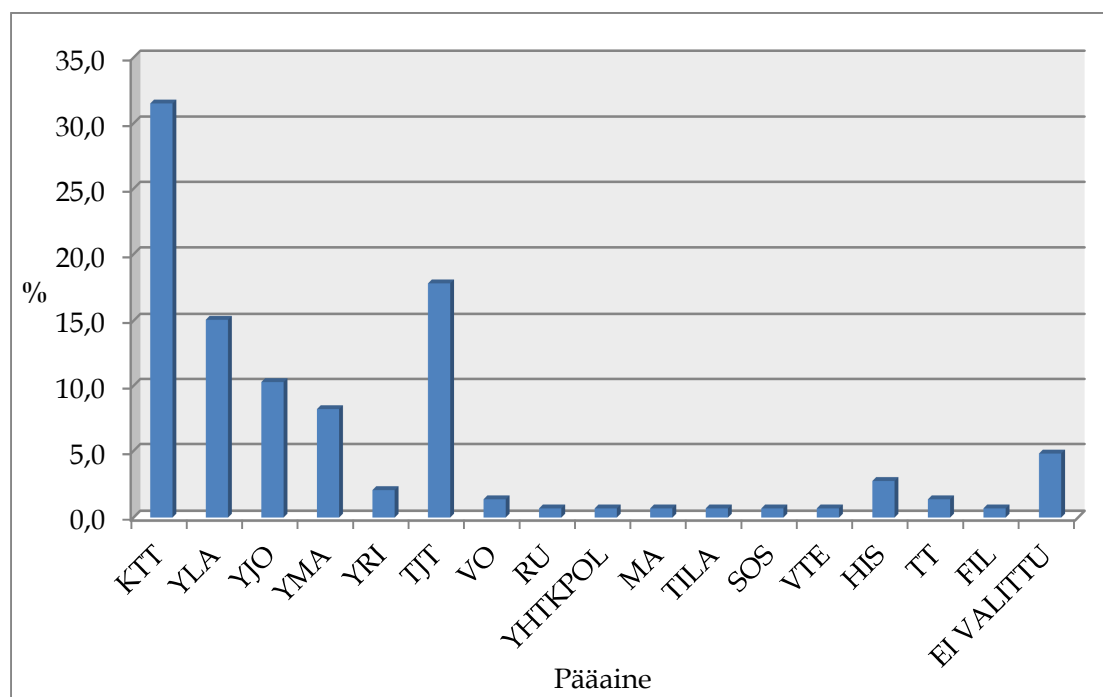
Opiskelijavastaajista 39,7 % oli naisia ja 60,3 % miehiä. Vastaajien keski-ikä oli reilut 24 vuotta, minimi 19 ja maksimi 51 vuotta. Taulukosta 2 nähdään opiskelijoiden jakautuminen ikäryhmittäin: suurin osa vastaajista oli 21 - 25 -vuotiaita.

⁷ Esimerkiksi seitsemän kuukauden piste-estimaatiksi oli annettu 95 ja luottamusväleiksi 10 ja viisi. Yläluottamusväleiksi korjattiin 105 ja alaluottamusväleiksi 90.

TAULUKKO 2 Opiskelijoiden jakautuminen ikäryhmiin

Ikäryhmä	N	%	Kumulatiivinen prosentti
<= 20	5	3,4	3,4
21 - 25	115	78,8	82,2
26 - 30	18	12,3	94,5
31 - 35	4	2,7	97,3
36 - 40	1	,7	97,9
41 - 45	1	,7	98,6
46 - 50	1	,7	99,3
51+	1	,7	100,0
Yhteensä	146	100,0	

Kuviosta 4 ja taulukosta 3 nähdään opiskelijoiden jakautuminen pääaineittain ja ryhmiteltynä tiedekunnittain. Kolme suurinta vastaajaryhmää pääaineittain mitattuna olivat kansantaloustieteen (31,5 %), tietojärjestelmätieteen (17,8 %) ja laskentatoimen (15,1 %) opiskelijat. Seuraavista pääaineista vastaajia oli vain yksi henkilö: ruotsi, yhteiskuntapolitiikka, matematiikka, tilastotiede, sosiologia, valmennus- ja testausoppi ja filosofia. Tiedekunnittain ryhmiteltynä vastaajissa oli ylivoimaisesti eniten kauppatieteilijöitä (105 henkeä), kuten taulukosta 3 voidaan nähdä.

KUVIO 4 Opiskelijoiden jakautuminen pääaineittain⁸

⁸ Kuviossa käytetyt lyhenteet on selitetty alaviitteessä 6. Lisäksi EI VALITTU tarkoittaa yrityksen taloustieteiden opiskelijoita, jotka eivät olleet vielä valinneet pääainettaan.

TAULUKKO 3 Opiskelijoiden jakautuminen tiedekunnittain

	N	%	Kumulatiivinen prosentti
Kauppatieteilijät	105	71,9	71,9
Humanistit	5	3,4	75,3
Yhteiskuntatieteilijät	5	3,4	78,8
Informaatioteknologit	28	19,2	97,9
Matemaattisluonnontieteilijät	2	1,4	99,3
Liikuntatieteilijät	1	,7	100,0
Yhteensä	146	100,0	

Taulukosta 4 puolestaan nähdään, että suurimmalla osalla vastaajista oli pohjakoulutuksena ylioppilastutkinto, eikä vastaajissa ollut mukana kuin muutama ammattikorkeakoulututkinnon suorittanut.

TAULUKKO 4 Opiskelijoiden pohjakoulutus

	N	%	Kumulatiivinen prosentti
Ylioppilastutkinto	130	89,0	89,0
Ammattikorkeakoulu-tutkinto	3	2,1	91,1
Aikaisempi yliopistotutkinto	9	6,2	97,3
Kaksoistutkinto	4	2,7	100,0
Yhteensä	146	100,0	

Opiskelijoiden suorittamat opintopisteet vaihtelivat välillä 0-500. Keskimäärin vastaajilla oli kyselyn teettämisen aikaan opintopisteitä n. 115. Taulukossa 5 on esitetty opiskelijoiden jakautuminen luokkiin suoritettujen opintopisteiden perusteella. Taulukosta nähdään, että suurimmalla osalla vastaajista on ollut kyselyn teettämisen aikaan 100 opintopistettä tai alle (82 henkeä).

TAULUKKO 5 Opiskelijoiden jakautuminen luokkiin suoritettujen opintopisteiden perusteella

Opintopisteet yhteensä	N	%	Validi prosentti	Kumulatiivinen prosentti
<= 50	33	22,6	22,9	22,9
51 - 100	49	33,6	34,0	56,9
101 - 150	25	17,1	17,4	74,3
151 - 200	13	8,9	9,0	83,3
201 - 250	12	8,2	8,3	91,7
251 - 300	5	3,4	3,5	95,1
301 - 350	3	2,1	2,1	97,2
351 - 400	2	1,4	1,4	98,6
401 - 450	1	,7	,7	99,3
451+	1	,7	,7	100,0
Yhteensä	144	98,6	100,0	
Puuttuu	2	1,4		
Yhteensä	146	100,0		

Keskimäärin kyselyyn vastanneet opiskelijat olivat suorittaneet 4,68 opintopistettä rahoituksen kurssseja. Vain 12 opiskelijalla niitä oli suoritettuna 20 opinto-

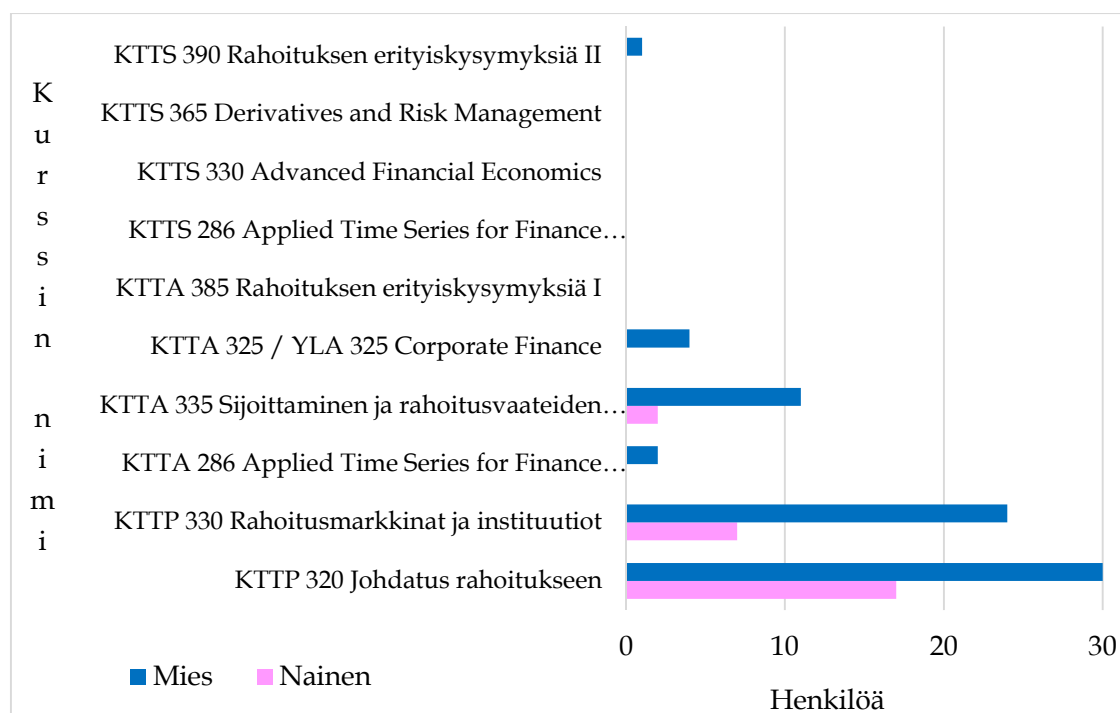
pistettä tai enemmän. Rahoituksen opintojen maksimiopintopistemäärä olisi ollut 54 opistopistettä. Kyselylomakkeeseen valitut kurssit ja niistä laskettu rahoituksen kurssien maksimiopintopistemäärä saatiin ottamalla huomioon Jyväskylän yliopiston kauppakorkeakoulun opinto-oppaassa lukuvuodelle 2010 - 2013 ilmoitetut kurssit sekä ne kurssit, jotka uuden opinto-oppaan (lukuvuosille 2013 - 2016) mukaan luennottiin syksyllä 2013.

Taulukko 6 esittää opiskelijoiden jakautumisen luokkiin suoritettujen rahoituksen kurssien perusteella. Valtaosalla (65,8 %) vastaajista ei ollut suoritettuna yhtään rahoituksen kursseja kyselyn teettämisen aikoihin. Tähän vaikuttaa osaltaan myös se, että kyselylomake ei ottanut huomioon keskeneräisiä kursseja ja esimerkiksi Johdatus rahoitukseen -kurssi oli kyselyn teettämisen aikoihin loppusuoralla.

TAULUKKO 6 Opiskelijoiden jakautuminen ryhmiin suoritettujen rahoituksen kurssien perusteella

Rahoituksen kurssien opintopisteet	N	%	Kumulatiivinen prosentti
0	96	65,8	65,8
1 - 6	3	2,1	67,8
7 - 12	17	11,6	79,5
13 - 18	17	11,6	91,1
19 - 24	8	5,5	96,6
25+	5	3,4	100,0
Yhteensä	146	100,0	

Kuviossa 5 on esitetty suoritettujen rahoituksen kurssit sukupuolittain.



KUVIO 5 Suoritettujen rahoituksen kurssit sukupuolittain

Kuviosta 5 nähdään, että miesopiskelijat olivat suorittaneet enemmän rahoituksen kursseja kuin naisopiskelijat ja että vain yksi opiskelija oli kyselyn teettämisen aikoihin ehtinyt suorittaa syventäviä opintoja (KTTS 390 Rahoituksen erityiskysymyksiä II -kurssin).

Taulukoissa 7 ja 8 on esitetty, kuinka hyviä arvosanoja opiskelijat olivat saaneet rahoituksen kursseista. Taulukon 7 tyydyttävästi tarkoittaa, että opiskelijan rahoituksen kursseista saamat arvosanat ovat olleet keskimäärin 1-2, hyvin vastaa arvosanaa 3 ja erittäin hyvin vastaa arvosanoja 4-5. Taulukosta 8 nähdään, että miehet olivat oman ilmoituksensa perusteella saaneet enemmän erinomaisia arvosanoja kuin naiset.

TAULUKKO 7 Opiskelijoiden ilmoittama opintomenestys rahoituksen kursseissa

	N	%	Kumulatiivinen prosentti
Tyydyttävästi	12	8,2	8,2
Hyvin	19	13,0	21,2
Erittäin hyvin	19	13,0	34,2
En ole vielä suorittanut rahoituksen opintoja	96	65,8	100,0
Yhteensä	146	100,0	

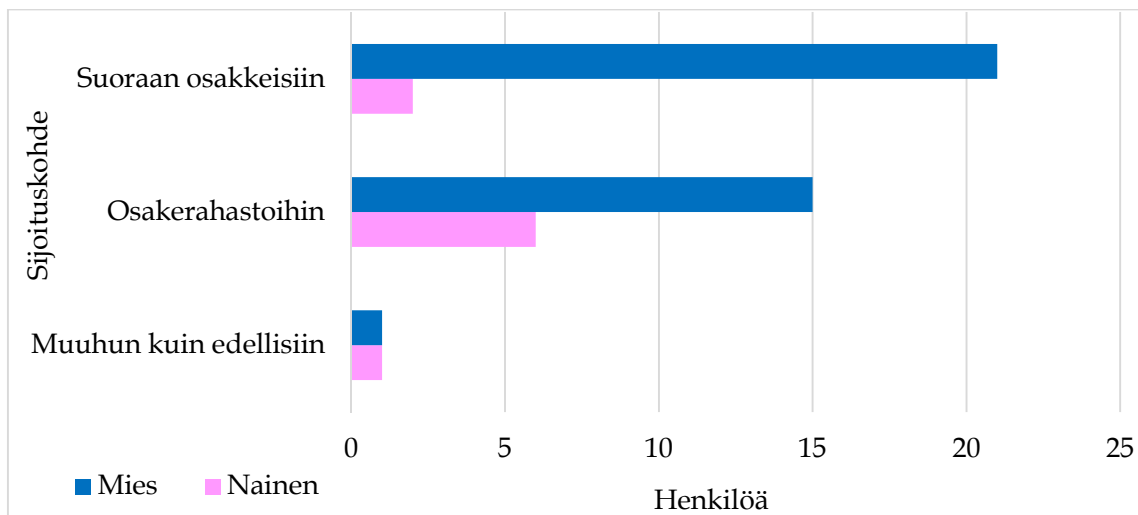
TAULUKKO 8 Rahoituksen kurssien opintomenestys sukupuolittain

		Sukupuoli		Yhteensä	
		Nainen	Mies		
Opintomenestys	Tyydyttävästi	N	5	7	12
		%	41,7 %	58,3 %	100,0 %
	Hyvin	N	7	12	19
		%	36,8 %	63,2 %	100,0 %
	Erittäin hyvin	N	5	14	19
		%	26,3 %	73,7 %	100,0 %
	En ole vielä suorittanut rahoituksen opintoja	N	41	55	96
		%	42,7 %	57,3 %	100,0 %
	Yhteensä	N	58	88	146
		%	39,7 %	60,3 %	100,0 %

Kyselylomakkeessa kysyttiin myös vastaajien sijoitusharrastusta. Vastaajista 38 ilmoitti harrastavansa sijoittamista. Heistä seitsemän oli naisia (12,1 % kaikista naisista) ja 31 miehiä (35,2 % kaikista miehistä). Lisäksi kyselylomakkeessa kysyttiin sijoitusharrastusta suhteessa vanhempiin ja parhaimpiin ystäviin. Sijoittamista harrastavien opiskelijoiden vanhemmista puolet harrasti sijoittamista ja puolet eivät. Ei-sijoittavien opiskelijoiden vanhemmista taas 64,8 % ei harrastanut sijoittamista ja 35,2 % harrasti sijoittamista. Sijoittamista harrastavien opiskelijoiden parhaimmista ystäväistä 78,9 % harrasti sijoittamista ja 21,1 % taas ei. Ei-sijoittavien opiskelijoiden parhaimmista ystäväistä sijoittamista harrasti 46,3 % ja 53,7 % ei harrastanut sijoittamista.

Kuviosta 6 nähdään opiskelijoiden sijoituskohteet sukupuolittain. Kuvion mukaan miehistä 21 ja naisista kaksi sijoittivat suoraan osakkeisiin. Osakerahas-

toihin taas sijoitti 15 miestä ja kuusi naista. Yksi mies ja yksi nainen ilmoitti sijoittavansa muuhun kuin suoraan osakkeisiin tai osakerahastoihin: mies ilmoitti sijoittavansa jalometalleihin ja nainen indeksiin.



KUVIO 6 Sijoituskohteet sukupuolittain.

Kyselylomakkeessa kysyttiin sijoitusharrastuksen yhteydessä myös talousuutisten seuraamista ja niiden vaikutusta sijoituspäätöksiin. Ristiintaulukoimalla uutisten seuraamista sukupuolittain huomattiin, että miehet olivat aktiivisempia seuraamaan päivittäin talousuutisia: sijoittavista miehistä 24 ilmoitti seuraavansa talousuutisia päivittäin, naisista vain yksi. Viikoittain talousuutisia seurasi seitsemän miestä ja kolme naista. Sijoittavista miesvastaajista kukaan ei seurannut uutisia tätä harvemmin, naisistakin vain kolme. Suurin osa vastaajista (71,7 %) ilmoitti talousuutisten vaikuttavan vain hieman sijoituspäätöksiin. 15,8 % vastaajista taas katsoi, että talousuutiset vaikuttavat sijoituspäätöksiin huomattavasti.

Sijoittavat opiskelijat arvioivat keskimäärin oman salkkunsu tulevan 12 kuukauden tuoton olevan n. 6,92 %, kun taas vastaavan markkinatuoton arvioitiin olevan 6,86 %. Arvioitu minimi- ja maksimituotto sekä omalle salkulle että markkinoille oli -5 % ja 30 %.

Kyselylomakkeen ensimmäisen osan viimeisenä kysymyksenä kysyttiin vanhempien sosioekonomista asemaa. Taulukoista 9 ja 10 nähdään, miten opiskelijat vastasivat. Taulukkojen perusteella voidaan sanoa, että yleisimmin opiskelijoiden isät ja äidit olivat työntekijöitä.

TAULUKKO 9 Äidin sosioekonominen asema

	N	%	Validi prosentti	Kumulatiivinen prosentti
Yrittäjä	6	4,1	4,2	4,2
Ylempi toimihenkilö	18	12,3	12,5	16,7
Alempi toimihenkilö	25	17,1	17,4	34,0
Työntekijä	83	56,8	57,6	91,7
Opiskelija	2	1,4	1,4	93,1
Eläkeläinen	8	5,5	5,6	98,6
Muu	2	1,4	1,4	100,0
Yhteensä	144	98,6	100,0	
Puuttuu	2	1,4		
Yhteensä	146	100,0		

TAULUKKO 10 Isän sosioekonominen asema

	N	%	Validi prosentti	Kumulatiivinen prosentti
Yrittäjä	28	19,2	20,0	20,0
Ylempi toimihenkilö	34	23,3	24,3	44,3
Alempi toimihenkilö	17	11,6	12,1	56,4
Työntekijä	47	32,2	33,6	90,0
Eläkeläinen	13	8,9	9,3	99,3
Muu	1	,7	,7	100,0
Yhteensä	140	95,9	100,0	
Puuttuu	6	4,1		
Yhteensä	146	100,0		

Ristiintaulukoinnilla selvitettiin lisäksi opiskelijoiden jakautumista tiedekunnittain suhteessa vanhempien sosioekonomiseen asemaan. Seuraavaksi raportoidaan vain muutamia tärkeimpiä tuloksia. Kauppatieteilijöiden äideistä suurin osa eli 58,3 % oli työntekijöitä ja kaksi toiseksi suurinta ryhmää olivat ylempät ja alemmat toimihenkilöt (15,5 %). Informaatioteknologian opiskelijoiden äideistä kaksi suurinta ryhmää sosioekonomisen aseman perusteella olivat työntekijät (60,7 %) ja alemmat toimihenkilöt (17,9 %). Humanistien äideistä suurin osa (60 %) oli työntekijöitä ja yhteiskuntatieteilijöiden taas alempia toimihenkilöitä (60 %).

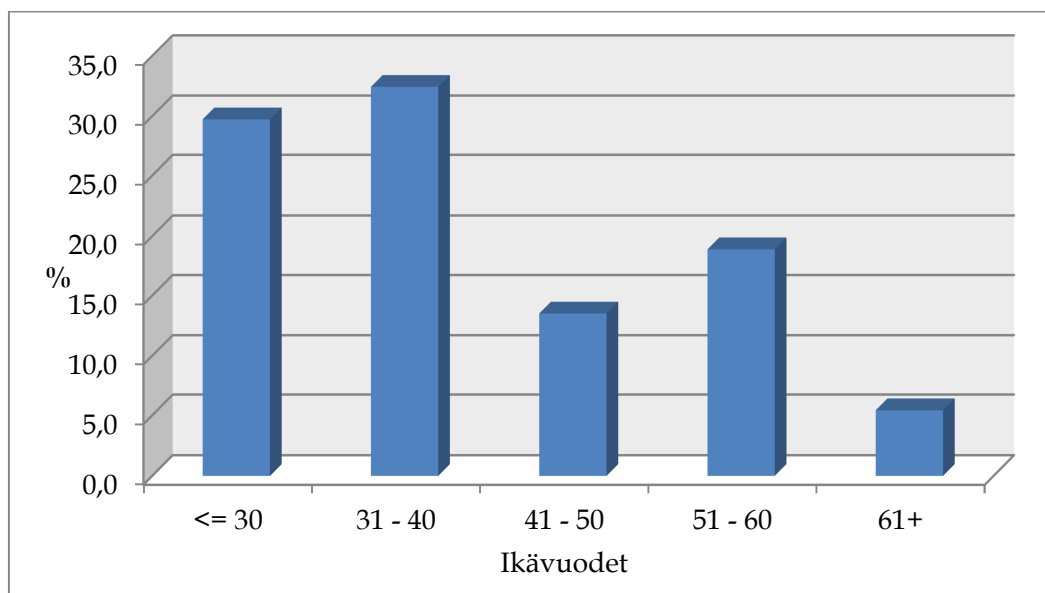
Kauppatieteilijöiden iseistä kolme suurinta ryhmää sosioekonomisen aseman perusteella olivat työntekijät (34,7 %), ylempät toimihenkilöt (25,7 %) ja yrittäjät (23,8 %). Informaatioteknologian opiskelijoiden iseistä kaksi suurinta ryhmää olivat työntekijät (37 %) ja ylempät toimihenkilöt (29,6 %). Humanistien iseistä suurin osa oli työntekijöitä (50 %) ja yhteiskuntatieteilijöiden puolestaan yrittäjiä ja alempia toimihenkilöitä (40 %).

Liitteessä 3 on esitetty opiskelijoiden osake-ennusteiden perustunnusluvut⁹.

⁹ SPSS:ssä ennusteet koodattiin seuraavasti: A1, A2, A3 jne., jossa A viittaa osakekuvaajaan yksi ja numero taas itse osake-ennusteeseen (1 = seitsemän kuukauden piste-estimaatti, 2 = 13 kuukauden piste-estimaatti, 3 = seitsemän kuukauden yläluottamusväli jne.).

4.6 Osuuspankin aineiston kuvailu

Osuuspankin sijoitusasiantuntijoille suunnattuun kyselyyn vastaisi 37 henkilöä, joista naisia oli 17 ja miehiä 20. Vastaajien keski-ikä oli reilut 39 vuotta, minimi-ikä 27 vuotta ja maksimi-ikä 62. Seuraavasta pylväskuvaajasta 7 nähdään pankin vastaajien jakautuminen ikäryhmiin. Reilut 60 % vastaajista oli 40 - vuotiaita tai sen alle, eli vastaajajoukko oli melko nuorta.



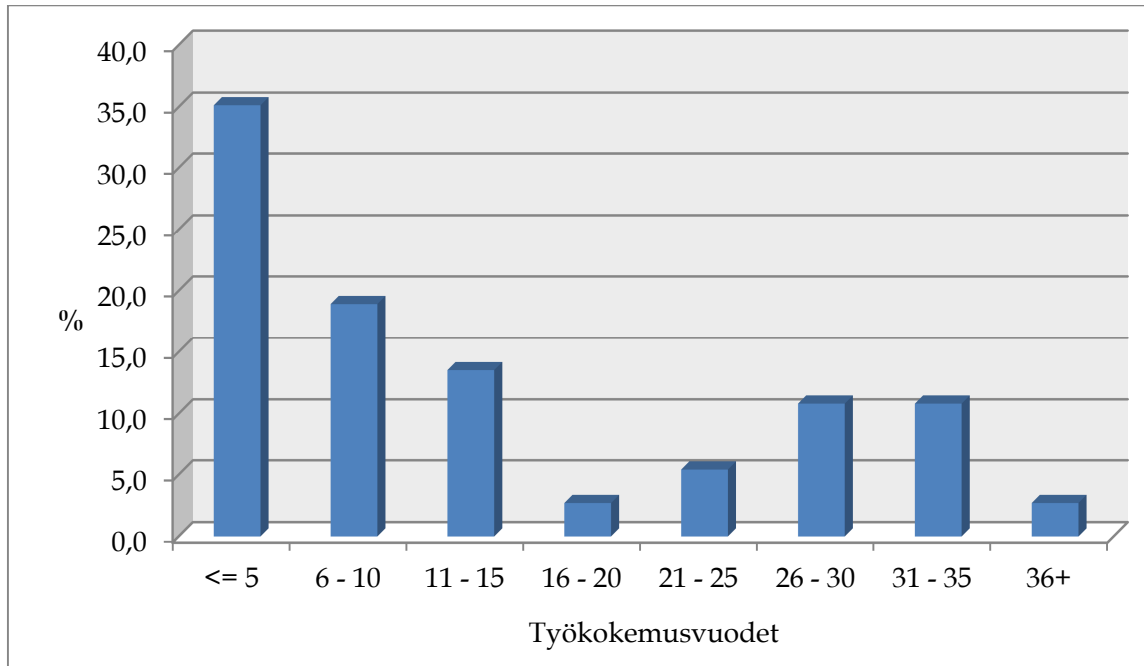
KUVIO 7 Pankin vastaajien jakautuminen ikäryhmiin

Taulukossa 11 on esitetty pankin vastaajien jakautuminen ryhmiin pohjakoulutuksen perusteella. Suurimmalla osalla vastaajista oli pohjakoulutuksena ammattikorkeakoulututkinto.

TAULUKKO 11 Pankin vastaajien jakautuminen pohjakoulutuksen mukaan

	N	%	Validi prosentti	Kumulatiivinen prosentti
Ylioppilastutkinto	1	2,7	2,7	2,7
Ammatillinen koulutus	8	21,6	21,6	24,3
Ammattikorkeakoulututkinto	15	40,5	40,5	64,9
Yliopistotutkinto	13	35,1	35,1	100,0
Yhteensä	37	100,0	100,0	

Pankin vastaajilla rahoitusalan työkokemuksen pituus oli keskimäärin noin 14 vuotta, minimi 1 vuosi ja maksimi 40 vuotta. Seuraava kuvio 8 esittää pankin vastaajien jakautumisen ryhmiin rahoitusalan työkokemuksen perusteella.



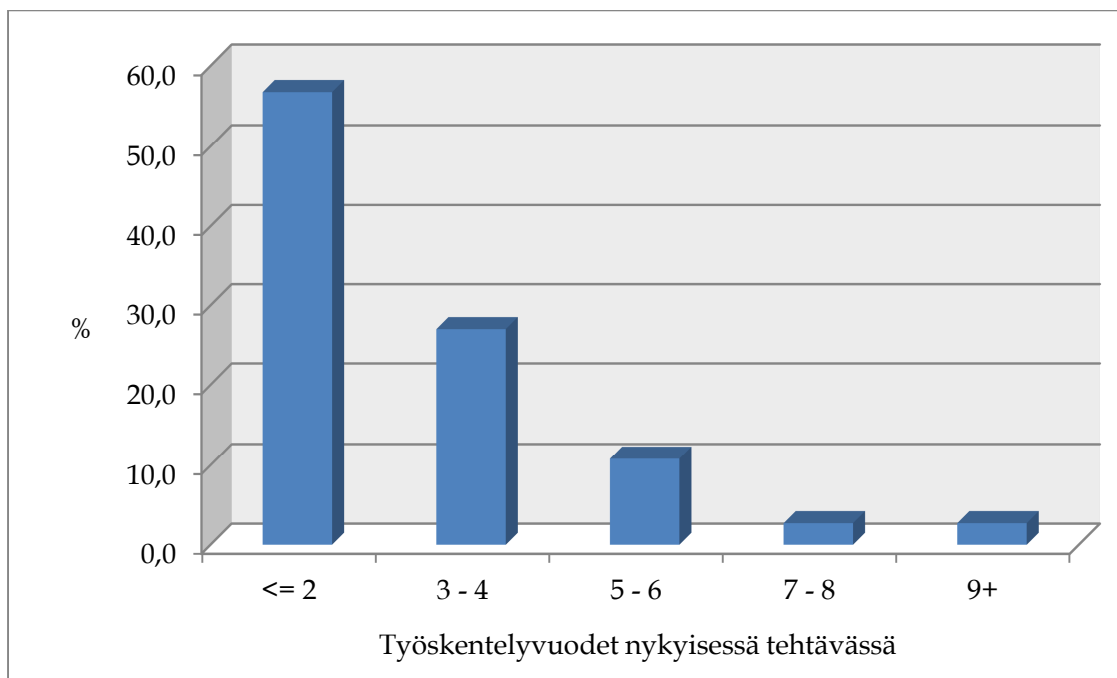
KUVIO 8 Rahoitusalan työkokemus (vuosina)

Pankin vastaajien jakautuminen tehtävänimikkeittäin on kuvattu taulukossa 12. Kuten taulukosta huomataan, kyselyyn vastasi myös muita pankin toimihenkilöitä kuin kohderyhmäksi määritellyt sijoitusneuvojat, -johtajat ja -päälliköt.

TAULUKKO 12 Jakautuminen nykyisten tehtävänimikkeiden mukaan

	N	%	Kumulatiivinen prosentti
Sijoitusneuvoja	13	35,1	35,1
Sijoituspäällikkö	11	29,7	64,9
Sijoitusjohtaja	4	10,8	75,7
Pankinjohtaja	1	2,7	78,4
Sijoitusasiantuntija	1	2,7	81,1
Varainhoitaja	3	8,1	89,2
Asiakkuusasiantuntija	1	2,7	91,9
Rahoitusneuvoja	1	2,7	94,6
OP-yhteyshenkilö	1	2,7	97,3
Talusasiantuntija	1	2,7	100,0
Yhteensä	37	100,0	

Pankin vastaajilla oli työkokemusta nykyisestä työtehtävästä keskimäärin vajaat kolme vuotta. Vähimmillään työkokemusta nykyisestä tehtävästä oli 0 vuotta ja enimmillään 10 vuotta. Seuraavassa kuviossa 9 on jaoteltu vastaajat nykyisessä työtehtävässä toimitun ajan mukaan. Valtaosa vastaajista (83,8 %) oli työskennellyt nykyisessä tehtävässään neljä vuotta tai sen alle.



KUVIO 9 Työskentelyvuodet nykyisessä tehtävässä

Pankin sijoitusasiantuntijoille suunnatussa kyselyssä esitettiin myös tehokkaiden markkinoiden hypoteesiin liittyviä väitteitä. Vain kolme vastaajaa ilmoitti, ettei kyseinen hypoteesi ollut heille tuttu. Liitteessä 4 olevissa taulukoissa on kuvattu pankin vastaajien suhtautuminen hypoteesiin liittyviin väitteisiin passiivisesta sijoitusstrategiasta, osakekurssien ennustettavuudesta ja satunnaiskulusta sekä markkinoiden seuraamisesta. Väitteiden tarkempi sisältö on nähtävissä liitteenä olevasta pankin kyselystä (liite 2, kysymys nro 8). Tässä tuodaan esille vain liitteenä 4 olevien taulukkojen keskeisin sisältö.

Liitteessä 4 olevasta taulukosta a) nähdään, että vain kaksi pankin vastaajaa oli täysin eri mieltä siitä, että asiakkaiden kannattaisi noudattaa passiivista salkun hajautusstrategiaa. Kukaan vastaajista ei ollut väitteen kanssa täysin samaa mieltä. Yllättävän moni (32,4 %) oli myös väitteen kanssa osittain samaa mieltä, mikä osoittaa, että pankin vastaajat näyttivät pitävän tehokkaiden markkinoiden hypoteesia jossain määrin totena. Taulukko b) kuvaa pankin vastaajien suhtautumista osakekurssien ennustettavuuteen. Suurin osa vastaajista (47,1 %) oli osittain samaa mieltä siitä, että osakekurseja voi ennustaa. Yllättävää on, ettei yhtään vastaajaa ollut väitteen kanssa täysin samaa mieltä, sillä luulisi työkseen sijoitusneuvoja antavan uskovan täysin osakkeiden ennustettavuuteen. Vastausten hajoaminen myös "Osittain eri mieltä" -sektorille osoittaa, että he uskoivat markkinoiden olevan jossain määrin tehokkaat. Taulukon c) perusteella huomataan, että suurin osa vastaajista (yhteensä 61,7 %) oli joko täysin eri mieltä tai osittain eri mieltä siitä, että asiakkaiden ei kannata kuluttaa aikaansa markkinoiden seuraamiseen. Tämä on hiukan ristiriidassa edellisten vastausten kanssa, sillä jos markkinoiden uskotaan olevan tehokkaat, niin silloinhan passiivinen sijoitusstrategia olisi hyvä vaihtoehto. Tehokkaiden markkinoiden hypoteesiin

liittyvässä väitepatterissa viimeinen kohta liittyi osakekurssien satunnaiskulkuun. Kuten taulukosta d) nähdään, tämän väitteen kohdalla vastaukset hajosivat "Osittain eri mieltä" ja "Osittain samaa mieltä" -akselille. Näyttäisi siltä, vastaajat uskoivat osakkeiden hinnoissa olevan jonkin verran ennustettavuutta.

Pankin vastaajilta kysyttiin myös talousuutisten vaikutusta sijoitusneuvoihin. Vaihtoehtoa "Ei mitenkään" ei ollut valinnut kukaan. Vastaajista yli 54 % arvioi talousuutisten vaikuttavan sijoitusneuvoihin huomattavasti, ja lähes 46 % hieman. Naiset ja miehet arvioivat talousuutisten vaikutukset lähes samalla tavalla.

Valtaosalle pankin vastaajista (81,1 %) työnantajan tarjoama materiaali oli pääasiallinen tietolähde sijoitusneuvoja annettaessa. Tämä oli odotettua, sillä Osuuspankissa sijoitusasiantuntijat eivät anna asiakkaille neuvoja omien analyysien perusteella, vaan OP-Pohjola ryhmän suosituksen mukaisesti. Yksi henkilö oli merkinnyt pääasialliseksi tietolähteekseen muu, ja tarkensi vastaustaan, ettei mikään yksittäinen asia ole pääasiallinen tietolähde, vaan tiedot koostuvat monista eri asioista. Kuitenkin lähes 60 % pankin vastaajista katsoi, että kollegakeskusteluilla olisi hieman vaikutusta annettuihin sijoitusneuvoihin ja lähes 38 % katsoi, että niillä oli huomattava merkitys. Tämä on aika yllättävää, jos kerran pääasiallinen tietolähde sijoitusneuvoja annettaessa on työnantajan tarjoama valmis materiaali.

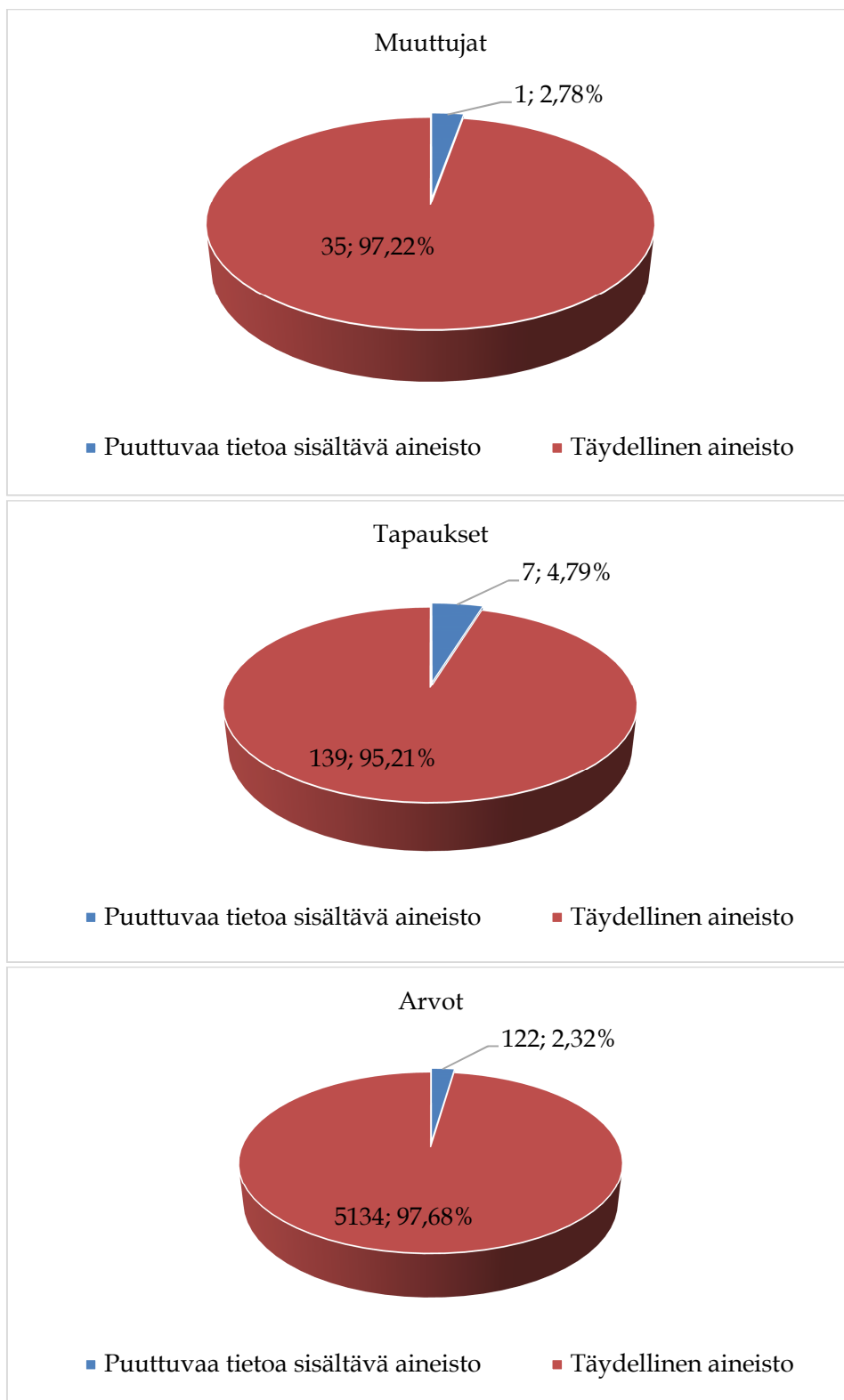
Liitteessä 3 on esitetty pankin vastaajien osake-ennusteiden perustunnusluvut.

4.7 Puuttuvan tiedon analyysi, opiskelijoiden osake-ennusteet

Aineiston analyysin näkökulmasta tärkeä huomioon otettava seikka on opiskelijoiden osake-ennusteissa esiintyvät puuttuvat tiedot. Kun asiaa tarkasteltiin muuttujittain, huomattiin, että eniten puuttuvia arvoja oli viimeisten ennusteiden kohdalla, joten jonkinlaista vastausväsymystä oli varmaankin ilmennyt. Puuttuvia ennusteita oli maksimissaan 3,4 % kaikista osake-ennusteista seuraavien ennusteiden kohdalla: D4-D6, E3-E6 ja F3-F6¹⁰.

Seuraavasta kuvioista 10 selviää, että puuttuvia tietoja oli 35 muuttujassa ja seitsemällä opiskelijalla (4,8 % vastaajista). Kaikkiaan puuttuvia arvoja oli 122 eli yhteensä 2,3 %, mikä oli verrattain vähän.

¹⁰ Kirjain viittaa osakekuvaajaan ja numero osake-ennusteeseen, esim. D4 = osakekuvaajaan neljä liittyvä seitsemän kuukauden alaluottamusväli, kts. liite 8.



KUVIO 10 Puuttuvan tiedon kokonaistarkastelu

Little's MCAR-testitulokset on 0,426 eli $> 0,05$, mikä tarkoittaa, että opiskelijoiden osake-ennusteista puuttuva tieto oli luonteeltaan satunnaista, Missing Completely at Random, MCAR. Käytännössä tämä tarkoittaa sitä, että puuttuvat ha-

vainnot olivat satunnaisesti jakautuneita kaikkien havaintojen kesken (Graham 2012, 13). MCAR-testin oletuksena on, että puuttuvaa tietoa sisältävä aineisto olisi normaalisti jakautunut, mikä ei tässä tapauksessa toteudu, kuten myöhemmin huomataan. Näin ollen testin tulos saattoi olla hiukan harhainen.

Tässä tutkimuksessa tehtyjä herkkyysanalyyskejä varten puuttuvan tietoa käsiteltiin Case Deletion -menetelmällä¹¹ (jatkossa menetelmästä käytetään lyhennettä CD-menetelmä). Menetelmässä poistetaan analyysistä ne muuttujat, joissa puuttuvaa tietoa on. Tämän menetelmän etuna on selkeys ja tehokkuus, etenkin jos puuttuvaa tietoa on vähän, kuten tässä opinnäytetyössä. Case deletion -menetelmän on osoitettu johtavan harhattomiin estimaatteihin, kun puuttuva tieto on luonteeltaan MCAR. (Schafer & Graham 2002, 156–157.) Menetelmän käyttö kuitenkin heikentää tilastollisten testien voimaa ja on hyväksyttävissä vain silloin, kun odotettavissa on vähäinen tiedon hävikki (Little & Rubin 2002, 41–42).

