



Vaasan yliopisto
UNIVERSITY OF VAASA

Timi Lankinen

Vedonlyöntimarkkinoiden tehokkuus

Vedonlyöntimarkkinoiden tehokkuus suomalaisessa jääkiekossa:
empiirinen analyysi Liigan ottelukertoimista

Laskentatoimen ja rahoituksen akateeminen yksikkö
Taloustieteen maisteriohjelma
Pro gradu -tutkielma

Vaasa 2026

VAASAN YLIOPISTO**Laskentatoimen ja rahoituksen akateeminen yksikkö**

Tekijä:	Timi Lankinen		
Tutkielman nimi:	Vedonlyöntimarkkinoiden tehokkuus: Vedonlyöntimarkkinoiden tehokkuus suomalaisessa jääkiekossa: empiirinen analyysi Liigan ottelukertoimista		
Tutkinto:	Kauppätieteiden maisteri		
Koulutusohjelma:	Taloustiede		
Työn ohjaaja:	Teemu Pekkarinen		
Valmistumisvuosi:	2026	Sivumäärä:	85

TIIVISTELMÄ:

Tässä pro gradu -tutkielmassa tarkastellaan suomalaisen jääkiekkoliiga Liigan vedonlyöntimarkkinoiden tehokkuutta ja kertoimissa mahdollisesti esiintyviä systemaattisia vinoumia. Tutkimuksen lähtökohtana on markkinatehokkuuden hypoteesi, jonka mukaan tehokkailla markkinoilla hinnat heijastavat kaikkea saatavilla olevaa informaatiota eikä systemaattista ylituottoa voida saavuttaa julkisesti saatavilla olevan datan perusteella. Urheiluedonlyöntimarkkinat tarjoavat poikkeuksellisen selkeän ympäristön markkinatehokkuuden testaamiselle, koska tapahtumien lopputulokset ovat yksiselitteisiä ja realisoituvat nopeasti.

Tutkimuksen teoreettinen viitekehys rakentuu markkinatehokkuuden teorian ja käyttäytymistaloustieteen varaan. Keskeisiä käsitteitä ovat implisiittinen todennäköisyys, vedonvälittäjän marginaali, kalibraatio sekä suosikki-altavastaja-harha. Tutkimusaineisto koostuu Liigan runkosarjoittelusta kymmeneltä kaudelta vuodesta 2015 vuoteen 2025, sisältäen yli 4400 ottelua.

Aineistoa analysoidaan usealla tilastollisella menetelmällä. Kalibraatioanalyysissä verrataan kertoimista johdettuja implisiittisiä todennäköisyyksiä toteutuneisiin voittoosuuksiin. Kertoimien ennustetarkkuutta mitataan Brier Scorella ja ennustevoimaa multinomiaalisella logistisella regressiolla. Suosikki-altavastaja-harhaa sekä kotiedun hinnoittelua testataan strategiasimuloinnein ja tilastollisin testein. Kehittyneempien strategioiden osalta toteutetaan in-sample- ja out-of-sample-testaus luotettavamman arvion saamiseksi.

Tutkimuksen keskeinen havainto on, että Liigan vedonlyöntimarkkinat ovat kohtuullisen tehokkaat. Kertoimet ovat pääosin hyvin kalibroituneet ja sisältävät merkittävää ennustevoimaa, eikä yksikään testatuista vedonlyöntistrategioista tuota positiivista odotusarvoa. Viitteitä lievistä suosikki-altavastaja-harhasta havaitaan, mutta havaitut vinoumat ovat pieniä eivätkä tarjoa riittävää näyttöä systemaattisesta virhehinnoittelusta, sillä ne voivat selittyä myös marginaalin jakautumisella, satunnaisvaihtelulla tai muilla ottelukohtaisilla tekijöillä. Tulosten perusteella markkinatehokkuus ei ole yksinomaan suurten markkinoiden ominaisuus, vaan myös pienemmät markkinat voivat toimia kohtuullisen tehokkaasti. Tutkimus täydentää vedonlyöntimarkkinoiden tehokkuuskirjallisuutta tarjoamalla empiirisen analyysin pienemmän jääkiekkoliigan kertoimista.

AVAINSANAT: Urheiluedonlyönti, markkinatehokkuus, vedonlyöntikertoimet, implisiittinen todennäköisyys, jääkiekko

Sisällys

1	Johdanto	6
2	Teoreettinen viitekehys	8
2.1	Markkinatehokkuus	8
2.2	Käyttäytymistaloustiede ja systemaattiset vinoumat	10
3	Aikaisempi kirjallisuus	14
3.1	NHL- ja jääkiekkotutkimus	14
3.2	Vedonlyöntimarkkinat rahoitusteorian laboratoriona	18
3.3	Suosikki-altavastaja-harha	19
3.4	Metodologiset haasteet ja kotiedun hinnoittelu	22
3.5	Yhteenveto	25
4	Vedonlyöntimarkkinoiden institutionaalinen tausta	27
4.1	Online-vedonlyöntimarkkinoiden kehitys	28
4.2	Bookmaker-markkinat ja vedonlyöntipörssit	30
4.3	Kertoimien muodostuminen ja vedonvälittäjän marginaali	32
4.4	Vedonlyöjien motiivit ja käyttäytyminen	35
4.5	Liiga vedonlyöntikohteena	38
5	Aineisto	41
5.1	Otteluaineisto	41
5.2	Vedonlyöntikertoimet	44
5.3	Aineiston käsittely	46
6	Empiirinen analyysi	50
6.1	Vedonvälittäjän marginaali	50
6.2	Kalibraatioanalyysi	52
6.3	Suosikki-altavastaja-analyysi	58
6.4	Multinomiaalinen logistinen regressio	62
6.5	Kotietu	65
6.6	Vedonlyöntistrategioiden simulointi	69
6.6.1	Perusstrategiat	70