Case Deletion -menetelmää suositellumpia puuttuvan tiedon käsittelymenetelmiä olisivat Full Information Maximum Likelihood (FIML)- ja Multiple Imputation (MI) -menetelmät. Kahta jälkimmäistä pidetään parhaimpana puuttuvan aineiston paikkausmenetelmistä (Rubin 1996; Schafer & Graham 2002). SPSS 20-ohjelma ei kuitenkaan kykene FIML-estimointiin, ja MI:n toteuttamiseenkin vain tietyin rajoituksin. Normaalisti MI perustuu ensiksi laskettuun EM-algoritmiin (Expectation Maximization), josta saadaan kovarianssimatriisin estimaatit aineiston augmentaatiota varten. Grahamin mukaan ei ole kuitenkaan varmuutta, tekeekö SPSS 20-ohjelma tätä, sillä vaiheen dokumentointi on ohjelmassa puutteellista. SPSS 20-ohjelma ei myöskään tarjoa MI:hin liittyvää diagnostiikkaa, esimerkiksi kuvioita parametrien estimaateista ja autokorrelaatiosta, jotta pystyttäisiin arvioimaan MI-ratkaisun hyväksyttävyyttä. (Graham 2012, 116–117.) On kuitenkin huomattava, että vaikka EM-algoritmin mukana olo SPSS 20-ohjelman MI-prosessissa ei ole selvästi tiedossa, löytyy EM-analyysi kuitenkin ohjelmasta Missing Value Analysis -kohdasta, josta siis voidaan laskea EM-parametrien estimaatit.

Yllä mainitut rajoitteet huomioon ottaen päädyttiin tässä opinnäytetyössä valitsemaan puuttuvan tiedon käsittelymenetelmäksi CD-menetelmä. Kaiken kaikkiaan Case Deletion koski seitsemää vastausta, joten herkkyysanalyysjä tehtäessä aineiston koko oli 139 vastausta.

4.8 Aineistojen normalisuuden tarkistukset ja normalisointi

Aineistojen analyysin kannalta olennaista on tarkistaa, onko opiskelijoiden ja pankin vastaajien antamat osake-ennusteet normaalijakautuneita. SPSS-ohjelmassa esimerkiksi t-testit, ANOVA ja lineaarinen regressioanalyysi perustuvat aineiston normaalijakautuneisuusoletukseen.

¹¹ Usein menetelmästä käytetään myös nimeä Listwise Deletion.

Kummankaan alkuperäisen perusaineiston osake-ennusteet eivät olleet normaalisti jakautuneita. Liitteestä 5 nähdään opiskelijoiden ja pankin vastaajien osake-ennusteaineiston normaalisuustestien p-arvot ja muuttujien symmetrisyystarkastelut. P-arvoja tulkittaessa tulee muistaa, että SPSS:n käyttämät tilastolliset testit paljastavat herkästi pienetkin erot tilastollisesti merkitseviksi, joten useissa tapauksissa on järkevämpää katsoa jakauman symmetrisyyttä. Jos muuttujan jakauman vinousluku (skewness) jaettuna keskivirheellään (standard error) on < -2 tai > 2 , niin jakauma ei ole symmetrinen. Liitteestä 5 huomataan, että opiskelijoiden osake-ennustemuuttujista 18 (50 %) ei ollut symmetrisiä. Pankin osake-ennustemuuttujista 13 (n. 36 %) ei ollut symmetrisiä.

Osake-ennusteet muodostavat aikasarja-aineiston, joka on ei-stationäärinen ja heteroskedastinen. Ei-stationäärisuus tarkoittaa sitä, että aineiston tilastolliset ominaisuudet, kuten keskiarvo ja keskihajonta, vaihtelevat ajassa. Heteroskedastisuus taas tarkoittaa sitä, että varianssit eivät ole vakioita ja että osakesarjojen volatiilisuus voi vaihdella melko huomattavastikin. Tämänkaltaisen aikasarja-aineiston normalisointi perinteisillä metodeilla ei aina onnistu. (Ogasawara, Martinez, de Oliveira, Zimbrão, Pappa & Mattoso 2010.) Tässä opinnäytetyössä käytettiin samaa normalisointitekniikkaa kuin De Bondtin (1993) tutkimuksessa. Odotetun hintamuutoksen (EPC) ja luottamusvälien vinouden parametrit normalisoitiin jakamalla ne toteutuneen S&P 500 indeksin kuukausituoton keskihajonnalla (σ) neljän kuukauden ajalta ennen ennusteen tekemistä. Keskihajonnan laskemiseen käytettiin Excel-
taulukkolaskennan STDEVS-kaavaa.

5 HAVAINTOJA ENNUSTEISTA

Tässä luvussa esitetään analyysien tulokset. Aineistot jaoteltiin seuraavasti:

1. Aineisto: opiskelijoiden perusaineisto
2. Aineisto: opiskelijoiden normalisoitu perusaineisto (odotetun hintamuutoksen ja luottamusvälien vinouden parametrit normalisoitu, puuttuvia tietoja ei poistettu)
3. Aineisto: opiskelijoiden täydellinen aineisto (puuttuvia ennusteita sisältävät vastaukset poistettu CD-menetelmän mukaisesti)
4. Aineisto: opiskelijoiden täydellinen ja normalisoitu aineisto (puuttuvia ennusteita sisältävät vastaukset poistettu CD-menetelmän mukaisesti sekä odotetun hintamuutoksen ja luottamusvälien vinouden parametrit normalisoitu)
5. Aineisto: pankin perusaineisto
6. Aineisto: pankin normalisoitu perusaineisto (odotetun hintamuutoksen ja luottamusvälien vinouden parametrit normalisoitu).

Analyysit tehtiin hypoteesien mukaisessa järjestyksessä aineistopareittain edeten. Aineistoparit olivat seuraavat:

- Aineistot 1. ja 5.
- Aineistot 2. ja 6.
- Aineistot 3. ja 5.
- Aineistot 4. ja 6.

Aineistot 1. ja 5. muodostivat tämän tutkimuksen varsinaiset perusaineistot ja muille aineistoille tehdyt analyysit olivat herkkyysanalyysijä. Huomion arvoista on lisäksi, että päätetty normalisointiratkaisu koski vain kolmea hypoteesia, eli hypoteeseja H_2 , H_3 ja H_4 . Muiden hypoteesien testaukseen normalisointi ei vaikuttanut.

Analyysimenetelminä käytettiin t-testejä, varianssianalyysijä, korrelaatiotarkasteluja sekä lineaarista ja logistista regressioanalyysijä. Parametrittomia menetelmiä, kuten Mann-Whitneyn U-testijä (jatkossa testistä käytetään lyhennettä MWU-testi), Kruskal-Wallis testijä ja Wilcoxonin Signed-Rank -testijä (jatkossa testistä käytetään lyhennettä WSR-testi) käytettiin parametristen testien rinnalla, koska osake-ennusteet ja niistä muodostetut muuttujat eivät useimmissa tapauksissa noudattaneet normaalijakaumaa. Analysointiin käytettiin IBM SPSS Statistics 20-ohjelmaa.

Tässä luvussa vain perusaineistojen analyysien tulokset esitetään taulukoina, lukuun ottamatta parametrittomien testien tuloksia. Parametrisijä testeijä pidetään yleisesti voimakkaampina kuin parametrittomia, eli niiden mukaan

mahdollisuus väärän nollahypoteesin hylkäämiseen on suurempi. Perusaineistojen osalta kukin hypoteesi on kerrattu analyysien yhteydessä. Hypoteesit on lisäksi esitetty liitteessä 6. Analyyseissä käytetyt muuttujat on pyritty esittämään mahdollisimman tarkasti kunkin analyysin kohdalla. Liitteessä 8 esitetään lisäksi kattava muuttujaluettelo.

5.1 Perusaineistot

Aineiston analyysi aloitettiin perusaineistoilla 1. ja 5. Opiskelijoiden perusaineistosta puuttuvia osake-ennusteita sisältäviä vastauksia ei tässä vaiheessa poistettu. On huomattava kuitenkin, että esimerkiksi t-testit vaativat määrittämään puuttuvan tiedon käsittelyn ennen testin tekemistä. Opiskelijoiden perusaineistossa parametrisissä testeissä käytettiin Exclude analysis by analysis -komentoa, mikä tarkoittaa, että testiä ajettaessa käytetään kaikkia niitä tapauksia, joissa on validia tietoa testattavia muuttujapareja varten. ”Otoksen” koko saattoi tämän vuoksi vaihdella testistä toiseen. Parametrittomien testien yhteydessä käytettiin vastaavanlaista Exclude cases test-by-test -komentoa.

5.1.1 Ankkuroituivatko osake-ennusteet keskiarvotuottoa kohti?

Ensimmäinen tämän opinnäytetyön hypoteeseista liittyy ennustettujen osakehintojen konvergoitumiseen esitetyn osakekuvaajan keskiarvoa kohti:

H_1 : Osake-ennusteet ankkuroituvat osakkeen hintakehityksen keskiarvoa kohti.

Tällöin kyse olisi toisinajattelijastrategiasta (contrarian investor strategy), jossa nousukauden jälkeen ennustetaan osaketuottojen laskevan ja laskukauden jälkeen taas nousevan. Kumpi vastaajaryhmistä, opiskelijat vain pankin sijoitusasiantuntijat, oli keskimäärin enemmän ankkuroitunut osakekuvaajan keskiarvoja kohti ennusteilla mitattuna?

Hypoteesin H_1 testausta varten yhdistettiin perusaineistot. Tämän jälkeen luotiin uusi muuttuja R (ryhmä), joka sai arvon 1, kun kyseessä oli opiskelija ja arvon 2, kun kyseessä oli pankin sijoitusasiantuntija. Sen jälkeen luotiin uusi muuttuja EROKA, joka kuvaa ennustettua osaketuottoa osakekuvaajan keskiarvoon verrattuna. Osakekuvaajan keskiarvo laskettiin kussakin kuvaajassa näkyvältä 48 kuukaudelta.

Seuraavaksi tutkittiin riippumattomien ryhmien t-testillä, oliko pankin sijoitusasiantuntijoiden ja opiskelijoiden EROKA-muuttujissa eroa, ja kumman ryhmän tuottoennusteet olivat lähempänä osakekuvaajien keskiarvoja. Se ryhmä, jonka muuttujan EROKA keskiarvot kunkin osakkeen kohdalla olivat pienemmät, oli konvergoitunut ennusteissaan enemmän kohti kyseisen osakekuvaajan keskiarvotuottoa. Erot olivat tilastollisesti melkein merkitseviä seuraavien osake-ennusteiden tapauksissa: B2, C1, C2, D1, E2, F1 (taulukko 13). Kaikki-

en edellä mainittujen ennusteiden, paitsi B2:n kohdalla, pankin vastaajat olivat enemmän konvergoituneita osakekuvaajien keskiarvotuottoja kohti.

TAULUKKO 13 Ennustettu osaketuotto osakekuvaajan keskiarvoon verrattuna

	t	Vapausasteet	P-arvo
EroKaA1	-,751	181	,454
EroKaA2	-1,324	180	,187
EroKaB1	-,667	180	,505
EroKaB2	-2,611(*)	179	,010
EroKaC1	-2,440(*)	180	,016
EroKaC2	-2,022(*)	180	,045
EroKaD1	2,473(*)	179	,014
EroKaD2	,683	178	,496
EroKaE1	-1,732	177	,085
EroKaE2	-2,443(*)	177	,016
EroKaF1	2,065(*)	177	,040
EroKaF2	-,040	177	,968

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)¹².

Riippumattomien ryhmien t-testin edellytyksenä on, että muuttujat ovat normaalisti jakautuneet ja varianssit ovat homogeenisia. Varianssien yhtäsuuruus varmistui testin yhteydessä olleen Levenen testin perusteella, mutta muuttujat eivät olleet normaalisti jakautuneita. Tämän vuoksi tehtiin vielä MWU-testi. MWU-testi tutkii, ovatko kaksi riippumatonta ostosta samoin jakautuneesta perusjoukosta. Sen tulokset olivat hyvin samansuuntaisia kuin yllä: ennusteiden B2, C1, D1, E2 ja F1 osalta muuttujien EROKA sijalukujen keskiarvoissa (mean rank) oli tilastollisesti melkein merkitsevää tai merkitsevää eroa (U-testisuureet 1916,5; 1967,0; 1976,0; 2031,5 ja 1989,5 sekä vastaavat p-arvot 0,008; 0,012; 0,015; 0,033 ja 0,021). Sijalukujen keskiarvoja verrattaessa saatiin kuitenkin t-testistä poikkeava tulos: pankin vastaajat olivat enemmän konvergoituneita osakekuvaajien keskiarvotuottoja kohti vain D1- ja F1-ennusteiden tapauksissa, muissa taas opiskelijat. Tulos siis poikkesi jonkin verran t-testillä saaduista tuloksista: tilastollisesti merkitseviä eroja saatiin kummallakin testillä lähes samoihin ennusteisiin (poikkeuksena C2), mutta testin tulkinta oli erilainen testistä riippuen. Sekä parametrinen että parametrittoman testin mukaan D1- ja F1-ennusteiden kohdalla pankin vastaajat olivat enemmän konvergoituneita kyseisten osakekuvaajien keskiarvotuottoa kohti, ja B2-ennusteiden kohdalla taas opiskelijat. C1- ja E2-ennusteiden kohdalla testien tulokset menivät kuitenkin ristiin: t-testin mukaan pankin vastaajat olivat niissä enemmän konvergoituneita kohti osakekuvaajien keskiarvoa, mutta MWU-testin mukaan taas opiskelijat.

Entä oliko vastaajaryhmien välillä eroa toisinajattelijastrategian mukaisessa käyttäytymisessä? Toisinajattelijastrategiaa noudattavan oletettiin nousukausien jälkeen (osakkeet 1., 4. ja 6., eli A, D ja F) ennustavan siten, että $P_0 > F_7 > F_{13}$ ja laskukauden jälkeen (osakkeet 2., 3. ja 5., eli B, C ja E) siten, että $P_0 < F_7 < F_{13}$ ¹³.

¹² (*) = tilastollisesti melkein merkitsevää, (**) = tilastollisesti merkitsevää ja (***) = tilastollisesti erittäin merkitsevää.

¹³ P_0 viittaa tässä kyselylomakkeessa näytettyjen osakekuvaajien viimeiseen arvoon, F_7 seitsemän kuukauden piste-estimaattiin ja F_{13} 13 kuukauden piste-estimaattiin.

Vastaajat jaettiin yllämainitun käyttäytymisen perusteella ryhmään 1. ja muun käyttäytymisen perusteella ryhmään 2.

Tuloksena saatiin uudet ryhmittelymuuttujat CONT_A, CONT_B, CONT_C, CONT_D, CONT_E ja CONT_F (CONT viittaa sanaan contrarian). Seuraavaksi tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla muuttujien välistä riippuvuutta. Ristiintaulukoinnissa selitettäviksi eli riippuviksi muuttujiksi asetettiin ensin nousukausien muuttujat CONT_A, CONT_D ja CONT_F ja selittäväksi eli riippumattomaksi muuttujaksi R, joka siis erotteli pankin vastaajat opiskelijoista. Toisessa analyysissä selitettäviksi muuttujiksi asetettiin laskukausien muuttujat CONT_B, CONT_C ja CONT_E.

Osakkeen A osalta 26 opiskelijaa (17,8 % kaikista opiskelijoista) ja kaksi pankin vastaajaa (5,4 % kaikista pankin vastaajista) kuului ryhmään 1. eli noudatti toisinajattelijastrategiaa. Osakkeen D tapauksessa taas 26 opiskelijaa (18,1 %) ja 6 pankin vastaajaa (16,2 %) kuului ryhmään 1. F-osakkeen osalta 10 opiskelijaa (7 %) ja neljä pankin vastaajaa (10,8 %) kuului ryhmään 1. Loput vastaajista siis kuuluivat ryhmään 2. Huomattiin siis, että A-osakkeen osalta opiskelijat olivat enemmän toisinajattelijoita. D- ja F-osakkeiden osalta suhteelliset osuudet olivat melko lailla yhtä suuria. Huomion arvoista on kuitenkin se seikka, että kokonaismäärällä mitattuna toisinajattelijastrategian noudattaminen nousukausien jälkeen oli aika heikkoa.

Laskukausien tuloksia tutkittaessa havaittiin, että osakkeen B osalta 55 opiskelijaa (37,9 %) ja 24 pankin vastaajaa (64,9 %) kuului ryhmään 1. Osakkeen C osalta taas 40 opiskelijaa (27,6 %) ja 17 pankin vastaajaa (45,9 %) kuului ryhmään 1. E-osakkeen tapauksessa taas ei yhtään opiskelijaa kuulunut ryhmään 1. ja pankin vastaajistakin vain yksi (2,7 %). Loput vastaajista kuuluivat ryhmään 2. B-osakkeen osalta tulosten tulkintaa hankaloittaa kuvaajan muoto. Jos tarkkoja ollaan, niin B-osakekuvaaja on sellainen, että osakkeen hinta ensin laskee, mutta kääntyy aikajakson loppupuolella nousuun. Siispä näkemys, jossa B-osakekuvaajaa pidettäisiin täysin laskukauden kuvaajana, ei välttämättä pidä paikkaansa, eli $P_0 < F_7 < F_{13}$ määrittely tarkoittaisikin B:n tapauksessa vahvaa (nousu-) trendinseuraajaa, eikä toisinajattelijastrategian noudattamista. Tämä kuitenkin osoittaa sen, kuinka herkkiä vastaajien ennusteet olivat varsinkin kuvaajan loppuosan muutoksille. C-osakkeen osalta taas pankin vastaajat olivat suhteellisesti enemmän taipuvaisia toisinajattelijastrategiaan. B- ja C-osakkeiden perusteella voidaan sanoa, että toisinajattelijastrategian "kannattaminen" oli yleisempää kuin nousukausien osakkeilla. E-osakkeen tapauksessa strategian noudattaminen oli heikkoa.

Seuraava taulukko 14 perustelee tulokset tilastollisesti. Taulukosta huomataan, että tulokset olivat tilastollisesti merkitseviä vain osakkeiden B ja C kohdalla. Voidaan todeta, että molemmissa tapauksissa vastaajaryhmällä ja toisinajattelijaluokittelulla oli riippuvuutta ja että pankin vastaajat noudattivat niissä enemmän toisinajattelijastrategiaa. Kontingenssikertoimista nähdään, että riippuvuudet olivat kuitenkin melko pieniä. Kontingenssikerroin mittaa kahden nominaalisen muuttujan välistä riippuvuutta. Suurempi kertoimen arvo merkitsee voimakkaampaa riippuvuutta muuttujien välillä. E- ja F-osakkeiden

kohdalla analyysit perustuivat Fisherin tarkan testin tuloksiin. Fisherin tarkkaa testiä voidaan käyttää χ^2 -testin sijaan, jos otoskoko on pieni ja odotetut frekvenssit liian pieniä.

TAULUKKO 14 Toisinajattelijastrategian testaus, yhdistetty perusaineisto

	χ^2	Vapausasteet	P-arvo	Kontingenssikerroin
CONT_A	3,504	1	,061	,137
CONT_B	8,727(**)	1	,003	,214
CONT_C	4,619(*)	1	,032	,157
CONT_D	,068	1	,794	,019
CONT_E			,207	
CONT_F			,492	

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

E- ja F-osakkeelle tehtiin Fisherin tarkka testi.

5.1.2 Erot odotetussa hintamuutoksessa

Hypoteesit H_{20} ja H_{21} liittyvät odotettuun hintamuutokseen:

$$H_{20}: \overline{EPC}_{\text{Bull}} = \overline{EPC}_{\text{Bear}}$$

$$H_{21}: \overline{EPC}_{\text{Bull}} > \overline{EPC}_{\text{Bear}}$$

Tässä tutkimuksessa oltiin kiinnostuneita siitä, oliko odotetun hintamuutoksen keskiarvo suurempi nousukausien jälkeen kuin laskukausien jälkeen. Aluksi tarkastellaan opiskelijavastauksia, sitten pankin vastauksia ja lopuksi yhdistettyjä vastauksia. Hypoteesin H_{20} testausta varten muodostettiin ensin kaikkiin perusaineistoihin osakekohtaiset muuttujat EPC_7 ja EPC_{13} , jotka kuvaavat odotettua hintamuutosta ($F_7 - P_0$ ja $F_{13} - P_0$). Sen jälkeen muodostettiin uudet muuttujat $EPC_{7\text{bull}}$, $EPC_{7\text{bear}}$, $EPC_{13\text{bull}}$ ja $EPC_{13\text{bear}}$, jotka kuvaavat nousu- ja laskukausien odotettujen hintamuutosten keskiarvoja.

Opiskelijoiden perusaineistolla hypoteesia testattiin verrannollisten pariien t-testillä. Seuraavasta taulukosta 15 nähdään, että keskimääräinen odotettu hintamuutos opiskelijoilla oli suurempi nousu- kuin laskukausien jälkeen, ja tulokset olivat tilastollisesti erittäin merkitseviä. T-testin tulosten perusteella H_{20} hylättiin.

TAULUKKO 15 Odotettujen hintamuutosten vertailu ja testitulokset, opiskelijoiden perusaineisto

	Keskiarvo	N	Keskihajonta	t	Vapausasteet	P-arvo
Pari 1	$EPC_{7\text{bull}}$,5190	146	3,29418	9,395	,000
	$EPC_{7\text{bear}}$	-2,2929	146	1,59111	(***)	
Pari 2	$EPC_{13\text{bull}}$	1,5557	146	5,29277	6,736	,000
	$EPC_{13\text{bear}}$	-1,8187	146	2,73586	(***)	

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

Koska normaalijakaumaoletus ei täytynyt muuttujien EPC13_{bull} ja EPC13_{bear} kohdalla, testattiin jälkimmäistä paria myös parametrittomalla WSR-testillä, jonka tulokset olivat samanlaisia kuin yllä: H₂₀-hypoteesi hylättiin merkitsevyystasolla $p < 0,001$ (Z-testisuure -5,905).

Pankin perusaineistolla hypoteesia testattiin ensin verrannollisten parien t-testillä, aivan kuten opiskelijoiden perusaineistollakin. Taulukosta 16 nähdään, että myös pankin vastaajilla keskimääräinen odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen ja tulokset olivat tilastollisesti melkein merkitseviä. T-testin tulosten perusteella H₂₀ hylättiin molempien parien osalta, mutta tulokset eivät kuitenkaan olleet tilastollisesti niin merkitseviä kuin opiskelija-aineistolla saadut tulokset.

TAULUKKO 16 Odotettujen hintamuutosten vertailu ja testitulokset, pankin perusaineisto

		Keskiarvo	N	Keskihajonta	t	Vapausasteet	P-arvo
Pari 1	EPC7bull	,0560	37	3,63520	2,426	36	,020
	EPC7bear	-1,4069	37	2,28770	(*)		
Pari 2	EPC13bull	2,1101	37	4,10258	2,385	36	,022
	EPC13bear	-,0015	37	4,42388	(*)		

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

Sama testaus tehtiin jälleen parametrittomalla WSR-testillä, sillä mikään muuttujista ei ollut normaalisti jakautunut. Tulokset olivat samansuuntaiset kuin verrannollisten parien t-testinkin tulokset. Ensimmäisen parin osalta H₂₀ hylättiin p-arvolla 0,028 ja toisen parin osalta p-arvolla 0,013 (Z-testisuure ensimmäiselle parille -2,196 ja toiselle -2,497). Huomattiin jälleen, etteivät parametrittomallakaan testillä saadut tulokset antaneet niin vahvaa todistetta nollihypoteesin hylkäämiseksi kuin opiskelijoiden perusaineistolla.

Lopuksi hypoteesia testattiin vielä yhdistetyllä perusaineistolla. Taulukko 17 esittää tulokset.

TAULUKKO 17 Odotettujen hintamuutosten vertailu ja testitulokset, yhdistetty perusaineisto

		Keskiarvo	N	Keskihajonta	t	Vapausasteet	P-arvo
Pari 1	EPC7bull	,4254	183	3,36068	9,392	182	,000
	EPC7bear	-2,1138	183	1,78311	(***)		
Pari 2	EPC13bull	1,6678	183	5,06927	7,113	182	,000
	EPC13bear	-1,4513	183	3,22025	(***)		

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

Tulosten mukaan H₂₀ hylättiin molempien parien osalta merkitsevyystasoilla $p < 0,001$. Muuttujien normaalijakautuneisuus ei ollut voimassa, joten tehtiin vielä parametriton WSR-testi. Tulokset olivat samanlaiset kuin yllä: H₂₀ hylättiin molempien parien osalta merkitsevyystasoilla $p < 0,001$ (Z-testisuure ensimmäiselle parille -7,777 ja toiselle parille -6,380). Opiskelijoiden ja pankin yhdiste-

tyllä perusaineistolla saadut tulokset olivat siis tilastollisesti erittäin merkitseviä: odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen.

Vastaajat siis osoittautuivat optimistisemmiksi nousukausien jälkeen kuin laskukausien jälkeen. Koska kaikkien aineistojen perusteella keskimääräinen odotettu hintamuutos nousukausien jälkeen oli positiivinen ja laskukausien jälkeen negatiivinen, kertoo se siitä, että vastaajat odottivat menneen trendin jatkuvan.

5.1.3 Luottamusvälien vinous

Hypoteesit H_{30} ja H_{31} liittyvät vastaajien ennustamien luottamusvälien vinouteen:

$$H_{30}: \bar{\Delta}_{\text{Bull}} = \bar{\Delta}_{\text{Bear}}$$

$$H_{31}: \bar{\Delta}_{\text{Bull}} < \bar{\Delta}_{\text{Bear}}$$

Haluttiin siis tutkia, oliko keskimääräinen vinous pienempi nousu- kuin laskukausien jälkeen. Aluksi hypoteesia H_{30} testattiin opiskelijoilla, sitten pankin aineistolla ja lopuksi yhdistetyllä aineistolla. Hypoteesin testausta varten muodostettiin kaikkiin perusaineistoihin ensin uudet muuttujat UCI_{7kk} , UCI_{13kk} , LCI_{7kk} ja LCI_{13kk} , jossa UCI on vastaajan antaman todennäköisyysjakauman ylä-arvo (H) vähennettynä piste-estimaatilla (F), ja LCI on piste-estimaatti (F) vähennettynä todennäköisyysjakauman ala-arvolla (L). Sen jälkeen muodostettiin uudet muuttujat Delta7 ja Delta13, jotka ovat $UCI_{7kk:n}$ ja $LCI_{7kk:n}$ sekä $UCI_{13kk:n}$ ja $LCI_{13kk:n}$ erotuksia. Lopuksi luotiin vielä uudet muuttujat $\Delta_{7\text{bull}}$, $\Delta_{7\text{bear}}$, $\Delta_{13\text{bull}}$ ja $\Delta_{13\text{bear}}$, jotka kuvaavat nousu- ja laskukausien luottamusvälien vinouksien keskiarvoja.

Seuraava taulukko 18 esittää verrannollisten pariin t-testin tulokset opiskelijoiden perusaineistolle. Taulukosta nähdään, että molemmissa luottamusväliennusteissa keskimääräiset vinoudet nousukausien jälkeen olivat pienempiä kuin laskukausien jälkeen, ja erot olivat tilastollisesti erittäin merkitseviä.

TAULUKKO 18 Luottamusvälien vinouden vertailu ja testitulokset, opiskelijoiden perusaineisto

		Keskiarvo	N	Keskihajonta	t	Vapausasteet	P-arvo
Pari 1	Delta7bull	-2,9528	143	4,59660	-7,623	142	,000
	Delta7bear	,1071	143	2,53814	(***)		
Pari 2	Delta13bull	-3,1324	143	5,69015	-5,867	142	,000
	Delta13bear	,0672	143	4,23700	(***)		

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Nousukausien jälkeen keskimääräinen vinous oli negatiivinen molempien luottamusvälien kohdalla, mikä tarkoittaa, että luottamusvälit olivat vinoja vasemmalle. Näin ollen näytetyn osakekuvaajan keskiarvo veti vastaajien ylä- ja ala-

luottamusvälejä alas, eli vastaajat suojautuivat mahdolliseen hintojen romahdukseen. Laskukausien jälkeen keskimääräinen vinous oli molemmissa tapauksissa positiivinen, mikä siis tarkoittaa, että luottamusvälit olivat vinoja oikealle. Näin ollen näytetyn osakekuvaajan keskiarvo veti ylä- ja alaluottamusvälejä ylös. Voimme siis todeta, että ankkurointiefektiin perustuva, De Bondtin luoma teoria luottamusvälien suojauksesta piti täydellisesti paikkansa opiskelijoiden perusaineistossa.

Koska vain muuttuja $\Delta_{13\text{bull}}$ oli normaalisti jakautunut, tehtiin vielä parametrin WSR-testi. Sen mukaan H_{30} hylättiin ensimmäisen parin osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure $-6,829$) ja jälkimmäisen parin osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure $-5,768$).

Seuraava taulukko 19 esittää keskiarvovertailujen tulokset pankin perusaineistolle.

TAULUKKO 19 Luottamusvälien vinouden vertailu, pankin perusaineisto

		Keskiarvo	N	Keskihajonta
Pari 1	Delta7bull	-2,9369	37	4,58987
	Delta7bear	-,5766	37	2,90981
Pari 2	Delta13bull	-3,1712	37	5,87032
	Delta13bear	-1,2613	37	3,96056

Taulukosta 19 nähdään tässäkin tapauksessa, että luottamusvälien vinouksien keskiarvot olivat nousukausien jälkeen jälleen pienemmät kuin laskukausien jälkeen. Huomataan myös, että molemmissa luottamusväliennusteissa keskimääräinen vinous nousu- ja laskukausien jälkeen oli negatiivinen, mikä tarkoittaa, että luottamusvälit olivat vinoja vasemmalle. Tässä tapauksessa De Bondtin luoma teoria luottamusvälien suojauksesta ei toteutunut niin vahvasti, sillä keskimääräiset vinoudet laskukausien jälkeen eivät olleet positiivisia (eli oikealle vinoja).

Verrannollisten parien t-testiä ei voitu käyttää tilastollisen merkitsevyyden tutkimiseen, sillä parit $\Delta_{7\text{bull}}$ ja $\Delta_{7\text{bear}}$ sekä $\Delta_{13\text{bull}}$ ja $\Delta_{13\text{bear}}$ korreloivat keskenään Spearmanin järjestykskorrelaatiokertoimella mitattuna. Ensimmäisen parin välinen korrelaatio oli merkitsevää (kerroin 0,531 ja p-arvo 0,001) ja toisen parin melkein merkitsevää (kerroin 0,403 ja p-arvo 0,013). Tästä syystä hypoteesia tutkittiin WSR-testillä. Testin mukaan H_{30} hylättiin molempien parien osalta (Z-testisuureksi ja p-arvoksi ensimmäiselle parille saatiin $-3,433$ ja 0,001 sekä toiselle parille $-2,019$ ja 0,043). Luottamusvälien keskimääräinen vinous oli näin ollen pienempi nousukausien jälkeen.

Yhdistetyllä perusaineistolla saadut tulokset olivat yhteneväisiä pankin perusaineistolla saatujen tulosten kanssa. Taulukon 20 perusteella huomataan, että luottamusvälien keskimääräinen vinous yhdistetyssä perusaineistossa oli pienempi nousukausien jälkeen molempien parien kohdalla. Tulosten perusteella voidaan sanoa, että De Bondtin teoria luottamusvälien suojauksesta toteutui tässäkin tapauksessa, mutta ei niin vahvasti kuin opiskelijoiden perusaineistossa, sillä luottamusvälit olivat vinoja vasemmalle myös laskukausien jälkeen.

TAULUKKO 20 Luottamusvälien vinouden vertailu, yhdistetty perusaineisto

		Keskiarvo	N	Keskihajonta
Pari 1	Delta7bull	-2,9495	180	4,58239
	Delta7bear	-,0334	180	2,62491
Pari 2	Delta13bull	-3,1404	180	5,71104
	Delta13bear	-,2059	180	4,20547

Koska molemmissa pareissa muuttujien välillä oli riippuvuutta, tutkittiin hypoteesia WSR-testillä (Spearmanin järjestyskorrelaatiokertoimen arvo ja p-arvo ensimmäiselle parille 0,151 ja 0,044 sekä jälkimmäiselle parille 0,151 ja 0,043). Testin mukaan H_{30} hylättiin molempien pariin osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure ensimmäiselle parille -7,634 ja toiselle -6,055).

Näin ollen De Bondtin kehittämä teoria luottamusvälien suojauksesta piti paikkansa tämän tutkimuksen perusaineistojen perusteella. Opiskelijoiden antamat luottamusvälit nousukausien jälkeen eivät olleet symmetrisiä, vaan vasemmalle vinoja. Laskukausien jälkeen luottamusvälit olivat vinoja oikealle. Pankin perusaineiston ja yhdistetyn perusaineiston perusteella voidaan sanoa, että luottamusvälien vinous oli nousukausien jälkeen pienempi kuin laskukausien jälkeen.

5.1.4 Regressioanalyysit luottamusvälien vinoudelle

Tämän työn keskeisimpiä tuloksia ovat tähän mennessä olleet trendinseuraaminen ja subjektiivisten todennäköisyysjakaumien vinous. De Bondtin luoma teoria luottamusvälien suojauksesta piti paikkansa myös tämän tutkimuksen perusaineistojen osalta. H_4 liittyy kysymykseen, voidaanko vinouden parametreja selittää vastaavilla odotetuilla hintamuutoksilla ja trendinseurantaluokituksella.

H_4 : Luottamusvälit kasvavat enemmän oikealle vinoiksi, kun suurempaa hinnanlaskua ennustetaan ja toisinpäin.

Hypoteesia tutkittiin regressioanalyysillä, jota varten perusaineistoihin muodostettiin joukko uusia muuttujia. Trendinseuraamiseen liittyen yhdistettiin ennustetun hintamuutoksen vahvuuden tutkimista varten nousu- ja laskumarkkinoiden osakekuvaajat uusiksi muuttujiksi TF_bull ja TF_bear. Yhdistäminen suoritettiin laskemalla keskiarvo kolmen nousu- ja laskuosakekuvaajan $F_{7:n}$ ja $P_{0:n}$ sekä $F_{13:n}$ ja $F_{7:n}$ välisistä prosentuaalisista muutoksista. Näin nousu- ja laskukausille saatiin vain yksi "aggregaattimuuttuja", joka kuvaa ennustetun hintamuutoksen vahvuutta ja siihen liittyvää käyttäytymistä (TF tulee sanoista trend follower). Muuttujat TF_bull ja TF_bear saivat arvoja 1-4 seuraavasti:

1. Jos $P_0 < F_7 < F_{13}$, niin kyseessä oli vahva nousun seuraaja
2. Jos $P_0 > F_7 > F_{13}$, niin kyseessä oli vahva laskun seuraaja

3. Jos $P_0 < F_7 > F_{13}$, niin kyseessä oli heikko nousun seuraaja
4. Jos $P_0 > F_7 < F_{13}$, niin kyseessä oli heikko laskun seuraaja.

Tämän jälkeen luotiin muuttujat TF_bear_CO ja TF_bull_CO, jotka jaottelivat vastaajat trendinseuraajiin ja toisinajattelijoihin yhdistämällä muuttujien TF_bull ja TF_bear luokat, ottaen huomioon nousu- ja laskukaudet. Toisin sanoen nousukausien muuttujan TF_bull osalta luokat 1. ja 3. saivat uudessa muuttujassa TF_bull_CO arvon 1, mikä merkitsee vahvaa tai heikkoa trendinseuraajaa. Vastaavasti luokkien 2. ja 4. osalta TF_bull_CO sai arvon 0, mikä tarkoittaa vahvaa tai heikkoa toisinajattelijaa. Laskukausien muuttujan TF_bear osalta luokat menivät päinvastoin: luokat 2. ja 4. saivat muuttujassa TF_bear_CO arvon 1 (trendinseuraaja) ja luokat 1. ja 3. arvon 0 (toisinajattelijat). Edellä esitettyjen muuttujien luonnissa käytetyt syntaksikomennot on esitetty liitteessä 7.

Vastaavalla tavalla muodostettiin muuttujat TF_bull_COstrong ja TF_bear_COstrong, jotka ottivat huomioon vain trendinseuraaja- ja toisinajattelijaluokkien vahvat muodot. Mikäli vastaaja ei kuulunut kumpaakaan vahvaan luokkaan, ei uusi muuttuja saanut mitään arvoa.

Edellä mainittujen muuttujien lisäksi luotiin uudet muuttujat EPC_bull ja EPC_bear, jotka laskettiin nousu- ja laskukausien odotettujen hintamuutosten keskiarvoina, sekä Delta_bull ja Delta_bear, jotka laskettiin nousu- ja laskukausien luottamusvälien vinouksien keskiarvoina.

Seuraavat havainnot tukivat ajatusta valita regressioanalyysin selittäjämuuttujiksi vain odotettu hintamuutos ja trendinseuraamis- tai toisinajattelijaluokittelu: opiskelijoiden perusaineistolla huomattiin, että suoritettujen rahoituksen kurssien määrä, sukupuoli, ikä, pääaine tai sijoitusharrastus eivät korreloineet muuttujien Delta_bull tai Delta_bear kanssa. Pankin perusaineiston kohdalla huomattiin puolestaan, etteivät sukupuoli, ikä tai rahoituksen alan työkokemus korreloineet muuttujien Delta_bull tai Delta_bear kanssa. Näin ollen regressionanalyysiä ei olisi ollut järkevää tehdä näiden muuttujien pohjalta.¹⁴

Opiskelijoiden perusaineistossa saatiin korrelaatiotarkastelujen yhteydessä selville, että muuttuja Delta_bull korreloi tilastollisesti melkein merkitsevästi muuttujien EPC_bull ja TF_bull_CO kanssa. Muuttujien EPC_bull ja TF_bull_CO välinen korrelaatio oli puolestaan tilastollisesti merkitsevää. Muuttuja Delta_bear taas korreloi merkitsevästi sekä EPC_bear:in kanssa että TF_bear_COstrong:in kanssa. Lisäksi huomattiin, että muuttujat EPC_bear ja TF_bear_COstrong korreloivat merkitsevästi keskenään.

Korrelaatiotarkastelujen perusteella päätettiin ajaa neljä regressiota. Ensimmäinen regressio suhteeseen muuttujaa Delta_bull erikseen muuttujien EPC_bull ja TF_bull_CO suhteen (mallit 1. ja 2.). Seuraavaksi regressio suhteeseen muuttujaa Delta_bear erikseen muuttujien EPC_bear ja TF_bear_COstrong suhteen (mallit 3. ja 4.). Puuttuvat tiedot poistettiin analyysistä komennolla Exclude cases pairwise.

¹⁴ Regressioanalyysissä edeltäneissä korrelaatiotarkasteluissa käytettiin Spearmanin järjestyskorrelaatiokerrointa.

Seuraavasta taulukosta 21 nähdään tulokset. Malleilla 1. ja 2. selitysasteet olivat huonoja, eivätkä ne varianssianalyysin mukaan sopineet aineistoon (tosin mallin 2. tapauksessa F-testin p-arvo oli 0,054, eli malli oli lähellä sopia aineistoon). Mallien 3. ja 4. selitysasteet olivat kohtalaisia, ja varianssianalyysin mukaan kyseiset mallit sopivat aineistoon.

TAULUKKO 21 Regressioanalyysien tulokset, opiskelijoiden perusaineisto

	Delta_bull	Delta_bull	Delta_bear	Delta_bear
Vakio	-2,961 (-7,726)	-1,787 (-2,422)	-1,265 (-3,654)	-3,407 (3,605)
EPC_bull	-,058 (-,621)			
EPC_bear			-,648 (-5,331 ^{***})	
TF_bull_CO		-1,665 (-1,945)		
TF_bear_COstrong				4,281 (4,053 ^{***})
Selitysaste R ²	,003	,027	,168	,218
F-testisuure	,385	3,784	28,419 ^{***}	16,424 ^{***}
Durbin-Watson	2,188	2,158	1,796	1,858
N	143	138	142	60

T-testisuure esitetty suluissa regressiokertoimen alla.

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

Mallien 3. ja 4. perusteella voidaan kirjoittaa regressioyhtälöt:

$$y = -1,265 - 0,648 * EPC_bear$$

$$y = -3,407 + 4,281 * TF_bear_COstrong$$

Sekä muuttuja EPC_bear että muuttuja TF_bear_COstrong vaikuttivat merkittävästi muuttujaan Delta_bear. H₄:n kannalta ensimmäiseen yhtälöön sisältyvä tärkeä tulos: kun EPC_bear pienenee, kasvaa Delta_bear:in arvo 0,648* EPC_bear:in muutos, eli vinous oikealle kasvaa. Jälkimmäisestä yhtälöstä nähdään, että kun kyse on vahvasta trendinseuraajasta (muuttujan TF_bear_COstrong saadessa arvon 1), niin Delta_bear saa arvo 0,87, mikä tarkoittaa, että luottamusvälit ovat vinoja oikealle. Kun kyseessä on vahva toisinajattelija (muuttujan TF_bear_COstrong saadessa arvon 0), saa Delta_bear arvon -3,4 eli luottamusvälit ovat vinoja vasemmalle.

Suoritettujen regressioanalyysien perusteella voidaan sanoa, että H₄ sai tukea mallin 3. tapauksessa: luottamusvälit kasvoivat enemmän oikealle vinoiksi, kun suurempaa hinnanlaskua ennustettiin. Mallissa 1. regressiokerroin muuttujalla EPC_bull oli myös hypoteesin mukainen eli negatiivinen, mutta ei tilastollisesti merkitsevä.

Regressioanalyysin tulosten tulkinnassa on muistettava tutkia analyysimenetelmän melko vahvojen oletusten voimassaolo. Seuraavaksi oletuksia tutkitaan vain mallien 3. ja 4. osalta. Regressionanalyysin yksi perusoletus on, että residuaalien jakauma on normaalijakauman kaltainen. Mallin 3. kohdalla histogrammi standardisoitujen jäännöstermien jakaumasta oli huipukkaampi kuin

normaalijakauma. Molempien mallien jäännösten kumulatiivinen prosenttijakauma oli toisen asteen käyrän muotoinen, muistuttaen painofunktioiden kuvaajien muotoa kuviossa 3 (s. 19), kun sen pitäisi asettua tasaisesti 45 asteen diagonaalille. Mallin 3. hajontakuvio standardoiduista residuaaleista ja ennusteista osoitti, että jäännökset olivat kohtalaisen hyvin hajaantuneet ennusteen eri arvoilla, mikä kertoo, ettei suuria poikkeamia varianssien homegeenisuudesta ollut. Suurta poikkeamaa lineaarisuuden oletuksesta ei hajontakuvion muodon perusteella myöskään ollut nähtävissä.

Regressioanalyysi edellyttää myös, että jatkuvat muuttujat olisivat likimain normaalisti jakautuneita. Kolmogorov-Smirnovin normaaliustestin mukaan tämä oletus toteutui vain muuttujan Delta_bull osalta. Selittäjämuuttujat eivät saisi korreloida keskenään (multikollinearisuus), mikä ei ollut tässä ongelma. Havaintojen virhetermien tulisi olla toisistaan riippumattomia ja tätä voidaan testata Durbin-Watsonin testillä. Malleissa 3. ja 4. D-W testin arvo oli lähellä kakkosta, mikä on hyvä asia. Autokorrelaatiota ei näin ollen esiintynyt.