6.6.2	Kehittyneemmät strategiat	72
6.6.3	Yhteenveto strategioista	74
6.7	Yhteenveto	76
7	Johtopäätökset	79
	Lähteet	83
8	Liitteet	85
	Liite 1. Tekoälyn käyttö tutkielmassa	85

Kuviot

Kuvio 1. Kalibraatioanalyysi	57
Kuvio 2. Suosikki- ja altavastajastrategioiden kumulatiivinen tuotto	62

Taulukot

Taulukko 1. Liigan runkosarjaotteluiden jakautuminen kausittain	43
Taulukko 2. B365:n 1X2-kertoimien tunnusluvut koko tarkastelujakson aineistossa	45
Taulukko 3. Vedonvälittäjän marginaalin tunnusluvut kausittain (%)	51
Taulukko 4. Kotivoiton todennäköisyyksien kalibraatio	53
Taulukko 5. Tasapelien todennäköisyyksien kalibraatio	55
Taulukko 6. Vierasvoiton todennäköisyyksien kalibraatio	55
Taulukko 7. Suosikki- ja altavastajastrategioiden tuotot	59
Taulukko 8. Suosikki- ja altavastajastrategioiden ROI kausittain (%)	60
Taulukko 9. Multinomiaalinen logistinen regressio	64
Taulukko 10. Kotivoitto- ja vierasvoitto% sekä kotiedun hinnoitteluero kausittain	66
Taulukko 11. Kotiedun muutos koronapandemian aikana	67
Taulukko 12. Suosikkistrategiat eri raja-arvoilla	70
Taulukko 13. Altavastajastrategiat eri raja-arvoilla	71
Taulukko 14. Kehittyneempien strategioiden tuotot	74
Taulukko 15. Kaikkien strategioiden yhteenveto	75

Lyhenteet

B365

Bet365

1 Johdanto

Rahoitusmarkkinoiden tehokkuus on yksi modernin taloustieteen keskeisimmistä tutkimusaiheista. Markkinatehokkuuden hypoteesin (Efficient Market Hypothesis, EMH) mukaan hinnat heijastavat kaikkea saatavilla olevaa tietoa, eikä systemaattista ylituottoa voida saavuttaa ilman lisäriskiä. (Fama, 1970). Hypoteesia on testattu perinteisillä osake- ja valuuttamarkkinoilla, mutta viime vuosikymmeninä urheiluedonlyöntimarkkinat ovat nousseet merkittäväksi vaihtoehtoiseksi tutkimusympäristöksi. Vedonlyöntimarkkinoilla hinnat ilmenevät kertoimina, jotka voidaan tulkita implisiittisiksi todennäköisyyksiksi tietyn tapahtuman toteutumiselle. Tämä tekee urheiluedonlyöntimarkkinoista poikkeuksellisen selkeän tutkimusympäristön, jossa hinnoittelun tarkkuutta voidaan arvioida vertaamalla ennustettuja todennäköisyyksiä toteutuneisiin lopputuloksiin.

Urheiluedonlyöntimarkkinoiden erityispiirteenä on, että lopputulos on yksiselitteinen ja realisoituu lyhyessä ajassa. Toisin kuin rahoitusmarkkinoilla, joissa "todellinen arvo" voi olla teoreettinen ja vaikeasti mitattava. Urheilussa tapahtuma joko toteutuu tai ei toteudu. Tämän vuoksi vedonlyöntikertoimien kalibraatiota voidaan arvioida suoraan havaittujen voittoprosenttien avulla. Mikäli kertoimista johdetut implisiittiset todennäköisyydet vastaavat systemaattisesti toteutuneita voittotodennäköisyyksiä eikä yksinkertaisilla strategioilla voida saavuttaa positiivista odotusarvoa, voidaan markkinoita pitää tehokkaina. Moskowitz (2021) on korostanut, että vedonlyöntimarkkinat myös ratkaisevat rahoitusmarkkinoita vaivaavan yhteishypoteesiongelman, koska vetojen lopputulos on eksogeeninen markkinatoiminnalle ja realisoituu tarkasti määriteltynä hetkenä, jolloin mahdolliset hinnoitteluvirheet voidaan havaita suoraan ilman hinnoittelumallin samanaikaista testaamista.

Tässä tutkimuksessa tarkastellaan vedonlyöntimarkkinoiden tehokkuutta suomalaisen jääkiekkoliigan Liigan ottelukertoimissa kausien 2014–2025 aikana. Liiga edustaa kansainvälisesti tarkasteltuna keskikokoista ja suhteellisen rajattua vedonlyöntimarkkinaa, sillä se on ottelumäärältään sekä arvioidulta

vedonlyöntivolyymltaan huomattavasti pienempi kuin esimerkiksi jääkiekon NHL tai Euroopan suuret jalkapalloliigat. Tämä tekee siitä erityisen kiinnostavan tutkimuskohteen. Vedonlyöntimarkkinoiden koko ja likviditeetti voivat vaikuttaa informaation heijastumiseen kertoimiin, joten tuloksia ei voida suoraan yleistää suurempiin urheilusarjoihin. Tästä näkökulmasta Liiga tarjoaa mahdollisuuden arvioida, kuinka tehokkaasti informaatio välittyy pienemmässä markkinaympäristössä.

Tutkimuksen päätavoitteena on arvioida, heijastavatko Liigan 1X2-vedonlyöntikertoimet otteluiden todellisia voittotodennäköisyyksiä vai esiintyykö kertoimissa systemaattisia vinoumia. Lisäksi tutkimuksessa analysoidaan, onko yksinkertaisilla vedonlyöntistrategioilla mahdollista saavuttaa positiivinen odotusarvo sekä ovatko mahdolliset vinoumat muuttuneet tarkastelujakson aikana.

2 Teoreettinen viitekehys

Tämän tutkimuksen teoreettinen viitekehys rakentuu kahden toisiaan täydentävän taloustieteen teoria-alueen ympärille. Ensimmäinen on markkinatehokkuuden teoria, joka tarjoaa lähtökohdan sille, miten tehokkaiden markkinoiden tulisi toimia ja miten kertoimien tulisi heijastaa saatavilla olevaa informaatiota. Toinen on käyttäytymistaloustiede, joka tarjoaa selitysmallin sille, miksi markkinat voivat poiketa tehokkuudesta ja millaisia systemaattisia vinoumia markkinaosapuolten päätöksenteossa voidaan havaita. Yhdessä nämä teoriat muodostavat kehyksen, jonka avulla Liigan vedonlyöntimarkkinoiden tehokkuutta ja mahdollisia vinoumia arvioidaan empiirisessä analyysissä.

2.1 Markkinatehokkuus

Tehokkaiden markkinoiden hypoteesin (Efficient Market Hypothesis, EMH) mukaan tehokkailla markkinoilla hinnat heijastavat kaikkea saatavilla olevaa informaatiota (Fama, 1970). Faman mukaan tämän kaltaisilla markkinoilla yksittäinen toimija ei voi saavuttaa systemaattisesti ylituottoa hyödyntämällä julkista tai historiallista tietoa ja dataa. Markkinatehokkuuden hypoteesi on yksi rahoitusteorian keskeisimmistä teorioista.

Fama jakoi teorian kolmeen eri malliin informaation kattavuuden mukaan. Ensimmäisessä, heikon tehokkuuden -mallissa hinnat heijastavat kaikkea historiallista tietoa, kuten kaupankäyntihintoja ja -volyymeja. Tällöin tekninen analyysi ei mahdollista ylituottoa. Puoli vahvassa mallissa hinnat heijastelevat historiallisen tiedon lisäksi kaiken julkisesti saatavilla olevan tiedon, kuten tilinpäätökset, uutiset ja suhdanne-ennusteet. Vahvassa mallissa hinnat sisältävät näiden lisäksi myös yksityisen informaation, jolloin edes sisäpiiritieto ei tuottaisi systemaattista etua. Empiirisessä tutkimuksessa tarkastelu kohdistuu usein heikkoon tai puolivahvaan malliin, koska vahvan muodon testaaminen käytännössä on haastavaa.

Myöhemmin markkinatehokkuuden teoria kohtasi yhteishypoteesi eli niin sanotun joint hypothesis -ongelman. Yhteishypoteesi ongelmalla tarkoitetaan, että ei voida testata suoraan markkinatehokkuutta ilman että testataan samanaikaisesti jotakin hinnoittelumallia. Jos huomataan poikkeamia, ei voida yksiselitteisesti päätellä johtuuko havainto markkinoiden tehottomuudesta vai siitä, että käytetty hinnoittelumalli on puutteellinen. (Fama, 1991). Urheiluedonlyöntimarkkinat tarjoavat tähän ongelmaan poikkeuksellisen ratkaisun. Koska vetokohteiden arvo realisoituu eksogeenisesti ottelun lopputuloksessa, mahdolliset hinnoitteluvirheet voidaan havaita suoraan ilman samanaikaista hinnoittelumallin testausta. Tällöin yhteishypoteesiongelma ei koske vedonlyöntimarkkinoita samalla tavoin kuin perinteisiä rahoitusmarkkinoita (Moskowitz, 2021).

Urheiluedonlyöntimarkkinat tarjoavat kuitenkin poikkeuksellisen selkeän ympäristön markkinatehokkuuden arviointiin. Rahoitusmarkkinoilla omaisuuserän fundamentaalinen arvo voi olla teoreettinen ja vaikeasti mitattava, kun taas urheilutapahtumissa lopputulos on yksiselitteinen sekä havaittavissa huomattavasti lyhyemmän ajan kuluessa. Kertoimet voidaan tulkita tapahtuman implisiittiseksi todennäköisyyksiksi, jolloin tehokkuutta voidaan arvioida vertaamalla näitä todennäköisyyksiä toteutuneisiin voittoprosentteihin. Mikäli kertoimista johdetut todennäköisyydet ovat systemaattisesti kalibroituja eikä yksinkertaisilla vedonlyöntistrategioilla saavuteta positiivista odotusarvoa, voidaan pitää markkinoita tehokkaina ainakin Faman heikon ja puolivahvan mallin näkökulmasta.