Malleihin 3. ja 4. vaikuttaneita tekijöitä tutkittiin vielä standardoitujen residuaalien, vipuvaikutuksen (leverage), Cookin etäisyysluvun (Cook's distance) ja DFBeta-lukujen avulla. Standardoitujen residuaalien avulla voidaan tutkia poikkeavien havaintojen määrää mallissa. Tällöin verrataan normalisoitujen residuaalien arvoa sekä määrää normaalijakauman arvoihin. Vipuvaikutus tutkii, kuinka kaukana selittävän muuttujan yksittäinen havainto on muista saman muuttujan havainnoista. Cookin etäisyysluvussa puolestaan tutkitaan yksittäisen poikkeavan havainnon vaikutusta malliin. Suuren arvon saanut havainto vaikuttaa regressioyhtälön kulmakertoimeen enemmän kuin muut havainnot, mikä voi vääristää mallin tulosta. DFBeta-lukujen avulla on mahdollista löytää tapaukset, joilla on suuri vaikutus regressiomallien parametreihin.

Mallissa 3. standardoitujen residuaalien arvoista kolme oli yli raja-arvon (3) ja yli 1 % arvoista ylitti 2,5:n, mikä kertoo poikkeavista havainnoista. Cookin etäisyysluvun perusteella mallissa 3. ei ollut yhtään sallitun rajan ylittäviä arvoja (yli 1:n). Vipuvakutusluvussa seitsemän arvoa oli yli 2 x keskiarvon ja yksi arvo yli 3 x keskiarvon. Sekä vakion että selittäjämuuttujan DFBeta-luvuissa oli yksi arvo yli sallitun (yli 1:n).

Mallissa 4. standardoitujen residuaalien arvoista kaksi oli yli raja-arvon (3) ja yli 1 % arvoista ylitti 2,5:n, mikä kertoo poikkeavista havainnoista tässäkin tapauksessa. Cookin etäisyysluvun perusteella mallissa 4. oli yksi sallitun rajan ylittävä arvo (yli 1:n). Vipuvakutusluvussa 10 arvoa oli yli 2 x keskiarvon, mutta ei yhtään yli 3 x keskiarvon. DFBeta-luvuissa oli yksi arvo yli sallitun (yli 1:n) sekä vakion että selittäjämuuttujan kohdalla.

Edellä mainittujen takastelujen perusteella havaittiin, että mallit 3. ja 4. sisältivät virheitä. Poikkeavien havaintojen poistaminen saattaisi parantaa malleja, mutta tässä ei tätä muokkausta enää tehty.

Seuraavaksi tehtiin regressioanalyysi pankin perusaineistolle. Korrelaatiotarkastelujen yhteydessä huomattiin, että muuttuja Delta_bull ei korreloinut tilastollisesti merkittävästi minkään $H_4:n$ testaukseen liittyvän relevantin selittäjämuuttujan kanssa. Muuttuja Delta_bear korreloi tilastollisesti merkittävästi

vain muuttujan EPC_bear kanssa. Näin ollen päätettiin ajaa vain yksi regressio, jossa muuttujaa Delta_bear regressioitiin EPC_bear:in suhteen. Taulukko 22 esittää tulokset.

TAULUKKO 22 Regressioanalyysin tulokset, pankin perusaineisto

	Delta_bear
Vakio	-1,098 (-2,179)
EPC_bear	-,254 (-1,640)
Selitysaste R ²	,071
F-testisuure	2,688
Durbin-Watson	1,733
N	37

T-testisuure esitetty suluissa regressiokertoimen alla.

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Taulukosta 22 nähdään, että mallin selitysaste oli melko huono, eikä malli varianssianalyysin mukaan sopinut aineistoon. Muuttuja EPC_bear ei vaikuttanut merkitsevästi muuttujaan Delta_bear. H₄:n kannalta tarkasteltuna muuttujan EPC_bear regressiokerroin oli hypoteesin mukainen eli negatiivinen, mutta tulos ei ollut tilastollisesti merkitsevä, joten tämän enempää ei tässä tulkita tuloksia tai pohdita analyysimenetelmän oletusten voimassaoloa.

Yhdistetyn perusaineiston korrelaatiotarkasteluissa saatiin selville, että muuttuja Delta_bull korreloi merkitsevästi muuttujien EPC_bull ja TF_bull_CO kanssa. Muuttujat EPC_bull ja TF_bull_CO puolestaan korreloivat merkitsevästi keskenään. Delta_bear taas korreloi merkitsevästi muuttujien EPC_bear ja TF_bear_COstrong kanssa. Lisäksi muuttuja EPC_bear korreloi merkitsevästi muuttujan TF_bear_COstrong kanssa.

Korrelaatiotarkastelujen perusteella päätettiin ajaa neljä regressiota. Ensin regressioitiin Delta_bull:ia erikseen muuttujien EPC_bull ja TF_bull_CO suhteen (mallit 2. ja 3.). Sen jälkeen regressioitiin Delta_bear:ia erikseen muuttujien EPC_bear ja TF_bear_COstrong suhteen (mallit 3. ja 4.). Puuttuvat tiedot poistettiin analyysistä komennolla Exclude cases pairwise. Seuraava taulukko 23 esittää tulokset.

TAULUKKO 23 Regressioanalyysien tulokset, yhdistetty perusaineisto

	Delta_bull	Delta_bull	Delta_bear	Delta_bear
Vakio	-2,917 (-8,673)	-1,686 (-2,733)	-1,013 (-3,747)	-3,361 (-4,692)
EPC_bull	-,107 (-1,258)			
EPC_bear			-,501 (-5,417***)	
TF_bull_CO		-1,855 (-2,557*)		
TF_bear_COstrong				3,997 (5,025***)

(jatkuu)

TAULUKKO 23 (jatkuu)

	Delta_bull	Delta_bull	Delta_bear	Delta_bear
Selitysaste R ²	,009	,036	,142	,260
F-testisuure	1,583	6,538(*)	29,348(***)	25,250(***)
Durbin-Watson	2,132	2,130	1,802	1,903
N	180	175	179	73

T-testisuure esitetty suluissa regressiokertoimen alla.

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Taulukosta 23 nähdään, että malleilla 1. ja 2. selitysasteet olivat huonoja. Malli 1. ei varianssianalyysin mukaan sopinut aineistoon. Mallien 3. ja 4. selitysasteet olivat kohtalaisia, ja varianssianalyysin mukaan kyseiset mallit sopivat aineistoon.

Mallien 3. ja 4. perusteella voidaan kirjoittaa regressioyhtälöt:

$$y = -1,013 - 0,501 * EPC_bear$$

$$y = -3,361 + 3,997 * TF_bear_COstrong$$

Sekä muuttuja EPC_bear että muuttuja TF_bear_COstrong vaikuttivat merkitsevästi muuttujaan Delta_bear. H₄:n kannalta ensimmäiseen yhtälöön sisältyvä tärkeä tulos: kun EPC_bear pienenee, kasvaa Delta_bear:in arvo 0,501*EPC_bear:in muutos eli vinous oikealle kasvaa. Muutos Delta_bear:ssa on puolet EPC_bear:in muutoksesta. Jälkimmäisestä yhtälöstä nähdään, että kun kyse on vahvasta trendinseuraajasta (muuttujan TF_bear_COstrong saadessa arvon 1), niin Delta_bear saa arvo 0,64, mikä tarkoittaa, että luottamusvälit ovat vinoja oikealle. Kun kyseessä on vahva toisinajattelija (muuttujan TF_bear_COstrong saadessa arvon 0), saa Delta_bear arvon -3,36 eli luottamusvälit ovat vinoja vasemmalle.

Suoritettujen regressioanalyysien perusteella voidaan sanoa, että H₄ sai tukea yhdistettyjen perusaineistojen perusteella mallin 3. tapauksessa: luottamusvälit kasvoivat enemmän oikealle vinoiksi, kun suurempaa hinnanlaskua ennustettiin. Mallin 1. osalta regressiokerroin muuttujalla EPC_bull oli hypoteesin mukainen eli negatiivinen, mutta tulos ei ollut tilastollisesti merkitsevää.

Lopuksi tutkitaan regressioanalyysin oletusten voimassaoloa mallien 3. ja 4. kautta. Molempien mallien histogrammit standardisoitujen jäännöstermien jakaumasta olivat huipukkaampia kuin normaalijakauma. Molemmissa malleissa jäännösten kumulatiivinen prosenttijakauma oli jälleen toisen asteen käyrän muotoinen, muistuttaen painofunktioiden kuvaajien muotoa kuviossa 3 (s. 19), mutta mallissa 4. jakauma hiukan suurempi. Mallin 3. kohdalla hajontakuvio standardoiduista residuaaleista ja ennusteista osoitti, että jäännökset olivat kohtalaisesti kasaantuneet kuvion oikeaan reunaan, mikä kertoo, että varianssit eivät olleet homogeeniset.

Kolmogorov-Smirnovin testin mukaan muuttuja EPC_bull ei ollut normaalista jakautunut, mikä on vastoin regressioanalyysin oletuksia. Malleissa 3. ja 4. Durbin-Watsonin testin arvot olivat lähellä kakkosta, eli autokorrelaatiota ei esiintynyt, mikä on hyvä asia.

Malleihin 3. ja 4. vaikuttaneita tekijöitä tutkittiin tässäkin standardoitujen residuaalien, vipuvaikutuksen (leverage), Cookin etäisyysluvun (Cook's distance) ja DFBeta-lukujen avulla. Mallissa 3. standardoitujen residuaalien arvoista viisi oli yli raja-arvon (3) ja yli 1 % arvoista ylitti 2,5:n, mikä kertoo poikkeavista havainnoista. Cookin etäisyysluvun perusteella mallissa 3. oli yksi sallitun rajan (yli 1:n) ylittävä arvo. Vipuvakutusluvussa neljä arvoa oli yli 2 x keskiarvon ja kaksi arvoa yli 3 x keskiarvon. Sekä vakion että selittäjämuuttujan DFBeta-luvuissa oli yksi arvo yli sallitun (yli 1:n).

Mallin 4. kohdalla standardoitujen residuaalien arvoista kaksi oli yli raja-arvon (3) ja yli 1 % arvoista ylitti 2,5:n. Cookin etäisyysluvun perusteella mallissa 4. oli yksi sallitun rajan (yli 1:n) ylittävä arvo. Vipuvakutusluvussa 14 arvoa oli yli 2 x keskiarvon, mutta ei yhtään yli 3 x keskiarvon. Sekä vakion että selittäjämuuttujan DFBeta-luvuissa oli yksi arvo yli sallitun (yli 1:n).

Edellä esitetyn perusteella molemmat mallit siis sisälsivät virheitä. Malleja voisi parantaa poistamalla poikkeavat havainnot, mutta tässä malleja ei enää muokattu.

Kaikilla perusaineistoilla ajettiin siis neljä samanlaista regressiomallia, joissa selittäjämuuttujia oli vain yksi. Koemielessä tehtiin myös opiskelijoiden perusaineistolle kaksi monimuuttujaregressiomallia, jossa ensimmäisessä Delta_bull:ia selitettiin muuttujilla EPC_bull ja TF_bull_CO ja toisessa Delta_bear:ia muuttujilla EPC_bear ja TF_bear_COstrong, vaikka tiedettiin multikollinearisuongelmasta. Tuloksista nähtiin, että ensimmäinen malli ei varianssianalyysin mukaan sopinut aineistoon ja malli selitysaste oli huono. Jälkimmäisen mallin selitysaste oli kohtalainen (21,8 %) ja malli sopi varianssianalyysin mukaan aineistoon. T-testin tulosten mukaan kumpikaan selittäjistä, EPC_bear ja TF_bear_COstrong, eivät näyttäneet kuitenkaan vaikuttavan merkittävästi muuttujaan Delta_bear. Multikollinearisuutta tutkivat Tolerance- ja VIF-luku olivat yli sallittujen rajojen (Tolerance-luku molemmille selittäjille 0,253 ja VIF-luku 3,946).

5.1.5 Yliluottamus

Seuraavat hypoteesit liittyvät yliluottamukseen:

H₅: Pankin sijoitusasiantuntijat ovat yliluottavaisempia kuin opiskelijat.

H₉: Miehet ovat yliluottavaisempia kuin naiset.

Yliluottavaiset henkilöt arvioivat luottamusvälit kapeammiksi ja arvioivat tulevia osaketuottoja optimistisemmin (eli suuremmiksi) kuin muut. Opiskelija-aineistossa yliluottamusta voidaan mitata myös sen kautta, miten sijoittamista harrastavat näkivät oman salkun tuoton suhteessa markkinatuottoon¹⁵. Seuraavaksi keskitytään jälleen ensin opiskelijoiden perusaineistoon, sitten pankin perusaineistoon ja lopuksi tehdään analyysyjä yhdistetylle perusaineistolle.

¹⁵ Markkinatuotto tarkoittaa tässä siis opiskelijoiden antamaa keskiarvotuottoennustetta tulevalle 12 kuukaudelle.

Aluksi ryhmiteltiin opiskelijat kolmeen ryhmään luottamuksen mukaan seuraavasti:

1. Oman salkun tuotto > markkinatuotto = ylliluottavainen
2. Oman salkun tuotto = markkinatuotto = neutraali
3. Oman salkun tuotto < markkinatuotto = aliluottavainen.

Näin muodostettiin uusi luottamusryhmittelymuuttuja LUOTTO, joka siis sai arvoja 1-3. Tämän jälkeen testattiin ristiintaulukoinnin ja χ^2 -testin avulla sukupuolen ja edellä olevan luottamusryhmittelyn merkitsevyyttä: haluttiin siis tutkia, oliko sukupuolella vaikutusta luottamusluokitteluun. Selitettäväksi eli riippuvaksi muuttujaksi valittiin luottamusryhmittely ja selittäväksi eli riippumattomaksi muuttujaksi sukupuoli.

χ^2 -testin edellytykset eivät kuitenkaan täytyneet, joten tehtiin Fisherin tarkka testi. Tilastollisesti merkitseviä eroja ei sen perusteella saatu, p-arvo oli 0,737. Sukupuolella ei siis ollut tilastollisesti merkitsevää vaikutusta luottamusluokitteluun.

Luottamusvälien kapeuden laskemiseksi luotiin apumuuttuja LUOTTAMUSVÄLI, joka saatiin vähentämällä ylliluottamusvälin arvosta alaluottamusvälin arvo, esimerkiksi A3-A4 jne. Sen jälkeen testattiin riippumattomien ryhmien t-testillä luottamusvälien keskiarvojen eroja sukupuolittain. Taulukko 24 esittää testin tulokset. Testin mukaan luottamusvälien keskiarvoissa oli tilastollisesti merkitsevää eroa miesten ja naisten välillä vain neljässä tapauksessa: C- ja E-osakkeiden 13 kuukauden luottamusvälissä sekä F-osakkeen seitsemän ja 13 kuukauden luottamusvälissä. Opiskelijamiehet eivät siis näyttäneet olevan naisia ylliluottavaisempia kuin näissä neljässä tapauksessa, joten muilta osin H_0 ei saanut tukea.

TAULUKKO 24 Luottamusvälien kapeuden testaus, opiskelijoiden perusaineisto

	t	Vapausasteet	P-arvo
Aluottamusväli7kk	-,472	142	,637
Aluottamusväli13kk	-,703	142	,483
Bluottamusväli7kk	-1,765	140	,080
Bluottamusväli13kk	-1,829	140	,069
Cluottamusväli7kk	-1,806	141	,073
Cluottamusväli13kk	-2,328(*)	141	,021
Dluottamusväli7kk	-,726	139	,469
Dluottamusväli13kk	-,329	139	,743
Eluottamusväli7kk	-1,799	139	,074
Eluottamusväli13kk	-2,072(*)	139	,040
Fluottamusväli7kk	-2,763(**)	139	,006
Fluottamusväli13kk	-3,932(***)	139	,000

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Riippumattomien ryhmien t-testin lisäksi testattiin luottamusvälien kapeuksien eroja parametrittömällä MWU-testillä, sillä t-testin oletukset eivät kaikilta osin täytyneet. Parametrittoman testin mukaan luottamusvälien sijalukujen keskiarvoissa oli sukupuolien välillä tilastollisesti merkitsevää eroa vain neljässä

tapauksessa: B- ja C-osakkeen 13 kuukauden luottamusvälissä sekä F-osakkeen seitsemän ja 13 kuukauden luottamusvälissä (U-testisuureet 1933,5; 1890,5; 1793,0 ja 1613,5 sekä vastaavat p-arvot 0,047; 0,024; 0,15 ja 0,001). H_0 sai siis tukea vain näiden tapausten osalta, mutta muilta osin ei. Tulos poikkesi jonkin verran t-testin tuloksista.

Entä ennustivatko opiskelijanaiset tulevia osaketuottoja varovaisemmin kuin miehet? Tätä kysymystä testattiin riippumattomien ryhmien t-testillä, jossa verrattiin osake-ennusteiden piste-estimaattien keskiarvojen eroja sukupuolittain. Tilastollisesti merkitseviä tuloksia saatiin vain C1-, D1- ja D2 -osake-ennusteiden kohdalla, kuten taulukosta 25 nähdään: vain kyseisten ennusteiden kohdalla miesten ja naisten keskiarvojen välillä oli tilastollisesti melkein merkitsevää eroa, eli naiset ennustivat niissä osaketuottoja varovaisemmin. Näin ollen H_0 sai tukea vain kolmessa tapauksessa 12:sta (ennustettavia piste-estimaatteja oli kaikkiaan 12).

TAULUKKO 25 Osake-ennusteiden varovaisuuden testaus, opiskelijoiden perusaineisto

	t	Vapausasteet	P-arvo
A1	-1,161	93,181	,249
A2	-1,658	80,616	,101
B1	-,114	143	,910
B2	-1,481	142	,141
C1	2,295(*)	143	,023
C2	1,841	143	,068
D1	-2,306(*)	142	,023
D2	-2,412(*)	91,517	,018
E1	-1,882	140	,062
E2	-1,540	140	,126
F1	-1,147	140	,254
F2	-1,914	140	,058

Merkitsevyystasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Naisopiskelijoiden varovaisuutta testattiin myös MWU-testillä. Sen mukaan muuttujien sijalukujen keskiarvot poikkesivat tilastollisesti merkitsevästi vain seuraavien neljän osake-ennusteen kohdalla: C1, D1, D2 ja F2 (U-testisuureet 1876,5; 1880,0; 1843,0 ja 1916,0 sekä vastaavat p-arvot 0,010; 0,014; 0,012 ja 0,039), joten melko samankaltainen tulos saatiin kuin t-testilläkin. Parametrittoman testin perusteella H_0 sai tukea vain näiden neljän osake-ennusteen osalta ja muilta osin ei.

Seuraavaksi tutkittiin asettivatko miesvastaajat luottamusvälit kapeammiksi kuin naisvastaajat pankin perusaineistossa? Kysymystä varten luotiin pankin aineistolle ensin apumuuttuja LUOTTAMUSVÄLI, aivan kuten opiskelija-aineistossakin. Riippumattomien ryhmien t-testillä testattiin luottamusvälien keskiarvojen eroja sukupuolittain, ja tulosten mukaan tilastollisesti merkitseviä eroja ei ollut sukupuolten välillä, joten H_0 ei saanut tukea. Vastaajien vähyys saattoi tosin vaikuttaa siihen, että tilastollisesti merkitseviä tuloksia ei saatu ($N = 37$). Koska kaikki LUOTTAMUSVÄLI-muuttujat eivät olleet normaalisti jakautuneita, tehtiin lisäksi MWU-testi. Tulosten mukaan ryhmien välillä ei sijalukujen keskiarvojen suhteen ollut tilastollisesti merkitseviä eroja muissa tapa-

uksissa kuin F-osakkeen seitsemän kuukauden luottamusvälin kohdalla (U-testisuure 106,5 ja p-arvo 0,026)¹⁶. F-osakkeen seitsemän kuukauden luottamusvälin kohdalla pankin naisvastaajat arvioivat luottamusvälin kapeammaksi kuin miehet. Parametrittoman testin tulos poikkesi siis hiukan t-testin tuloksista.

Entä ennustivatko naisvastaajat tulevia osaketuottoja varovaisemmin kuin miehet pankin perusaineistossa? Tätä testattiin riippumattomien ryhmien t-testillä ja MWU-testillä, mutta tilastollisesti merkitseviä tuloksia ei saatu. Voidaan siis todeta, että pankin perusaineistolla H_9 ei saanut juuri ollenkaan tukea. MWU-testin mukaan ainoastaan yhden luottamusvälin osalta saatiin hypoteesia tukeva tulos.

Lopuksi testattiin ylikuottamusta pankin vastaajien ja opiskelijoiden kesken yhdistetyllä perusaineistolla. Riippumattomien ryhmien t-testillä testattiin ensin hypoteesia H_5 : arvioivatko pankin sijoitusasiantuntijat luottamusvälit kapeammiksi kuin opiskelijat ja ennustivatko opiskelijat tulevia tuottoja varovaisemmin kuin pankin vastaajat. Taulukosta 26 nähdään, että luottamusvälien kapeuksien perusteella ei voida sanoa, että pankin sijoitusasiantuntijat olisivat ylikuottavaisempia kuin opiskelijat. Tilastollisesti merkitsevä tulos saatiin vain C-osakkeen seitsemän kuukauden luottamusvälin kohdalla ja pankin vastaajat ennustivat siinä luottamusvälin kapeammaksi. Muilta osin siis H_5 ei saanut tukea.

TAULUKKO 26 Luottamusvälien kapeuden testaus, yhdistetty perusaineisto

	t	Vapausasteet	P-arvo
Aluottamusvali7kk	,552	179	,582
Aluottamusvali13kk	-,623	179	,534
Bluottamusvali7kk	,735	177	,463
Bluottamusvali13kk	-1,152	177	,251
Cluottamusvali7kk	2,137(*)	178	,034
Cluottamusvali13kk	,705	178	,482
Dluottamusvali7kk	,478	176	,633
Dluottamusvali13kk	,372	176	,710
Eluottamusvali7kk	1,301	176	,195
Eluottamusvali13kk	-,627	176	,531
Fluottamusvali7kk	,704	176	,483
Fluottamusvali13kk	,077	176	,938

Merkitsevyytastot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***)

Sama testaus tehtiin myös parametrittomalla MWU-testillä, sillä muuttujien normaalijakaumaoletus ei täyttnyt. MWU-testin mukaan tilastollisesti merkitsevä tulos saatiin vain C-osakkeen seitsemän kuukauden luottamusvälin kohdalla (U-testisuure 2050,5 ja p-arvo 0,035), joten aivan samansuuntainen tulos saatiin kuin riippumattomien ryhmien t-testilläkin.

Riippumattomien ryhmien t-testillä tutkittiin lisäksi, olivatko pankin vastaajat ylikuottavaisempia kuin opiskelijat. Taulukon 27 tuloksista nähdään, että osake-ennusteiden varovaisuudella mitattuna pankin sijoitusasiantuntijat olivat ylikuottavaisempia osake-ennusteiden B2, C1, C2, D1, E2 ja F1 kohdalla, mutta

¹⁶ P-arvo oli tässä Exact sig. arvo, jossa sidoksia ei korjattu.

niiden osalta tulokset olivat vain tilastollisesti melkein merkitseviä. Muilta osin H_5 ei saanut tukea.

TAULUKKO 27 Osake-ennusteiden varovaisuuden testaus, yhdistetty perusaineisto

	t	Vapausasteet	P-arvo
A1	-,751	181	,454
A2	-1,324	180	,187
B1	-,667	180	,505
B2	-2,611(*)	179	,010
C1	-2,440(*)	180	,016
C2	-2,022(*)	180	,045
D1	2,473(*)	179	,014
D2	,683	178	,496
E1	-1,732	177	,085
E2	-2,443(*)	177	,016
F1	2,065(*)	177	,040
F2	-,040	177	,968

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

MWU-testin mukaan opiskelijat ennustivat varovaisemmin vain seuraavien osake-ennusteiden kohdalla: B2, C1, D1, E2 ja F1 (U-testisuureet 1916,5; 1967,0; 1976,0; 2031,5 ja 1989,5 sekä vastaavat p-arvot 0,008; 0,012; 0,015; 0,033 ja 0,021). Muilta osin H_5 ei saanut tukea. Tässäkin tapauksessa saatiin hyvin samankaltainen tulos kuin riippumattomien ryhmien t-testillä.

Verrattaessa saatuja tuloksia teoriataustaan, voidaan todeta, että saadut tulokset vahvistavat teoriaa vain osittain. Opiskelijoiden perusaineiston perusteella miehet eivät näyttäneet kaikkien osake-ennusteiden kohdalla aliarvioivan riskiä. Yhdistetyllä aineistolla saadut tulokset taas osoittivat, että ammattilaisuus ei tarkoittanut yliluottamusta kaikkien osake-ennusteiden kohdalla.

5.1.6 Tekivätkö koulutetut henkilöt täsmällisempiä ennusteita?

Kuudes tämän opinnäytetyön hypoteeseista liittyy koulutuksen ja sijoittajien käyttäytymisessä ilmenevien virheiden yhteyteen:

H_6 : Paremmin koulutetut henkilöt tekevät vähemmän virheitä kuin heikommin koulutetut.

Tähän mennessä on käynyt ilmi, että kyselyyn vastaajat ennustivat menneen trendin jatkuvan. Tämä on vastoin perinteistä näkemystä sijoittajien rationaalisuudesta. Trendinseuraamista pidetään rahoituksen käyttäytymistieteessä virheenä, ja hypoteesia H_6 tutkittiin ensin tältä pohjalta.

Ensimmäiseksi haluttiin selvittää, oliko opiskelijoiden trendinseuraamis-käyttäytymisessä eroja sen suhteen, kuinka paljon rahoituksen kursseja opiskelijoilla oli suoritettuina. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla opiskelijoiden perusaineistolla. Analyysissä käytettiin "aggregaattimuuttujia" TF_bull ja TF_bear, jotka kuvaavat ennustetun hintamuutoksen vahvuutta ja siihen liittyvää käyttäytymistä. Muuttujat luotiin hypoteesin H_4 tutkimista var-

ten, ja liitteessä 7 on esitetty niiden luomisperiaatteet. Jotta hypoteesia H_6 voitiin järkevästi tutkia, suoritettujen rahoituksen opintojen perusteella luotiin uusi ryhmittelymuuttuja, jonka perusteella opiskelijat jaettiin kahteen ryhmään: niihin, jotka eivät olleet suorittaneet yhtään rahoituksen opintoja ja niihin, jotka olivat suorittaneet enemmän kuin yhden opintopisteen rahoituksen opintoja. Vaikka jaottelu vaikuttaa epäloogiselta, niin tällä tavalla saatiin riittävästi opiskelijoita kumpaakin ryhmään, jotta χ^2 -testi oli mielekästä tehdä.

Käyttäytymistä tutkittiin ensin nousukausien muuttujan TF_bull avulla ja sitten laskukausien muuttujan TF_bear avulla. Nämä muuttujat asetettiin selitettäviksi muuttujiksi ja suoritettujen rahoituksen kurssit selittäviksi muuttujaksi. Kummassakaan tapauksessa ei kuitenkaan saatu tilastollisesti merkitseviä tuloksia. Nousukausien analyysin tulokset olivat: $\chi^2(3) = 0,139$, p-arvo 0,987 ja kontingenssikerroin 0,031. Laskukausien analyysin osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi, jonka p-arvo oli 0,690. Näin ollen opiskelijoiden perusaineiston perusteella ei voida sanoa, että enemmän rahoituksen kursseja käyneet opiskelijat olisivat käyttäytyneet trendinseuraamisen suhteen eri tavalla kuin vähemmän rahoituksen kursseja käyneet.

Entä oliko opiskelijoiden trendinseuraamiskäyttäytymisessä eroja sen suhteen, kuinka paljon kaikkia opintoja oli suoritettuina? Tätäkin tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Analyysiä varten luotiin uusi muuttuja, jolla jaettiin opiskelijat kahteen ryhmään suoritettujen opintopisteiden perusteella: alle ja yli 100 opintopistettä suorittaneet. Muuten χ^2 -testin asetukset olivat kuten edellisessä analyysissä. Nousukausien analyysin tulokset eivät olleet tilastollisesti merkitseviä: $\chi^2(3) = 1,359$, p-arvo 0,715 ja kontingenssikerroin 0,099. Laskukausien osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Testin p-arvo oli 0,938, eli tulokset eivät olleet tilastollisesti merkitseviä. Opiskelijoiden perusaineiston perusteella ei siis voida sanoa, että trendinseuraamiskäyttäytymisessä olisi ollut eroja sen suhteen, kuinka paljon kaikkia opintoja oli suoritettuina.

Trendinseuraamiseen liittyen haluttiin vielä tutkia, oliko trendinseuraamiskäyttäytymisessä eroja sen suhteen, kuinka hyviä arvosanoja vastaajat olivat saaneet rahoituksen kursseilta. Tässäkin analyysi tehtiin ristiintaulukoinnin ja χ^2 -testin avulla. Kummankaan analyysin kohdalla χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Nousukausien osalta p-arvo oli 0,473, ja laskukausien osalta 0,936. Tilastollisesti merkitseviä tuloksia ei saatu kummassakaan tapauksessa. Trendinseuraamiskäyttäytymisessä ei siis ollut eroja rahoituksen kurssien arvosanojen suhteen.

Hypoteesia H_6 tutkittiin vielä ennustevirheen kautta. Haluttiin siis selvittää, tekivätkö enemmän rahoituksen kursseja opiskelleet tarkempia ennusteita kuin vähemmän rahoitusta opiskelleet. Lisäksi haluttiin tutkia, tekivätkö ne opiskelijat, joilla oli kaikki opintopisteet mukaan luettuna enemmän opintopisteitä koossa, tarkempia ennusteita kuin vähemmän opintopisteitä suorittaneet, tai tekivätkö parempia arvosanoja rahoituksen kursseista saaneet opiskelijat tarkempia ennusteita.

Suoritettujen rahoituksen kurssien opintopisteiden ja ennustevirheen yhteyttä tutkittiin yksisuuntaisella varianssianalyysillä, jolla voidaan tutkia yhden ryhmittelymuuttujan vaikutusta jatkuvan muuttujan vaihteluun. Tässä on otettava huomioon, että muuttujat eivät olleet normaalisti jakautuneita. Opiskelijat jaettiin ryhmiin suorittujen rahoituksen kurssien perusteella samalla tavalla kuin edellä tehtiin. Lisäksi luotiin uusi muuttuja ENNVIRHE, joka laskettiin kaikkien osake-ennusteiden virheen neliösummana.

Levenen varianssien yhtäsuuruustestin mukaan varianssit eri ryhmissä olivat yhtä suuret (p-arvo 0,290). F-testin mukaan ryhmäkeskiarvoissa ei ollut kuitenkaan eroa ryhmien välillä: $F(1) = 0,169$ ja p-arvo 0,682. Näin ollen ryhmien välillä ei ollut eroa ennustetarkkuudessa, jos käytetään 5 %:n merkitsevyystasoa, mutta 10 %:n merkitsevyystasolla mitattuna eroja löytyi: enemmän rahoituksen kurseja suorittaneilla ennustevirhe oli keskimäärin pienempi kuin niillä, joilla ei ollut yhtään rahoituksen kurseja suoritettuina. Täytyy kuitenkin muistaa, että yksisuuntaisen varianssianalyysin oletus muuttujien normaalijakautuneisuudesta ei täyttynyt. Sen vuoksi tehtiin lisäksi parametriton Kruskal-Wallis testin. Kruskal-Wallis testillä tutkitaan, ovatko k riippumatonta otosta samoin jakautuneesta perusjoukosta. Suoritettujen rahoituksen kurssien määrällä ei näyttänyt testin mukaan olevan vaikutusta ennustevirheeseen: $\chi^2(1) = 1,766$ ja p-arvo 0,184.

Suoritettujen rahoituksen kurssien määrän yhteyttä ennustevirheeseen tutkittiin vielä korrelaatiomatriisin avulla. Spearmanin korrelaatiokerrointa ja sen merkitsevyyden testiä käytetään järjestysasteikollisille sijaluvuille tai jos muuttujat eivät ole normaalisti jakautuneet. Testin mukaan ennustevirheen ja suorittujen rahoituksen kurssien määrän välinen korrelaatiokerroin oli -0,153 ja p-arvo 0,073, eli muuttujat eivät korreloineet keskenään, mikäli käytetään tiukempaa 5 %:n merkitsevyystasoa.

Seuraavaksi tarkasteltiin kaikkien suorittujen opintopisteiden vaikutusta ennustevirheisiin yksisuuntaisella varianssianalyysillä. Opiskelijat jaettiin kahteen ryhmään suorittujen opintopisteiden perusteella samalla tavoin kuin edellä tehtiin trendinseuraamista tutkittaessa. Levenen varianssien yhtäsuuruustestin mukaan ryhmien varianssit olivat tässä tapauksessa yhtä suuret (p-arvo 0,122), mutta oletus muuttujien normaalijakautuneisuudesta ei pätenyt. Varianssianalyysin mukaan suorittujen opintopisteiden määrällä ei ollut tilastollisesti merkitsevää vaikutusta ennustevirheeseen: ryhmien välinen $F(1) = 1,347$ ja p-arvo 0,248. Samansuuntainen tulos saatiin myös Kruskal-Wallis testillä: suorittujen opintopisteiden määrällä ei ollut testin mukaan vaikutusta ennustevirheeseen: $\chi^2(1) = 0,324$ ja p-arvo 0,569.

Kokonaisopintopistemäärän ja ennustevirheen yhteyttä tutkittiin niin ikään Spearmanin korrelaatiokerroimen avulla. Kerroimen perusteella muuttujat eivät korreloineet keskenään (korrelaatiokerroin 0,058 ja p-arvo 0,501).

Yksisuuntaisella varianssianalyysillä yritettiin vielä tutkia, oliko rahoituksen kurssien arvosanoilla vaikutusta ennustevirheisiin. Eri arvosanaryhmiin jakautui niin vähän opiskelijoita (alle 20 opiskelijaa), että analyysiä ei kuitenkaan ollut mielekäästä viedä loppuun.

Opiskelijoiden perusaineiston perusteella ei siis voida sanoa, että rahoituksen alaan liittyvä sofistikoituneisuus olisi ollut yhteydessä lievempään virhekäyttäytymiseen, kun trendinseuraamista pidetään yhtenä virhekäyttäytymisen muotona. Ei myöskään voida sanoa, että rahoituksen alaan liittyvä sofistikoituneisuus olisi ollut vahvasti yhteydessä ennustetarkkuuteen. Tässä mielessä tulosten voidaan nähdä osoittavan, että osakekurssija on vaikea ennustaa menneiden tuottojen perusteella.

5.1.7 Sijoituskokemus ja virheet

Seitsemäs hypoteesi liittyy sijoituskokemuksen ja virheiden väliseen yhteyteen:

H_7 : Sijoituskokemus vähentää sijoittajien tekemiä virheitä.

Ensimmäiseksi haluttiin tutkia, oliko sijoittamista harrastavien ja ei-harrastavien opiskelijoiden välillä eroa trendinseuraamisessa. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Selitettäväksi eli riippuvaksi muuttujaksi valittiin vuorollaan TF_bull ja TF_bear ja selittäväksi eli riippumattomaksi muuttujaksi sijoitusharrastus. Kummankaan muuttujan kohdalla tulokset eivät kuitenkaan olleet tilastollisesti merkitseviä. Muuttujalle TF_bull tulokset olivat: $\chi^2(3) = 1,591$, p-arvo 0,661 ja kontingenssikerroin 0,107. Muuttujan TF_bear osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi, jonka mukaan p-arvo oli 0,652. Tällä perusteella H_7 ei siis saanut tukea, eli sijoittamista harrastavat opiskelijat käyttäytyivät trendinseuraamisen suhteen samalla tavalla kuin ei-sijoittajat.

Hypoteesia H_7 voidaan tutkia myös ennustevirheen kautta. Riippumattomien ryhmien t-testillä testattiin opiskelijoiden perusaineistolla sijoittamisharrastuksen yhteyttä ennustevirheeseen. Tulosten mukaan $t(137) = -0,590$ ja p-arvo 0,556, joten H_7 ei saanut tukea tältä osin: ennustevirheellä ei ollut tilastollisesti merkitsevää eroa sijoittamista harrastavien ja ei-harrastavien välillä. MWU-testillä saatiin samanlainen tulos (U-testisuure 1918,0 ja p-arvo 0,996), eli ennustevirheiden sijalukujen keskiarvoissa ei ollut tilastollisesti merkitsevää eroa.

Seuraavaksi tutkittiin pankin perusaineistolla, oliko rahoitusalan työkokemuksella yhteyttä trendinseuraamiskäyttäytymiseen. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Analyysiä varten muodostettiin rahoitusalan työkokemuksesta kaksikategorinen muuttuja, joka jakoi pankin vastaajat kahteen ryhmään: niihin, joilla rahoitusalan työkokemusta oli alle 10 vuotta ja niihin, joilla työkokemusta oli yli 10 vuotta. Analyysissä selitettäväksi eli riippuvaksi muuttujaksi valittiin vuorollaan TF_bull ja TF_bear. Selittäväksi eli riippumattomaksi muuttujaksi valittiin edellä mainittu rahoitusalan työkokemus. Kummankaan analyysin kohdalla χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Nousukausien osalta p-arvo oli 0,869, ja laskukausien osalta 0,422. Tilastollisesti merkitseviä tuloksia ei siten saatu. Rahoitusalan työkokemuksella ei siis ollut vaikutusta trendinseuraamiskäyttäytymiseen.

Pankin perusaineistolla yritettiin tutkia rahoitusalan työkokemuksen yhteyttä ennustevirheeseen lineaarisen regression avulla. Tässä oletettiin, että rahoitusalan työkokemus olisi yhtä kuin sijoituskokemus. Korrelaatiotarkastelussa saatiin kuitenkin selville, etteivät työkokemus rahoitusosalta ja ennustevirhe korreloineet keskenään, joten regressioanalyysiä ei tehty.

5.1.8 Lähipiiri ja virheet

Hypoteesi H_8 liittyy rahoitusalan sofistikoituneisuuteen ja sijoittamiseen liittyviin virheisiin:

H_8 : Ne vastaajat, joiden vanhemmilla on korkea sosioekonominen asema ja joiden vanhemmat ja lähipiiri harrastavat sijoittamista tekevät vähemmän sijoittamiseen liittyviä virheitä.

Ensimmäiseksi haluttiin selvittää, oliko opiskelijoiden trendinennustamiskäyttäytymisessä eroja sen suhteen, harrastivatko heidän vanhempansa sijoittamista vai eivät. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla opiskelijoiden perusaineistolla. Selitettäviksi muuttujiksi asetettiin vuorollaan TF_bull ja TF_bear ja selittäväksi muuttujaksi vanhempien sijoitusharrastus. Nousukausien analyysien osalta tilastollisesti merkitseviä tuloksia ei saatu: $\chi^2(3) = 1,808$, p-arvo 0,613 ja kontingenssikerroin 0,113. Laskukausien osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Sen mukaan p-arvo oli 0,014, eli tulos oli tilastollisesti melkein merkitsevä.

Nousukausien osalta ei siis voida sanoa, että ne opiskelijat, joiden vanhemmat harrastivat sijoittamista, käyttäytyisivät trendinseuraamisen suhteen eri tavalla kuin ne opiskelijat, joiden vanhemmat eivät harrastaneet sijoittamista. Laskukausien kohdalla suurin osa opiskelijoista, joiden vanhemmat harrastivat sijoittamista, käyttäytyivät joko vahvan (53,7 %) tai heikon (37 %) laskun seuraajan tavoin. Niistä opiskelijoista, joiden vanhemmat eivät harrastaneet sijoittamista, suurin osa käyttäytyi heikon (59 %) tai vahvan (26,5 %) laskun seuraajan tavoin. Vaikka ryhmien välillä oli tilastollisesti melkein merkitsevää eroa, molemmissa ryhmissä yli 85 % vastaajista käyttäytyi trendinseuraajan tavoin. Jos oletetaan aikaisempien tutkimusten mukaisesti, että vanhempien rahoitusalan sofistikoituneisuus välittyisi lapsille, ja sofistikoituneisuuden osoituksena pidetään sijoittamisharrastusta, pitäisi tämän vähentää opiskelijoiden trendinseuraamiskäyttäytymistä. Tässä tapauksessa kävi kuitenkin päinvastoin, sijoittamista harrastavien vanhempien lapset olivat enemmän trendinseuraajia kuin ei-sijoittavien vanhempien lapset. H_8 ei siten saanut tukea tässä tapauksessa.

Seuraavaksi haluttiin selvittää, oliko opiskelijoiden trendinennustamiskäyttäytymisessä eroja sen suhteen, harrastivatko jotkut heidän parhaimmista ystävistään sijoittamista vai eivät. χ^2 -testin edellytykset eivät tässä tapauksessa toteutuneet laskukausien osalta, joten käytettiin Fisherin tarkkaa testiä, jonka p-arvoksi saatiin 0,150. Myöskään nousukausien tulokset eivät olleet tilastollisesti merkitseviä: $\chi^2(3) = 2,191$, p-arvo 0,534 ja kontingenssikerroin 0,124. Opiskelijoiden perusaineiston perusteella ei siis voida sanoa, että ne opiskelijat,

joiden parhaimmat ystävät harrastivat sijoittamista, käyttäytyisivät trendinseurannan suhteen toisin kuin ne opiskelijat, joiden parhaimmat ystävät eivät harrastaneet sijoittamista. Tässäkin tapauksessa H_8 ei saanut tukea.

Trendinseuraamiseen liittyen haluttiin vielä selvittää, oliko opiskelijoiden trendinseurantaikäyttäytymisessä eroja sen suhteen, mikä vanhempien sosioekonominen asema oli. Analyysiä varten ryhmiteltiin vanhemmat uudelleen sosioekonomisen aseman mukaan muuttujan SOSEKAS alle. Muuttuja sai arvoja 1.- 4., jossa 1. tarkoitti yrittäjää, opiskelijaa, eläkeläistä tai muuta, 2. tarkoitti ylempää toimihenkilöä, 3. tarkoitti alempaa toimihenkilöä ja 4. tarkoitti työntekijää. Näin menetellen pyrittiin jokaiseen ryhmään saamaan riittävä määrä henkilöitä, jotta analyysi oli mielekästä tehdä. Tässä ei täysin onnistuttu, vaan pienimpään ryhmään saatiin vain 17 henkilöä. Ristiintaulukoinnissa ja χ^2 -testissä selittävinä muuttujina käytettiin vuorollaan muuttujia SOSEKAS_äiti ja SOSEKAS_isä. Tuloksista havaittiin kuitenkin, etteivät χ^2 -testin edellytykset täyttyneet kummankaan selitettävän muuttujan, eikä selittävän muuttujan osalta, joten käytettiin Fisherin tarkkaa testiä. Äidin osalta nousukausien p-arvo oli 0,196 ja laskukausien 0,366. Isän osalta nousukausien p-arvo oli puolestaan 0,844 ja laskukausien 0,116. Näin ollen tulokset eivät olleet tilastollisesti merkitseviä, eikä H_8 saanut tukea. Opiskelijoiden perusaineiston osalta ei siis voida sanoa, että vanhempien sosioekonomisella asemalla olisi vaikutusta opiskelijoiden trendinseuraamiskäyttäytymiseen.