Taloustieteellinen tutkimus vedonlyöntimarkkinoista on korostanut markkinarakenteen merkitystä. Raymond Sauerin (1998) mukaan vedonlyöntimarkkinat luovat luonnollisen laboratorion tehokkuuden testaamiselle, koska ne ovat kilpailullisia ja informaatio kulkee nopeasti markkinaosapuolten välillä. Samalla hän kuitenkin tuo esille huomion, että vedonvälittäjän rooli erottaa vedonlyöntimarkkinat perinteisistä rahoitusmarkkinoista. Vedonvälittäjä eli bookkeri ei välttämättä ainoastaan välitä vetoja tasapainottamalla riskinsä, vaan voi pyrkiä optimoimaan omat voittonsa asettamalla kertoimet strategisesti.

Hyun Song Shin (1993) puolestaan esitti teoreettisen mallin, jonka mukaan vedonlyöntimarkkinoilla voi esiintyä informaatioepäsymmetriaa. Hänen tutkimuksensa mukaan vedonvälittäjät voivat ottaa huomioon mahdollisen sisäpiiritiedon hinnoittelussaan, mikä vaikuttaa kertoimien marginaaleihin ja rakenteeseen. Tämä viittaa siihen, että kertoimet eivät ole pelkästään neutraaleja todennäköisysestimaatteja, vaan ne heijastavat myös markkinarakenteellisia tekijöitä ja riskienhallintaa.

Steven D. Levitt (2004) argumentoi, että vedonlyöntimarkkinat eroavat olennaisesti rahoitusmarkkinoista, koska vedonvälittäjän tavoitteena ei ole välttämättä saavuttaa tasapainotilaa, jossa riskit kumoutuvat täydellisesti vaan maksimoida voitto hyödyntämällä pelaajakäyttäytymistä. Tällöin kertoimet voivat heijastaa osittain kysyntäperusteisia vääristymiä, eivätkä ne välttämättä vastaa täysin todennäköisyyksiä.

Tässä tutkimuksessa markkinatehokkuudella tarkoitetaan sitä, että vedonlyöntikertoimista johdetut implisiittiset todennäköisyydet vastaavat toteutuneita voittotodennäköisyyksiä siten, ettei systemaattista positiivista odotusarvoa voida saavuttaa yksinkertaisilla ja toistettavilla strategioilla. Mikäli kertoimissa havaitaan systemaattisia ja taloudellisesti merkittäviä poikkeamia, voidaan tulkita, että markkinat eivät ole täysin tehokkaat. Jos taas poikkeamat ovat satunnaisia, tulokset tukevat käsitystä siitä, että vedonlyöntimarkkinat toimivat lähellä tehokasta tasapainoa.

2.2 Käyttäytymistaloustiede ja systemaattiset vinoumat

Perinteinen markkinatehokkuuden teoria perustuu oletukseen rationaalisista toimijoista, jotka päivittävät odotuksiaan johdonmukaisesti uuden informaation perusteella rationaalisten odotusten mukaisesti. Käyttäytymistaloustiede on kuitenkin osoittanut, että yksilöiden päätöksenteko poikkeaa systemaattisesti rationaalisesta. Erityisesti epävarmuuden ja riskin vallitessa havaitaan toistuvia vinoumia, jotka voivat vaikuttaa

markkinahintoihin. Vedonlyöntimarkkinat tarjoavat otollisen ympäristön näiden ilmiöiden tarkasteluun, sillä päätökset tehdään selkeästi määritellyissä riskitilanteissa ja lopputulos realisoituu nopeasti.

Prospektiteorian mukaan yksilöt eivät arvioi vaihtoehtoja odotetun hyödyn perusteella, vaan suhteessa vertailupisteeseen sekä painottavat todennäköisyyksiä epälineaarisesti. Yleensä pieniä todennäköisyyksiä yliarvioidaan ja suuria taas aliarvioidaan. Tämän lisäksi usein tappiot koetaan voimakkaampina kuin vastaavan suuruiset voitot. (Kahneman, 2012). Urheiluvedonlyönnin näkökulmasta tämä tarkoittaa, että epätodennäköiset, mutta suuria voittoja tarjoavat kohteet houkuttelevat suhteettoman paljon panoksia, vaikka niiden odotusarvo olisi negatiivinen.

Eräs tunnetuimmista vedonlyöntimarkkinoilla havaituista ilmiöistä on niin sanottu suosikki–altavastajaharha. Perinteisessä muodossaan se tarkoittaa, että altavastajien kertoimet ovat usein suhteessa liian matalia niiden todelliseen voittotodennäköisyyteen nähden, minkä seurauksena altavastajiin panostaminen tuottaa keskimäärin heikomman tuoton kuin suosikkeihin panostaminen (Snowberg & Wolfers, 2010). Ilmiötä on havaittu lukuisissa lajeissa ja markkinoissa. Snowberg ja Wolfers (2010) argumentoivat, että harhan ensisijainen selitys on todennäköisyyksien vääristyminen prospektiteorian mukaisesti eikä riskihakuisuus sinänsä. Vedonlyöjät siis vain uskovat virheellisesti altavastajan voiton olevan todennäköisempää kuin se todellisuudessa on. On kuitenkin tärkeää huomata, että harhan suunta vaihtelee markkinarakenteen mukaan. Newall ja Cortis (2021) osoittavat kirjallisuuskatsauksessaan, että kolmen lopputuloksen markkinoilla, kuten jalkapallon ja jääkiekon 1X2-markkinoilla, harha ilmenee tyypillisesti perinteisessä muodossaan, mutta kaksiulotteisilla moneyline-markkinoilla, kuten NHL:ssä, on havaittu käänteistä harhaa, jossa altavastajia aliarvioidaan ja suosikkeja yliarvioidaan (Woodland & Woodland, 2001).

Suosikki-altavastaja-harhalle on esitetty useita selityksiä. Käyttäytymistaloustieteen näkökulmasta ilmiö voi johtua siitä, että vedonlyöjät yliarvioivat pienten

todennäköisyyksien toteutumista ja hakevat suuria voittoja pienellä panoksella. Tällöin altavastajaat saavat paljon kysyntää, mikä taas painaa niiden kertoimia alaspäin. Ilmiö voidaan tulkita myös rationaalisten riskipreferenssien seuraukseksi, jos osa vedonlyöjistä on riskihakuisia suurten tuottojen suhteen. Tällöin altavastajien korkeampi tuottovarianssi voi tehdä niistä houkuttelevia, vaikka odotusarvo on matalampi. Shin (1993) selittää harhan vedonvälittäjän reaktiona sisäpiiritiedon uhkalle, kun taas Whelan (2024) osoittaa sen voivan syntyä kilpailullisilla markkinoilla vedonvälittäjien riskin kaihtamisen seurauksena ilman irrationaalisuusoletuksia.

Suosikki-altavastaja-harhan lisäksi toinen yleisesti havaittu ilmiö on yli- tai alireagointi informaatioon. Markkinaosapuolet voivat reagoida liiallisesti tuoreeseen informaatioon. Vaihtoehtoisesti markkinaosapuolet voivat alireagoida hitaasti kertyvään informaatioon. (Barberis, Shleifer & Vishny, 1998). Urheilussa tämän kaltaisia informaatioita voisi olla esimerkiksi voittoputki, kokoonpanomuutokset tai pitkän aikavälin tulokset. Urheiluviedonlyöntimarkkinoilla tämä näkyy esimerkiksi siten, että lyhyen voittoputken jälkeen joukkueen kertoimet muuttuvat enemmän kuin mitä todellinen todennäköisyysmuutos edellyttäisi. Tämänkaltaiset dynamiikat voivat aiheuttaa tilapäisiä hinnoitteluvirheitä, mikäli markkina ei korjaa niitä riittävän nopeasti.

Käyttäytymistekijät voivat myös kytkeytyä markkinarakenteeseen. Jos vedonvälittäjä asettaa kertoimet osittain kysynnän perusteella eikä pelkästään objektiivisten todennäköisyysarvioiden pohjalta, pelaajakäyttäytyminen voi heijastua suoraan hintoihin (Levitt, 2004). Tällaisessa tilanteessa kertoimet eivät ole puhtaita todennäköisysestimaatteja, vaan ne sisältävät komponentin, joka heijastaa kysyntäperusteisia vääristymiä. Tämä on keskeinen näkökulma arvioitaessa, ovatko havaitut vinoumat seurausta irrationaalisesta käyttäytymisestä vai markkinarakenteellisista tekijöistä.

Tässä tutkimuksessa käyttäytymistaloustieteellinen viitekehys toimii mahdollisena selitysmallina, mikäli Liigan vedonlyöntikertoimissa havaitaan systemaattisia poikkeamia

tai tilastollisesti merkitseviä tuottokuvioita eri todennäköisyysluokissa. Erityinen huomio kiinnittyy siihen, esiintyykö suosikki-altavastaaja-harhaa ja onko se ajassa pysyvä vai muuttuva ilmiö. Mikäli vinoumat ovat systemaattisia ja taloudellisesti merkittäviä, tulokset voivat viitata siihen, että käyttäytymistekijät tai markkinarakenteelliset piirteet estävät täyden informaation toteutumisen. Jos taas poikkeamat ovat pieniä, yksittäisiä tai katoavat ajan kuluessa, tukee tämä näkemystä siitä, että markkina korjaa käyttäytymisperäiset vääristymät suhteellisen nopeasti.

3 Aikaisempi kirjallisuus

Vedonlyöntimarkkinoiden tehokkuutta on tutkittu taloustieteessä jo useiden vuosikymmenten ajan. Tutkimuskenttä on kehittynyt hevoskilpailuiden parimutuelmarkkinoiden analyysistä kohti monipuolisempaa tarkastelua, joka kattaa nykypäivänä lukuisia urheilulajeja, markkinarakenteita ja erilaisia menetelmiä. Vaikka tutkimuksen painopiste on pitkään ollut suurissa ja likvideissä markkinoissa, kuten yhdysvaltalaisissa ammattilaisliigasarjoissa ja eurooppalaisessa jalkapallossa, on jääkiekko noussut omaksi tutkimuskohteikseen erityisesti NHL:ää käsittelevien tutkimusten kautta. Suomalaista Liigaa ei ole aiemmin tutkittu vedonlyöntimarkkinoiden tehokkuuden näkökulmasta, minkä vuoksi tämä tutkimus nojaa sekä NHL-spesifiseen kirjallisuuteen että laajempaan kansainväliseen tutkimukseen muista lajeista.

Aiempi kirjallisuus voidaan jakaa karkeasti neljään toisiaan täydentävään osaan. Ensimmäinen osa käsittää jääkiekkomarkkinoiden tehokkuutta suoraan tarkastelevat tutkimukset, joista keskeisimpiä ovat Woodland ja Woodlandin pitkä tutkimussarja sekä Paul ja Weinbachin NHL-analyysi. Toinen osa tarkastelee vedonlyöntimarkkinoita laajemmasta rahoitusteoreettisesta näkökulmasta, jota edustaa erityisesti Moskowitz (2021). Kolmas osuus koostuu suosikki–altavastaaajaharhan teoreettisesta perustasta ja empiiristä esiintymistä käsittelevistä tutkimuksista, erityisesti Newall ja Cortisin (2021) kirjallisuuskatsauksesta sekä Whelanin (2024) teoreettisesta mallista. Neljäs osa kattaa metodologiset ja kotietua käsittelevät tutkimukset, joita edustavat Winkelmann ym. (2024) sekä Fischer ja Haucap (2022). Yhdessä nämä erilaiset tutkimukset ja analyysit muodostavat teoreettisen ja empiirisen perustan, johon tämä tutkimus nojaa.