Hypoteesia H_8 tutkittiin vielä ennustevirheen kautta opiskelijoiden perusaineistolla. Ensimmäiseksi tutkittiin riippumattomien ryhmien t-testin avulla vanhempien sijoittamisharrastuksen vaikutusta opiskelijoiden osakeennusteiden tarkkuuteen. Tulosten mukaan $t(96,223) = 2,053$ ja p-arvo 0,043, joten testin mukaan sijoittamista harrastavien ja ei-harrastavien osakeennusteiden virheissä oli tilastollisesti melkein merkitsevää eroa: niiden opiskelijoiden, joiden vanhemmat eivät harrastaneet sijoittamista, ennustevirhe oli keskimäärin pienempi kuin niiden, joiden vanhemmat harrastivat sijoittamista. Samansuuntainen tulos saatiin myös MWU-testillä (U-testisuureen arvo 1844,0 ja p-arvo 0,045). Saadut tulokset ovat melko yllättäviä. Aikaisempi tutkimus on painottanut rahoitusalan sofistikoituneisuuden välittymistä vanhemmilta lapsille. Jos tulkitaan ennustetarkkuuden olevan osa sofistikoituneisuutta, niin tämän työn aineiston perusteella aikaisemmat tutkimukset eivät saaneet vahvistusta.

Riippumattomien ryhmien t-testillä testattiin seuraavaksi, oliko opiskelijoiden ennustevirheiden keskiarvoissa eroa, jos otetaan huomioon parhaimpien ystävien sijoitusharrastus. T-testin tulostamien keskiarvovertailuiden perusteella nähtiin, että niiden opiskelijoiden, joiden parhaimmat ystävät harrastivat sijoittamista, ennustevirheiden keskiarvo oli pienempi kuin niiden, joiden ystävät eivät harrastaneet sijoittamista. Tulokset eivät kuitenkaan olleet tilastollisesti merkitseviä: $t(137) = -1,648$ ja p-arvo 0,102. Parametrittoman MWU-testin tulokset antoivat erilaisen tuloksen: U-testisuureeksi saatiin 1911,0 ja p-arvoksi 0,044, mikä tarkoittaa, että ennustevirheiden sijalukujen keskiarvoissa oli eroa. Niiden opiskelijoiden, joiden parhaimmat ystävät harrastivat sijoittamista, en-

nustevirheet olivat keskimäärin pienempiä kuin niiden, joiden parhaimmat ystävät eivät harrastaneet sijoittamista. Tämä antoi tukea teorialle: jos sofistikoituneisuus merkitsee tarkempia ennusteita, niin ystävien sijoitusharrastuksella saattoi olla vaikutusta tähän tarkkuuteen.

Yllä saatujen tulosten perusteella nousee esiin kysymys, oliko ennustevirhe kaikkein pienin niillä opiskelijoilla, joiden vanhemmat eivät harrastaneet sijoittamista ja parhaimmat ystävät harrastivat. Tätä tutkittiin kaksisuuntaisella varianssianalyysillä. Kaksisuuntaisella varianssianalyysillä voidaan tutkia kahden ryhmittelevän muuttujan vaikutusta jatkuvan muuttujan vaihteluun. Tässä ryhmittelevinä muuttujina olivat vanhempien ja parhaimpien ystävien sijoittamisharrastus ja jatkuvana muuttujana ennustevirhe. Keskiarvoja verrattaessa ennustevirhe oli todellakin pienin silloin, kun vanhemmat eivät harrastaneet sijoittamista ja parhaimmat ystävät harrastivat, mutta yhdysvaikutus ei ollut tilastollisesti merkitsevä, kuten taulukon 28 tulokoosteesta nähdään.

TAULUKKO 28 Kaksisuuntaisen varianssianalyysin tulokset

Lähde ¹⁷	F-testisuure	Vapausasteet	P-arvo
K12	2,542	1	,113
K11	4,688(*)	1	,032
K12 * K11	1,286	1	,259

Merkitsevyytasot $p < ,05$ (*), $p < ,01$ (**) ja $p < ,001$ (***).

Seuraavaksi tutkittiin yksisuuntaisella varianssianalyysillä, oliko opiskelijoiden ennustevirheissä eroja suhteessa siihen, mikä äidin tai isän sosioekonominen asema oli. Analyysissä käytettiin edellä luotuja muuttujia SOSEKAS_äiti ja SOSEKAS_isä. Tilastollisesti merkitseviä tuloksia ei saatu, eli vanhempien sosioekonomisella asemalla ei ollut vaikutusta ennustevirheiden keskiarvojen eroihin.

Hypoteeseihin H₇ ja H₈ liittyen päätettiin vielä tutkia logistisen regression avulla, mitkä tekijät selittävät sitä, harrastivatko opiskelijat sijoittamista vai eivät. Tätä tutkittiin opiskelijoiden perusaineistolla. Korrelaatiotarkasteluissa havaittiin, että sukupuoli ja parhaimpien ystävien sijoittamisharrastusta kuvaava muuttuja korreloivat merkitsevästi selitettävän muuttujan kanssa (Spearman). Lisäksi huomattiin, että opiskeltujen rahoituksen kurssien määrä sekä äidin ja isän luokiteltu sosioekonomisen aseman muuttuja SOSEKAS korreloivat selitettävän muuttujan kanssa melkein merkitsevästi. Selittävien muuttujien korrelaatiotarkasteluista nähtiin, että sukupuoli ja parhaimpien ystävien sijoittamisharrastusta kuvaava muuttuja korreloivat merkitsevästi keskenään. Opiskeltujen rahoituksen kurssien määrä ja parhaimpien ystävien sijoittamisharrastusta kuvaava muuttuja korreloivat melkein merkitsevästi keskenään. Myös SOSEKAS_äiti ja SOSEKAS_isä korreloivat keskenään merkitsevästi. Koska selittäjämuuttujien välinen korrelaatio on ongelma myös logistisessa regressiossa, niin

¹⁷ K11 ja K12 viittaavat tässä opiskelijoille suunnatun kyselylomakkeen vastaaviin kysymyksiin: K11 = Harrastavatko vanhempasi sijoittamista, K12 = Harrastavatko jotkut parhaimmista ystävistäsi sijoittamista.

päätettiin muodostaa kaksi mallia. Mallissa 1. selittäjinä eli kovariaatteina olivat sukupuoli, opiskeltujen rahoituksen kurssien määrä ja SOSEKAS_äiti¹⁸. Mallissa 2. selittäjinä olivat parhaimpien ystävien sijoittamisharrastusta kuvaava muuttuja ja SOSEKAS_äiti.

Ensimmäisen mallin muuttujan SOSEKAS_äiti kontrastiksi valittiin Indicator ja muuttujan ensimmäinen arvo (yrittäjät, eläkeläiset ja opiskelijat) asetettiin referenssikategoriaksi. Havaittiin kuitenkin, ettei muuttuja SOSEKAS_äiti parantanut mallia. P-arvot olivat suuria kaikissa kyseisen muuttujan luokissa, mikä tarkoittaa, että muuttujalla ei ollut tilastollisesti merkitsevää vaikutusta ennusteeseen, joten muuttuja poistettiin mallista. Uusi regressio tehtiin sukupuolen ja opiskeltujen rahoituksen kurssien määrää kuvaavan muuttujan avulla. Taulukko 29 esittää tulokset.

TAULUKKO 29 Logistisen regressioanalyysin tulokset, opiskelijoiden perusaineisto, malli 1.

	B	Keski- virhe	Waldin testi- suure	Vapaus- asteet	P-arvo	Exp(B)	95 %:n luottamusväli Exp(B)	
							Ala	Ylä
K2	1,266	,467	7,331 (**)	1	,007	3,545	1,418	8,863
K6 ¹⁹	,054	,025	4,702 (*)	1	,030	1,055	1,005	1,108
Vakio	-3,463	,846	16,747	1	,000	,031		

$\chi^2(2) = 15,254$ (***)

Pseudo R² = ,145

N = 146

Selitettävä muuttuja Harrastaako itse sijoittamista koodattiin seuraavasti: kyllä = 1, ei = 0.

Merkitsevyystasot p < ,05 (*), p < ,01 (**), ja p < ,001 (***)

Pseudo R² tarkoittaa Nagelkerken selitystasetta.

Taulukosta 29 nähdään, että malli oli tilastollisesti erittäin merkitsevä. Nagelkerken selitystasote oli kohtalainen. Waldin testin tuloksien perusteella voidaan sanoa, että molemmat selittäjämuuttujista lisäsivät mallin ennustekykä. Muuttujan K2 ristitulo-suhteen, exp(B), perusteella voidaan tulkita, että toisella sukupuolella on 3,5-kertainen riski ryhtyä harrastamaan sijoittamista. Muuttujan K6 ristitulo-suhteen perusteella voidaan puolestaan sanoa, että yksi opintopiste lisää rahoituksen kurssia kasvattaa 5,5 % riskiä ryhtyä harrastamaan sijoittamista. Malli pystyi kokonaisuudessaan ennustamaan 74,7 % siitä, sijoittaako henkilö vai ei.

Mallin perusteella voidaan kirjoittaa ennusteyhtälö:

$$P(Y) = \frac{1}{1 + e^{-(1,266 \cdot K2 + 0,054 \cdot K6 - 3,463)}}$$

Tarkasteltaessa mallin residuaaleja kiinnitettiin huomiota Cookin etäisyysluokun (Cook's distance), vipuvaikutukseen (leverage), DFBeta-lukuihin ja stan-

¹⁸ Päätös keskittyä pelkästään äidin sosioekonomisen aseman vaikutukseen perustui aikaisempaan tutkimukseen, esim. Lusardi, Mitchell & Curto 2010.

¹⁹ K2 tarkoittaa sukupuolta ja K6 opiskeltujen rahoituksen kurssien määrää.

dardoituihin residuaaleihin. Cookin etäisyysluku oli kaikissa tapauksissa alle yhden, vipuvaikutusluvusta 18 oli suurempia kuin 2 x keskiarvo ja 12 suurempia kuin 3 x keskiarvo. DFBeta-arvot olivat alle yhden, mikä on hyvä asia. Mallin standardoituja residuaaleja tutkittaessa havaittiin, että viidessä tapauksessa arvot ylittivät raja-arvona pidetyn 3:n ja yli 1 % arvoista oli yli raja-arvon 2,5, mikä saattaa kertoa poikkeavista havainnoista. Mallin parantamiseksi poikkeavat havainnot tulisi poistaa aineistosta.

Toisessa mallissa muuttujan SOSEKAS_äiti kontrastiksi valittiin jälleen Indicator ja muuttujan ensimmäinen arvo (yrittäjät, eläkeläiset ja opiskelijat) asetettiin referenssikategoriaksi. Havaittiin kuitenkin, ettei muuttuja SOSEKAS_äiti parantanut mallia, eli sillä ei ollut tilastollisesti merkitsevää vaikutusta ennusteeseen, joten kyseinen muuttuja poistettiin mallista. Taulukko 30 esittää tulokset.

TAULUKKO 30 Logistisen regressioanalyysin tulokset, opiskelijoiden perusaineisto, malli 2.

	B	Keski- virhe	Waldin testisuure	Vapaus- asteet	P-arvo	Exp(B)	95 %:n luottamusväli Exp(B)	
							Ala	Ylä
K12 ²⁰	-1,470	,442	11,052 (**)	1	,001	,230	,097	,547
Vakio	-,511	,231	4,893	1	,027	,600		

$\chi^2(1) = 12,814 (***)$
Pseudo R² = ,123
N = 146

Selitettävä muuttuja Harrastaako itse sijoittamista koodattiin seuraavasti: kyllä = 1, ei = 0. Merkitsevyytasot p < ,05 (*), p < ,01 (**) ja p < ,001 (***). Pseudo R² tarkoittaa Nagelkerken selitystasetta.

Taulukosta 30 nähdään, että malli oli tilastollisesti erittäin merkitsevä. Nagelkerken selitystasote oli kohtalainen. Waldin testin tuloksien perusteella voidaan sanoa, että selittäjämuuttuja lisäsi mallin ennustekykä. Muuttujan K12 ristitulo-suhteen, exp(B), perusteella voidaan sanoa, että jos parhaimmat ystävät harrastavat sijoittamista, niin se kasvattaa 23 % riskiä ryhtyä harrastamaan sijoittamista. Malli pystyi kokonaisuudessaan ennustamaan 74 % siitä, ryhtyykö henkilö sijoittamaan vai ei.

Mallin perusteella voidaan kirjoittaa ennusteyhtälö:

$$P(Y) = \frac{1}{1 + e^{-(1,470 \cdot K12 - 0,511)}}$$

Tarkasteltaessa mallin residuaaleja kiinnitettiin jälleen huomiota Cookin etäisyyslukuun (Cook's distance), vipuvaikutukseen (leverage), vakion DFBeta-lukuihin ja standardoituihin residuaaleihin. Cookin etäisyysluku oli kaikissa tapauksissa alle yhden, vipuvaikutusluku alle 2 x keskiarvon ja DFBeta-arvot

²⁰ K12 tarkoittaa kyselylomakkeen kysymystä Harrastavatko jotkut parhaimmista ystävästäsi sijoittamista.

alle yhden, mikä on hyvä asia. Mallin standardoituja residuaaleja tutkittaessa saatiin selville, että yhtään arvoa ei ollut yli kriittisenä rajana pidetyn 3:n, mutta yli 1 % oli yli rajana pidetyn 2,5, mikä saattaa kertoa poikkeavista havainnoista. Kaiken kaikkiaan tässä mallissa poikkeavia havaintoja oli selvästi vähiten tässä tutkimuksessa tehtyjen regressioanalyysien joukossa.

5.1.9 Ikä ja virheet

Hypoteesi 10 liittyy iän ja sijoittamiseen liittyvien virheiden väliseen yhteyteen:

H_{10} : Keski-ikäiset tekevät kaikista vähiten virheitä osake-ennusteissa.

Analyyseissä käytettiin yhdistettyä perusaineistoa. Ensimmäiseksi haluttiin tutkia, oliko trendinseuraamiskäyttäytymisessä eroja sen suhteen, mikä oli vastaajan ikä. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Järkevien ikäluokkien luominen osoittautui hankalaksi, koska yhdistetyssä vastaajajoukossa keski-ikäisiä oli niin vähän. Suurin osa vastaajista oli kolmikymppisiä tai sen alle (80,5 % kaikista vastaajista). Analyysiä varten olisi kuitenkin hyvä saada jokaiseen selittäjänä toimivaan ikäryhmään riittävästi vastaajia. Ikään liittyen päätettiin luoda apumuuttuja, joka luokitteli kaikki vastaajat kahteen ikäluokkaan: alle 30-vuotiaat ja yli 30-vuotiaat. Vaikka tällainen jaottelu ei suoraan erottelekaan keski-ikäisiä omaksi ikäryhmäkseen, oli tämän tyylinen jaottelu kuitenkin analyysia ajatellen järkevää tehdä. Selitettäviksi muuttujiksi valittiin jälleen vuorollaan TF_bull ja TF_bear. Tuloksista havaittiin, että muuttujan TF_bear osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. P-arvoksi saatiin 0,436. Muuttujan TF_bull osalta tulokset olivat: $\chi^2(3) = 3,297$, p-arvo 0,348 ja kontingenssikerroin 0,135. Saadut tulokset eivät olleet tilastollisesti merkitseviä. Näin ollen yhdistetyn perusaineiston perusteella ei voida sanoa, että eri-ikäiset olisivat käyttäytyneet trendinseuraamisen suhteen eri tavoilla, eli H_{10} ei saanut tukea.

Seuraavaksi tutkittiin hypoteesia H_{10} ennustevirheen kautta. Ensin vertailtiin yksisuuntaisen varianssianalyysin avulla ennustevirheen keskiarvoja eri ikäryhmien välillä. Ikäryhmät muodostettiin samalla tavoin kuin edellä. Levenen varianssitestin p-arvoksi saatiin 0,420, joten tämä tuki oletusta, että varianssit olisivat eri ikäryhmissä yhtä suuret. Yksisuuntaisen varianssianalyysin tulosten mukaan eri ikäryhmien ennustevirheiden ryhmäkeskiarvoissa ei ollut eroa: $F(1) = 0,098$ ja p-arvo 0,755. Näin ollen H_{10} ei saanut yhdistetyn perusaineiston perusteella tukea. Yhdistetyn perusaineiston perusteella ei siis voida sanoa, että eri ikäryhmien välillä olisi eroja ennustetarkkuudessa.

Sama testaus suoritettiin myös parametrittömällä Kruskal-Wallis testillä, sillä muuttujat eivät olleet normaalisti jakautuneita. Tulokset olivat samanlaiset, eli ikäryhmien välillä ei ollut eroa ennustetarkkuudessa: $\chi^2(1) = 0,372$ ja p-arvo 0,542. H_{10} ei siis saanut tukea parametrittömän testin perusteella.

Lopuksi tutkittiin vielä korrelaatiomatriisin avulla, oliko iän ja ennustevirheen välillä lineaarista riippuvuutta. Spearmanin korrelaatiokertoimen ar-

voksi saatiin -0,096 ja p-arvoksi 0,207, joten ikä ei korreloinut ennustevirhe-muuttujan kanssa.

5.1.10 Opiskelijatko häirikkösijoittajia?

Hypoteesi 11 liittyy häirikkösijoittamiseen eli trendinseuraamiseen:

H_{11} : Opiskelijat käyttäytyvät häirikkösijoittajan tavoin.

Olivatko opiskelijat siis enemmän trendinseuraajia suhteessa pankin sijoitusasi-antuntijoihin? Hypoteesia tutkittiin muuttujien TF_bull_COstrong ja TF_bear_COstrong avulla. Haluttiin selvittää χ^2 -yhteensopivuustestillä, oliko vahvassa trendinseuraamis- ja toisinajattelijakäyttäytymisessä eroa opiskelija-vastaaajien ja pankin sijoitusasiantuntijoiden välillä. Tässä toisen luokitteluas-teikollisen muuttujan muodosti ensimmäisen hypoteesin yhteydessä luotu muuttuja R, joka siis sai arvon 1, kun kyseessä oli opiskelija ja arvon 2, kun ky-seessä oli pankin sijoitusasiantuntija. Analyysissä selitettäväksi muuttujaksi valittiin vuorollaan TF_bull_COstrong sekä TF_bear_COstrong ja selittäväksi muuttujaksi ryhmä (R).

Huomattiin kuitenkin, että sekä nousu- että laskukausien analyysin osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Nousu-kausien osalta p-arvoksi saatiin 1,000 ja laskukausien osalta 0,253. Tulokset ei-vät siten olleet tilastollisesti merkitseviä, joten H_{11} ei saanut tämän perusteella tukea.

Häirikkösijoittamista tutkittiin vielä χ^2 -testillä sukupuolittain. Haluttiin tutkia, erosiko naisten ja miesten vahva trendinseuraamis- ja toisinajattelijakäyt-täytyminen opiskelijoiden ja pankin aineistoissa. Ensiksi kysymystä testattiin opiskelijoiden perusaineistolla. Analyyseissä selitettäväksi muuttujaksi asetetiin vuorollaan TF_bull_COstrong ja TF_bear_COstrong ja selittäväksi muuttu-jaksi valittiin sukupuoli. Tuloksista nähtiin tässäkin tapauksessa, etteivät χ^2 -testin edellytykset täyttyneet kummankaan analyysin kohdalla, joten tehtiin Fisherin tarkka testi. Nousukausien kohdalla saatiin tilastollisesti merkitsevä tulos: p-arvo oli 0,005. Laskukausien kohdalla tulokset taas eivät olleet tilastolli-sesti merkitseviä: p-arvo oli 1,000. Nousukausien tulosten mukaan vahvoja trendinseuraajia oli opiskelijamiehistä 71,2 % ja -naisista 28,8 %. Tämän perus-teella miehet siis olivat opiskelijoiden perusaineistossa enemmän häirikkösijoit-tajia.

Myöskään pankin perusaineiston tapauksessa χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Sekä nousu- että laskukausien osalta p-arvoksi saatiin 1,000, joten tilastollisesti merkitseviä eroja sukupuolten välillä ei ollut.

Tässä yhteydessä haluttiin selvittää, mitkä tekijät selittivät sitä, oliko vas-taaja vahva trendinseuraaja vai vahva toisinajattelija. Tätä tutkittiin yhdistetyllä perusaineistolla logistisen regression avulla. Luvun 5.1.4 perusteella ainakin luottamusvälien vinoudet ja odotetut hintamuutokset voisivat muodostaa selit-täjämuuttujat. Korrelaatiotarkasteluissa (Spearman) saatiin selville, että muut-

tuja TF_bull_COstrong korreloi merkitsevästi muuttujien EPC_bull ja sukupuoli kanssa, mutta ei korreloinut muuttujien ikä, R (ryhmä) tai Delta_bull kanssa. Sukupuoli ja EPC_bull taas korreloivat merkitsevästi keskenään. Lisäksi havaittiin, että muuttajat Delta_bear ja EPC_bear korreloivat merkitsevästi selitettävän muuttujan TF_bear_COstrong kanssa, mutta niiden keskinäinen korrelaatio oli merkitsevää.

Korrelaatiotarkasteluiden perusteella päätettiin koemielessä muodostaa kaksi mallia ennustamaan nousu- ja laskukausien käyttäytymistä. Selitettäväksi muuttujaksi nousukausien analyysissä valittiin siis TF_bull_COstrong ja selittäjäksi EPC_bull (malli 1.) ja sukupuoli (malli 2.). Laskukausien analyysissä selitettävänä muuttujana oli siis TF_bear_COstrong ja selittäjinä EPC_bear (malli 3.) sekä Delta_bear (malli 4.). Monimuuttajamalleja ei tehty selittävien muuttujien merkitsevän keskinäisen korrelaation vuoksi.

Mallit 1., 2. ja 4. olivat tilastollisesti erittäin merkitseviä tai merkitseviä. Tässä ei kuitenkaan enempää raportoida tuloksia, koska kyse on samasta ilmiöstä, jota tutkittiin jo lineaarisen regression avulla luvussa 5.1.4. Mallista 3. ei saatu järkeviä tuloksia Quasi-Complete separation -ongelman vuoksi, mikä tarkoittaa, ettei suurimman uskottavuuden estimaattia ollut olemassa.

Edellä yritettiin tutkia, oliko vastaajaryhmillä eroa trendinseuraamiskäyttäytymisen ja toisinajattelijastrategian noudattamisessa, mutta Fisherin tarkan testin tulokset eivät olleet tilastollisesti merkitseviä. Jos kuitenkin pidetään pankin vastaajia rationaalisina tai informoituina sijoittajina, jotka tuntevat tehokkaiden markkinoiden hypoteesin, niin mitkä tekijät selittivät tätä tietämystä? Kysymystä tutkittiin logistisen regression avulla. Selittäjämuuttujiksi eli kovariaateiksi päätettiin valita työkokemus rahoitusosalta, pohjakoulutus, ikä ja sukupuoli. Kahden ensiksi mainitun ajateltiin erityisesti vaikuttavan sijoitusasian tuntijoiden rationaalisuuteen ja sofistikoituneisuuteen. Havaittiin kuitenkin, että kyseiset muuttajat eivät korreloineet selitettävän muuttujan kanssa ja lisäksi niiden keskinäinen korrelaatio oli tilastollisesti merkitsevää Spearmanin järjestyskorrelaatiokertoimen avulla mitattuna. Koska selittäjämuuttujien korrelaatio on ongelma myös logistisen regression tapauksessa, päätettiin tehdä kaksi regressiomallia. Mallissa 1. selittäjinä olivat pohjakoulutus, ikä ja sukupuoli. Mallissa 2. selittäjinä olivat taas työkokemus rahoitusosalta, ikä ja sukupuoli. Mallin 1. kategorisen selittäjämuuttujan (pohjakoulutus) kontrastiksi valittiin Indicator ja muuttujan viimeinen arvo (yliopistotutkinto) asetettiin referenssikategoriaksi.

Molempien mallien osalta saatiin kuitenkin selville, etteivät malleissa mukana olleet selittäjämuuttajat parantaneet malleja. P-arvot olivat suuria, eli niillä ei ollut tilastollisesti merkitsevää vaikutusta ennusteeseen, joten analyysiä ei viety enää pidemmälle.

5.2 Herkkyysanalyysit: normalisoidut perusaineistot

Analysoitaessa eri aikasarjahavainnoista yhdistettyä aineistoa tulisi normalisoinnilla eliminoida hintaeroista johtuvat vaihtelut. Tässä opinnäytetyössä päädyttiin kuitenkin normalisoimaan vain odotetun hintamuutoksen ja luottamusvälien vinouden parametrit, kuten De Bondtin (1993) tutkimuksessa. Tässä alaluvussa tutkitaan hypoteeseja H_{20} , H_{30} ja H_4 siten, että odotetun hintamuutoksen ja luottamusvälien vinouden parametrit on normalisoitu jakamalla ne toteutuneen S&P 500 indeksin kuukausituoton keskihajonnalla (σ) neljän kuukauden ajalta ennen kunkin osakekuvaajan viimeistä näkyvää arvoa. Opiskelijoiden osake-ennusteista puuttuvia tietoja ei tässä vaiheessa poistettu. Ensinnä tutkitaan yllä mainittuja hypoteeseja opiskelijoiden normalisoidulla perusaineistolla (aineisto 2.), sitten pankin normalisoidulla perusaineistolla (aineisto 6.) ja lopuksi näiden yhdistelmällä.

5.2.1 Erot odotetussa hintamuutoksessa

Hypoteesia H_{20} testattiin kaikilla yllä mainituilla aineistoilla muuttujien $EPC7_{bull_norm}$, $EPC7_{bear_norm}$, $EPC13_{bull_norm}$ ja $EPC13_{bear_norm}$ avulla. Kyseiset muuttujat siis kuvaavat nousu- ja laskukausien odotettujen hintamuutosten normalisoituja keskiarvoja.

Opiskelijoiden normalisoidussa perusaineistossa huomattiin, että muuttujaparit eivät korreloineet keskenään Spearmanin järjestyskorrelaatiokertoimien perusteella, joten verrannollisten parien t-testiä voitiin käyttää. Keskiarvovertailuista havaittiin, että kummankin parin kohdalla odotetun hintamuutoksen keskiarvo oli nousukausien jälkeen suurempi kuin laskukausien jälkeen. T-testin tulokset ensimmäiselle parille $EPC7_{bull_norm} - EPC7_{bear_norm}$ olivat $t(145) = 7,052$; $p < 0,001$ ja toiselle parille $EPC13_{bull_norm} - EPC13_{bear_norm}$ $t(145) = 5,719$; $p < 0,001$. Näin ollen H_{20} hylättiin molempien parien osalta, ja tulokset olivat tilastollisesti erittäin merkitseviä, mikä ei poikennut perusaineistolla saaduista tuloksista.

Koska normaalijakauma- tai symmetrisyysoletus ei täytynyt muuttujien $EPC13_{bull_norm}$ ja $EPC13_{bear_norm}$ kohdalla, päätettiin jälkimmäistä paria vielä testata parametrittömällä WSR-testillä. Sen tulokset olivat vastaavat kuin yllä: jälkimmäisen parin osalta H_{20} hylättiin merkitsevyystasolla $p < 0,001$ (Z-testisuure -5,278). Jälkimmäisen parin osalta sekä parametrisella että parametrittömällä testillä saatiin siis samankaltainen tulos kuin perusaineistollakin.

Pankin normalisoidussa perusaineistossa huomattiin muuttujaparin $EPC7_{bull_norm}$ ja $EPC7_{bear_norm}$ välillä olevan korrelaatiota (Pearsonin korrelaatiokerroin 0,345 ja p-arvo 0,037). Lisäksi muuttuja $EPC13_{bull_norm}$ ei ollut normaalisti jakautunut, joten hypoteesia tutkittiin suoraan parametrittömällä WSR-testillä. Tulosten mukaan ensimmäisen parin osalta H_{20} jäi voimaan p-arvolla 0,079 (Z-testisuure -1,758) ja toisen parin osalta se hylättiin p-arvolla 0,015 (Z-testisuure -2,421).

Huomattiin siis, että pankin aineistossa EPC parametrien normalisointi vaikutti siten, että ensimmäisen parin $EPC7_{\text{bull_norm}} - EPC7_{\text{bear_norm}}$ kohdalla ei voida sanoa, että odotettujen hintamuutosten keskiarvoissa olisi tilastollisesti merkitsevää eroa nousu- ja laskukausien välillä. Tämä tulos poikkesi perusaineistolla saadusta tuloksesta. Jälkimmäisen parin $EPC13_{\text{bull_norm}} - EPC13_{\text{bear_norm}}$ osalta H_{20} puolestaan hylättiin, joten kyseisen parin osalta tulokset taas vastasivat perusaineistolla saatuja tuloksia.

Lopuksi hypoteesia H_{20} testattiin vielä normalisoidulla ja yhdistetyllä perusaineistolla. Muuttujaparien $EPC7_{\text{bull_norm}} - EPC7_{\text{bear_norm}}$ ja $EPC13_{\text{bull_norm}} - EPC13_{\text{bear_norm}}$ välillä ei havaittu korrelaatioita, joten hypoteesia testattiin ensin verrannollisten parien t-testillä. Keskiarvovertailuista nähtiin jälleen, että odotetun hintamuutoksen keskiarvot nousukausille olivat suurempia kuin laskukausille. T-testin tulokset ensimmäiselle parille olivat $t(182) = 7,071$; $p < 0,001$ ja jälkimmäiselle parille $t(182) = 6,263$; $p < 0,001$, joten aivan samanlainen tulos saatiin kuin taulukossa 17 on esitetty. Molempien parien osalta H_{20} hylättiin, ja tulokset olivat tilastollisesti erittäin merkitseviä.

Koska muuttujat $EPC7_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$ eivät olleet normaalisti jakautuneet, testattiin hypoteesia vielä parametrittömällä WSR-testillä. Tulosten mukaan H_{20} hylättiin molempien parien osalta merkitsevyystasolla $p < 0,001$ (Z-testisuureen arvo ensimmäiselle parille $-6,512$ ja jälkimmäiselle parille $-5,803$). Voidaan siis todeta, että normalisointi ei muuttanut tuloksia suhteessa yhdistettyyn perusaineistoon.

5.2.2 Luottamusvälien vinous

Seuraavaksi tutkittiin, vaikuttiko normalisointi tuloksiin hypoteesin H_{30} kohdalla. Aluksi keskityttiin jälleen opiskelijoihin, sitten pankin vastaajiin ja lopuksi yhdistettyyn aineistoon. Kaikissa aineistoissa hypoteesia tutkittiin muuttujien $\Delta 7_{\text{bull_norm}}$, $\Delta 7_{\text{bear_norm}}$, $\Delta 13_{\text{bull_norm}}$ ja $\Delta 13_{\text{bear_norm}}$ avulla. Muuttujat kuvaavat siis nousu- ja laskukausien luottamusvälien vinouksien normalisoituja keskiarvoja.

Opiskelijoiden normalisoidussa perusaineistossa muuttujien välillä ei havaittu korrelaatiota (Spearman), joten hypoteesia H_{30} testattiin ensin verrannollisten parien t-testillä. Keskiarvovertailuiden mukaan luottamusvälien vinoudet nousukausien jälkeen olivat negatiivisia ja laskukausien jälkeen positiivisia molempien parien kohdalla. T-testin tulokset parille $\Delta 7_{\text{bull_norm}} - \Delta 7_{\text{bear_norm}}$ olivat $t(142) = -7,538$; $p < 0,001$ ja parille $\Delta 13_{\text{bull_norm}} - \Delta 13_{\text{bear_norm}}$ $t(142) = -5,912$; $p < 0,001$. Näin ollen H_{30} hylättiin molempien parien osalta, ja tulokset olivat tilastollisesti erittäin merkitseviä: keskimääräinen luottamusvälien vinous nousukausien jälkeen oli pienempi kuin laskukausien jälkeen. Normalisointi ei siis muuttanut tuloksia suhteessa perusaineistoon.

Koska muuttujat $\Delta 7_{\text{bear_norm}}$ ja $\Delta 13_{\text{bear_norm}}$ eivät olleet normaalisti jakautuneita, tutkittiin hypoteesia vielä parametrittömällä WSR-testillä. Sen tulosten mukaan H_{30} hylättiin molempien parien osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure ensimmäiselle parille $-6,776$ ja jälkimmäiselle parille $-5,681$). Tuloksissa ei siis ollut muutosta perusaineistolla saatuihin tuloksiin nähden.

Seuraavaksi hypoteesia tutkittiin pankin normalisoidulla perusaineistolla. Verrannollisten parien t-testiä ei voitu käyttää hypoteesin testaukseen, sillä muuttujaparit korreloivat merkitsevästi keskenään Spearmanin järjestyskorrelaatiokerroinlaatiokertoimella mitattuna. Muuttujien $\Delta_{7_{\text{bull_norm}}}$ ja $\Delta_{7_{\text{bear_norm}}}$ välinen korrelaatiokerroin oli 0,557 (p-arvo < 0,001) ja muuttujien $\Delta_{13_{\text{bull_norm}}}$ ja $\Delta_{13_{\text{bear_norm}}}$ välinen 0,344 (p-arvo < 0,001). Havaittiin kuitenkin, että keskimääräinen luottamusvälien vinous nousu- ja laskukausien jälkeen oli molempien parien kohdalla negatiivista, mutta nousukausien jälkeen keskimääräinen vinous oli pienempi. Parametrittoman WSR-testin mukaan H_{30} hylättiin ensimmäisen parin osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure -3,616) ja jälkimmäisen parin osalta merkitsevyystasolla $p = 0,033$ (Z-testisuure -2,136). Tulokset olivat siis hyvin samankaltaisia kuin perusaineistollakin saadut tulokset, eli luottamusvälien keskimääräinen normalisoitu vinous pankin aineistossa oli pienempi nousukausien jälkeen.

Lopuksi hypoteesia tutkittiin normalisoidulla ja yhdistetyllä perusaineistolla. Muuttujaparit korreloivat keskenään Spearmanin järjestyskorrelaatiokerroimen perusteella. Muuttujien $\Delta_{7_{\text{bull_norm}}}$ ja $\Delta_{7_{\text{bear_norm}}}$ välinen korrelaatiokerroin oli 0,188 ja p-arvo 0,011, kun taas muuttujien $\Delta_{13_{\text{bull_norm}}}$ ja $\Delta_{13_{\text{bear_norm}}}$ välinen korrelaatiokerroin oli 0,163 ja p-arvo 0,028. Molempien parien kohdalla luottamusvälien keskimääräinen vinous oli negatiivinen sekä nousu- että laskukausien jälkeen, mutta nousukausien jälkeen vinous oli pienempää. Hypoteesia testattiin parametrittomalla WSR-testillä ja sen tulosten mukaan H_{30} hylättiin molempien parien osalta merkitsevyystasolla $p < 0,001$ (Z-testisuure ensimmäiselle parille -7,649 ja jälkimmäiselle parille -6,037). Näin ollen myös normalisoidulla ja yhdistetyllä aineistolla saadut tulokset vahvistavat, että luottamusvälien keskimääräiset vinoudet nousukausien jälkeen olivat pienemmät kuin laskukausien jälkeen.

5.2.3 Regressioanalyysit luottamusvälien vinoudelle

Seuraavaksi tehtiin regressioanalyysit, joissa pyrittiin selvittämään H_4 :n mukaisesti, mitkä tekijät vaikuttivat muuttujiin $\Delta_{\text{bull_norm}}$ ja $\Delta_{\text{bear_norm}}$. Aivan kuten luvussa 5.1.4, selittäjämuuttujat pyrittiin ensin löytämään korrelaatiotarkastelujen perusteella (Spearman). Regressiot tehtiin totuttuun tapaan ensin opiskelija-aineistolle, sitten pankin aineistolle ja lopuksi näiden yhdistelmälle.

Opiskelijoiden normalisoidun perusaineiston kohdalla nähtiin, että $\Delta_{\text{bull_norm}}$ korreloi melkein merkitsevästi muuttujan $\text{EPC}_{\text{bull_norm}}$ kanssa ja merkitsevästi muuttujan $\text{TF}_{\text{bull_CO}}$ kanssa. $\Delta_{\text{bear_norm}}$ taas korreloi merkitsevästi sekä muuttujan $\text{EPC}_{\text{bear_norm}}$ että muuttujan $\text{TF}_{\text{bear_COstrong}}$ kanssa. Lisäksi havaittiin, että muuttujat $\text{EPC}_{\text{bull_norm}}$ ja $\text{TF}_{\text{bull_CO}}$ korreloivat merkitsevästi keskenään, ja että $\text{EPC}_{\text{bear_norm}}$ korreloi merkitsevästi $\text{TF}_{\text{bear_COstrong}}$:in kanssa. Tämän vuoksi päätettiin tehdä neljä regressiota samalla periaatteella kuin luvussa 5.1.4 tehtiin. Ensin regressoitiin muuttujaa $\Delta_{\text{bull_norm}}$ erikseen muuttujien $\text{EPC}_{\text{bull_norm}}$ ja $\text{TF}_{\text{bull_CO}}$ suhteen (mallit 1. ja 2.). Sen jälkeen regressoitiin muuttujaa $\Delta_{\text{bear_norm}}$ erikseen muuttujien $\text{EPC}_{\text{bear_norm}}$ ja $\text{TF}_{\text{bear_COstrong}}$ suhteen (mallit 3. ja 4.).

Mallin 1. osalta tulokset olivat samansuuntaiset kuin perusaineistolla saadut tulokset: mallin selitysaste oli heikko, eikä malli varianssianalyysin mukaan sopinut aineistoon. Muuttuja $EPC_{bull_{norm}}$ ei vaikuttanut tilastollisesti merkitsevästi muuttujaan $\Delta_{bull_{norm}}$. H_4 :n kannalta tarkasteltuna huomattiin, että muuttujan $EPC_{bull_{norm}}$ regressiokerroin oli hypoteesin mukainen, eli negatiivinen, mutta tulos ei siis ollut tilastollisesti merkitsevä, joten tässä ei tulkita tuloksia enempää. Mallin 2. osalta tulokset olivat niin ikään samansuuntaiset kuin perusaineistolla saadut. Varianssianalyysin mukaan malli oli lähellä sopia aineistoon (F-testisuure 3,866 ja p-arvo 0,051).

Mallin 3. selitysaste R^2 oli 15,4 %, ja varianssianalyysin tulokset $F(1) = 25,640$ ja p-arvo $< 0,001$, kun taas mallin 4. selitysaste R^2 oli 21,7 %, ja varianssianalyysin tulokset $F(1) = 16,358$ ja p-arvo $< 0,001$. Molemmat mallit siis sopivat aineistoon. T-testisuureen arvojen ja merkitsevyytasojen perusteella sekä muuttuja $EPC_{bear_{norm}}$ että muuttuja $TF_{bear_COstrong}$ vaikuttivat merkitsevästi muuttujaan $\Delta_{bear_{norm}}$ (t-testisuureen arvo ja merkitsevyytaso muuttujalle $EPC_{bear_{norm}}$ -5,064; $p < 0,001$ sekä muuttujalle $TF_{bear_COstrong}$ 4,044; $p < 0,001$). Mallien 3. ja 4. perusteella voidaan kirjoittaa regressioyhtälöt:

$$y = -0,312 - 0,631 * EPC_{bear_{norm}}$$

$$y = -0,774 + 0,928 * TF_{bear_COstrong}$$

Ensimmäiseen yhtälöön sisältyy H_4 :n kannalta keskeinen tulos: kun $EPC_{bear_{norm}}$ pienenee, kasvaa muuttujan $\Delta_{bear_{norm}}$ arvo $0,631 * EPC_{bear_{norm}}$:n muutos, eli vinous oikealle kasvaa. Jälkimmäisestä yhtälöstä nähdään, että kun kyse on vahvasta trendinseuraajasta (muuttujan $TF_{bear_COstrong}$ saadessa arvon 1), niin $\Delta_{bear_{norm}}$ saa arvon 0,15, mikä tarkoittaa, että luottamusvälit ovat vinoja oikealle. Kun kyseessä on vahva toisinajattelijä (muuttujan $TF_{bear_COstrong}$ saadessa arvon 0), saa $\Delta_{bear_{norm}}$ arvon -0,77 eli luottamusvälit ovat vinoja vasemmalle. Malli 3. antoi tukea hypoteesille H_4 : luottamusvälit kasvoivat enemmän oikealle vinoiksi, kun suurempaa hinnan laskua ennustettiin. Mallin 1. kohdallakin regressiokerroin muuttujalla $EPC_{bull_{norm}}$ oli hypoteesin mukainen, mutta ei ollut tilastollisesti merkitsevä.

Residuaalien jakaumia tutkittaessa saatiin selville, että varsinkin mallissa 3. histogrammi standardisoitujen jäännöstermien jakaumasta oli huipukkaampi kuin normaalijakauma. Molemmissa malleissa (mallit 3. ja 4.) jäännösten kumulatiivinen prosenttijakauma oli loivan toisen asteen käyrän muotoinen, muistuttaen painofunktioiden kuvaajien muotoa kuviossa 3 (s. 19), mutta mallissa 4. jakauma hiukan suurempi. Mallissa 3. hajontakuviosta standardoiduista residuaaleista ja ennusteista osoitti, että jäännökset eivät olleet kovin tasaisesti hajaantuneet ennusteen eri arvoilla. Tämä paljastaa, että varianssit eivät olleet täysin homogeenisia.

Kolmogorov-Smirnovin testin mukaan vain muuttuja $\Delta_{bull_{norm}}$ oli normaalisti jakautunut, mikä rikkoi regressioanalyysin oletuksia. Durbin-Watsonin testin arvo kolmannessa mallissa oli 1,789 ja neljännessä 1,878. Auto-

korrelaatiota ei siis esiintynyt. Muita malleihin 3. ja 4. vaikuttavia tekijöitä ei tässä enää tutkittu.

Seuraavaksi tehtiin regressioanalyysi pankin normalisoidulle perusaineistolle. Korrelaatiotarkasteluiden perusteella päätettiin ajaa yksi regressio, jossa muuttujaa $\Delta_{bear_{norm}}$ regressoitiin muuttujan $EPC_{bear_{norm}}$ suhteen. Saadut tulokset olivat hyvin lähellä perusaineistolla saatuja tuloksia. Mallin selitysaste oli melko huono, eikä malli sopinut varianssianalyysin perusteella aineistoon. Muuttuja $EPC_{bear_{norm}}$ ei näyttänyt vaikuttavan merkittävästi muuttujaan $\Delta_{bear_{norm}}$. Muuttujan $EPC_{bear_{norm}}$ regressiokerroin oli H_4 :n mukainen, eli negatiivinen, mutta tulos ei ollut tilastollisesti merkitsevää.

Normalisoidun ja yhdistetyn aineiston korrelaatiotarkasteluiden perusteella päätettiin ajaa neljä regressiota, kuten yhdistetyllä perusaineistollakin tehtiin. Ensin regressoitiin muuttujaa $\Delta_{bull_{norm}}$ erikseen muuttujien $EPC_{bull_{norm}}$ ja TF_{bull_CO} suhteen (mallit 1. ja 2.). Sen jälkeen regressoitiin muuttujaa $\Delta_{bear_{norm}}$ erikseen muuttujien $EPC_{bear_{norm}}$ ja $TF_{bear_COstrong}$ suhteen (mallit 3. ja 4.). Puuttuvat tiedot käsiteltiin komennolla `Exclude cases pairwise`.

Tulosten mukaan malleilla 1. ja 2. oli huonot selitysasteet, eivätkä mallit varianssianalyysin mukaan sopineet aineistoon. Muuttujat $EPC_{bull_{norm}}$ ja TF_{bull_CO} eivät vaikuttaneet tilastollisesti merkittävästi muuttujaan $\Delta_{bull_{norm}}$.

Varianssianalyysien mukaan mallit 3. ja 4. sopivat aineistoon. Mallissa 3. selitysaste R^2 oli 12,7 % ja varianssianalyysin tulokset $F(1) = 25,850$ ja p -arvo $< 0,001$. Mallissa 4. taas selitysaste R^2 oli 26,0 % ja varianssianalyysin tulokset $F(1) = 25,341$ ja p -arvo $< 0,001$. T-testisuureen arvojen ja merkitsevyytasojen perusteella sekä muuttuja $EPC_{bear_{norm}}$ että $TF_{bear_COstrong}$ vaikuttivat merkittävästi muuttujaan $\Delta_{bear_{norm}}$ (t-testisuureen arvo ja merkitsevyytaso muuttujalle $EPC_{bear_{norm}}$ -5,084; $p < 0,001$ sekä muuttujalle $TF_{bear_COstrong}$ 5,034; $p < 0,001$). Mallien 3. ja 4. perusteella voidaan kirjoittaa regressioyhtälöt:

$$y = -0,255 - 0,480 * EPC_{bear_{norm}}$$

$$y = -0,837 + 0,973 * TF_{bear_COstrong}$$

Ensimmäiseen yhtälöön sisältyy H_4 :n kannalta keskeinen tulos: kun $EPC_{bear_{norm}}$ pienenee, kasvaa muuttujan $\Delta_{bear_{norm}}$ arvo $0,480 * EPC_{bear_{norm}}$:n muutos eli vinous vasemmalle kasvaa. Jälkimmäisestä yhtälöstä nähdään, että kun kyse on vahvasta trendinseuraajasta ($TF_{bear_COstrong} = 1$), niin $\Delta_{bear_{norm}}$ saa arvon 0,14, mikä tarkoittaa, että luottamusvälit ovat vinoja oikealle. Kun kyseessä on vahva toisinajattelija ($TF_{bear_COstrong} = 0$), saa $\Delta_{bear_{norm}}$ arvon -0,84 eli luottamusvälit ovat vinoja vasemmalle. Malli 3. antoi tukea hypoteesille H_4 : luottamusvälit kasvoivat enemmän oikealle vinoiksi, kun suurempaa hinnan laskua ennustettiin. Myös mallissa 1. muuttujan $EPC_{bull_{norm}}$ regressiokerroin oli hypoteesin mukainen, mutta tulos ei ollut tilastollisesti merkitsevää.

Residuaalien jakaumia tutkittaessa huomattiin, että varsinkin mallissa 3. histogrammi standardisoitujen jäännöstermien jakaumasta oli huipukkaampi

kuin normaalijakauma. Molemmissa malleissa (mallit 3. ja 4.) jäännösten kumulatiivinen prosenttijakauma oli loivan toisen asteen käyrän muotoinen, muistuttaen painofunktioiden kuvaajien muotoa kuviossa 3 (s. 19), mutta mallissa 4. jakauma hiukan suurempi. Mallin 3. hajontakuviota standardoiduista residuaaleista ja ennusteista osoitti, että jäännökset olivat melko vahvasti kertyneet rykelmäksi kuvion oikeaan reunaan, mikä osoittaa, että varianssien homogeenisuusoletus rikkoontui.

Kolmogorov-Smirnovin normaaliustestin mukaan regressioanalyysin oletus jatkuvien muuttujien normaalijakautuneisuudesta toteutui vain muuttujien $\Delta_{\text{bull}_{\text{norm}}}$ ja $\text{EPC}_{\text{bull}_{\text{norm}}}$ osalta. Mallissa 3. Durbin-Watsonin testin tulos oli 1,790 ja 1,892 mallissa 4., mikä on hyvä asia ja kertoo, ettei autokorrelaatiota esiintynyt. Muita malleihin vaikuttavia tekijöitä ei tässä enää tutkittu.

5.3 Herkkyysanalyysit: puuttuvat tiedot poistettu opiskelija-aineistosta

Tässä alaluvussa käydään läpi, miten puuttuvia osake-ennusteita sisältävien vastausten poistaminen opiskelija-aineistosta vaikutti tuloksiin. Analyysyjä tehtiin vastaavalla tavalla kuin luvussa 5.1, pelkästään pankin perusaineistoa koskevat analyysit pois lukien, opiskelijoiden täydelliselle aineistolle (aineisto 3.) ja yhdistetylle aineistolle (aineistot 3. + 5.). Opiskelija-aineiston puuttuvat tiedot poistettiin siis Case Deletion -menetelmällä, jolloin opiskelijavastauksia jäi jäljelle 139 kappaletta. Tässä alaluvussa tuloksia ei enää esitetä taulukoiden muodossa, elleivät tulokset poikkea merkittävästi aiemmasta.

5.3.1 Ankkuroituivatko osake-ennusteet keskiarvotuottoa kohti?

Ensimmäisen hypoteesin tutkimista varten yhdistettiin opiskelijoiden täydellinen aineisto ja pankin perusaineisto. Kumpi vastaajaryhmistä oli enemmän ankkuroitunut ennusteissaan osakekuvaajan keskiarvoja kohti? Yhdistetylle aineistolle luotiin muuttujat R ja EROKA, kuten luvussa 5.1.1 tehtiin. Riippumattomien ryhmien t-testin tulokset ja p-arvot olivat samansuuntaiset kuin taulukossa 13 on esitetty: ennusteiden B2, C1, D1, E2 ja F1 kohdalla muuttujien EROKA keskiarvoissa oli tilastollisesti melkein merkitsevää tai merkitsevää eroa. Kaikkien yllä mainittujen ennusteiden, paitsi B2:n kohdalla, pankin vastaajat olivat enemmän konvergoituneita osakekuvaajien keskiarvotuottoja kohti. Näin ollen vain ennusteen C2 tulos poikkesi taulukon 13 tuloksista: täydellisellä, yhdistetyllä aineistolla ennusteen C2 tulos ei ollut tilastollisesti merkitsevää.

Koska muuttujat eivät olleet normaalisti jakautuneita, tehtiin vielä MWU-testi. Tulosten mukaan ennusteiden B2, C1, D1, E2 ja F1 kohdalla sijalukujen keskiarvoissa oli tilastollisesti merkitseviä eroja (U-testisuureet ja p-arvot samaa suuruusluokkaa kuin perusaineistolla). Huomattavaa siis on, että parametritestien tulokset olivat melkein samat kuin luvussa 5.1.1, eli pankin vastaajat olivat tässä enemmän konvergoituneita osakekuvaajien keskiarvotuottoja kohti.

vain D1-ennusteen kohdalla, muissa tapauksissa opiskelijat. Tulos oli ristiriidassa yllä olevien t-testin tulosten kanssa.

Entä oliko vastaajaryhmien välillä eroa toisinajattelijastrategian mukaisessa käyttäytymisessä? Ristiintaulukoinnin ja χ^2 -testin valmistelut olivat vastaavat kuin luvussa 5.1.1. Tuloksista huomattiin, että osakekuvaajien E ja F osalta χ^2 -testin edellytykset eivät täytyneet, joten niille tehtiin Fisherin tarkka testi. Tilastollisesti merkitseviä tuloksia saatiin vain osakkeiden A, B ja C kohdalla, joten yhdistetyllä perusaineistolla saatuihin tuloksiin nähden nyt A-osakkeellekin saatiin merkitsevä tulos (χ^2 -testin tulos A-osakkeelle oli $\chi^2(1) = 3,864$, p-arvo 0,049 ja kontingenssikerroin 0,147). Kyseisten kolmen osakkeen kohdalla vastaajaryhmän ja toisinajattelijaluokittelun välillä oli riippuvuutta, mutta riippuvuus oli heikkoa. Muut tulokset olivat samansuuntaisia kuin taulukossa 14 on esitetty. Ristiintaulukoinnin avulla saatuja jakaumia ei tässä enää kerrata, sillä ne olivat lähes identtiset luvussa 5.1.1 esitettyjen jakaumien kanssa. Johtopäätöksenä voidaan kuitenkin todeta, että osakkeen A osalta opiskelijat noudattivat enemmän toisinajattelijastrategiaa kuin pankin vastaajat, kun taas B- ja C-osakkeiden osalta pankin vastaajat.

5.3.2 Erot odotetussa hintamuutoksessa

Seuraavaksi tutkittiin hypoteesia H_{20} opiskelijoiden täydellisellä aineistolla verrannollisten parien t-testillä samojen muuttujien avulla kuin perusaineistossa. Tulosten mukaan keskimääräinen odotettu hintamuutos opiskelijoilla oli nousukausien jälkeen suurempi kuin laskukausien jälkeen molempien parien kohdalla ja erot olivat tilastollisesti erittäin merkitseviä, aivan kuten luvussa 5.1.2. T-testin tulokset ja p-arvot molemmille pareille olivat samansuuruiset kuin taulukossa 15 on esitetty. Muuttujien välillä ei ollut korrelaatiota, joten t-testiä voitiin käyttää.

Parametritonta testausta ei tässä tehty opiskelijavastauksille, sillä muuttujat olivat riittävän symmetrisiä. Vinousluku jaettuna keskivirheellään oli jokaisella muuttujalla $[-2, 2]$, joten normaalijakaumaoletus säilyi.

Yhdistetyllä aineistolla verrannollisten parien t-testin tulokset olivat samanlaiset kuin taulukossa 17 on esitetty: molempien parien osalta H_{20} hylättiin. T-testin tulokset ja p-arvot molemmille pareille vastasivat yhdistetyllä perusaineistolla saatuja arvoja. Keskimääräinen odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen.

Koska yhdistetyn aineiston muuttujista EPC_7 , EPC_{13} , $EPC_{7\text{bull}}$, $EPC_{7\text{bear}}$, $EPC_{13\text{bull}}$ ja $EPC_{13\text{bear}}$ kaikki eivät olleet normaalisti jakautuneita, eivätkä edes symmetrisiä, tehtiin vielä WSR-testi. Saatujen tulosten mukaan H_{20} hypoteesi hylättiin. Z-testisuureet ja p-arvot molemmille pareille olivat samansuuruiset kuin yhdistetyllä perusaineistolla, joten tuloksissa ei ilmennyt tässä suhteessa mitään poikkeavaa.

Voimme todeta siis, että puuttuvia osake-ennusteita sisältävien vastausten poistaminen opiskelija-aineistosta ei vaikuttanut hypoteesia H_{20} koskeviin tuloksiin. Vastaajat odottivat täydellisenkin aineiston perusteella menneen tren-

din jatkuvan. Keskimääräinen odotettu hintamuutos nousukausien jälkeen oli positiivinen ja laskukausien jälkeen negatiivinen.

5.3.3 Luottamusvälien vinous

Voidaanko H_{30} hylätä myös täydellisen opiskelija-aineiston tapauksessa? Tätä tutkittiin verrannollisten pariin t-testillä. Testiä varten opiskelijoiden täydelliseen aineistoon luotiin samat muuttujat kuin luvussa 5.1.3. Testiä ei voitu kuitenkaan käyttää parille $\Delta_{7\text{bull}}$ ja $\Delta_{7\text{bear}}$, koska muuttujien välinen korrelaatio oli melkein merkitsevää (Spearman).

Parin $\Delta_{13\text{bull}}$ ja $\Delta_{13\text{bear}}$ välillä oli tilastollisesti erittäin merkitsevää eroa. H_{30} hylättiin tämän parin osalta merkitsevyystasolla $p < 0,001$, $t(138) = -5,601$. Parille $\Delta_{7\text{bull}}$ ja $\Delta_{7\text{bear}}$ tehtiin parametriton WSR-testi ja sen mukaan H_{30} hylättiin merkitsevyystasolla $p < 0,001$ (Z-testisuureen arvo $-6,662$).

Koska vain muuttuja $\Delta_{13\text{bull}}$ oli normaalisti jakautunut, tehtiin myös parille $\Delta_{13\text{bull}}$ ja $\Delta_{13\text{bear}}$ WSR-testi ja senkin mukaan H_{30} hylättiin merkitsevyystasolla $p < 0,001$ (Z-testisuureen arvo $-5,559$). Näin ollen opiskelijoiden täydellisellä aineistolla edellä saadut tulokset olivat samoja kuin perusaineistolla saadut. Luottamusvälien keskimääräinen vinous nousukausien jälkeen oli negatiivinen, eli luottamusvälit olivat vinoja vasemmalle ja laskukausien jälkeen positiivinen, mikä tarkoittaa, että luottamusvälit olivat vinoja oikealle.

Yhdistetyllä aineistolla saadut tulokset olivat samansuuntaisia kuin luvussa 5.1.3 on esitetty. Molempien pariin muuttujien välillä oli merkitsevyystasoltaan samantasoista korrelaatiota kuin perusaineistolla, joten H_{30} hypoteesia tutkittiin vain WSR-testillä. Sen mukaan H_{30} hylättiin molempien pariin osalta merkitsevyystasolla $p < 0,001$ (Z-testisuureen arvo ensimmäiselle parille $-7,481$ ja toiselle parille $-5,855$). Yhdistetyn perusaineiston tapaan myös tässä keskimääräinen luottamusvälien vinous oli sekä nousu- että laskukausien jälkeen negatiivinen, mutta nousukausien jälkeen vinous oli pienempää.

5.3.4 Regressioanalyysit luottamusvälien vinoudelle

Hypoteesiin H_4 liittyen tehtiin regressioanalyysit samalla periaatteella kuin luvussa 5.1.4.

Regressioanalyysien tekoa edelsivät tässäkin korrelaatiotarkastelut (Spearman). Opiskelijoiden täydellisellä aineistolla nähtiin, että Δ_{bull} korreloi tilastollisesti melkein merkitsevästi muuttujan EPC_{bull} kanssa ja merkitsevästi muuttujan $\text{TF}_{\text{bull_CO}}$ kanssa. Δ_{bear} taas korreloi merkitsevästi muuttujien EPC_{bear} ja $\text{TF}_{\text{bear_COstrong}}$ kanssa. Lisäksi nähtiin, että muuttujat EPC_{bear} ja $\text{TF}_{\text{bear_COstrong}}$ korreloivat merkitsevästi keskenään, samoin kuin muuttujat Δ_{bull} ja $\text{TF}_{\text{bull_CO}}$. Päätettiin tehdä neljä regressioanalyysiä, aivan kuten perusaineistonkin kohdalla tehtiin. Mallissa 1. regressioitiin muuttujaa Δ_{bull} muuttujan EPC_{bull} suhteen ja mallissa 2. muuttujaa Δ_{bull} muuttujan $\text{TF}_{\text{bull_CO}}$ suhteen. Seuraavaksi regressioitiin muuttujaa Δ_{bear} erikseen muuttujien EPC_{bull} ja $\text{TF}_{\text{bear_COstrong}}$ suhteen (mallit

3. ja 4.). Nyt puuttuvien tietojen käsittelykomennolla ei ollut merkitystä, koska opiskelijoiden täydellinen aineisto ei sisältänyt puuttuvia tietoja.

Ajettujen regressioiden tulokset olivat samansuuntaiset kuin perusaineistolla ajettujen regressioiden tulokset. Regressiokertoimien arvot olivat samaa suuruusluokkaa, kuten myös F- ja t-testisuureiden arvot sekä merkitsevyystasot. H_4 sai tukea mallin 3. tapauksessa, ja mallin 1. osaltakin regressiokerroin muuttujalle EPC_bull oli hypoteesin mukainen, mutta tulos ei ollut tilastollisesti merkitsevä. Tässä ei käydä tuloksia tämän tarkemmin läpi, sillä mitään merkittävää muutosta suhteessa perusaineistolla saatuihin tuloksiin ei ilmennyt. Tässä ei myöskään enää tutkita muita regressioanalyysin oletuksia tai malleihin vaikuttaneita tekijöitä.

Yhdistetyllä aineistolla saadut korrelaatiotarkastelujen tulokset olivat samansuuntaisia kuin yhdistetyllä perusaineistolla, joten tehtiin samanlaiset regressioanalyysit kuin luvussa 5.1.4. Puuttuvien tietojen käsittelyä ei tarvinnut tässäkään tapauksessa erikseen määritellä.

Ajettujen regressioanalyysien tulokset olivat aivan samansuuntaiset kuin yhdistetyllä perusaineistolla saadut tulokset. Selitysasteet, regressiokertoimet, F- ja t-testisuureiden arvot, merkitsevyystasot ja Durbin-Watsonin testin arvot olivat samaa tasoa kuin yhdistetyllä perusaineistolla saadut, joten H_4 sai tukea mallin 3. tapauksessa. Tässä ei enää esitetä tuloksia tarkemmin, eikä tutkita muita regressioanalyysin oletuksia, koska tulokset ovat samoja kuin luvussa 5.1.4 esitetyt. Myöskään malleihin vaikuttaneita tekijöitä ei enää tutkita tarkemmin.

Johtopäätöksenä voidaan todeta lyhyesti, että puuttuvien tietojen poistaminen ei muuttanut regressioanalyysien tuloksia suhteessa perusaineistoihin.

5.3.5 Yliluottamus

Yliluottamusta testattiin samalla tavalla kuin luvussa 5.1.5, lukuun ottamatta omien sijoitusten tuottoon perustuvaa χ^2 - testiä. Analyysit tehtiin ensin opiskelijoiden täydelliselle aineistolle ja lopuksi yhdistetylle aineistolle.

Hypoteesiin H_9 liittyen haluttiin tutkia, olivatko miehet yliluottavaisempia kuin naiset. Hypoteesia testattiin aluksi riippumattomien ryhmien t-testillä, jossa verrattiin luottamusvälien keskiarvoja sukupuolittain opiskelijoiden täydellisessä aineistossa. Sitä varten luotiin muuttuja LUOTTAMUSVÄLI, joka siis saatiin vähentämällä yläluottamusvälin arvosta alaluottamusvälin arvo. Tulokset olivat samansuuntaiset kuin perusaineistolla saadut tulokset: luottamusväleissä oli tilastollisesti merkitsevää eroa C- ja E-osakkeiden 13 kuukauden sekä F-osakkeen seitsemän ja 13 kuukauden luottamusväliennusteiden kohdalla. T-testin tulokset ja p-arvot luottamusväleittäin olivat samansuuruiset kuin taulukossa 24 esitetyt. Miehet eivät siis opiskelijoiden täydellisen aineiston perusteella näyttäneet olevan naisia yliluottavaisempia kuin edellä mainitussa neljässä tapauksessa, joten muilta osin H_9 ei saanut tukea. On kuitenkin huomattava, että F-osakkeen seitsemän ja 13 kuukauden luottamusväliennusteiden kohdalla t-testin toinen oletus rikkoontui: varianssit eivät olleet Levenen testin mukaan yhtä suuret.

Luottamusvälien kapeutta testattiin vielä parametrittomalla MWU-testillä, sillä normaalijakautuneisuusoletus ei toteutunut millään muuttujalla. Testin tulosten mukaan luottamusvälien sijalukujen keskiarvot eivät olleet samoja C-osakkeen 13 kuukauden sekä F-osakkeen seitsemän ja 13 kuukauden luottamusväliennusteissa, eli tulokset olivat hiukan erilaiset verrattuna t-testillä saatuihin ja myös perusaineistolla saatuihin tuloksiin nähden. H_9 sai siis tukea vain näiden kolmen muuttujan osalta (U-testisuureet 1794,0; 1786,0 ja 1583,5 sekä vastaavat p-arvot 0,026; 0,023 ja 0,002), mutta muuten ei.

Yliluottamusta tutkittiin lisäksi ennustettujen tulevien tuottojen perusteella. Ennustivatko opiskelijanaiset tulevia tuottoja varovaisemmin kuin miehet? Riippumattomien ryhmien t-testin oletus varianssien yhtäsuuruudesta rikkoon tui ennusteiden A1, A2 ja D2 kohdalla, aivan kuten perusaineistonkin kohdalla. Tilastollisesti merkitseviä tuloksia saatiin ennusteiden C1, D1, D2 ja F2 kohdalla, joten melko samankaltainen tulos saatiin kuin perusaineistolla, joka ei tosin näyttänyt ennusteissa F2 olevan tilastollisesti merkitsevää eroa. Tässä t-testin tulos ja p-arvo F2-ennusteille oli $t(137) = -2,012; 0,046$. Muilta osin tulokset vastasivat taulukon 25 tasoa. Voimme siis todeta, että opiskelijoiden täydellisen aineiston perusteella vain C1-, D1-, D2- ja F2-ennusteiden kohdalla miesten ja naisten ennusteiden keskiarvoissa oli tilastollisesti merkitsevää eroa, joten H_9 sai tukea vain näiltä osin. Näiden ennusteiden kohdalla opiskelijamiehet olivat yliluottavaisempia kuin opiskelijanaiset.

Sukupuolittaisia eroja osake-ennusteissa testattiin myös MWU-testillä, sillä ennusteet eivät olleet normaalisti jakautuneet. Testin tulokset olivat täysin samat kuin perusaineistolla saadut tulokset: naisten ja miesten ennusteiden sijalukujen keskiarvot poikkesivat tilastollisesti merkitsevästi C1-, D1-, D2- ja F2-ennusteiden kohdalla (U-testisuureet ja p-arvot samaa tasoa kuin perusaineistolla saaduissa tuloksissa). Näin ollen H_9 sai tukea vain näiden neljän ennusteen kohdalla, mutta ei muuten.

Lopuksi testattiin yliluottamusta H_5 :een liittyen yhdistetyllä aineistolla. Riippumattomien ryhmien t-testillä testattiin, arvioivatko pankin sijoitusasiantuntijat luottamusvälit kapeammiksi kuin opiskelijat ja ennustivatko opiskelijat tulevia tuottoja varovaisemmin kuin pankin vastaajat. T-testin tulokset luottamusvälien kapeudelle olivat samat kuin taulukossa 26 esitetyt, eli tilastollisesti merkitsevä tulos saatiin vain C-osakkeen seitsemän kuukauden luottamusvälin osalta (t-testisuure ja p-arvo samaa tasoa kuin taulukossa 26). Koska luottamusvälimuuttujat eivät olleet normaalisti jakautuneet, niin tehtiin vielä MWU-testi. Testin tulos oli aivan samanlainen kuin t-testinkin: luottamusvälien sijalukujen keskiarvot eivät olleet samat C-osakkeen seitsemän kuukauden luottamusvälin osalta (U-testisuure ja p-arvo samaa tasoa kuin perusaineistolla saadussa tuloksessa), joten vain kyseisen luottamusvälin osalta H_5 sai tukea. Pankin sijoitusasiantuntijat olivat sen osalta yliluottavaisempia kuin opiskelijat.

Opiskelijoiden ennusteiden varovaisuutta suhteessa pankin ennusteisiin testattiin sekä riippumattomien ryhmien t-testillä että MWU-testillä, sillä osakeennusteet eivät olleet normaalisti jakautuneita. T-testin mukaan pankin vastaajat olivat yliluottavaisempia ennusteiden B2, C1, D1, E2 ja F1 kohdalla. T-testin

tulokset ja p-arvot osake-ennusteittain olivat samaa tasoa kuin taulukossa 27 on esitetty. Muilta osin H_5 ei saanut tukea tämän yhdistetyn aineiston perusteella. Tulokset olivat siis lähes samanlaisia kuin perusaineistolla saadut tulokset, sillä poikkeuksella, että C2-ennusteen kohdalla piste-estimaattien keskiarvoissa ei ollut nyt tilastollisesti merkitsevää eroa. MWU-testin mukaan sijalukujen keskiarvoissa oli tilastollisesti merkitseviä eroja ennusteissa B2, C1, D1, E2 ja F1 (U-testisuureet ja p-arvot samaa tasoa kuin perusaineistolla saaduissa tuloksissa). Vain kyseisten ennusteiden osalta voidaan sanoa, että opiskelijoiden ennusteet olisivat varovaisempia kuin pankin ennusteet. Parametrittoman testin tulos oli näin ollen sama kuin perusaineistolla saatu.

5.3.6 Tekivätkö koulutetut henkilöt täsmällisempiä ennusteita?

Opiskelijoiden täydelliselle aineistolle tehtiin samat analyysit liittyen kuudenteen hypoteesiin kuin luvussa 5.1.6.

Ensimmäiseksi haluttiin selvittää, oliko opiskelijoiden trendinseuranta-käyttäytymisessä eroja sen suhteen, kuinka paljon rahoituksen kursseja oli suoritettuina. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Analyysin valmistelut tehtiin kuten luvussa 5.1.6, eikä nousu- tai laskukausien analyysien osalta saatu tilastollisesti merkitseviä tuloksia: χ^2 -testisuureen arvo, p-arvot ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. Tässäkin tapauksessa voidaan siis todeta, etteivät enemmän rahoituksen kursseja suorittaneet käyttäytyneet trendinseuraamisen suhteen eri tavalla kuin vähemmän rahoituksen kursseja käyneet. H_6 ei siis saanut tukea, eli tulokset vastasivat perusaineistolla saatuja tuloksia.

Seuraavaksi tutkittiin, oliko opiskelijoiden trendinseuraamiskäyttäytymisessä eroja sen suhteen, kuinka paljon kaikkia opintoja oli suoritettuina. Tätä tutkittiin jälleen ristiintaulukoinnin ja χ^2 -testin avulla. Testin valmistelut tehtiin kuten luvussa 5.1.6. Tulokset olivat samansuuntaiset kuin perusaineistolla saadut tulokset. Laskukausien analyysin osalta χ^2 -testin edellytykset eivät täyttyneet, ja Fisherin tarkan testin p-arvo oli samaa tasoa kuin perusaineistolla saatu, eli tulos ei ollut tilastollisesti merkitsevää. Nousukausien analyysin osalta ei myöskään saatu tilastollisesti merkitseviä tuloksia: χ^2 -testisuureen arvo, p-arvo ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. Trendinseuraamiskäyttäytymisessä ei siis ollut eroja sen suhteen, kuinka paljon kaikkia opintoja oli suoritettuna. Tulokset eivät siis poikenneet perusaineistolla saaduista tuloksista.

Trendinseuraamiskäyttäytymiseen liittyen haluttiin vielä tutkia, aivan kuten perusaineistollakin, oliko käyttäytymisessä eroja sen suhteen, kuinka hyviä arvosanoja rahoituksen kursseista oli saatu. Tätä tutkittiin jälleen kerran ristiintaulukoinnin ja χ^2 -testin avulla. Kummankaan analyysin osalta χ^2 -testin edellytykset eivät täyttyneet, joten tehtiin Fisherin tarkka testi. Nousu- ja laskukausien p-arvot olivat samaa tasoa kuin perusaineistolla, joten tilastollisesti merkitseviä tuloksia ei saatu kummassakaan tapauksessa. Trendinseuraamiskäyttäytymisessä ei siis ollut eroja rahoituksen kurssien arvosanojen suhteen.

Seuraavaksi hypoteesia H_6 tutkittiin vielä ennustevirheen kautta. Ensiksi tutkittiin yksisuuntaisella varianssianalyysillä suorittujen rahoituksen kurssien määrän yhteyttä ennustevirheeseen. Levenen varianssien yhtäsuuruustestin mukaan varianssit olivat yhtä suuret (p-arvo 0,290), ja F-testin tulokset olivat identtiset perusaineistolla saatujen tulosten kanssa. Koska oletus muuttujien normaalijakautuneisuudesta ei täytynyt, tehtiin vielä Kruskal-Wallis testin. Sen tulokset olivat niin ikään samat kuin perusaineistolla saadut, eli suoritetun rahoituksen kurssien määrällä ei ollut vaikutusta ennustevirheeseen.

Ennustevirheen ja suoritettujen rahoituksen kurssien välisen korrelaatiotarkastelun tulokset eivät poikenneet perusaineistolla saaduista tuloksista, eli korrelaatiota ei esiintynyt tässäkään tapauksessa.

Seuraavaksi tutkittiin kaikkien suoritettujen opintopisteiden yhteyttä ennustevirheeseen yksisuuntaisella varianssianalyysillä. Opiskelijat jaettiin ryhmiin suorittujen opintopisteiden perusteella samalla tavalla kuin perusaineiston analyysin yhteydessä tehtiin. Tulokset olivat identtiset perusaineistolla saatuihin tuloksiin nähden: suoritettujen opintopisteiden määrällä ei ollut tilastollisesti merkitsevää vaikutusta ennustevirheeseen. Koska oletus muuttujien normaalijakautuneisuudesta ei pätenyt, tehtiin vielä Kruskal-Wallis parametrin testi. Sen tulokset olivat jälleen samat kuin perusaineistolla saadut tulokset. Muuttujan ENNVIRHE sija lukujen keskiarvoissa ei eri opintopisteluokissa ollut tilastollisesti merkitsevää eroa.

Kokonaisopintopistemäärän ja ennustevirheen korrelaatiotarkastelut toivat niin ikään samat tulokset kuin perusaineiston analyysillä, eli korrelaatiota ei esiintynyt.

5.3.7 Sijoituskokemus ja virheet

Seitsemänten hypoteesiin liittyen tutkittiin ensiksi, oliko sijoittamista harrastavien ja ei-harrastavien opiskelijoiden välillä eroa trendinseuraamiskäyttäytymisessä. Tätä tutkittiin opiskelijoiden täydellisellä aineistolla ristiintaulukoinnin ja χ^2 -testin avulla. Selitettäväksi muuttujiksi valittiin vuorollaan TF_bull ja TF_bear sekä selittäväksi muuttujaksi sijoitusharrastus. Tässäkään tapauksessa, laskukausien analyysin osalta, χ^2 -testin edellytykset eivät täytyneet, joten tehtiin Fisherin tarkka testi, josta ei saatu tilastollisesti merkitseviä tuloksia. Myöskään nousukausien analyysin osalta tulokset eivät olleet tilastollisesti merkitseviä: χ^2 -testisuureen arvo, p-arvo ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. Voidaan siis todeta, että opiskelijoiden täydellisen aineiston perusteella sijoittamista harrastavat käyttäytyivät trendinseuraamisen suhteen samoin kuin ei-sijoittajat, eli H_7 ei saanut tukea. Tulos siis vastasi perusaineistolla saatua tulosta.

Hypoteesia H_7 tutkittiin vielä ennustevirheen kautta. Riippumattomien ryhmien t-testillä testattiin, oliko opiskelijoiden sijoitusharrastuksella vaikutusta ennustevirheeseen. T-testisuureen arvo ja p-arvo olivat täsmälleen samat kuin perusaineistolla, joten H_7 ei saanut tukea. Koska muuttuja ENNVIRHE ei ollut normaalisti jakautunut, tehtiin vielä parametrin MWU-testi, jonka tulokset olivat niin ikään samanlaiset kuin perusaineistolla saadut. Johtopäätöksenä

voidaan täydellisen aineiston kohdallakin todeta, että ennustevirheissä ei ollut tilastollisesti merkitsevää eroa sijoittamista harrastavien ja ei-harrastavien opiskelijoiden välillä.

5.3.8 Lähipiiri ja ennustevirheet

Hypoteesiin H_8 liittyen haluttiin tietää, oliko opiskelijoiden trendinseuraamiskäyttäytymisessä eroja sen suhteen, harrastivatko heidän vanhempansa sijoittamista vai eivät. Tätä tutkittiin opiskelijoiden täydellisellä aineistolla ristiintaulukoinnin ja χ^2 -testin avulla. Testin valmistelut tehtiin kuten luvussa 5.1.8. Tuloksista havaittiin kuitenkin, että laskukausien analyysin osalta testin edellytykset eivät täytyneet, joten tehtiin Fisherin tarkka testi. P-arvo oli samaa tasoa kuin perusaineistolla, ja tulos siis oli tilastollisesti melkein merkitsevä. Ristiintaulukoinnin avulla saatuja jakaumia ei tässä enää käydä läpi, sillä jakaumat olivat samoja kuin luvussa 5.1.8.

Nousukausien analyysin osalta tulokset eivät olleet tilastollisesti merkitseviä: χ^2 -testisuureen arvo, p-arvo ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. Hypoteesi H_8 ei siis saanut tukea nousu- eikä laskukausien osalta.

Seuraavaksi samaa asiaa tutkittiin suhteessa siihen, harrastivatko vastaajien parhaimmat ystävät sijoittamista vai eivät. Testin valmistelut tehtiin kuten luvussa 5.1.8. Tulokset olivat samansuuntaiset kuin perusaineistolla saadut tulokset. Laskukausien analyysin osalta χ^2 -testin edellytykset eivät täytyneet, ja Fisherin tarkan testin p-arvo ei ollut tilastollisesti merkitsevä. Nousukausien analyysin tulokset eivät olleet tilastollisesti merkitseviä: χ^2 -testisuureen arvo, p-arvo ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. H_8 ei siis saanut tukea, ja tulokset vastasivat perusaineistolla saatuja tuloksia.

Trendinseuraamiseen liittyen haluttiin vielä selvittää, oliko käyttäytymisessä eroja suhteessa siihen, mikä oli vanhempien sosioekonominen asema. Tätä tutkittiin jälleen ristiintaulukoinnin ja Fisherin tarkan testin avulla. Tulokset olivat samoja kuin perusaineistolla saadut, eli tilastollisesti merkitseviä eroja ei löydetty. H_8 ei saanut tukea, eli opiskelijoiden trendinseuraamiskäyttäytymisessä ei ollut eroja suhteessa vanhempien sosioekonomiseen asemaan.

Hypoteesia H_8 tutkittiin vielä ennustevirheen kautta. Ensimmäiseksi haluttiin selvittää, oliko opiskelijoiden ennustevirheissä eroa riippuen siitä, harrastivatko heidän vanhempansa sijoittamista vai eivät. Keskiarvovertailuista nähtiin, että ennustevirhe oli keskimäärin pienempi niillä opiskelijoilla, joiden vanhemmat eivät harrastaneet sijoittamista. T-testin tulos ja p-arvo olivat täsmälleen samat kuin perusaineistolla saadut. Perusaineiston tuloksia vastaava tulos saatiin myös MWU-testillä. Kummankaan testin perusteella H_8 ei siis saanut tukea.

Entä oliko opiskelijoiden ennustevirheissä eroa riippuen siitä, harrastivatko jotkut heidän parhaimmista ystävästään sijoittamista vai eivät? Riippumattomien ryhmien t-testin tulokset olivat täysin samanlaiset kuin perusaineistolla saadut tulokset. Opiskelijoiden ennustevirheiden keskiarvoissa ei ollut tilastol-

lisesti merkitsevää eroa sen suhteen, harrastivatko heidän parhaimmat ystävänä sijoittamista vai eivät. T-testisuureen arvo ja p-arvo olivat samat kuin perusaineistolla. Yhtä lailla MWU-testin tulokset olivat samanlaiset kuin perusaineistolla (U-testisuureen arvo ja p-arvo samat): ennustevirheiden sijalukujen keskiarvoissa oli tilastollisesti merkitsevää eroa ja niiden opiskelijoiden, joiden parhaimmat ystävät harrastivat sijoittamista, ennustevirheiden sijalukujen keskiarvot olivat pienemmät kuin niiden, joiden parhaimmat ystävät eivät harrastaneet sijoittamista. Parametrittoman testin tulokset antoivat siis tukea H_8 :lle.

Lopuksi tutkittiin kaksisuuntaisella varianssianalyysillä, oliko ennustevirhe kaikkein pienin niillä opiskelijoilla, joiden vanhemmat eivät harrastaneet sijoittamista, mutta parhaimmat ystävät harrastivat. Jatkuvaksi muuttujaksi asetettiin ennustevirhe ja ryhmitteleviksi muuttujiksi vanhempien ja parhaimpien ystävien sijoittamisharrastus. Keskiarvoja verrattaessa ennustevirhe oli pienin silloin, kun vanhemmat eivät harrastaneet sijoittamista ja parhaimmat ystävät harrastivat, mutta yhdysvaikutus ei ollut tilastollisesti merkitsevää. Muutenkin F-testin tulokset olivat täsmälleen samat kuin perusaineistolla saadut tulokset.

Lopuksi tutkittiin vielä yksisuuntaisella varianssianalyysillä ennustevirheiden eroja suhteessa siihen, mikä oli äidin tai isän sosioekonominen asema. Vanhemmat ryhmiteltiin samalla tavalla sosioekonomisen aseman mukaan, kuten luvussa 5.1.8 tehtiin. Tässäkin tapauksessa ei saatu tilastollisesti merkitseviä tuloksia, eli vanhempien sosioekonomisella asemalla ei ollut vaikutusta ennustevirheiden keskiarvojen eroihin.

Hypoteeseihin H_7 ja H_8 liittyen tutkittiin vielä opiskelijoiden täydellisellä aineistolla, mitkä tekijät selittävät sijoitusharrastusta. Korrelaatiotarkasteluiden perusteella päätettiin muodostaa tässäkin tapauksessa kaksi logistista regressiomallia: mallissa 1. selittäjinä eli kovariaatteina olivat sukupuoli, opiskeltujen rahoituksen kurssien määrä ja SOSEKAS_äiti. Mallissa 2. selittäjinä olivat parhaimpien ystävien sijoittamista kuvaava muuttuja ja SOSEKAS_äiti. Mallin asetukset pidettiin samoina kuin luvussa 5.1.8.

Mallista 1. päätettiin kuitenkin poistaa muuttuja SOSEKAS_äiti, sillä se ei parantanut mallia: p-arvot olivat suuria kaikissa muuttujan kategoriaissa. Uuden mallin tulokset olivat hyvin samanlaiset kuin perusaineistolla saadut tulokset, joten tässä tyydytään suppeampaan esitykseen kuin luvussa 5.1.8. Malli oli tilastollisesti merkitsevä: $\chi^2(2) = 14,881$; p-arvo 0,001, ja Nagelkerken selitysaste oli 14,7 %. Waldin testin perusteella voidaan sanoa, että molemmat selittäjämuuttujat lisäsivät mallin ennustekykyä, ja p-arvot olivat samaa luokkaa kuin taulukossa 29 on esitetty. Ristitulosuhteiden arvot olivat niin ikään samansuuntaisia kuin taulukossa 29. Mallin residuaaleja ei enää tutkittu tässä yhteydessä.

Myös mallin 2. asetukset olivat vastaavat kuin luvussa 5.1.8. Muuttuja SOSEKAS_äiti poistettiin mallista, sillä se ei parantanut mallia. Lopulta ainoana selittäjämuuttujana mallissa oli siis parhaimpien ystävien sijoittamista kuvaava muuttuja. Regression tulokset olivat samansuuntaisia kuin taulukossa 30 on esitetty, joten tässä ei enää tarkemmin selosteta niitä.

5.3.9 Ikä ja ennustevirheet

Hypoteesia H_{10} tutkittiin yhdistetyllä aineistolla. Ensimmäiseksi haluttiin selvittää, oliko trendinseuraamiskäyttäytymisessä eroja suhteessa vastaajan ikään. Tätä tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla. Testin valmistelut tehtiin kuten luvussa 5.1.9. Laskukausien analyysin osalta testin edellytykset eivät kuitenkaan täyttyneet, joten tehtiin Fisherin tarkka testi. Kyseisen testin p-arvo oli samaa tasoa kuin perusaineistolla, eli ei tilastollisesti merkitsevä. Myöskään nousukausien analyysien tulokset eivät olleet tilastollisesti merkitseviä: χ^2 -testisuureen arvo, p-arvo ja kontingenssikertoimen arvo olivat samaa tasoa kuin perusaineistolla saaduissa tuloksissa. Näin ollen H_{10} ei saanut tukea: yhdistetyn aineiston perusteella ei voida sanoa, että trendinseuraamiskäyttäytymisessä olisi ollut eroja sen suhteen, mikä vastaajan ikä oli. Tulokset vastasivat yhdistetyllä perusaineistolla saatuja tuloksia.

Seuraavaksi tutkittiin hypoteesia H_{10} ennustevirheen kautta. Oliko iällä ja ennustevirheellä yhteyttä? Varianssianalyysin tulokset olivat tässäkin tapauksessa samanlaiset kuin yhdistetyllä perusaineistolla saadut tulokset: eri ikäryhmien ennustevirheiden keskiarvoissa ei ollut eroa (F-testisuureen arvo ja p-arvo samat). Myös Kruskal-Wallis testin tulokset olivat yhtä pitäviä luvussa 5.1.9 esitettyjen tulosten kanssa (χ^2 -testisuureen arvo ja p-arvot samat). H_{10} ei siis saanut näiden testien tulosten mukaan tukea.

5.3.10 Opiskelijatko häirikkösijoittajia?

Lopuksi hypoteesia H_{11} tutkittiin yhdistetyllä aineistolla. Haluttiin siis tietää, olivatko opiskelijat enemmän trendinseuraajia suhteessa pankin sijoitusasiantuntijoihin. Aluksi luotiin aineistolle samat muuttujat kuin luvussa 5.1.4 tehtiin. Sen jälkeen tutkittiin ristiintaulukoinnin ja χ^2 -testin avulla, oliko vahvassa trendinseuraamis- ja toisinajattelijakäyttäytymisessä eroa vastaajaryhmien välillä. Selitettäväksi muuttujaksi valittiin vuorollaan TF_bull_COstrong sekä TF_bear_COstrong ja selittäväksi muuttujaksi ryhmä (R). Molempien analyysien kohdalla havaittiin kuitenkin, etteivät χ^2 -testin edellytykset täyttyneet, joten tehtiin Fisherin tarkka testi, jonka p-arvot olivat samaa tasoa kuin yhdistetyllä perusaineistolla. Tulokset eivät siten olleet tilastollisesti merkitseviä, joten H_{11} ei saanut tämän perusteella tukea.

Häirikkösijoittamista tutkittiin vielä sukupuolittain opiskelijoiden täydellisellä aineistolla ristiintaulukoinnin ja χ^2 -testin avulla, kuten luvussa 5.1.10 tehtiin. Tulosten mukaan χ^2 -testin edellytykset eivät täyttyneet kummankaan analyysin kohdalla, joten tehtiin Fisherin tarkka testi. Nousukausien osalta saatiin tilastollisesti merkitsevä tulos, laskukausien osalta tulokset taas eivät olleet tilastollisesti merkitseviä. Tulosten perusteella miehet olivat enemmän häirikkösijoittajia, kuten perusaineistollakin.