3.1 NHL- ja jääkiekkotutkimus

Jääkiekon vedonlyöntimarkkinoita on tutkittu akateemisessa kirjallisuudessa huomattavasti vähemmän kuin esimerkiksi amerikkalaisen jalkapallon tai hevoskilpailuiden vedonlyöntimarkkinoita. Jääkiekon NHL:n vedonlyöntimarkkinat ovat kuitenkin tarjonneet hyvän tutkimusympäristön erityisesti siksi, että jääkiekko poikkeaa

rakenteellisesti monista muista lajeista, sillä tulokset ovat suhteellisen tasaisia ja satunnaisempia kuin esimerkiksi jalkapallossa. Myös maalimäärien jakaumat eroavat merkittävästi ja pelimuoto on muuttunut olennaisesti siirryttäessä tasapeleihin päättyivistä otteluista jatkoaika- ja rangaistuslaukausratkaisuihin. Nämä piirteet tekevät jääkiekosta erityisen kiinnostavan kohteen vedonlyöntimarkkinoiden tehokkuuden kannalta.

Merkittävin jääkiekkomarkkinoiden tehokkuutta käsittelevä tutkimussarja on Linda ja Bill Woodlandin tutkimukset, jotka kattavat useita vuosikymmeniä ja useita eri lajeja. Woodland ja Woodland (1994) käsittelevät alun perin baseball-markkinoita ja havaitsivat yllättäen, että aiemmin hevoskilpailuissa dokumentoitu suosikki-altavastaaja-harha esiintyi baseball-markkinoilla käänteisessä muodossa. Tämä löydös osoitti ensimmäistä kertaa systemaattisesti, että käänteinen suosikki-altavastaajaharha on mahdollinen moneyline-markkinoilla ja loi perustan myöhemmälle jääkiekkotutkimukselle.

Woodland ja Woodland (2001) laajensivat tarkastelunsa NHL:ään kausien 1990–1996 kattavalla aineistoilla. Tutkimuksessa havaittiin, että NHL:n moneyline-markkinat ovat osittain tehottomat, ja yksinkertaisia vedonlyöntistrategioita tunnistettiin, jotka tuottivat positiivisia tuottoja. Baseball- ja jalkapallotutkimuksen tapaan havaittiin, että jääkiekkovedonlyöjät ylipelaavat suosikkeja suhteessa heidän havaittuun voittotodennäköisyyteensä. Kiinnostavaa on lisäksi se, että markkinat eivät näyttäneet konvergoituvan tehokkuuteen kauden edetessä. Tutkimuksen metodologinen lähestymistapa perustui vedonvälittäjän marginaalin laskemiseen moneyline-kertoimista ja implisiittisten voittotodennäköisyyksien johtamiseen, joita verrattiin toteutuneisiin voitto-osuuksiin.

Woodland ja Woodlandin (2001) tulosten merkittävyyttä arvioidessa on otettava huomioon, että heidän laskentamenetelmänsä sai myöhemmin kritiikkiä. Gandar, Zuber ja Johnson (2004) osoittivat, että Woodlandien tapa laskea yksikköpanokset suosikille ja altavastaajalle oli metodologisesti puutteellinen ja korjasivat laskelman. Tärkeää on

kuitenkin se, että vaikka korjauksen jälkeen havaittu harha heikkeni, se säilyi tilastollisesti merkitsevä. Tämä vahvisti Woodlandien perushavainnon, että NHL:ssä esiintyy käänteinen suosikki–altavastajaharha, jossa altavastajat ovat systemaattisesti alihinnoiteltuja suhteessa heidän todelliseen voittotodennäköisyyteensä.

Woodland ja Woodlandin (2010) maalimäärämarkkinoita tarkasteleva tutkimus laajensi kuvaa NHL:n tehottomuuksista uuteen suuntaan. Tutkimuksessa analysoitiin yli 5 000 ottelun aineistolla NHL:n maalimäärien yli tai alle -markkinoita ja havaittiin merkittävä tehottomuus. Alle-vetojen voittoprosentti oli 54,20 prosenttia, ja tehokkaan markkinan hypoteesi hylättiin erittäin vahvalla tilastollisella merkitsevyydellä. Tehottomuus oli erityisen selkeä korkeissa maalimäärätasoissa, joissa yli 5,5 maalia sisältävissä otteluissa alle-vedot olivat systemaattisesti alihinnoiteltuja. Tulos täydentää moneyline-tutkimuksia osoittamalla, että NHL:n markkinatehottomuudet eivät rajoitu voittajamarkkinoihin.

Woodland ja Woodland (2011) päivittivät NHL-moneyline-tutkimuksensa uudemmalla aineistolla kysyen suoraan otsikossaan, onko käänteinen suosikki–altavastajaharha edelleen olemassa. Vastaus oli myöntävä, sillä harha säilyi tilastollisesti merkitsevä uudemmassa aineistossa. Tämä ajallinen pysyvyys on teoreettisesti merkittävä löydös, sillä se viittaa siihen, että kyse ei ole satunnaisesta poikkeamasta vaan rakenteellisesta ilmiöstä, joka ei poistu markkinoiden oppimisen myötä.