Luvussa 5.1.10 tuotiin esiin neljä logistista regressiomallia, joissa muuttujia TF_bull_COstrong ja TF_bear_COstrong pyrittiin selittämään. Tässä alaluvussa käsitellylle yhdistetylle aineistolle samanlaisia logistisia regressioita ei enää tehty.

5.4 Herkkyysanalyysit: puuttuvat tiedot poistettu opiskelija-aineistosta ja normalisoinnit tehty

Tässä aluvussa tutkitaan hypoteeseja H_{20} , H_{30} ja H_4 siten, että odotetun hintamuutoksen ja vinouden parametrit on normalisoitu jakamalla ne toteutuneen S&P 500 indeksin kuukausituoton keskihajonnalla (σ) neljän kuukauden ajalta ennen kunkin osakekuvaajan viimeistä näkyvää arvoa. Yllä mainittuja hypoteeseja tutkitaan opiskelijoiden täydellisellä ja normalisoidulla aineistolla (aineisto 4.) sekä yhdistetyllä aineistolla 4. + 6. Pankin normalisoidulle aineistolle (aineisto 6.) analyysijä ei enää tässä luvussa tehdä.

5.4.1 Erot odotetussa hintamuutoksessa

Vaikuttivatko odotetun hintamuutoksen parametrien normalisointi ja puuttuvien tietojen poistaminen hypoteesin H_{20} hylkäämispäätökseen? Kysymystä tutkittiin ensin opiskelijavastausten ja sen jälkeen yhdistettyjen vastausten kautta. Hypoteesin tutkimista varten molemmille aineistoille luotiin totuttuun tapaan muuttujat $EPC7_{\text{bull_norm}}$, $EPC7_{\text{bear_norm}}$, $EPC13_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$.

Täydellisellä ja normalisoidulla opiskelija-aineistolla huomattiin, että muuttujaparit $EPC7_{\text{bull_norm}}$ ja $EPC7_{\text{bear_norm}}$ sekä $EPC13_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$ eivät korreloineet keskenään Pearsonin korrelaatiokertoimilla mitattuna, joten verrannollisten pariin t-testiä voitiin käyttää. Keskiarvovertailuista nähtiin, että molempien pariin osalta keskimääräinen odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen. T-testin tulokset ja p-arvot olivat molemmille pareille samansuuruiset kuin perusaineistolla saadut. Näin ollen H_{20} hylättiin molempien pariin osalta, ja tulokset olivat tilastollisesti erittäin merkitseviä. Keskimääräinen odotettu hintamuutos nousukausien jälkeen oli positiivinen ja laskukausien jälkeen negatiivinen, eli tämänkin aineiston perusteella vastaajat osoittautuivat optimistisemmiksi nousukausien jälkeen.

Kaikki muuttujista noudattivat normaalijakaumaa (tosin $EPC13_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$ niukasti: Kolmogorov-Smirnovin testin p-arvot 0,053 ja 0,055), joten parametritonta testiä ei tehty.

Seuraavaksi hypoteesia tutkittiin yhdistetyllä aineistolla. Muuttujapariin $EPC7_{\text{bull_norm}}$ ja $EPC7_{\text{bear_norm}}$ sekä $EPC13_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$ välillä ei ollut korrelaatiota Spearmanin järjestyskorrelaatiokertoimen perusteella, joten hypoteesia testattiin ensin verrannollisten pariin t-testillä. Keskiarvovertailuista nähtiin jälleen, että keskimääräinen odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen. T-testin tulokset ja p-arvot olivat samansuuntaiset kuin yhdistetyllä perusaineistolla saadut. Näin ollen H_{20} hylättiin, ja tulokset olivat tilastollisesti erittäin merkitseviä.

Koska muuttujat $EPC7_{\text{bull_norm}}$ ja $EPC13_{\text{bear_norm}}$ eivät olleet normaalisti jakautuneita, suoritettiin hypoteesin testaus myös parametrittomalla WSR-testillä. Sen mukaan H_{20} hylättiin molempien pariin osalta merkitsevyydellä $p < 0,001$. Näin ollen tulokset eivät poikenneet yhdistetyllä perusaineistolla saa-

duista tuloksista. Vastaajat osoittautuivat optimistisemmiksi nousukausien jälkeen, eli keskimääräinen odotettu hintamuutos oli nousukausien jälkeen positiivinen ja laskukausien jälkeen negatiivinen.

5.4.2 Luottamusvälien vinous

Seuraavaksi tutkittiin, vaikuttivatko normalisointi ja puuttuvia osakeennusteita sisältävien vastausten poistaminen tuloksiin hypoteesin H_{30} kohdalla. Aluksi keskityttiin totuttuun tapaan opiskelijoihin ja lopuksi yhdistettyyn aineistoon. Analyysin kohteena olivat muuttujat $\Delta_{7\text{bull_norm}}$, $\Delta_{7\text{bear_norm}}$, $\Delta_{13\text{bull_norm}}$ ja $\Delta_{13\text{bear_norm}}$.

Opiskelijoiden täydellisellä ja normalisoidulla aineistolla havaittiin, että muuttujaparien välillä ei ollut korrelaatiota Spearmanin järjestyskorrelaatiokerrotoimien perusteella. Keskiarvovertailujen perusteella luottamusvälien keskimääräiset normalisoidut vinoudet olivat nousukausien jälkeen pienempiä kuin laskukausien jälkeen. Nousukausien jälkeen luottamusvälit olivat vinoja vasemmalle, eli vinouden parametrien arvot olivat negatiivisia ja laskukausien jälkeen vinoja oikealle, eli vinouden parametrien arvot olivat positiivisia. T-testin tulokset ja p-arvot olivat samantasoisia kuin perusaineistolla saadut, joten molempien parien osalta H_{30} hylättiin, ja tulokset olivat tilastollisesti erittäin merkitseviä.

Koska muuttujat $\Delta_{7\text{bear_norm}}$ ja $\Delta_{13\text{bear_norm}}$ eivät olleet normaalisti jakautuneet, testattiin hypoteesia vielä parametrittomalla WSR-testillä. Tulosten mukaan H_{30} hylättiin molempien parien osalta merkitsevyystasolla $p < 0,001$: nousukausien jälkeen luottamusvälien vinous oli keskimäärin pienempää kuin laskukausien jälkeen. Tulokset olivat siis samat kuin perusaineistolla saadut.

Yhdistetyn aineiston kohdalla havaittiin, että muuttujapari $\Delta_{7\text{bull_norm}}$ ja $\Delta_{7\text{bear_norm}}$ korreloi keskenään, kuten myös muuttujapari $\Delta_{13\text{bull_norm}}$ ja $\Delta_{13\text{bear_norm}}$ (Spearmanin järjestyskorrelaatiokerrotoimen arvo ja p-arvo ensimmäiselle parille 0,188 ja 0,013 sekä jälkimmäiselle parille 0,164 ja 0,030). Lisäksi ainoastaan muuttuja $\Delta_{13\text{bull_norm}}$ oli normaalisti jakautunut. Keskiarvovertailuiden perusteella nähtiin, että luottamusvälien keskimääräiset normalisoidut vinoudet olivat nousukausien jälkeen pienempiä kuin laskukausien jälkeen. Nousu- ja laskukausien jälkeen luottamusvälit olivat molempien parien kohdalla vinoja vasemmalle (vinouden parametrien arvot olivat negatiivisia). Hypoteesia testattiin parametrittomalla WSR-testillä. Testin tulosten mukaan H_{30} hylättiin molempien parien osalta merkitsevyystasolla $p < 0,001$. Näin ollen tulokset eivät muuttuneet perusaineistoilla tehdyistä tuloksista.

5.4.3 Regressioanalyysi luottamusvälien vinoudelle

Tutkimuksen viimeiset analyysit tehtiin hypoteesiin H_4 liittyen. Molempien aineistojen kohdalla regressioanalyysijä edelsivät korrelaatiotarkastelut (Spearman), joiden tulokset olivat samoja kuin luvussa 5.1.4 on esitetty. Näin ollen molemmilla aineistolla tehtiin neljä regressiota: ensin regressoitiin muuttujaa $\Delta_{\text{bull_norm}}$ erikseen muuttujien $\text{EPC}_{\text{bull_norm}}$ ja $\text{TF}_{\text{bull_CO}}$ suhteen (mallit

2. ja 3.) ja sen jälkeen muuttujaa $\Delta_{bear_{norm}}$ erikseen muuttujan $EPC_{bear_{norm}}$ ja $TF_{bear_COstrong}$ suhteen (mallit 3. ja 4.).

Opiskelijoiden täydellisellä ja normalisoidulla aineistolla ajettujen regressioanalyysien tulokset olivat hyvin samankaltaiset kuin perusaineistolla ja normalisoidulla perusaineistolla saadut tulokset. Selitysasteet, regressiokertoimet, F- ja t-testisuureiden arvot, merkitsevyystasot ja D-W testin arvot olivat samansuuntaiset kaikkien mallien kohdalla. H_4 sai tukea mallin 3. tapauksessa. Tässä ei enää esitetä tuloksia tarkemmin, eikä tutkita muita regressioanalyysin oletuksia, koska tulkinnat ovat samansuuntaisia kuin aikaisemmin tehtyjen analyysien kohdalla. Myöskään malleihin vaikuttaneita tekijöitä ei tässä enää tutkita tarkemmin.

Yhdistetyllä aineistolla suoritettujen regressioanalyysien perusteella saatiin selville, että kaikkien mallien selitysasteet, regressiokertoimet, F- ja t-testisuureiden arvot, merkitsevyystasot ja D-W testin arvot olivat samansuuntaiset kuin luvuissa 5.1.4 ja 5.2.3 esitetyt. H_4 sai tukea mallin 3. tapauksessa. Tässä ei enää esitetä tuloksia tarkemmin, eikä tutkita muita regressioanalyysin oletuksia, koska tulkinnat ovat samansuuntaisia kuin luvussa 5.1.4 esitetyt. Malleihin vaikuttaneita tekijöitä ei tässäkään enää tutkita tarkemmin.

Johtopäätöksenä voidaan todeta, että puuttuvien tietojen poistaminen ja normalisointi ei siis vaikuttanut tuloksiin.

6 TUTKIMUKSEN LUOTETTAVUUDESTA

Tutkimuksen luotettavuutta arvioidaan usein validiteetin, eli pätevyyden ja reliabiliteetin, eli pysyvyyden käsitteiden avulla. Validiteetti liittyy kysymyksiin, kuinka hyvin hankittu aineisto pystyy kuvaamaan sitä, mitä sen on tarkoituskin kuvata, tai kuinka täsmällisesti tutkimus pyrkii ratkaisemaan juuri ne ongelmat, joita tutkimuksessa asetettiin. Reliabiliteetti puolestaan liittyy kysymykseen, onko samanlainen aineisto saatavissa toistetussa tutkimuksessakin. Kun tutkimuksen reliabiliteetti on hyvä, tulokset eivät johdu satunnaisista tekijöistä. (Erätuuli, Leino & Yli-Luoma 1996, 19.) Tässä luvussa arvioidaan tämän tutkimuksen luotettavuutta muun muassa aineiston riittävyyden, tutkimusmenetelmien, mittaamisen onnistumisen, aineiston analyysitapojen, puuttuvan tiedon käsittelyn ja tulosten yleistettävyyden kautta.

Luotettavuuden näkökulmasta tässä tutkimuksessa ongelmallista on Osuuspankin sijoitusasiantuntijoilta saatujen vastausten vähyys. Vastauksia saatiin vain 37, vaikka kyselylinkki oli Webropol-ohjelman mukaan avattu 88 kertaa. Ilmeisesti kysely oli liian pitkä tai vaikea, koska niin monelta oli jäänyt vastaamatta. On myös mahdollista, että verkkokysely jäi Keski-Suomen Osuuspankissa kyselyn teettämisen aikoihin meneillään olleiden YT-neuvotteluiden jalkoihin, mikä saattoi alentaa vastausprosenttia. Joka tapauksessa pankin vastaajien vähyys rajoitti joissain tapauksissa analyysin tekoa.

Opiskelijavastausten määrää voi pitää opinnäytetyöhön sopivana. Opiskelijavastausten osalta ongelmaksi muodostui suoritettujen rahoituksen kurssien määrän vähyys. Tutkimuksen luotettavuuden kannalta olisi ollut parempi, jos useampia rahoituksen syventäviä kursseja suorittaneita opiskelijoita olisi saatu mukaan tutkimukseen.

Tutkimusmenetelmäksi valittu kyselytutkimus soveltuu hyvin osakeennusteiden tutkimista varten. De Bondt oli toteuttanut vastaavanlaisen tutkimuksen koeasetelman muodossa, jossa vastaajat saapuivat paikalle osallistukseen sijoittamiseen liittyvään peliin. Kummassakin tapauksessa törmätään kuitenkin samaan ongelmaan: aineiston hankkiminen on haasteellista ja vastausprosentit tahtovat jäädä pieniksi. Kenelläkään ei tunnu olevan aikaa vastata kyselyihin, eikä kokeellinen tutkimus houkuttele suuria vastaajakandidaattimääriä. Tässä mielessä kyselyn toteuttaminen luentojen yhteydessä on hyvä vaihtoehto, vaikka se saattaaakin kirvoittaa vastalauseita²¹. Verkkokyselyn etuna on suurempien vastaajajoukkojen tavoittaminen, mutta ongelmana on tällöinkin se, kuinka saada motivoitua kyselyyn vastaamiseen. Tässä tutkimuksessa käytettiin vastausmotivaation lisäajina kirjapalkintoja ja lahjakortteja, mutta on vaikea arvioida, kuinka paljon ne vaikuttivat vastaamishalukkuuteen. Testikyselyyn

²¹ Eräs opiskelijavastaaja kommentoi kyselylomakkeen reunaan tuskastuneen oloisena: "Epäreilua, kun aikaa menee tällaiseen!"

vastanneilta saadun palautteen perusteella sillä saattoi kuitenkin olla huomattava positiivinen vaikutus.

Siirrettäessä vastauksia SPSS-ohjelmaan havaittiin, että aika moni vastaaja oli ymmärtänyt luottamusvälit väärin. Tämä korostaa tutkijan vastuuta täsmällisten ohjeiden antamisessa ennen kyselyn alkua. Tässä tutkimuksessa vastausohjeisiin yritettiin panostaa siten, että tutkimuksen toteuttaja oli itse paikan päällä ohjeistamassa kyselyä ja että kyselylomakkeeseen oli vielä laitettu tarkat ohjeet vastaamiseen. Ilmeisesti tämä ei ollut kaikilta osin riittävää ohjeistusta. Voidaan myös pohtia, oliko luentojen alku sopiva hetki pitää kysely. Luennon alkuun liittyy lähes aina eriasteista liikehdintää. Osa luennoitsijoista oli informoinut opiskelijoita etukäteen kyselyn teettämisestä, osa taas ei. Jälkeenpäin tarkasteltuna voidaan todeta, että kyselyn teettäminen sujui paremmin niissä tapauksissa, joissa opiskelijoilla oli etukäteen tieto kyselyn teettämisestä, vaikka he eivät kyselyn tarkempaa sisältöä tienneetkään. Ennakkoon tiedon saaneet opiskelijat olivat keskittyneempiä kuin muut.

Miten hyvin kerätyt tiedot sitten vastasivat niitä asioita, joita oli tarkoitus tutkia? Entä miten mittaaminen onnistui? Kaiken kaikkiaan voidaan todeta, että kerätyt tiedot vastasivat suurelta osin sitä, mitä oli tarkoituskin tutkia. Poikkeuksia tästä olivat laumakäyttäytymiseen, talousuutisten seuraamiseen ja pääasialliseen tiedonlähteeseen liittyvät kysymykset pankin kyselyssä sekä talousuutisten seuraamiseen ja sijoituskohteisiin liittyvät kysymykset opiskelijakyselyssä. Niistä saadut vastaukset jäivät analysoimatta tarkemmin, koska kysymyksien ja tutkimusongelmien välinen yhteys jäi puutteellisesti määritellyksi. Lisäksi tutkimusta suunniteltaessa oli nimenomaan tarkoitus saada mukaan myös sellaisia opiskelijoita, joilla oli useita rahoituksen syventäviä kursseja suoritettuina, mutta tässä ei onnistuttu.

Aineisto päätettiin käsitellä SPSS-ohjelmalla, koska sitä käytetään yleisesti sosiaali- ja ihmistieteissä. Pitää muistaa, että tämä tutkimus liittyy rahoituksen käyttäytymistieteeseen, joka linkittää psykologian ilmiöt rahoituksen teoriaan. Aineiston analyysitavat määräytyivät kyselylomakkeiden kysymysten perusteella, ja parametristen sekä parametrittömien testien tekeminen rinnakkain vahvasti tulosten luotettavuutta, sillä useissa tapauksissa testit antoivat samansuuntaisen tuloksen. Puuttuvan tiedon käsittelyyn käytetty Case Deletion -menetelmä ei ollut paras mahdollinen, mutta hyväksyttävissä, koska puuttuvaa tietoa oli niin vähän ja se oli luonteeltaan MCAR.

Lopuksi on paikallaan vielä korostaa, että tulokset eivät ole yleistettävissä koskemaan perusjoukkoa, sillä aineistojen hankinnassa ei käytetty otantaasetelmaa. Tehtävät johtopäätökset rajoittuvat näin ollen vain kyselyyn vastanneisiin.

7 JOHTOPÄÄTÖKSET

Tässä opinnäytetyössä tutkittiin Jyväskylän yliopiston kauppakorkeakoulun järjestämille kursseille osallistuneiden opiskelijoiden ja Osuuspankin sijoitusasiantuntijoiden (lähinnä sijoitusneuvojat, -päälliköt ja -johtajat) käsityksiä tulevista osaketuotoista ja riskeistä. Tutkimuksen aineistot kerättiin kyselyjen avulla. Opiskelijavastaukset kerättiin syyslukukaudella 2013 muutamien luentojen alussa, ja pankin vastaukset sähköisesti Webropol-kyselyn avulla sama vuoden marras-joulukuussa.

Tämän opinnäytetyön aihe ideoitiin Werner De Bondtin (De Bondt 1993) tutkimusten pohjalta. Kyselyssä vastaajia pyydettiin laatimaan seitsemän ja 13 kuukauden hintaennusteet ja 80 %:n luottamusvälit osakekuvaajien perusteella. Osakekuvaajat esittivät S&P 500 indeksin kuvaajia kuudesta eri aikajaksosta ja ne oli skaalattu siten, ettei käytettyä indeksiä pystynyt identifioimaan. Osakeindeksin nimeä ei kerrottu vastaajille etukäteen. Työssä haluttiin selvittää, vastasivatko ennusteet rationaalisen sijoittajan odotuksia. Erityisesti tutkittiin, oliko opiskelijoilla ja pankin sijoitusasiantuntijoilla taipumusta ekstrapoloida mennyttä hintakehitystä, eli olivatko he optimistisia nousukausien jälkeen ja pessimistisiä laskukausien jälkeen sekä olivatko vastaajien subjektiiviset todennäköisyysjakaumat vinoja vasemmalle nousukausien jälkeen ja oikealle laskukausien jälkeen. Lisäksi selvitettiin taustatekijöiden, kuten iän, sukupuolen, opiskeltujen rahoituksen kurssien, työkokemuksen tai sijoituskokemuksen vaikutusta vastaajien trendinseuraamiskäyttäytymiseen ja ennustevirheisiin. Haluttiin myös selvittää, oliko vanhempien ja parhaimpien ystävien käyttäytymisellä vaikutusta vastaajien trendinseuraamiskäyttäytymiseen ja ennustevirheisiin.

Aineisto analysointiin IBM SPSS 20-ohjelmalla, ja analyysimenetelminä käytettiin t-testejä ja niiden parametrittomia vastineita, varianssianalyysijä, χ^2 -testiä sekä lineaarista ja logistista regressioanalyysiä. Analyyseissä edettiin aineistopareittain. Tutkimuksen tärkeimmät aineistot olivat opiskelijoiden ja pankin perusaineistot, joista odotetun hintamuutoksen ja luottamusvälien vinouden parametreja ei oltu normalisoitu, eikä puuttuvia osake-ennusteita oltu poistettu. Herkkyysanalyysijä tehtiin opiskelijoiden täydelliselle aineistolle, ja normalisoimalla perusaineistoista odotetun hintamuutoksen ja vinouden parametrit. Seuraavaksi raportoidaan pääasiassa vain perusaineistoilla saadut tulokset, sillä tulokset eivät merkittävästi muuttuneet herkkyysanalyysien myötä.

Tärkeimmät tutkimustulokset osoittivat De Bondtin luoman teorian luottamusvälien suojauksesta pitävän paikkaansa myös tämän tutkimuksen aineistojen perusteella. Vastaajat ennustivat menneen trendin jatkumista, eli keskimääräinen odotettu hintamuutos nousukausien jälkeen oli suurempi kuin laskukausien jälkeen. Näin ollen vastaajat olivat optimistisia nousukausien jälkeen

ja pessimistisiä laskukausien jälkeen. Tulokset olivat tilastollisesti erittäin merkitseviä. Ainoa poikkeus tästä oli pankin normalisoidulla perusaineistolla saatu tulos koskien seitsemän kuukauden keskimääräisiä odotettuja hintamuutoksia.

Vastaajat suojasivat ennusteitaan. Kaikilla aineistoilla mitattuna luottamusvälien keskimääräinen vinous oli pienempi nousukausien jälkeen kuin laskukausien jälkeen. Tulokset olivat tilastollisesti erittäin merkitseviä, pois lukien pankin aineistoilla saadut tulokset, jotka olivat tilastollisesti melkein merkitseviä. Opiskelija-aineistoilla saadut tulokset vastasivat vahvasti De Bondtin teoriaa siinä mielessä, että nousukausien jälkeen luottamusvälien vinouden parametrien arvot olivat negatiivisia, eli luottamusvälit olivat vinoja vasemmalle, ja laskukausien jälkeen positiivisia, eli luottamusvälit olivat vinoja oikealle. Näytetyn osakekuvaajan keskimääräinen hintataso veti ylä- ja alaluottamusväläjä ylöspäin laskukausien jälkeen ja alaspäin nousukausien jälkeen. Tämä tarkoitti nousukausien tapauksessa, että varauduttiin mahdolliseen hinnanpudotukseen. Myös muilla kuin pelkästään opiskelija-aineistoilla mitattuna De Bondtin luoma teoria toteutui, mutta ei aivan yhtä vahvasti: vinouden parametrien arvot olivat negatiivisia sekä nousu- että laskukausien jälkeen, mutta nousukausien vinous oli keskimäärin pienempää.

De Bondtin teoriaan liittyen osoitettiin myös lineaaristen regressioanalyysien perusteella, että opiskelija-aineistoilla ja yhdistetyillä aineistoilla luottamusvälit kasvoivat enemmän oikealle vinoiksi, kun suurempaa hinnanlaskua ennustettiin. Tulokset olivat tilastollisesti erittäin merkitseviä. Nousukausien tapauksessa myös muuttujan EPC_bull regressiokerroin oli hypoteesin H_4 mukainen, mutta ei ollut tilastollisesti merkitsevä.

Koska suurin osa vastaajista ennusti menneen trendin jatkuvan, se kertoo myös ylireagoinnista. Vaikka nousukausien jälkeen reagoitiinkin mahdolliseen hintojen laskuun luottamusvälien kautta, tulokset olivat yleisellä tasolla ristiriidassa sen ajatuksen kanssa, että osakekurssit ennemmin tai myöhemmin kääntyvät laskuun.

Tulokset osoittivat, etteivät vastaajat käyttäytyneet rationaalisesti. Odotukset sekä ennustettujen hintamuutosten että luottamusvälien vinouden perusteella mitattuina osoittautuivat pikemminkin adaptiiviksi, kuten Muradoğlu on esittänyt (Muradoğlu 2002).

Tässä opinnäytetyössä tutkittiin myös sitä, kumpi vastaajaryhmistä oli enemmän konvergoitunut kohti osakekuvaajien keskiarvoa. Parametrisellä ja parametrittomalla testillä saadut tilastollisesti merkitsevät tulokset menivät osittain ristiin: t-testin perusteella suurimmassa osassa ennusteista pankin vastaajat olivat enemmän konvergoituneita osaketuottojen keskiarvoja kohti, Mann-Whitneyn U-testin (MWU-testi) mukaan taas opiskelijat.

Toisinajattelijastrategian noudattaminen ei ollut kovin yleistä. Vain osakekuvaajien B ja C tapauksissa voitiin sanoa, että pankin vastaajat noudattivat suhteellisesti enemmän toisinajattelijastrategiaa ja kyseisten kuvaajien kohdalla erot olivat tilastollisesti merkitseviä. Tulos oli vain näiltä osin samansuuntainen kuin De Bondtin saamat tulokset. De Bondtin tutkimukset (1993) osoittivat, että osa ammattisijoittajista ennusti markkinoiden täyskäännöstä.

χ^2 -testillä yritettiin lisäksi selvittää, oliko *vahvalla* trendinseuraamis- tai toisinajattelijakäyttäytymisellä eroa opiskelijoiden ja pankin sijoitusasiantuntijoiden välillä. Testin edellytykset eivät kuitenkaan täyttyneet, joten tehtiin Fisherin tarkka testi. Testin tulokset eivät kuitenkaan olleet tilastollisesti merkitseviä.

Yliluottamusta tutkittiin luottamusvälien kapeuden ja ennustettujen osaketuottojen varovaisuuden perusteella. Luottamusvälien kapeuden perusteella ei voida kaikkien osake-ennusteiden kohdalla sanoa, että miehet ja pankin sijoitusasiantuntijat olisivat yliluottavaisempia, vaikka aikaisemmassa tutkimuksessa on paljon tämänsuuntaisia tuloksia raportoitukin. Ennusteiden varovaisuuteen perustuen ei myöskään voida sanoa yksiselitteisesti, että naiset ennustivat tulevia tuottoja varovaisemmin, tai että opiskelijat ennustivat tulevia tuottoja varovaisemmin kuin pankin sijoitusasiantuntijat.

Vastaajat ennustivat menneiden trendien jatkuvan, mikä on vastoin perinteistä näkemystä rationaalisesta päätöksenteosta riskitilanteissa. Trendinseuraamista tutkittiin tässä työssä virheenä, ja haluttiin tarkemmin selvittää, oliko virheellisessä käyttäytymisessä eroja tiettyjen taustamuuttujien suhteen. Tulosten perusteella voidaan todeta, ettei trendinseuraamiskäyttäytymisessä ollut eroja sen suhteen, kuinka paljon rahoituksen kurseja oli käytynä, kuinka paljon kaikkia opintoja oli yhteensä suoritettuina, tai millaisia arvosanoja rahoituksen kurseista oli saatu.

Lisäksi huomattiin, että suoritettujen rahoituksen kurssien määrällä ei ollut tilastollisesti merkitsevää vaikutusta ennustevirheeseen, jos merkitsevyystasona pidettiin viittä prosenttia. Kuitenkin kymmenen prosentin merkitsevyystasolla mitattuna vastaajaryhmien välillä oli eroa sen suhteen, kuinka paljon rahoituksen kurseja oli suoritettuna, mikä tarkoitti, että enemmän kurseja suorittaneiden ennustevirhe oli pienempi. Tämä antaa heikkoa tukea esimerkiksi Alexanderin ja McElreathin (1999) tutkimukselle. Tässä tutkimuksessa saatiin kuitenkin selville, että kaikkien suoritettujen opintopisteiden määrällä ei ollut vaikutusta ennustevirheeseen.

Aikaisempien tutkimusten mukaan sijoituskokemus saattaisi vähentää sijoittajien tekemiä virheitä. Tässä tutkimuksessa saatiin tulos, jonka mukaan opiskelijoiden sijoitusharrastus ei vaikuttanut trendinseuraamiskäyttäytymiseen tai ennustevirheeseen tilastollisesti merkitsevästi. Pankin perusaineiston analyysi osoitti puolestaan, ettei työkokemuksen pituudella ollut tilastollisesti merkitsevää vaikutusta trendinseuraamiskäyttäytymiseen. Lisäksi havaittiin, ettei ennustevirhe korreloinut rahoitusalan työkokemuksen pituuden kanssa.

Aikaisemmat tutkimukset ovat lisäksi tuoneet esille, että rahoitusalan sofistikoituneisuus ja käyttäytymismallit esimerkiksi sijoittamiseen liittyen voisivat välittyä vanhemmilta lapsille. Tässä suhteessa myös ystävien merkitys voi olla suuri. Tässä tutkimuksessa haluttiin tähän liittyen selvittää, oliko opiskelijoiden trendinseuraamiskäyttäytymisessä eroja sen suhteen, harrastivatko heidän vanhempansa tai parhaimmat ystävänsä sijoittamista, tai mikä oli vanhempien sosioekonominen asema. Näitä tutkittiin χ^2 - ja Fisherin tarkkojen testien avulla. Laskukausi tulosten perusteella niiden vastaajien, joiden vanhemmat

harrastivat sijoittamista, trendinseuraaminen oli oletuksista poiketen vahvempaa kuin niiden vastaajien, joiden vanhemmat eivät harrastaneet sijoittamista. Tulos oli tilastollisesti melkein merkitsevä. Nousukausien osalta tilastollisesti merkitseviä tuloksia ei saatu. Parhaimpien ystävien sijoitusharrastuksella ei puolestaan ollut tilastollisesti merkitsevää vaikutusta vastaajien trendinseuraamiskäyttäytymiseen.

Edelliseen liittyen tutkittiin riippumattomien ryhmien t-testillä lähipiirin, eli vanhempien ja parhaimpien ystävien, sijoitusharrastuksen vaikutusta vastaajien ennustevirheisiin. Tuloksista huomattiin yllättävästi, että niiden opiskelijoiden, joiden vanhemmat eivät harrastaneet sijoittamista, ennustevirhe oli pienempi kuin niiden, joiden vanhemmat harrastivat sijoittamista. Tulokset olivat tilastollisesti melkein merkitseviä. Parhaimpien ystävien sijoitusharrastuksella ei ollut tilastollisesti merkitsevää vaikutusta vastaajan ennustevirheeseen t-testin perustella. Parametrittoman MWU-testin tulokset osoittivat kuitenkin, että niiden opiskelijoiden, joiden parhaimmat ystävät harrastivat sijoittamista, ennustevirheet olivat keskimäärin pienempiä kuin niiden, joiden parhaimmat ystävät eivät harrastaneet sijoittamista. Tulos oli tilastollisesti melkein merkitsevä ja antoi heikkoa tukea teorialle. Vanhempien sosioekonomisella asemalla ei puolestaan ollut vaikutusta vastaajien ennustevirheiden keskiarvoihin.

Lisäksi havaittiin, ettei iällä ollut tilastollisesti merkitsevää vaikutusta trendinseuraamiskäyttäytymiseen, eikä ennustevirheeseen.

Tässä opinnäytetyössä selvitettiin vielä logistisen regressioanalyysin avulla, mitkä tekijät selittivät opiskelijoiden sijoitusharrastusta. Tulosten perusteella voidaan sanoa, että sukupuolella, opiskelujen rahoituksen kurssien määrällä ja parhaimpien ystävien sijoittamisharrastuksella oli tilastollisesti melkein merkitsevää tai merkitsevää vaikutusta regressiomallin ennustekykyyneen ja että näiden muuttujien avulla rakennetut mallit olivat tilastollisesti erittäin merkitseviä.

Herkkyysanalyysien tulokset osoittivat, että puuttuvien tietojen poistaminen opiskelija-aineistosta vaikutti vain mitättömän vähän tuloksiin. Näin tuli todistettua se seikka, että jos puuttuvia tietoja on vähän ja ne ovat luonteeltaan MCAR, voivat tulokset olla lähes harhattomia ilman, että puuttuville tiedoille tehdään mitään. Odotettujen hintamuutosten ja luottamusvälien vinouksien parametrien normalisointi vaikutti niin ikään mitättömän vähän tuloksiin.

Tätä tutkimusta voidaan kritisoida sillä perusteella, ettei vastaajille tarjottu muuta taustainformaatioita ennusteiden tekemiseen kuin pelkät kuvaajat menneistä hintakehityksistä neljän vuoden alalta. Vastaajilla ei myöskään ollut mahdollisuutta oppia tai saada palautetta ennusteistaan. Jos uskomme kuitenkin, että markkinat ovat täydellisen tehokkaat, voimme kysyä, mitä muuta tietoa osake-ennusteiden tekemiseen olisi tarvittu – eihän sillä olisi ollut merkitystä.

Kuten edellisessä luvussa tuotiin esille, tuloksia ei voida yleistää kokemaan koko perusjoukkoa. Ei voida olettaa, että tähän tutkimukseen osallistuneet edustaisivat tyypillisiä yksityissijoittajia tai sijoitusammattilaisia. Jotta saataisiin paremmin selville, kuinka relevantteja tulokset ovat suhteessa tyypillisiin sijoittajiin tai treidaajiin, salkunhoitajiin tai analyytikkoihin (joita useim-

missa tutkimuksissa pidetään *varsinaisina* sijoitusammattilaisina), tarvitaan lisää tutkimusta, jossa käytetään satunnaisotantaa. Tässä opinnäytetyössä pankin aineisto jäi liian pieneksi, eikä yhden pankin asiantuntijoiden vastausten perusteella voi kovin syvälle meneviä johtopäätöksiä tehdä.

Tämän tutkimuksen tuloksia voidaan kuitenkin peilata perinteisiä rahoitusteorioita vasten. Tulokset osoittavat omalta osaltaan, etteivät yksilöt aina toimi rationaalisesti, kun valinnat sisältävät riskejä. Epärationalisesta käyttäytymisestä seuraa se, etteivät markkinat ole täydellisen tehokkaat.

LÄHTEET

Kirjallisuus

- Agarwal, S., Driscoll, J., Gabaix, X. & Laibson, D. 2009. The age of reason: financial decision over the lifecycle and implications for regulation. *Brookings Papers on Economic Activity* 2, 51-117.
- Alexander, J. C. & McElreath, R. B. 1999. Does education affect how well students forecast the market? *Financial Services Review* 8, 253-260.
- Allais, M. 1953. Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica*, Vol. 21, No 4, 503-546.
- Baker, H. K. & Nofsinger, J. R. 2002. Psychological biases of investors. *Financial Services Review* 11, 97-116.
- Bange, M. M. 2000. Do the portfolios of small investors reflect positive feedback trading? *Journal of Financial and Quantitative Analysis* 35, 239-255.
- Barber, B. M. & Odean, T. 1999. The courage of misguided convictions. *Financial Analysts Journal*, Vol 55, 41-55.
- Barber, B. M. & Odean, T. 2001. Boys will be boys: gender, overconfidence and common stock investment. *Quarterly Journal of Economics*, Vol. 116, Issue 1, 261-292.
- Barberis, N. & Thaler, R. 2002. *A survey of behavioral finance*. NBER working paper series. (Viitattu 16.9.2013) Saatavana: <http://www.nber.org/papers/w9222>.
- Bernheim, B. D., Skinner, J. & Weinberg, S. 2001. What accounts for the variation in retirement wealth among US households? *American Economic Review* 91, 832-857.
- Breen, R., Luijkx, R., Müller, W. & Pollak, R. 2005. *Non-persistent inequality in educational attainment: evidence from eight European countries*. (Viitattu 23.10.2013) Saatavana: <http://www.ccpr.ucla.edu/publications/conference-proceedings/CP-05-001.pdf>.
- Brown, J., Ivkovic, Z., Smith, P. & Weisbenner, S. 2008. Neighbors matter: causal community effects and stock market participation. *Journal of Finance* 63, 1509-1531.
- Calvet, L. E., Campbell, J. Y. & Sodini, P. (2009) Measuring the financial sophistication of households. *American Economic Review: Papers & Proceedings* 99:2, 393-398.
- Camerer, C. F. & Hogarth, R. M. 1999. The effects on financial incentives in experiments: a review and capital-labor-production framework. *Journal of Risk and Uncertainty*, 19: 1-3, 7-42.

- Campbell, S. D. & Sharpe, S. A. 2009. Anchoring bias in consensus forecasts and its effect on market prices. *Journal of Financial and Quantitative Analysis*, Vol. 44, No. 2, 369-390.
- Cen, L., Hilary, G. & Wei, J. K. C. 2013. The role of anchoring bias in the equity market: evidence from analysts' earnings forecasts and stock returns. *Journal of Financial and Quantitative Analysis*, Vol. 48, No. 1, 47-76.
- Chen, G., Kim, K. A., Nofsinger, J. R. & Rui, O. M. 2007. Trading performance, disposition effect, overconfidence, representativeness bias and experience of emerging market investors. *Journal of Behavioral Decision Making* 20, 425-451.
- Chiteji, N. & Stafford, F. 1999. Portfolio choices of parents and their children as young adults: asset accumulation by African American families. *American Economic Review: Papers and Proceedings* 89, 377-380.
- Christiansen, C., Joensen, J. S. & Rangvid, J. (2008) Are economists more likely to hold stocks? *Review of Finance* 12, 465-496.
- Clement, M. B. 1999. Analyst forecast accuracy: Do ability, resources and portfolio complexity matter? *Journal of Accounting and Economics* 27, 285-303.
- Copeland, T. E., Weston, J. F. & Shastri, K. 2005. *Financial theory and corporate policy*. 4. painos. USA: Pearson Addison Wesley.
- De Bondt, W. F. M. 1993. Betting on trends: intuitive forecasts of financial risk and return. *International Journal of Forecasting* 9, 355-371.
- De Bondt, W. F. M., Muradoğlu, G., Shefrin, H. & Staikouras, S. K. 2008. Behavioral finance: quo vadis? *Journal of Applied Finance*, Vol. 18, Issue 2, 7-21.
- De Long, J. B., Shleifer, A., Summers, L. H. & Waldmann, R. J. 1990. Positive feedback investment strategies and destabilizing rational speculation. *Journal of Finance* 45, 379-395.
- Dhar, R. & Kumar, A. 2001. *A non-random walk down the main street: impact of price trends on trading decisions of individual investors*. Yale International Center for Finance, Yale School of Management. (Viitattu 23.9.2013) Saatavana: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.198.5736&rep=rep1&type=pdf>.
- Dorn, D. & Huberman, G. 2002. *Who trades?* EFA 2003 Annual Conference Paper, No. 645. (Viitattu 8.9.2013) Saatavana: <http://ssrn.com/abstract=423995>.
- Erätuuli, M., Leino, J. & Yli-luoma, P. 1996. *Kvantitatiiviset analyysimenetelmät ihmistieteissä*. Rauma: Kirjayhtymä.
- Fama, E. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance* 25, 383-417.
- Fama, E. 1991. Efficient capital markets II. *Journal of Finance* 46, 1575-1617.
- Felton, J., Gibson, B. & Sanbonmatsu D. M. 2003. Preference for risk in investing as a function of trait optimism and gender. *Journal of Behavioral Finance* 4, 33-40.

- Feng, L. & Seasholes, M. S. (2005) Do investor sophistication and trading experience eliminate behavioral biases in financial markets? *Review of Finance* 9, 305-351.
- Fujiwara, I. & Ichiue, H. & Nakazono, Y. & Shigemi, Y. 2013. Financial markets forecasts revisited: Are they rational, stubborn or jumpy? *Economics Letters* 118, 526-530.
- Glaser, M., Langer, T. & Weber, M. 2005. *Overconfidence of professionals and lay men: individual differences within and between tasks*. (Viitattu 10.9.2013) Saatavana: <http://www.wiwi.unifrankfurt.de/profs/schlag/dgf2009/Contribution168.pdf>.
- Graham, J. W. 2012. *Missing data: analysis and design. Statistics for social and behavioral sciences*. New York: Springer.
- Hibbert, A. M., Lawrence, E. R. & Prakash, A. J. 2013. Does knowledge of finance mitigate the gender difference in financial risk-aversion? *Global Finance Journal*. (Viitattu 4.9.2013) Saatavana: <http://dx.doi.org/10.1016/j.gfj.2013.07.002>.
- Hirsjärvi, S., Remes, P. & Sajavaara, P. 1997. *Tutki ja kirjoita*. 1.-2. painos. Helsinki: Kirjayhtymä.
- Hommes, C. 2005. *Heterogeneous agent models in economics and finance*. Tinbergen Institute Discussion Paper, TI 2005-056/1, CeNDEF, Department of Quantitative Economics, University of Amsterdam. (Viitattu 24.9.2013) Saatavana: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=742384.
- Hong, H., Kubik, J. & Stein, J. (2004) Social interaction and stock-market participation. *Journal of Finance* 59, 137-163.
- Jianakoplos, N. A. & Bernasek, A. 1998. Are women more risk averse? *Economic Inquiry*, Vol. 36, Issue 4, 620-630.
- Joo, S. & Grable, J. E. (2005) Employee education and the likelihood of having a retirement savings program. *Financial Counselling and Planning* 16, 37-49.
- Kahneman, D. & Tversky, A. 1974. Judgment under uncertainty: heuristics and biases. *Science*, New Series, Vol. 185, No. 4157, 1124-1131.
- Kahneman, D. & Tversky, A. 1979. Prospect theory: an analysis of decision under risk. *Econometrica* 47 (2), 263-291.
- Kahneman, D. & Tversky, A. 1992. Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297-323.
- Kaustia, M. 2003. *Essays on investor behavior and psychological reference prices*. Helsinki School of Economics.
- Kaustia, M., Alho, E. & Puttonen, V. 2008. How much does expertise reduce behavioral biases? The case of anchoring effects in stock returns estimates. *Financial Management*, 391-411.
- Kaustia, M., Lehtoranta, A. & Puttonen, V. 2013. *Does sophistication affect long-term return expectations? Evidence from financial advisers' exam scores*. SAFE Working Paper Series No. 3. (Viitattu 10.9.2013) Saatavana: <http://ssrn.com/abstract=2209649>.