Paul ja Weinbachin (2012) NHL-tutkimus täydentää Woodlandien tutkimuksia. Se lähestyy markkinatehokkuutta hieman erilaisesta näkökulmasta. Kun Woodland ja Woodland tarkastelivat tehokkuutta vedonlyöjän näkökulmasta, Paul ja Weinbach analysoivat vedonvälittäjän hinnoittelukäyttäytymistä käyttäen poikkeuksellista aineistoa. Heidän aineistonsa koostui todellisista vedonlyöntiosuuksista neljältä yhdysvaltalaiselta vedonvälittäjältä kolmelta eri NHL-kaudelta 2005–2008. Tämä mahdollisti sen testaamisen, asettaako vedonvälittäjä kertoimet tasapainottaakseen

vetokohteen panokset, maksimoidakseen voiton hyödyntämällä vedonlyöjien harhoja vai hinnoittellessaan ottelun tuloksen ennusteena.

Tutkimuksen keskeinen löydös oli, että NHL-vedonvälittäjät eivät pyri tasapainottamaan vetokohteen panoksia, sillä vedonlyöntiosuudet eivät ole proportionaalisia asetettuihin kertoimiin. Vedonlyöjillä havaittiin vahva preferenssi suosikkeja ja erityisesti vierassuosikkeja kohtaan. Tärkeintä on kuitenkin se, että vaikka vierassuosikkeja ylivetävä harha on selkeä ja tilastollisesti merkitsevä, vedonvälittäjät eivät näytä hyödyntävän tätä harhaa kertoimissaan, vaan hinnoittelu seuraa pikemminkin neutraalia ennustetta ottelun todellisesta lopputuloksesta. Paul ja Weinbach tulkitsevat tämän vedonvälittäjän strategisena valintana. Hinnoittelemalla kertoimet ennusteena pikemminkin kuin harhan eksplikoitina vedonvälittäjä estää informoituneita ammattivedonlyöjiä tunnistamasta ja hyödyntämästä hinnoitteluväristymiä. Pienillä markkinoilla, kuten NHL:ssä, tämä strategia suojaa vedonvälittäjää asiantunteilta vastapuolilta paremmin kuin voiton maksimointi lyhyellä aikavälillä.

Woodlandien tutkimussarjan sekä Paul ja Weinbachin tutkimuksen yhdistelmä tarjoaa tärkeän teoreettisen lähtökohdan tämän tutkimuksen kannalta kolmesta syystä. Ensinnäkin ne osoittavat, että jääkiekkomarkkinoilla harhan suunta voi erota muista lajeista. Siinä missä monissa lajeissa altavastaaajia yliarvioidaan, jääkiekon NHL:ssä on havaittu päinvastaista käyttäytymistä. Toiseksi ne paljastavat, että harha voi olla ajallisesti pysyvä eikä poistu markkinoiden kehittymisen myötä. Kolmanneksi ne muistuttavat tärkeästä rajoitteesta. Tässä tutkimuksessa käytetyt Liigan 1X2-kertoimet ovat kolmen lopputuloksen markkina, kun taas Woodlandien sekä Paul ja Weinbachin tutkimukset perustuvat kaksiulotteisiin moneyline-kertoimiin. Tämä rakenteellinen ero tarkoittaa, ettei NHL-tulosten suoraa soveltamista Liigaan voida pitää itsestään selvänä, ja tekee tästä tutkimuksesta itsenäisesti perustellun.

3.2 Vedonlyöntimarkkinat rahoitusteorian laboratoriona

Vaikka vedonlyöntimarkkinoita on tutkittu pitkään markkinatehokkuuden näkökulmasta, on sen rooli laajempaan rahoitusteoreettisena tutkimusympäristönä noussut keskeisemmäksi vasta viime vuosikymmeninä. Tähän kehitykseen on vaikuttanut erityisesti se, että vedonlyöntimarkkinat tarjoavat ratkaisun yhteen rahoitustieteen keskeisimmistä metodologisista ongelmista eli yhteishypoteesiongelmiaan. Perinteisillä osakemarkkinoilla havaittu poikkeama tehokkaista markkinoista voi johtua joko markkinoiden tehottomuudesta tai siitä, että käytetty hinnoittelumalli on puutteellinen, eikä näitä kahta selitystä voida empiirisesti erottaa toisistaan. Vedonlyöntimarkkinoilla tämä ongelma poistuu, koska vetosopimuksen arvo realisoituu eksogeenisesti ottelun lopputuloksessa, joka on täysin riippumaton markkinatoiminnasta.

Moskowitz (2021) hyödynsi tätä ominaisuutta omassa laajassa tutkimuksessaan. Tutkimuksen lähtökohtana oli testata rahoitusmarkkinoiden hinnoitteluanomaliaita vedonlyöntiympäristössä, jossa yhteishypoteesiongelmia ei vaikuta tulkintaan. Aineisto kattoi yli satatuhatta likvidiä vedonlyöntikohdetta lähes kolmen vuosikymmenen ajalta neljässä suuressa yhdysvaltalaisessa ammattiliigassa: NBA:ssa, NFL:ssä, MLB:ssä ja NHL:ssä. Tutkimus tarkasteli kolmea rahoitusmarkkinoilla tunnettua tekijää, momentumia, arvoa ja kokoa.