- Kendall, M. G. & Bradford Hill, A. 1953. The analysis of economic time series, part 1: prices. *Journal of the Royal Statistical Society, Series A (General)*, Vol. 116, Issue 1, 11-34.
- Kivinen, O., Hedman, J. & Kaipainen, P. 2012. Koulutusmahdollisuuksien yhdenvertaisuus Suomessa. *Yhteiskuntapolitiikka* 77 (5), 559-566.
- Lakonishok, J., Shleifer, A. & Vishny, R. W. 1994. Contrarian investment, extrapolation and risk. *The Journal of Finance*, Vol. 49, No. 5, 1541-1578.
- Li, G. 2009. *Information sharing and stock market participation: evidence from extended families*. Federal Reserve Board. (Viitattu 20.9.2013) Saatavana: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1383072.
- Little, R. J. A. & Rubin, D. B. 2002. *Statistical analysis with missing data*. Second edition. New Jersey: John Wiley & Sons, Inc.
- Lusardi, A. & Mitchell, O. S. 2006. *Financial literacy and planning: implications for retirement wellbeing*. Working paper, Pension Research Council, Wharton School, University of Pennsylvania. (Viitattu 18.9.2013) Saatavana: http://www.dartmouth.edu/~alusardi/Papers/Financial_Literacy.pdf.
- Lusardi, A. (toim.) 2008. *Overcoming the saving slump: how to increase the effectiveness of financial education and saving program*. Chicago: University Press.
- Lusardi, A., Mitchell, O. S. & Curto, V. 2010. Financial literacy among the young. *The Journal of Consumer Affairs*, Vol. 44, No. 2, 358-380.
- Mandell, L. 2008. Financial literacy in high school. Teoksessa Lusardi, A. (toim.) 2008. *Overcoming the saving slump: how to increase the effectiveness of financial education and saving program*. Chicago: University Press.
- Mata, R., Josef, A., K., Samanez-Larkin, G., R. & Hertwig, R. 2011. Age differences in risky choice: a meta-analysis. *Annals of the New York Academy of Sciences* 1235, 18-29. (Viitattu 10.9.2013) Saatavana: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3332530/>
- Muradoğlu, G. 2002. Portfolio managers' and novices' forecasts of risk and return: are there predictable forecasts errors? *Journal of Forecasting* 21, 395-416.
- Nakazono, Y. 2012. Heterogeneity and anchoring in financial markets. *Applied Financial Economics* 22, 1821-1826.
- Nguyen, T. & Schüßler, A. 2012. Investment and socio-demographic characteristics - empirical evidence from Germany. *International Journal of Economics and Finance*, Vol. 4, No. 9, 1-12.
- Odean, T. 1998. Volume, volatility, price and profit when all traders are above average. *Journal of Finance* 53, 1887-1934.
- Ogasawara, E., Martinez, L., C., de Oliveira, D., Zimbrão, G., Pappa, G., L. & Mattoso, M. 2010. *Adaptive normalization: a novel data normalization approach for non-stationary time series*. (Viitattu 20.1.2014) Saatavana: <http://homepages.dcc.ufmg.br/~glpappa/papers/Ogasawaraetal-2010-IJCNN.pdf>.

- Powell, M. & Ansic, D. 1997. Gender differences in risk behaviour in financial decision-making: An experimental analysis. *Journal of Economic Psychology* 18, 605-628.
- Riley, W. B. & Chow, K. V. (1992) Asset allocation and individual risk aversion. *Financial Analysts Journal* 48, 32-37.
- Rubin, D. B. 1996. Multiple Imputation After 18 + Years. *Journal of the American Statistical Association*, Vol. 91, No. 434, 473-489.
- Schafer, J., L. & Graham, J., W. 2002. Missing data: our view of the state of the art. *Psychological Methods*, Vol. 7, No. 2, 147-177.
- Schubert, R., Brown, M., Gysler, M. & Brachinger, H. W. 1999. Financial decision-making: are women really more risk-averse? *American Economic Review*, Vol. 89, Issue 2, 381-385.
- Shapira, Z. & Venezia, I. 2001. Patterns of behavior of professionally managed and independent investors. *Journal of Banking and finance*. Vol. 25, Issue 8, 1573-1587.
- Shavit, Y. & Blossfield, H-P. (toim.) 1993. *Persistent inequalities: a comparative study of educational attainment in thirteen countries*. Boulder Colorado: Westview Press.
- Shefrin, H. 2008. *A behavioral approach to asset pricing*. Second edition. Academic press advanced finance series. Elsevier.
- Shiller, R. J. 2003. From efficient market theory to behavioral finance. *Journal of Economic Perspectives*, Vol. 17, No. 1, 83-104.
- Shleifer, A. & Summers, L. H. 1990. The noise trader approach to finance. *Journal of Economic Perspectives*, Vol 4, No. 2, 19-33.
- Van de Venter, G. & Michayluk, D. 2008. An insight into overconfidence in the forecasting abilities of financial advisors. *Australian Journal of Management*, Vol. 32, No 3, 545-557.
- Van Rooij, M., Lusardi, A. & Alessie, R. 2011. Financial literacy and stock market participation. *Journal of Financial Economics* 101, 449-472.
- Vehkalahti, K. 2008. *Kyselytutkimuksen mittarit ja menetelmät*. Helsinki: Tammi.
- Vissing-Jørgensen, A. 2003. Perspectives on behavioral finance: does "irrationality" disappear with wealth? Evidence from expectations and actions. *NBER Macroeconomics Annual*, Vol. 8, 139-194. (Viitattu 20.9.2013) Saatavana: <http://www.nber.org/chapters/c11443.pdf>.
- Von Neumann, J. & Morgenstern, O. 1967. *Theory of games and economic behavior*. Third edition. New York: John Wiley & Sons.

Internet – lähteet

[Http://research.stlouisfed.org/fred2/series/SP500/downloaddata](http://research.stlouisfed.org/fred2/series/SP500/downloaddata). Aineisto osakekuvaajiin haettu 14.8.2013.

LIITTEET

Liite 1 Kyselylomake opiskelijoille

ARVOISA VASTAANOTTAJA

Teen pro gradu - tutkimusta Jyväskylän yliopiston kauppakorkeakoululle ja edessäsi on työhön liittyvä kyselylomake. Kyselyn tarkoituksena on selvittää opiskelijoiden käsityksiä tulevista osaketuotoista. Kyselyssä on kaksi osaa: ensimmäisessä osassa kartoitetaan perustietoja ja toisessa osassa sinua pyydetään tekemään ennusteita tulevista osaketuotoista.

Vastauksesi on arvokas opinnäytetyön onnistumisen kannalta. Kyselyn tuloksia käytetään vain pro gradu - työn lähdeaineistona, eikä niitä luovuteta kolmannelle osapuolelle. Vastaamiseen menee aikaa noin 20 minuuttia. Kyselyyn vastaaminen on vapaaehtoista.

Kaikkien kyselyyn vastanneiden ja sähköpostiosoitteensa antaneiden kesken arvotaan Paul Krugmanin kirjoittama kirja Lama. Tämän lisäksi tarkimmat ennusteet tehneelle vastaajalle luovutetaan 50 euron lahjakortti Intersportiin. Arvonnan tuloksista ilmoitetaan henkilökohtaisesti. Sähköpostiosoitteet jäävät vain tutkijan tietoon, eikä niitä julkaista tutkimuksessa. Jos haluat osallistua kirjan arvontaan ja tavoitella lahjakorttia, merkitse sähköpostiosoitteesi selvästi tähän:

_____.

Ystävällisin terveisin,
Anni Kinttu
anni.h.kinttu@student.jyu.fi

OSA 1: VASTAAJAN PERUSTIEDOT

1. Syntymävuotesi: _____

2. Sukupuolesi: (ympyröi oikea vaihtoehto):

nainen

mies

3. Pääaineesi: _____

4. Pohjakoulutuksesi (ympyröi sopiva vaihtoehto):

peruskoulu

ylioppilastutkinto

ammattillinen koulutus

ammattikorkeakoulututkinto

aikaisempi yliopistotutkinto

5. Tämän hetkinen opintopisteitteesi määrä (kaikki opinnot yhteensä):

n. _____ opintopistettä.

6. Tähän mennessä opiskelemiesi rahoituksen kurssien määrä opintopisteinä:

n. _____ opintopistettä.

7. Mitä rahoituksen kursseja olet suorittanut?

8. Kuinka hyvin olet tähän mennessä yleisesti ottaen suoriutunut rahoituksen opinnoista (kaikki rahoituksen opinnot mukaan luettuna, ympyröi sopiva vaihtoehto)?

tyydyttävästi (arvosanat keskimäärin 1-2)

hyvin (arvosanat keskimäärin 3)

erittäin hyvin (arvosanat keskimäärin 4-5)

en ole vielä suorittanut rahoituksen opintoja

9. Harrastatko itse sijoittamista (ympyröi sopiva vaihtoehto)? Jos vastauksesi on kielteinen, siirry kysymykseen numero 11.

kyllä

en

10. a) Jos vastasit edelliseen kysymykseen kyllä, kuinka paljon arvioit sijoitus-salkkusi tuottavan seuraavan 12 kuukauden aikana? _____ %

b) Kun ajatellaan Suomen osakemarkkinoita yleisesti, paljonko arvioit markkinatuoton olevan seuraavan 12 kuukauden aikana? _____ %

c) Teetkö suoria osakesijoituksia, sijoitatko osakerahastoihin vai johonkin muuhun kohteeseen (ympyröi sopiva vaihtoehto)?

sijoitan suoraan osakkeisiin

sijoitan osakerahastoihin

sijoitan muuhun kohteeseen kuin edellisiin, mihin: _____

d) Miten usein seuraat talousuutisia (ympyröi sopivin vaihtoehto)?

päivittäin

viikoittain

kerran kuukaudessa tai harvemmin

e) Arvioi, kuinka paljon talousuutiset vaikuttavat sijoituspäätöksiisi (ympyröi sopivin vaihtoehto).

ei mitenkään

hieman

huomattavasti

11. Harrastavatko vanhempasi sijoittamista (ympyröi sopiva vaihtoehto)?

kyllä

ei

12. Harrastavatko jotkut parhaimmista ystävästäsi sijoittamista (ympyröi sopiva vaihtoehto)?

kyllä

ei

13. Mikä on vanhempiesi sosioekonominen asema (ympyröi sopiva vaihtoehto)?

yrittäjät

ylemmät toimihenkilöt

alemmat toimihenkilöt

työntekijät

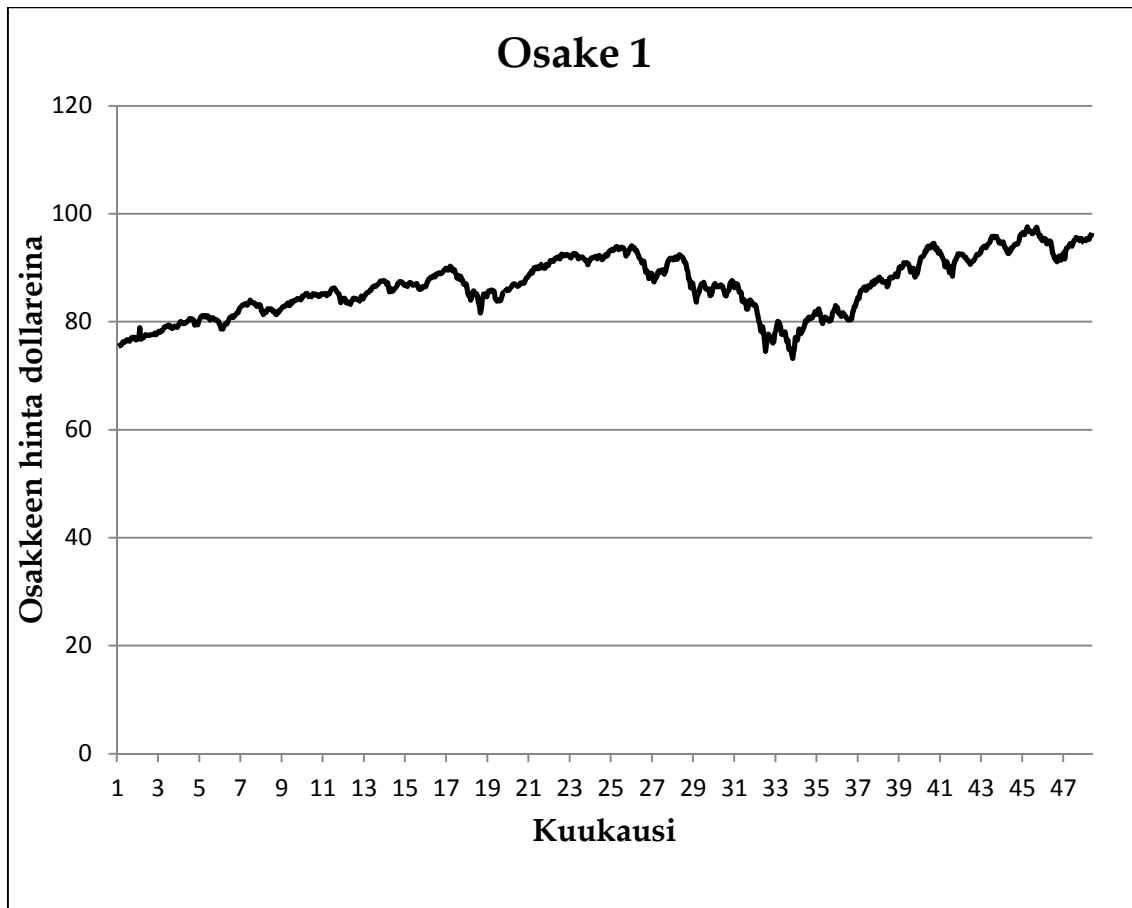
opiskelijat

eläkeläiset

muu, mikä _____

OSA 2: OSAKE-ENNUSTEET

Seuraavaksi näet kuusi kuvaajaa eräiden osakkeiden hintakehityksestä neljän vuoden ajalta. Tehtävänäsi on *ennustaa* kunkin osakkeen kohdalla, mikä on kyseisen osakkeen hinta *seitsemän* ja *13 kuukauden* kuluttua kussakin kuviossa esiintyvistä viimeisestä arvosta lukien. Tehtävänäsi on myös ennustaa *80 %:n luottamusväli*. Tämä tarkoittaa, että *ennustat kullekin osakkeelle matalan ja korkean hinnan* seitsemän ja 13 kuukauden päähän siten, että on *vain 10 %:n mahdollisuus, että toteutunut hinta on korkeampi kuin antamasi korkea ennuste ja vain 10 %:n mahdollisuus, että toteutunut hinta on matalampi kuin antamasi matala ennuste*.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

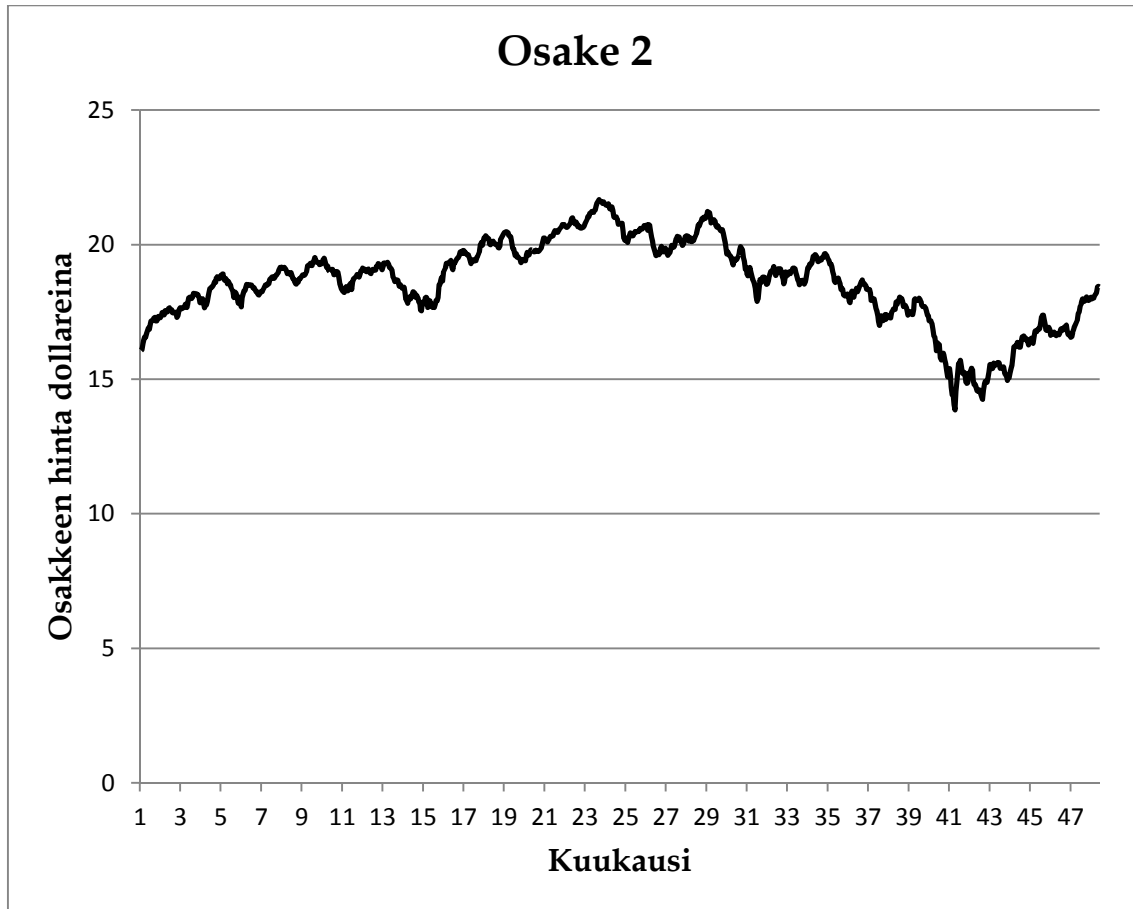
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

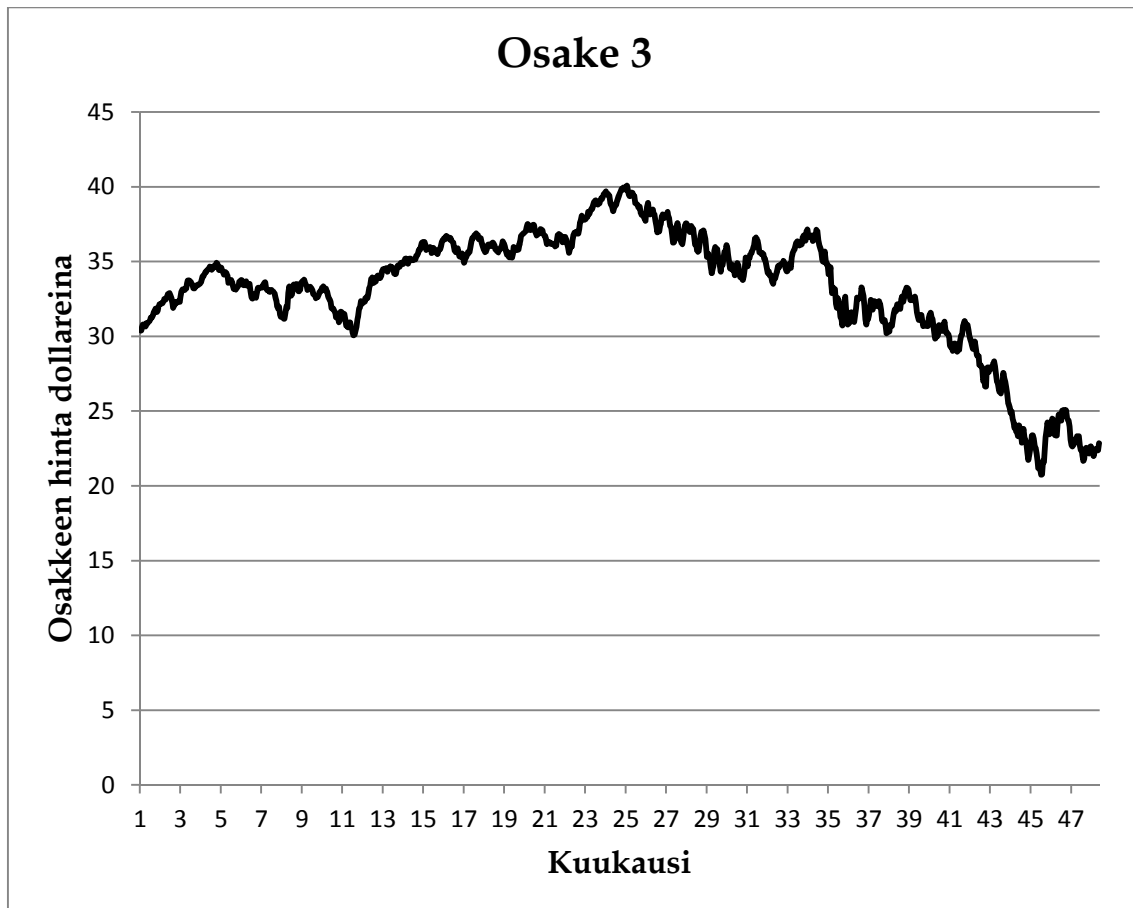
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

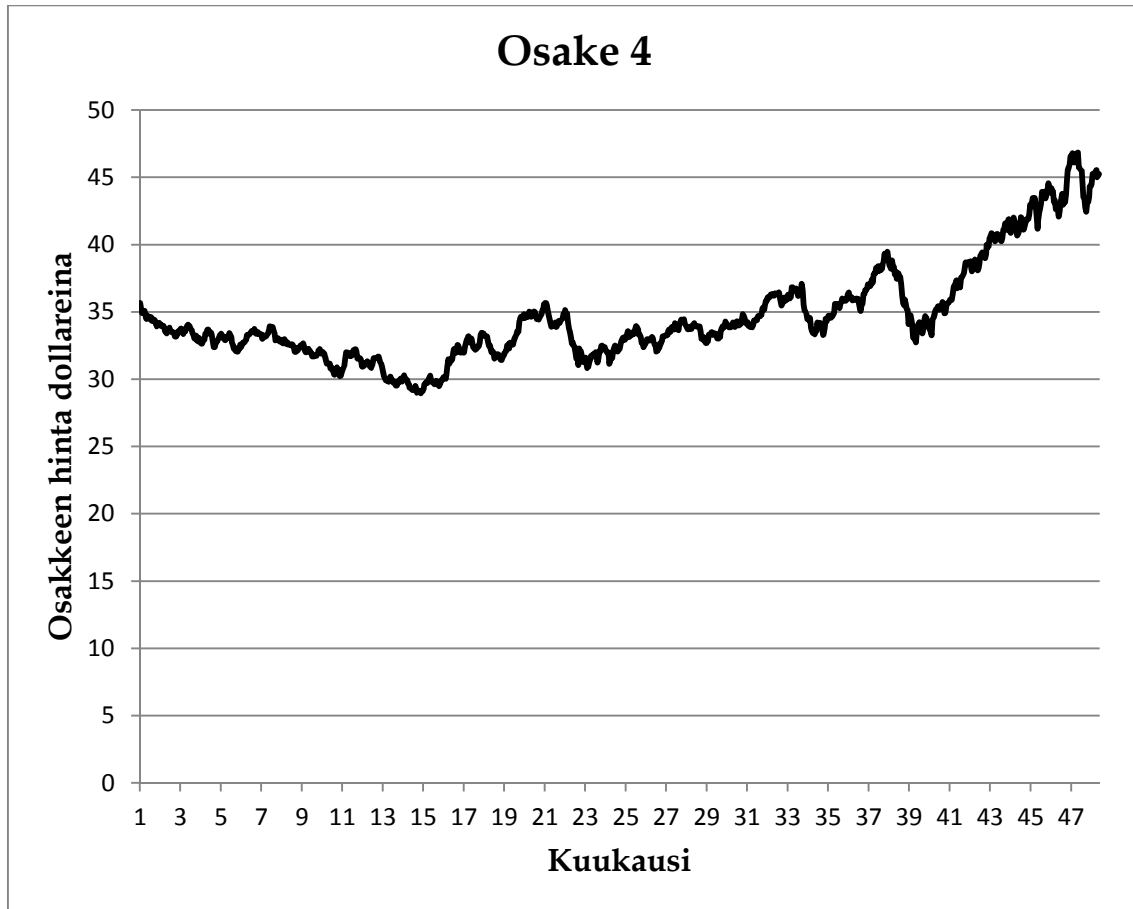
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

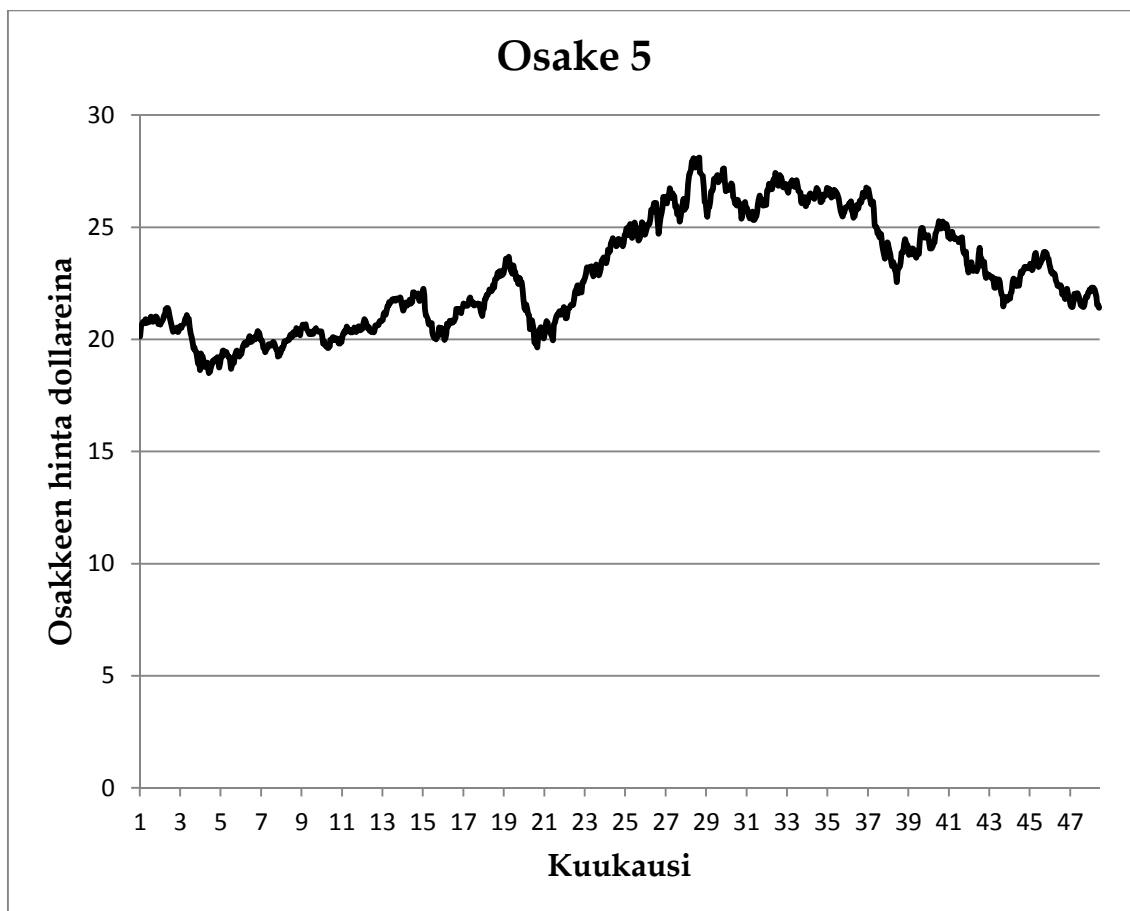
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

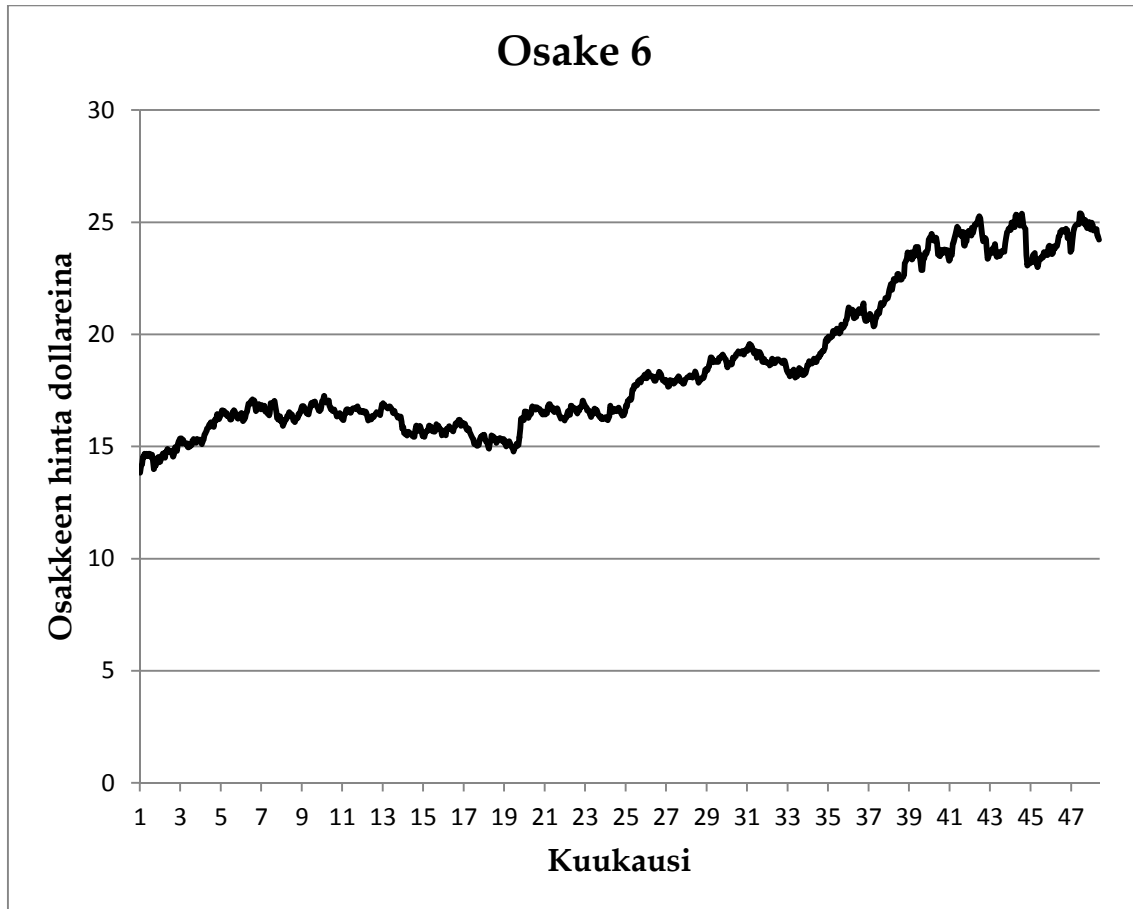
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Kiitos vastauksestasi! Muistithan laittaa etusivulle sähköpostiosoitteesi, niin olet mukana arvonnassa ja lahjakortin tavoittelussa.

Liite 2 Kyselylomake pankin sijoitusasiantuntijoille

ARVOISA VASTAANOTTAJA

Teen pro gradu - tutkimusta Jyväskylän yliopiston kauppakorkeakoululle ja edessäsi on työhön liittyvä kyselylomake. Kyselyn tarkoituksena on selvittää pankin sijoitusasiantuntijoiden käsityksiä tulevista osaketuotoista. Kyselyssä on kaksi osaa: ensimmäisessä osassa kartoitetaan perustietoja ja toisessa osassa sinua pyydetään tekemään ennusteita tulevista osaketuotoista.

Vastauksesi on arvokas opinnäytetyön onnistumisen kannalta. Kyselyn tuloksia käytetään vain pro gradu - työn lähdeaineistona, eikä niitä luovuteta kolmannelle osapuolelle. Kyselyyn vastaaminen on vapaaehtoista. Vastaamiseen menee aikaa noin 20 minuuttia.

Kaikkien kyselyyn vastanneiden, sähköpostiosoitteensa antaneiden kesken arvotaan Paul Krugmanin kirjoittama kirja Lama. Tämän lisäksi tarkimmat ennusteet tehneelle vastaajalle luovutetaan 50 euron lahjakortti Intersportiin. Arvonnan tuloksista ilmoitetaan henkilökohtaisesti. Sähköpostiosoitteet jäävät vain tutkijan tietoon, eikä niitä julkaista tutkimuksessa. Jos haluat osallistua kirjan arvontaan ja tavoitella lahjakorttia, merkitse sähköpostiosoitteesi selvästi tähän:

_____.

Ystävällisin terveisin,
Anni Kinttu
anni.h.kinttu@student.jyu.fi

OSA 1: VASTAAJAN PERUSTIEDOT

1. Syntymävuotesi: _____

2. Sukupuolesi: (ympyröi oikea vaihtoehto):

nainen

mies

3. Pohjakoulutuksesi (ympyröi sopiva vaihtoehto):

peruskoulu

ylioppilastutkinto

ammattillinen koulutus

ammattikorkeakoulututkinto

yliopistotutkinto

4. Kuinka pitkä työkokemus sinulla on rahoitusosalta (merkitse työkokemuksesi pituus vuosina tai kuukausina)?: _____

5. Nykyinen tehtävänimikkeesi: _____

6. Kuinka pitkään olet työskennellyt nykyisessä tehtävässäsi (merkitse työkokemuksesi pituus vuosina tai kuukausina)? _____

7. Onko tehokkaiden markkinoiden hypoteesi sinulle tuttu (ympyröi sopiva vaihtoehto)? Jos vastauksesi on kielteinen, siirry kysymykseen numero 9.

kyllä

ei

8. Jos vastasit edelliseen kysymykseen kyllä, ympyröi seuraavista väittämistä omaa käsitystäsi parhaiten vastaava vaihtoehto.

Vaihtoehdot ovat: 1. Täysin samaa mieltä, 2. Osittain samaa mieltä, 3. Ei samaa eikä eri mieltä, 4. Osittain eri mieltä, 5. Täysin eri mieltä.

a) Asiakkaiden kannattaa noudattaa passiivista salkun hajautusstrategiaa.

1 2 3 4 5

b) Osakekurseja voi ennustaa.

1 2 3 4 5

c) Asiakkaiden ei kannata kuluttaa aikaansa markkinoiden seuraamiseen.

1 2 3 4 5

d) Osakekurssit seuraavat satunnaiskulkua.

1 2 3 4 5

9. Arvioi, kuinka paljon talousuutiset vaikuttavat nykyisessä työssäsi antamiisi sijoitusneuvoihin tai -suosituksiin (ympyröi sopiva vaihtoehto).

ei mitenkään

hieman

huomattavasti

10. Arvioi, kuinka paljon keskustelut kollegoiden kanssa vaikuttavat nykyisessä työssäsi antamiisi sijoitusneuvoihin tai -suosituksiin (ympyröi sopiva vaihtoehto).

ei mitenkään

hieman

huomattavasti

11. Jos vastasit edelliseen kysymykseen huomattavasti, arvioi toimitko yleensä (ympyröi sopiva vaihtoehto):

samalla tavalla kuin muut

heidän näkemyksiään vastaan

joskus samalla tavalla, joskus päinvastoin

12. Kun annat asiakkaillesi sijoitusneuvoja tai -suosituksia, mikä on pääasiallinen tiedonlähdeesi, johon neuvosi perustuvat (ympyröi sopiva vaihtoehto)?

työnantajan tarjoama materiaali

rahoitukseen liittyvät uutiset

vallitsevat markkinahinnat ja niiden implikoimat odotukset

keskustelut kollegoiden kanssa

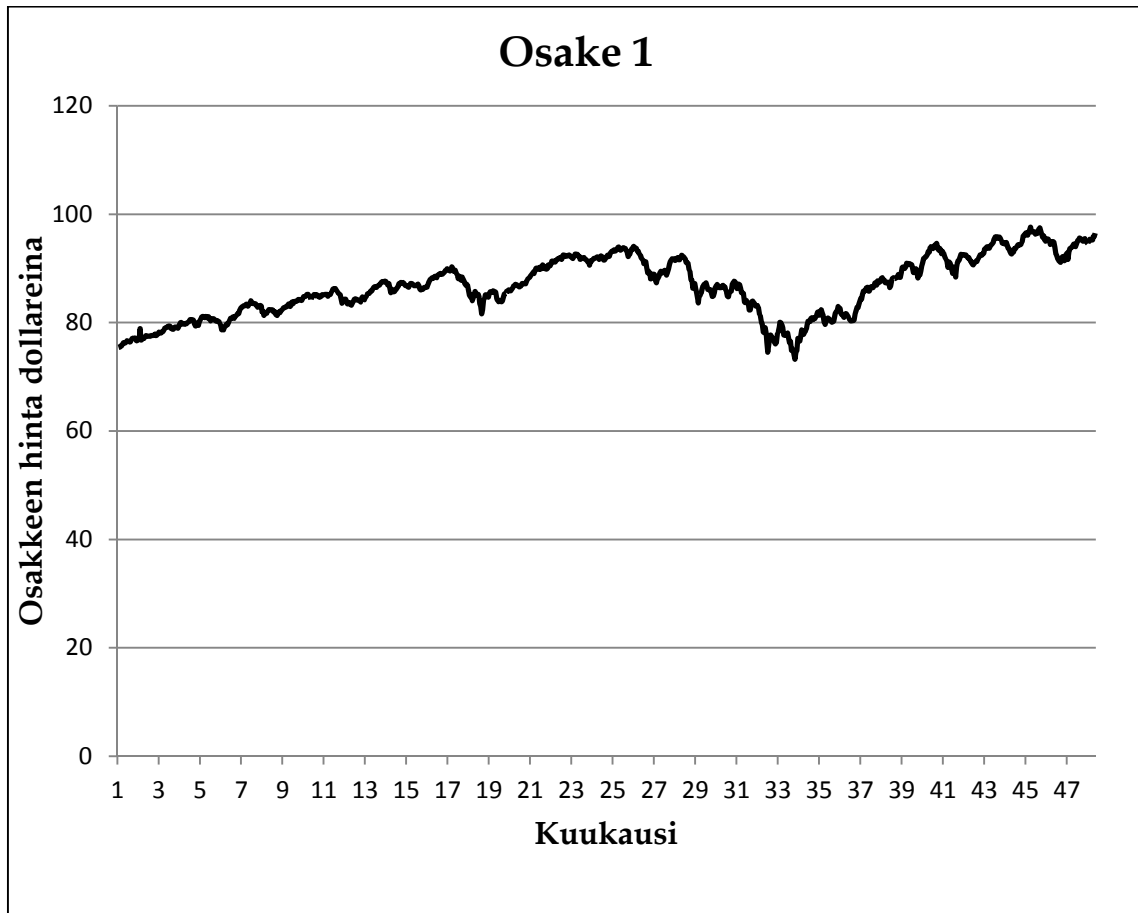
taloudelliset tilastot

historialliset tuotot

muu, mikä _____

OSA 2: OSAKE-ENNUSTEET

Seuraavaksi näet kuusi kuvaajaa eräiden osakkeiden hintakehityksestä neljän vuoden ajalta. Tehtävänäsi on *ennustaa* kunkin osakkeen kohdalla, mikä on kyseisen osakkeen hinta *seitsemän ja 13 kuukauden* kuluttua kussakin kuviossa esiintyvistä viimeisestä arvosta lukien. Tehtävänäsi on myös ennustaa *80 %:n luottamusväli*. Tämä tarkoittaa, että *ennustat kullekin osakkeelle matalan ja korkean hinnan seitsemän ja 13 kuukauden päähän siten, että on vain 10 %:n mahdollisuus, että toteutunut hinta on korkeampi kuin antamasi korkea ennuste ja vain 10 %:n mahdollisuus, että toteutunut hinta on matalampi kuin antamasi matala ennuste.*



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

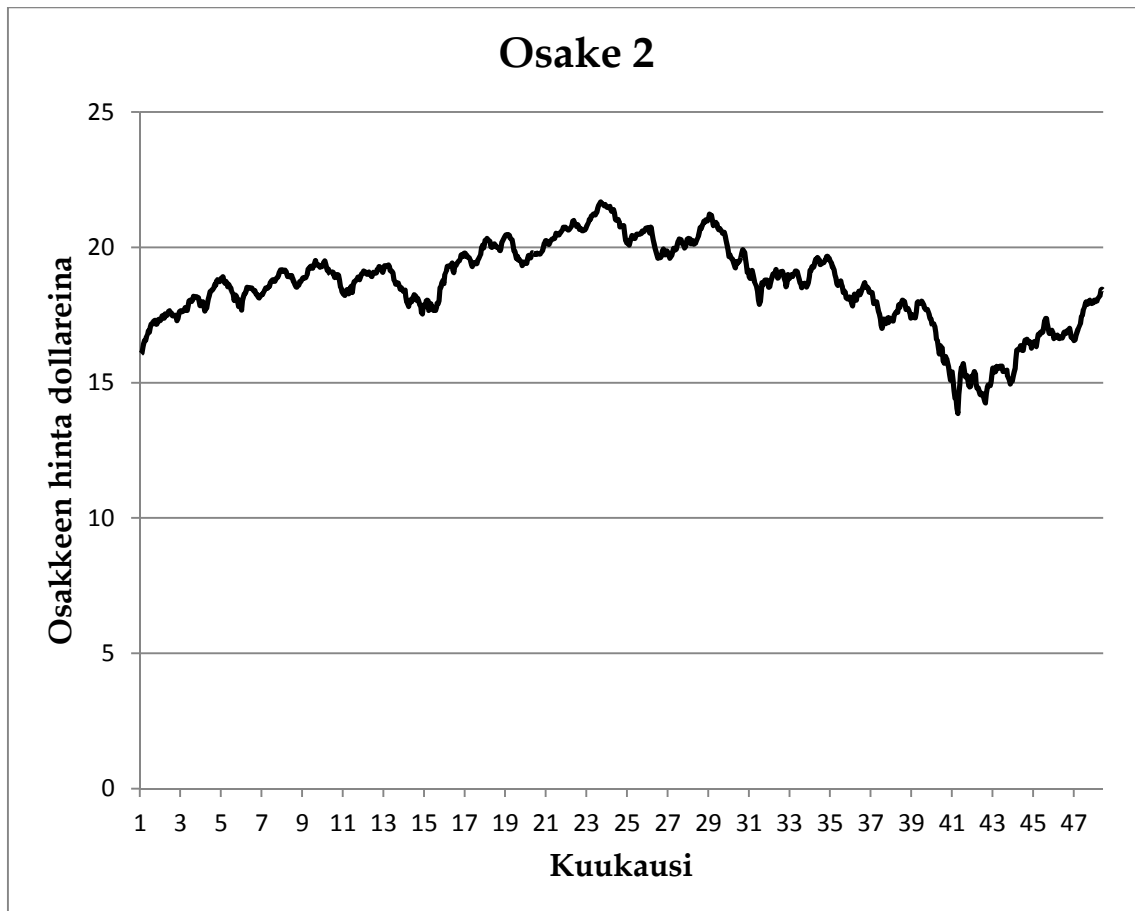
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

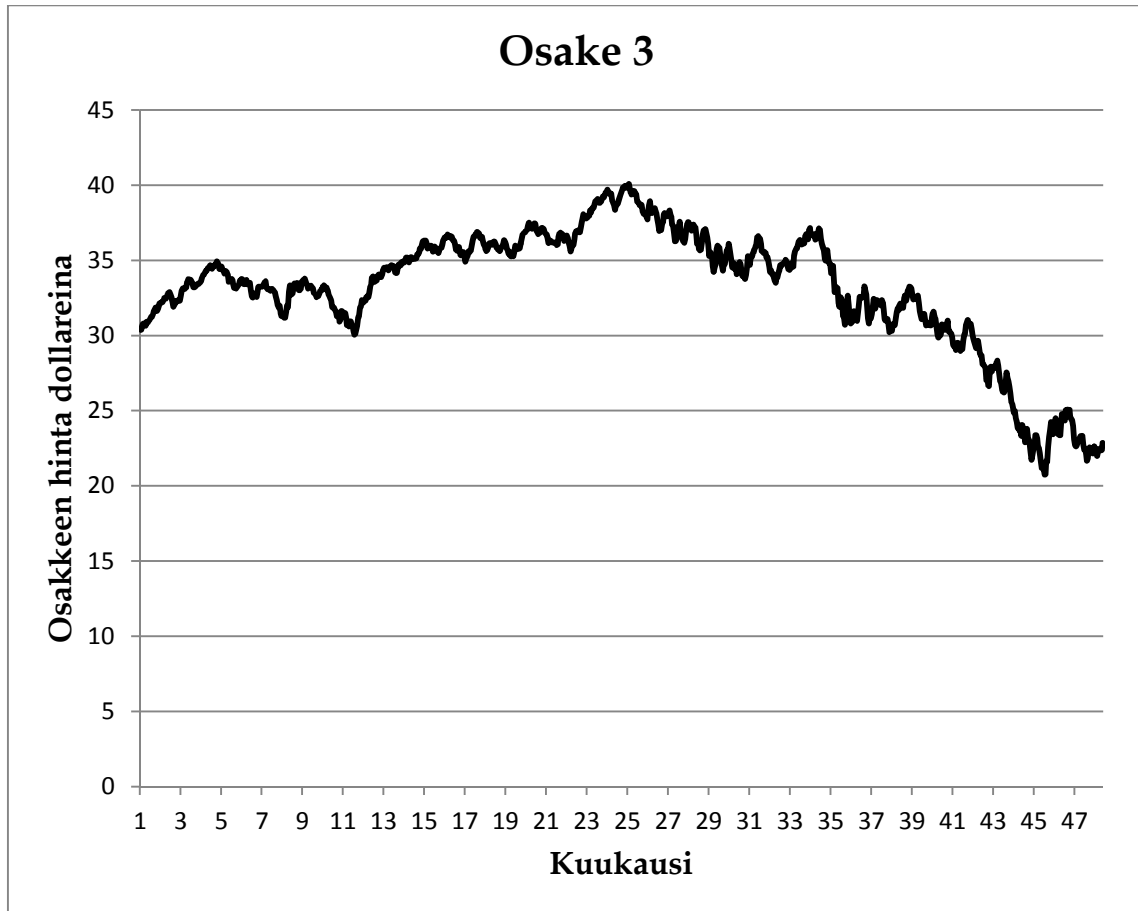
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

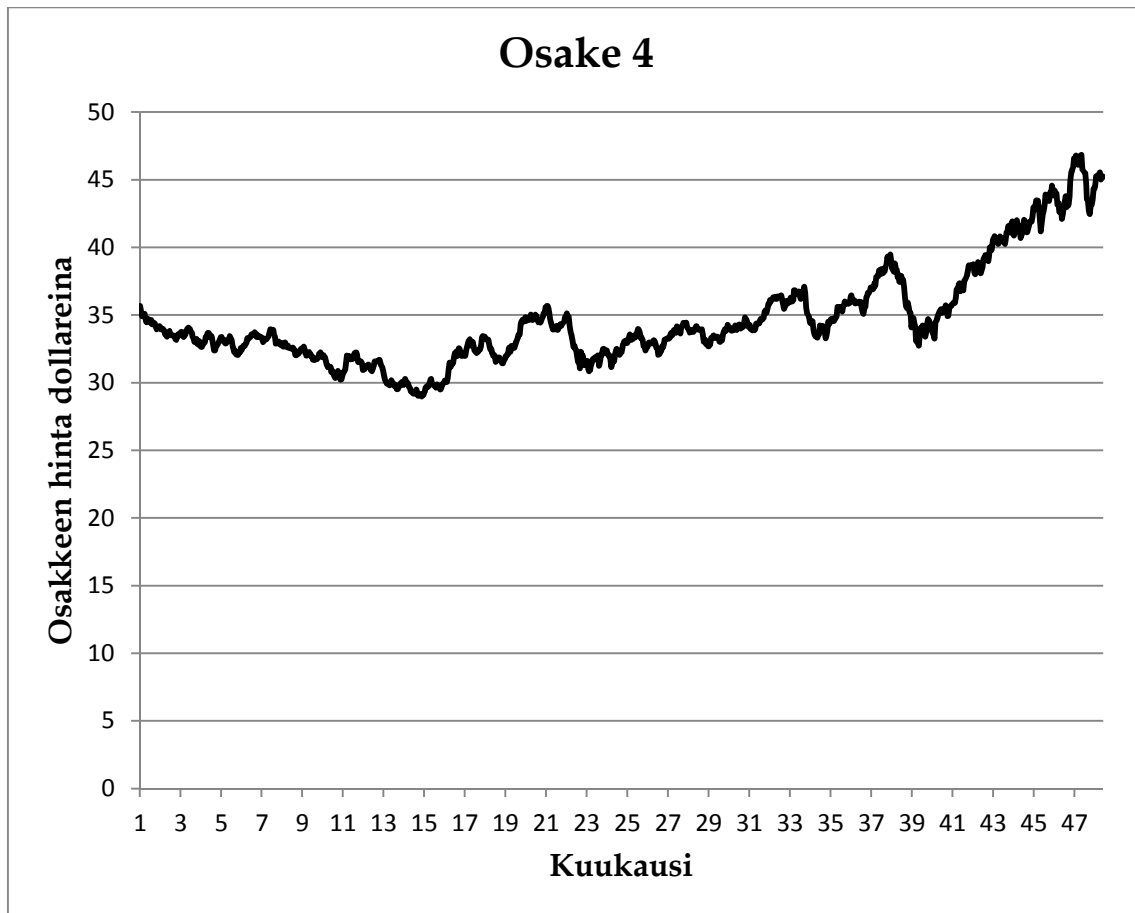
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

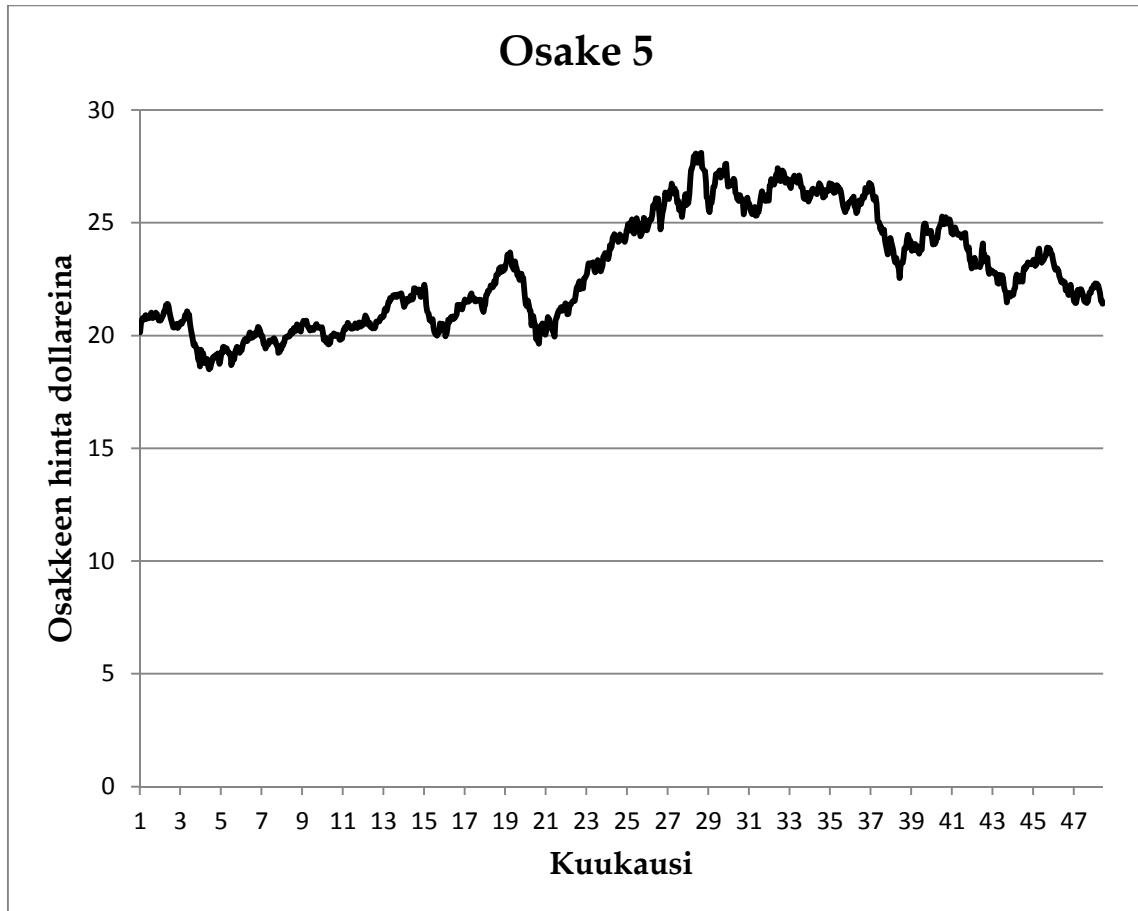
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

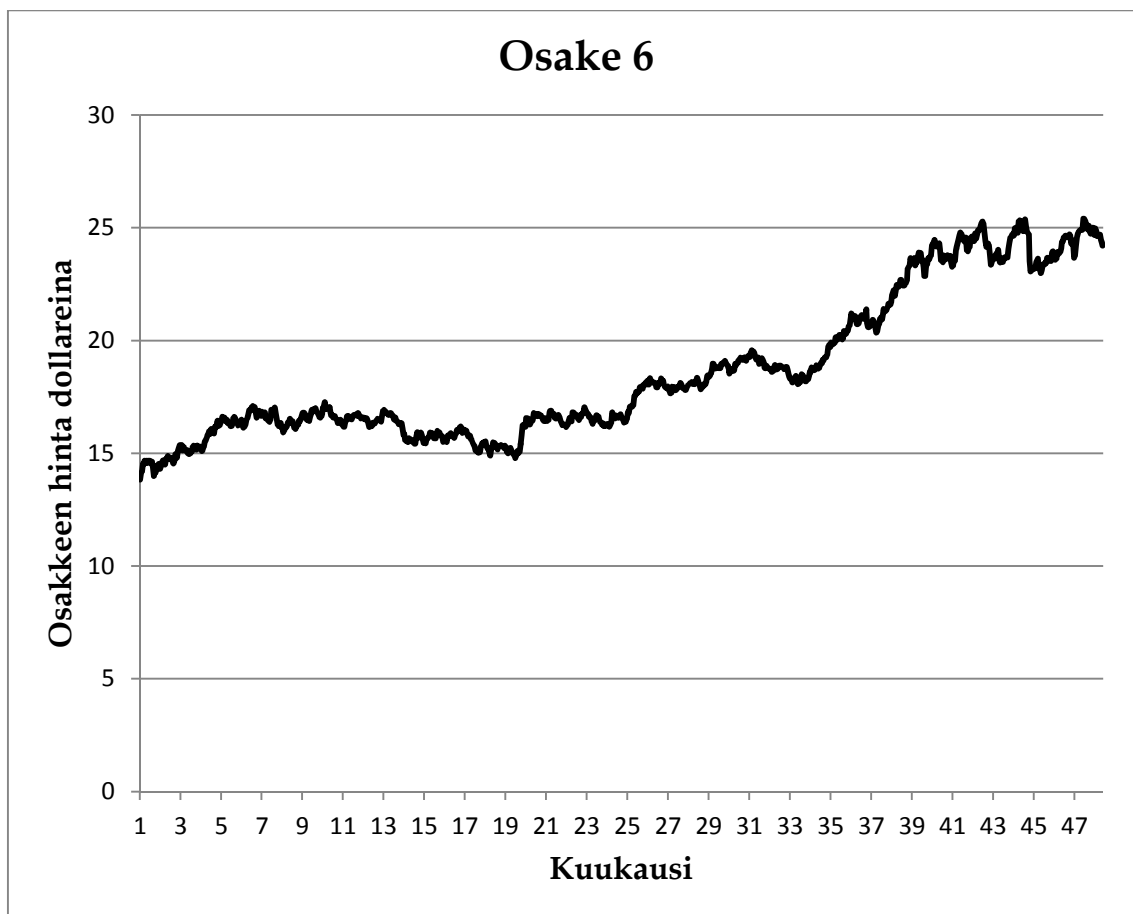
Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.



Ennustan, että kyseisen osakkeen hinta seitsemän kuukauden kuluttua on _____ dollaria.

Ennustan, että kyseisen osakkeen hinta 13 kuukauden kuluttua on _____ dollaria.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta seitsemän kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on korkeampi kuin _____ dollaria on 10 %.

Todennäköisyys sille, että osakkeen hinta 13 kuukauden kuluttua on matalampi kuin _____ dollaria on 10 %.

Kiitos vastauksestasi!

Muistithan laittaa etusivulle sähköpostiosoitteesi, niin olet mukana arvonnassa ja lahjakortin tavoittelussa.

Liite 3 Osake-ennusteiden perustunnusluvut

a) Opiskelija-aineisto

	N		Keski- arvo	Keski- hajonta	Va- rianssi	Vinous	Huipuk- kuus	Vaihte- luväli	Min.	Maks.
	Validi	Puut- tuu								
A1	146	0	95,10	7,34	53,94	-,30	,04	40	70	110
A2	145	1	96,40	10,09	101,91	-,63	1,66	60	60	120
A3	144	2	104,95	9,57	91,49	-,26	2,01	70	70	140
A4	144	2	80,63	10,87	118,21	-,78	1,09	60	40	100
A5	144	2	109,89	14,17	200,93	-,74	3,05	100	50	150
A6	144	2	78,34	14,38	206,66	-,60	1,30	90	20	110
B1	145	1	19,70	2,07	4,28	-,37	,90	13	12	25
B2	144	2	20,00	3,11	9,70	,24	-,05	15	13	28
B3	142	4	22,94	2,63	6,91	,33	1,50	16	15	31
B4	142	4	15,07	2,50	6,25	-,59	1,62	16	5	21
B5	142	4	24,39	4,11	16,93	1,81	11,91	40	10	50
B6	142	4	14,72	3,52	12,37	,07	1,42	20	5	25
C1	145	1	21,37	3,83	14,70	,15	-,42	19	12	31
C2	145	1	21,97	6,03	36,33	,34	,16	34	10	44
C3	143	3	27,64	4,97	24,68	,03	,08	25	15	40
C4	143	3	16,19	3,77	14,19	-,12	,83	23	5	28
C5	143	3	29,72	6,98	48,67	-,24	-,43	35	10	45
C6	143	3	14,87	5,80	33,67	,93	2,05	35	5	40
D1	144	2	46,70	3,51	12,32	-,50	,51	20	35	55
D2	143	3	47,55	6,53	42,70	-,09	,26	38	30	68
D3	142	4	51,90	6,29	39,51	-,49	10,56	65	15	80
D4	141	5	38,55	6,14	37,73	-1,26	3,98	40	10	50
D5	141	5	54,63	9,54	90,98	,36	3,94	75	25	100
D6	141	5	36,92	8,32	69,22	-,24	,98	55	5	60
E1	142	4	21,38	2,23	4,98	,74	1,32	13	17	30
E2	142	4	21,82	3,12	9,75	,10	,40	18	15	33
E3	141	5	25,91	3,20	10,24	1,23	4,58	22	18	40
E4	141	5	17,51	2,47	6,11	,54	5,08	22	8	30
E5	141	5	27,33	4,64	21,50	,44	3,27	35	10	45
E6	141	5	16,67	3,22	10,39	,22	2,15	23	7	30
F1	142	4	25,75	2,36	5,56	1,16	9,85	23	17	40
F2	142	4	26,65	3,48	12,13	-,13	,11	18	18	36
F3	141	5	29,57	4,19	17,52	3,26	19,74	38	22	60
F4	141	5	20,56	3,31	10,98	-1,95	6,68	23	3	26
F5	141	5	31,33	5,22	27,25	,93	3,00	35	15	50
F6	141	5	19,98	5,03	25,26	-,11	2,04	38	2	40

b) Pankin aineisto

	N		Keski- arvo	Keski- hajonta	Va- rianssi	Vinous	Huipuk- kuus	Vaihtelu- väli	Min.	Maks.
	Validi	Puut- tuu								
A1	37	0	96,14	7,91	62,57	-1,36	,87	30	75	105
A2	37	0	98,84	9,69	93,92	,19	,37	45	80	125
A3	37	0	105,00	10,35	107,17	-,58	,72	45	75	120
A4	37	0	82,22	10,56	111,62	-,46	-,38	40	60	100
A5	37	0	113,16	13,20	174,25	,58	,85	65	85	150
A6	37	0	79,32	14,66	214,95	-1,02	2,57	75	30	105
B1	37	0	19,95	1,87	3,50	,00	,70	9	16	25
B2	37	0	21,62	4,21	17,74	2,00	9,95	28	12	40
B3	37	0	22,95	3,45	11,89	1,24	3,99	20	15	35
B4	37	0	15,62	2,43	5,91	,48	,81	12	10	22
B5	37	0	25,78	5,55	30,84	2,94	10,65	31	19	50
B6	37	0	15,03	4,80	23,08	1,89	7,64	30	5	35
C1	37	0	23,08	3,71	13,80	-,09	,29	17	15	32
C2	37	0	24,27	6,74	45,48	,04	-,19	30	10	40
C3	37	0	27,49	4,83	23,31	,11	-,29	21	17	38
C4	37	0	18,38	3,85	14,85	-,32	,78	19	8	27
C5	37	0	30,54	7,91	62,64	,23	1,54	40	10	50
C6	37	0	16,81	6,33	40,05	,56	,70	30	5	35
D1	37	0	45,08	3,69	13,63	-,64	,21	17	35	52
D2	37	0	46,76	5,10	26,02	-,52	-,37	21	35	56
D3	37	0	50,24	4,95	24,47	,18	1,52	25	40	65
D4	37	0	37,57	5,46	29,86	-,42	,13	23	25	48
D5	37	0	53,86	6,63	43,95	,47	2,30	38	37	75
D6	37	0	36,89	7,52	56,49	-,13	-,37	30	20	50
E1	37	0	22,16	3,16	9,97	1,81	6,77	18	17	35
E2	37	0	23,51	5,59	31,26	2,86	13,79	35	15	50
E3	37	0	25,81	4,41	19,49	2,55	9,67	25	20	45
E4	37	0	18,41	2,79	7,80	,03	-,13	13	12	25
E5	37	0	29,11	6,76	45,71	3,03	12,31	40	20	60
E6	37	0	17,76	6,37	40,58	1,95	8,68	40	5	45
F1	37	0	24,89	1,84	3,38	-,77	1,11	8	20	28
F2	37	0	26,68	2,88	8,28	-,39	,30	12	20	32
F3	37	0	28,32	3,57	12,73	1,17	2,45	18	22	40
F4	37	0	20,05	3,37	11,33	-,57	,84	16	10	26
F5	37	0	30,62	4,72	22,24	,81	2,13	25	20	45
F6	37	0	19,38	4,64	21,58	-,52	1,31	23	5	28

Liite 4 Pankin vastaajien suhtautuminen tehokkaiden markkinoiden hypoteesin liittyviin väitteisiin

a)

		Tehokkaiden markkinoiden hypoteesin arviointi, strategia				Yhteensä	
		Osittain samaa mieltä	Ei samaa eikä eri mieltä	Osittain eri mieltä	Täysin eri mieltä		
Suku- puoli	Nainen	N	3	5	6	1	15
		%	20,0 %	33,3 %	40,0 %	6,7 %	100,0 %
Mies	Mies	N	8	5	6	0	19
		%	42,1 %	26,3 %	31,6 %	0,0 %	100,0 %
Yhteensä		N	11	10	12	1	34
		%	32,4 %	29,4 %	35,3 %	2,9 %	100,0 %

b)

		Tehokkaiden markkinoiden hypoteesin arviointi, ennustettavuus				Yhteensä	
		Osittain samaa mieltä	Ei samaa eikä eri mieltä	Osittain eri mieltä	Täysin eri mieltä		
Suku- puoli	Nainen	N	6	4	3	2	15
		%	40,0 %	26,7 %	20,0 %	13,3 %	100,0 %
Mies	Mies	N	10	0	9	0	19
		%	52,6 %	0,0 %	47,4 %	0,0 %	100,0 %
Yhteensä		N	16	4	12	2	34
		%	47,1 %	11,8 %	35,3 %	5,9 %	100,0 %

c)

		Tehokkaiden markkinoiden hypoteesin arviointi, markkinoiden seuraaminen					Yhteensä	
		Täysin samaa mieltä	Osittain samaa mieltä	Ei samaa eikä eri mieltä	Osittain eri mieltä	Täysin eri mieltä		
Suku- puoli	Nainen	N	0	1	3	7	4	15
		%	0,0 %	6,7 %	20,0 %	46,7 %	26,7 %	100,0 %
Mies	Mies	N	3	5	1	6	4	19
		%	15,8 %	26,3 %	5,3 %	31,6 %	21,1 %	100,0 %
Yhteensä		N	3	6	4	13	8	34
		%	8,8 %	17,6 %	11,8 %	38,2 %	23,5 %	100,0 %

d)

		Tehokkaiden markkinoiden hypoteesin arviointi, satunnaiskulku				Yhteensä
		Osittain samaa mieltä	Ei samaa eikä eri mieltä	Osittain eri mieltä	Täysin eri mieltä	
Sukupuoli	Nainen	N 4	3	6	2	15
		% 26,7 %	20,0 %	40,0 %	13,3 %	100,0 %
	Mies	N 8	0	10	1	19
		% 42,1 %	0,0 %	52,6 %	5,3 %	100,0 %
Yhteensä		N 12	3	16	3	34
		% 35,3 %	8,8 %	47,1 %	8,8 %	100,0 %

Liite 5 Osake-ennusteiden jakaumien normalisuuden testaus ja symmetrisyyden taskastelu

a) Opiskelija-aineisto

	Kolmogorov-Smirnov			Shapiro-Wilk			Vinous/ keskivirhe	Symmetri- nen?
	Testi- suure	Vapaus- asteet	P-arvo	Testi- suure	Vapaus- asteet	P-arvo		
A1	,113	146	,000	,973	146	,006	-1,480	on
A2	,108	145	,000	,959	145	,000	-3,110	ei ole
A3	,146	144	,000	,953	144	,000	-1,290	on
A4	,171	144	,000	,951	144	,000	-3,830	ei ole
A5	,114	144	,000	,944	144	,000	-3,670	ei ole
A6	,096	144	,003	,972	144	,005	-2,970	ei ole
B1	,145	145	,000	,966	145	,001	-1,850	on
B2	,094	144	,003	,979	144	,029	1,200	on
B3	,132	142	,000	,945	142	,000	1,610	on
B4	,164	142	,000	,958	142	,000	-2,910	ei ole
B5	,230	142	,000	,830	142	,000	8,900	ei ole
B6	,116	142	,000	,961	142	,000	,350	on
C1	,157	145	,000	,974	145	,007	,770	on
C2	,092	145	,004	,978	145	,019	1,680	on
C3	,135	143	,000	,966	143	,001	,130	on
C4	,153	143	,000	,965	143	,001	-,590	on
C5	,118	143	,000	,975	143	,009	-1,180	on
C6	,099	143	,001	,946	143	,000	4,600	ei ole
D1	,145	144	,000	,968	144	,002	-2,480	ei ole
D2	,122	143	,000	,976	143	,013	-,460	on
D3	,163	142	,000	,848	142	,000	-2,390	ei ole
D4	,168	141	,000	,910	141	,000	-6,170	ei ole
D5	,130	141	,000	,931	141	,000	1,760	on
D6	,093	141	,004	,975	141	,011	-1,150	on
E1	,146	142	,000	,947	142	,000	3,630	ei ole
E2	,105	142	,001	,968	142	,002	,490	on
E3	,186	141	,000	,898	141	,000	6,040	ei ole
E4	,134	141	,000	,900	141	,000	2,660	ei ole
E5	,169	141	,000	,936	141	,000	2,140	ei ole
E6	,139	141	,000	,956	141	,000	1,090	on
F1	,158	142	,000	,871	142	,000	5,700	ei ole
F2	,098	142	,002	,972	142	,006	-,660	on
F3	,261	141	,000	,752	141	,000	15,960	ei ole
F4	,256	141	,000	,834	141	,000	-9,560	ei ole
F5	,144	141	,000	,925	141	,000	4,570	ei ole
F6	,080	141	,026	,963	141	,001	-,550	on

b) Pankin aineisto

	Kolmogorov-Smirnov			Shapiro-Wilk			Vinous/ keskivirhe	Symmetri- nen?
	Testi- suure	Vapaus- asteet	P-arvo	Testi- suure	Vapaus- asteet	P-arvo		
A1	,263	37	,000	,807	37	,000	-3,519	ei ole
A2	,141	37	,059	,966	37	,302	,497	on
A3	,125	37	,150	,941	37	,050	-1,500	on
A4	,121	37	,192	,954	37	,134	-1,192	on
A5	,216	37	,000	,950	37	,095	1,495	on
A6	,167	37	,011	,927	37	,018	-2,627	ei ole
B1	,191	37	,002	,931	37	,025	,005	on
B2	,183	37	,003	,803	37	,000	5,151	ei ole
B3	,196	37	,001	,870	37	,000	3,186	ei ole
B4	,150	37	,034	,948	37	,084	1,242	on
B5	,313	37	,000	,678	37	,000	7,582	ei ole
B6	,178	37	,005	,846	37	,000	4,886	ei ole
C1	,138	37	,072	,975	37	,555	-,236	on
C2	,165	37	,013	,962	37	,232	,106	on
C3	,129	37	,123	,979	37	,689	,275	on
C4	,136	37	,080	,961	37	,221	-,835	on
C5	,151	37	,032	,944	37	,061	,605	on
C6	,153	37	,028	,951	37	,102	1,453	on
D1	,167	37	,011	,936	37	,033	-1,638	on
D2	,195	37	,001	,929	37	,021	-1,330	on
D3	,168	37	,010	,938	37	,039	,463	on
D4	,158	37	,020	,952	37	,115	-1,083	on
D5	,123	37	,169	,946	37	,074	1,222	on
D6	,120	37	,200	,973	37	,501	-,325	on
E1	,233	37	,000	,838	37	,000	4,660	ei ole
E2	,260	37	,000	,715	37	,000	7,381	ei ole
E3	,267	37	,000	,741	37	,000	6,584	ei ole
E4	,148	37	,039	,961	37	,211	,086	on
E5	,285	37	,000	,699	37	,000	7,819	ei ole
E6	,146	37	,044	,835	37	,000	5,039	ei ole
F1	,199	37	,001	,921	37	,012	-1,984	on
F2	,137	37	,077	,953	37	,123	-1,001	on
F3	,184	37	,003	,917	37	,009	3,009	ei ole
F4	,155	37	,025	,947	37	,076	-1,474	on
F5	,198	37	,001	,929	37	,021	2,095	ei ole
F6	,121	37	,191	,955	37	,137	-1,347	on

Liite 6 Tutkimuksen hypoteesit

H₁: Osake-ennusteet ankkuroituvat osakkeen hintakehityksen keskiarvoa kohti.

$$H_{20}: \overline{EPC}_{\text{Bull}} = \overline{EPC}_{\text{Bear}}$$

$$H_{21}: \overline{EPC}_{\text{Bull}} > \overline{EPC}_{\text{Bear}}.$$

$$H_{30}: \bar{\Delta}_{\text{Bull}} = \bar{\Delta}_{\text{Bear}}$$

$$H_{31}: \bar{\Delta}_{\text{Bull}} < \bar{\Delta}_{\text{Bear}}.$$

H₄: Luottamusvälit kasvavat enemmän oikealle vinoiksi, kun suurempaa hin-
narlaskua ennustetaan ja toisinpäin.

H₅: Pankin sijoitusasiantuntijat ovat yliluottavaisempia kuin opiskelijat.

H₆: Paremmin koulutetut henkilöt tekevät vähemmän virheitä kuin heikommin
koulutetut.

H₇: Sijoituskokemus vähentää sijoittajien tekemiä virheitä.

H₈: Ne vastaajat, joiden vanhemmilla on korkea sosioekonominen asema ja joi-
den vanhemmat ja lähipiiri harrastavat sijoittamista tekevät vähemmän sijoit-
tamiseen liittyviä virheitä.

H₉: Miehet ovat yliluottavaisempia kuin naiset.

H₁₀: Keski-ikäiset tekevät kaikista vähiten virheitä osake-ennusteissa.

H₁₁: Opiskelijat käyttäytyvät häirikkösijoittajan tavoin.

Liite 7 Trendinseuraaja- ja toisinajattelijaluokan määrittymuuttujan luonti

Muuttujien TF_bull ja TF_bear luonti SPSS syntaksilla:

IF((((A1-96.47)/96.47+(D1-45.25)/45.25+(F1-24.22)/24.22))/3>0 AND (((A2-A1)/A1+(D2-D1)/D1+(F2-F1)/F1))/3>0)TF_bull=1.

IF((((A1-96.47)/96.47+(D1-45.25)/45.25+(F1-24.22)/24.22))/3<0 AND (((A2-A1)/A1+(D2-D1)/D1+(F2-F1)/F1))/3<0)TF_bull=2.

IF((((A1-96.47)/96.47+(D1-45.25)/45.25+(F1-24.22)/24.22))/3>0 AND (((A2-A1)/A1+(D2-D1)/D1+(F2-F1)/F1))/3<0)TF_bull=3.

IF((((A1-96.47)/96.47+(D1-45.25)/45.25+(F1-24.22)/24.22))/3<0 AND (((A2-A1)/A1+(D2-D1)/D1+(F2-F1)/F1))/3>0)TF_bull=4.

IF((((B1-18.43)/18.43+(C1-22.85)/22.85+(E1-28.13)/28.13))/3>0 AND (((B2-B1)/B1+(C2-C1)/C1+(E2-E1)/E1))/3>0)TF_bear=1.

IF((((B1-18.43)/18.43+(C1-22.85)/22.85+(E1-28.13)/28.13))/3<0 AND (((B2-B1)/B1+(C2-C1)/C1+(E2-E1)/E1))/3<0)TF_bear=2.

IF((((B1-18.43)/18.43+(C1-22.85)/22.85+(E1-28.13)/28.13))/3>0 AND (((B2-B1)/B1+(C2-C1)/C1+(E2-E1)/E1))/3<0)TF_bear=3.

IF((((B1-18.43)/18.43+(C1-22.85)/22.85+(E1-28.13)/28.13))/3<0 AND (((B2-B1)/B1+(C2-C1)/C1+(E2-E1)/E1))/3>0)TF_bear=4.

EXECUTE.

Muuttujien TF_bull_CO ja TF_bear_CO luonti SPSS syntaksilla:

IF (TF_bull=1 OR TF_bull=3) TF_bull_CO=1.

IF (TF_bull=2 OR TF_bull=4) TF_bull_CO=0.

IF (TF_bear=2 OR TF_bear=4) TF_bear_CO=1.

IF (TF_bear=1 OR TF_bear=3) TF_bear_CO=0.

EXECUTE.

Liite 8 Muuttujaluettelo

Muuttuja	Selite
A	osakekuvaaja tai osake 1.
A1	osakekuvaajaan 1. liittyvä 7 kk:n piste-estimaatti
A2	osakekuvaajaan 1. liittyvä 13 kk:n piste-estimaatti
A3	osakekuvaajaan 1. liittyvä 7 kk:n yläluottamusväli
A4	osakekuvaajaan 1. liittyvä 7 kk:n alaluottamusväli
A5	osakekuvaajaan 1. liittyvä 13 kk:n yläluottamusväli
A6	osakekuvaajaan 1. liittyvä 13 kk:n alaluottamusväli
Aluottamusväli7kk	luottamusvälin kapeus, A-osakekuvaajaan liittyvä 7 kk:n luottamusväli
Aluottamusväli13kk	luottamusvälin kapeus, A-osakekuvaajaan liittyvä 13 kk:n luottamusväli
B	osakekuvaaja tai osake 2.
B1	osakekuvaajaan 2. liittyvä 7 kk:n piste-estimaatti
B2	osakekuvaajaan 2. liittyvä 13 kk:n piste-estimaatti
B3	osakekuvaajaan 2. liittyvä 7 kk:n yläluottamusväli
B4	osakekuvaajaan 2. liittyvä 7 kk:n alaluottamusväli
B5	osakekuvaajaan 2. liittyvä 13 kk:n yläluottamusväli
B6	osakekuvaajaan 2. liittyvä 13 kk:n alaluottamusväli
Bluottamusväli7kk	luottamusvälin kapeus, B-osakekuvaajaan liittyvä 7 kk:n luottamusväli
Bluottamusväli13kk	luottamusvälin kapeus, B-osakekuvaajaan liittyvä 13 kk:n luottamusväli
C	osakekuvaaja tai osake 3.
C1	osakekuvaajaan 3. liittyvä 7 kk:n piste-estimaatti
C2	osakekuvaajaan 3. liittyvä 13 kk:n piste-estimaatti
C3	osakekuvaajaan 3. liittyvä 7 kk:n yläluottamusväli
C4	osakekuvaajaan 3. liittyvä 7 kk:n alaluottamusväli
C5	osakekuvaajaan 3. liittyvä 13 kk:n yläluottamusväli
C6	osakekuvaajaan 3. liittyvä 13 kk:n alaluottamusväli
Cluottamusväli7kk	luottamusvälin kapeus, C-osakekuvaajaan liittyvä 7 kk:n luottamusväli
Cluottamusväli13kk	luottamusvälin kapeus, C-osakekuvaajaan liittyvä 13 kk:n luottamusväli
CONT_A	toisinajattelijastrategiaryhmä, osakekuvaaja A
CONT_B	toisinajattelijastrategiaryhmä, osakekuvaaja B
CONT_C	toisinajattelijastrategiaryhmä, osakekuvaaja C
CONT_D	toisinajattelijastrategiaryhmä, osakekuvaaja D
CONT_E	toisinajattelijastrategiaryhmä, osakekuvaaja E
CONT_F	toisinajattelijastrategiaryhmä, osakekuvaaja F
D	osakekuvaaja tai osake 4.
D1	osakekuvaajaan 4. liittyvä 7 kk:n piste-estimaatti

Muuttuja	Selite
D2	osakekuvaajaan 4. liittyvä 13 kk:n piste-estimaatti
D3	osakekuvaajaan 4. liittyvä 7 kk:n yläluottamusväli
D4	osakekuvaajaan 4. liittyvä 7 kk:n alaluottamusväli
D5	osakekuvaajaan 4. liittyvä 13 kk:n yläluottamusväli
D6	osakekuvaajaan 4. liittyvä 13 kk:n alaluottamusväli
Delta, Δ	vinous
Delta_bear	vinous, laskukausien osakekuvaajat
Delta_bear _{norm}	vinous, laskukausien osakekuvaajat, normalisoitu
$\bar{\Delta}_{\text{Bear}}$	keskimääräinen vinous, laskukaudet
Delta_bull	vinous, nousukausien osakekuvaajat
Delta_bull _{norm}	vinous, nousukausien osakekuvaajat, normalisoitu
$\bar{\Delta}_{\text{Bull}}$	keskimääräinen vinous, nousukaudet
Delta7 _{bear}	vinous, laskukausien 7 kk:n luottamusvälit
Delta7 _{bear_norm}	vinous, laskukausien 7 kk:n luottamusvälit, normalisoitu
Delta7 _{bull}	vinous, nousukausien 7 kk:n luottamusvälit
Delta7 _{bull_norm}	vinous, nousukausien 7 kk:n luottamusvälit, normalisoitu
Delta13 _{bear}	vinous, laskukausien 13 kk:n luottamusvälit
Delta13 _{bear_norm}	vinous, laskukausien 13 kk:n luottamusvälit, normalisoitu
Delta13 _{bull}	vinous, nousukausien 13 kk:n luottamusvälit
Delta13 _{bull_norm}	vinous, nousukausien 13 kk:n luottamusvälit, normalisoitu
Dluottamusväli7kk	luottamusvälin kapeus, D-osakekuvaajaan liittyvä 7 kk:n luottamusväli
Dluottamusväli13kk	luottamusvälin kapeus, D-osakekuvaajaan liittyvä 13 kk:n luottamusväli
E	osakekuvaaja tai osake 5.
E1	osakekuvaajaan 5. liittyvä 7 kk:n piste-estimaatti
E2	osakekuvaajaan 5. liittyvä 13 kk:n piste-estimaatti
E3	osakekuvaajaan 5. liittyvä 7 kk:n yläluottamusväli
E4	osakekuvaajaan 5. liittyvä 7 kk:n alaluottamusväli
E5	osakekuvaajaan 5. liittyvä 13 kk:n yläluottamusväli
E6	osakekuvaajaan 5. liittyvä 13 kk:n alaluottamusväli
Eluottamusväli7kk	luottamusvälin kapeus, E-osakekuvaajaan liittyvä 7kk:n luottamusväli
Eluottamusväli13kk	luottamusvälin kapeus, E-osakekuvaajaan liittyvä 13kk:n luottamusväli
ENNVIRHE	kaikkien osake-ennusteiden virheiden neliösumma
EPC_bear	odotettu hintamuutos, laskukaudet
EPC_bear _{norm}	odotettu hintamuutos, laskukaudet, normalisoitu
\bar{EPC}_{Bear}	odotetun hintamuutoksen keskiarvo, laskukaudet
EPC_bull	odotettu hintamuutos, nousukaudet

Muuttuja	Selite
EPC_{bull_norm}	odotettu hintamuutos, nousukaudet, normalisoitu
\overline{EPC}_{Bull}	odotetun hintamuutoksen keskiarvo, nousukaudet
EPC_7	odotettu hintamuutos, 7 kk:n piste-estimaatit
$EPC7_{bear}$	odotettu hintamuutos, 7 kk:n piste-estimaatit, laskukaudet
$EPC7_{bear_norm}$	odotettu hintamuutos, 7 kk:n piste-estimaatit, laskukaudet, normalisoitu
$EPC7_{bull}$	odotettu hintamuutos, 7 kk:n piste-estimaatit, nousukaudet
$EPC7_{bull_norm}$	odotettu hintamuutos, 7 kk:n piste-estimaatit, nousukaudet, normalisoitu
EPC_{13}	odotettu hintamuutos, 13 kk:n piste-estimaatit
$EPC13_{bear}$	odotettu hintamuutos, 13 kk:n piste-estimaatit, laskukaudet
$EPC13_{bear_norm}$	odotettu hintamuutos, 13 kk:n piste-estimaatit, laskukaudet, normalisoitu
$EPC13_{bull}$	odotettu hintamuutos, 13 kk:n piste-estimaatit, nousukaudet
$EPC13_{bull_norm}$	odotettu hintamuutos, 13 kk:n piste-estimaatit, nousukaudet, normalisoitu
EROKA	vastaajan antama piste-estimaatti vähennettynä kunkin osakekuvaajan keskiarvolla
EroKaA1	piste-estimaatti A1 vähennettynä osakekuvaajan A keskiarvolla
EroKaA2	piste-estimaatti A2 vähennettynä osakekuvaajan A keskiarvolla
EroKaB1	piste-estimaatti B1 vähennettynä osakekuvaajan B keskiarvolla
EroKaB2	piste-estimaatti B2 vähennettynä osakekuvaajan B keskiarvolla
EroKaC1	piste-estimaatti C1 vähennettynä osakekuvaajan C keskiarvolla
EroKaC2	piste-estimaatti C2 vähennettynä osakekuvaajan C keskiarvolla
EroKaD1	piste-estimaatti D1 vähennettynä osakekuvaajan D keskiarvolla
EroKaD2	piste-estimaatti D2 vähennettynä osakekuvaajan D keskiarvolla
EroKaE1	piste-estimaatti E1 vähennettynä osakekuvaajan E keskiarvolla
EroKaE2	piste-estimaatti E2 vähennettynä osakekuvaajan E keskiarvolla
EroKaF1	piste-estimaatti F1 vähennettynä osakekuvaajan F keskiarvolla

Muuttuja	Selite
EroKaF2	piste-estimaatti F2 vähennettynä osakekuvaajan F keskiarvolla
F	piste-estimaatti; osakekuvaaja tai osake 6.
F1	osakekuvaajaan 6. liittyvä 7 kk:n piste-estimaatti
F2	osakekuvaajaan 6. liittyvä 13 kk:n piste-estimaatti
F3	osakekuvaajaan 6. liittyvä 7 kk:n yläluottamusväli
F4	osakekuvaajaan 6. liittyvä 7 kk:n alaluottamusväli
F5	osakekuvaajaan 6. liittyvä 13 kk:n yläluottamusväli
F6	osakekuvaajaan 6. liittyvä 13 kk:n alaluottamusväli
F7	seitsemän kuukauden piste-estimaatti
F ₁₃	13 kuukauden piste-estimaatti
Fluottamusväli7kk	luottamusvälin kapeus, E-osakekuvaajaan liittyvä 7 kk:n luottamusväli
Fluottamusväli13kk	luottamusvälin kapeus, E-osakekuvaajaan liittyvä 13 kk:n luottamusväli
H	vastaajan ennustaman todennäköisyysjakauman ylä-arvo
IKÄ	vastaajan ikä vuosina
K2	opiskelijoille suunnatun kyselylomakkeen kysymys K2: Sukupuoli?
K6	opiskelijoille suunnatun kyselylomakkeen kysymys K6: Opiskeltujen rahoituksen kurssien määrä?
K11	opiskelijoille suunnatun kyselylomakkeen kysymys K11: Harrastavatko vanhempasi sijoittamista?
K12	opiskelijoille suunnatun kyselylomakkeen kysymys K12: Harrastavatko jotkut parhaimmista ystävästäsi sijoittamista?
P ₀	kyselylomakkeessa näytettyjen osakekuvaajien viimeinen arvo
L	yksilön ennustaman todennäköisyysjakauman ala-arvo
LCI	alempi luottamusväli, piste-estimaatti (F) vähennettynä todennäköisyysjakauman ala-arvolla (L)
LCI _{7kk}	alempaan luottamusvälin 7 kk:n ennusteen arvo
LCI _{13kk}	alempaan luottamusvälin 13 kk:n ennusteen arvo
LUOTTAMUSVÄLI	ylä- ja alaluottamusvälien erotus eli luottamusvälin kapeus
LUOTTO	oman salkun odotettuun tuottoon perustuva jaottelu
R	vastaajaryhmä
SOSEKAS	vanhempien sosioekonominen asema
TF_bear	trendinseuraamisen luokka, laskukaudet
TF_bull	trendinseuraamisen luokka, nousukaudet
TF_bear_CO	jako trendinseuraajiin ja toisinajattelijoihin, laskukaudet

Muuttuja	Selite
TF_bull_CO	jako trendinseuraajiin ja toisinajattelijoihin, nousukaudet
TF_bear_COstrong	jako vahvoihin trendinseuraajiin ja toisinajattelijoihin, laskukaudet
TF_bull_COstrong	jako vahvoihin trendinseuraajiin ja toisinajattelijoihin, nousukaudet
TS	trendinennustamislukokka ennusteen vahvuuden mukaan
UCI	ylempi luottamusväli, todennäköisyysjakauman yläarvo (H) vähennettynä piste-estimaatilla (F)
UCI _{7kk}	ylemmän luottamusvälin 7 kk:n ennusteen arvo
UCI _{13kk}	ylemmän luottamusvälin 13 kk:n ennusteen arvo