Vedonlyöntikontekstissa momentum tarkoittaa sitä, että viime aikoina hyvin menestyneet joukkueet hinnoitellaan suosituimmiksi kuin heidän pitkän aikavälin suorituksensa edellyttäisi. Arvo puolestaan viittaa siihen, onko joukkue alihinnoiteltu suhteessa objektiivisiin suoritustilastoihin, kuten pelaajien palkkabudjettiin tai muihin fundamentaalsiin tekijöihin. Koko mittaa joukkueen markkina-arvoa tai kotipaikkakunnan asukaslukua.

Moskowitz havaitsi vahvaa tilastollista näyttöä momentum-ilmioista. Vedonlyöntikertoimet liikkuvat avauksen ja sulkemisen välillä momentum-signaalien mukaisesti siten, että voittoputkessa olevat joukkueet hinnoitellaan suosituimmiksi kuin

heidän pitkän aikavälin tilastonsa tukisivat. Ratkaiseva löydös on kuitenkin se, että nämä kertoimien liikkeet kumoutuvat täysin ottelun lopputuloksessa. Momentum ei siis ennusta ottelun todellista lopputulosta paremmin kuin satunnaisesti. Tämä tarkoittaa, että vedonlyöjät reagoivat liialti joukkueiden viimeaikaiseen menestykseen, mikä on yhdenmukainen viivästyneen ylireagointimallin kanssa. Arvotekijä löytyi heikompana mutta suuntaa antavana, kun taas koon ei havaittu ennustavan tuottoja tilastollisesti merkitsevästi.

Tulostensa tulkinnassa Moskowitz on huolellinen. Hän korostaa, että vaikka momentum- ja arvoaikutukset ovat tilastollisesti merkitseviä, niiden suuruusluokka on noin viidesosa vastaavien ilmiöiden koosta osakemarkkinoilla. Tärkeintä on, että tuotot eivät riitä ylittämään transaktiokustannuksia, mikä estää arbitraasien eliminoimasta niitä. Markkinat ovat siis tehottomat, mutta tehottomuus on niin pieni, ettei se ole käytännössä hyödynnettävissä.

Tämän tutkimuksen kannalta Moskowitz (2021) on merkittävä kahdella eri tavalla. Teoreettisesti se tarjoaa vahvan perustelun sille, miksi vedonlyöntimarkkinoiden tutkiminen on taloustieteellisesti mielekästä muutenkin kuin urheilullisesta kiinnostuksesta. Nämä markkinat ovat ainutlaatuinen testi käyttäytymistaloudellisille teorioille. Empiirisesti se osoittaa, että momentum-tyyppinen ylireagointi esiintyy myös NHL:ssä osana laajempaa neljän lajin aineistoa. Mikäli Liigan kertoimissa havaitaan kausittaisia vaihteluita tai ajallisia muutoksia harhoissa, voidaan näitä verrata Moskowitzin löytämään dynamiikkaan. Lisäksi tutkimus muistuttaa tärkeästä tulkinnallisesta kehyksestä eli siitä, että vedonlyöntimarkkinoiden tehottomuudet ovat tyypillisesti niin pieniä, etteivät ne ylitä transaktiokustannuksia, ja tämä on otettava huomioon myös Liigan tulosten raportoinnissa.

3.3 Suosikki-altavastaja-harha

Suosikki-altavastaja-harha on yksi vedonlyöntimarkkinoilla eniten tutkittu ja dokumentoitu anomalia. Ilmiö on havaittu ensimmäistä kertaa hevoskilpailuissa jo 1940-

luvulla, ja sittemmin se on dokumentoitu lukuisissa lajeissa ja maissa. Harhan teoreettinen selittäminen on kuitenkin osoittautunut haastavaksi, ja kirjallisuudessa on esitetty useita kilpailevia ja toisiaan täydentäviä selitysmalleja. Tässä kappaleessa käydään läpi harhan empiirinen kirjallisuus ja sen keskeisimmät teoreettiset selitykset kahden viimeaikaisen tutkimuksen avulla.

Newall ja Cortis (2021) julkaisivat kattavan kirjallisuuskatsauksen suosikki-altavastaaja-harhasta, joka kokoaa yhteen laajan empiirisen tutkimuskentän ja tarjoaa sille uuden yhtenäistävän tulkinnan. Tutkimuksen lähtökohtana on ratkaista kirjallisuudessa pitkään vallinnut näennäinen ristiriita. Osa tutkimuksista löytää suosikkiharhaa, kun taas toiset altavastaajaharhaa, ja kokonaisuudessaan tutkimuskenttä on näyttänyt hajanaiselta ilman selkeää lopputulemaa.

Newall ja Cortis esittävät, että tutkimukset viittaavat siihen, että molemmat harhat voivat esiintyä samanaikaisesti vedonlyöntimarkkinoilla. Heidän keskeinen havaintonsa koskee markkinarakennetta. Kolmen lopputuloksen markkinoilla (1X2), joissa yksittäisen lopputuloksen todennäköisyys on tyypillisesti alle 50 prosenttia, harha ilmenee useimmiten perinteisessä muodossaan eli altavastaajia yliarvioidaan. Kaksiulotteisilla markkinoilla (moneyline), joissa suosikin voittotodennäköisyys ylittää 50 prosenttia, harha on tyypillisesti käänteinen eli suosikkeja yliarvioidaan. Rajoitettu kertoimien vaihteluväli yksittäisessä markkinassa selittää sen, miksi yksittäiset tutkimukset tukevat korkeintaan toista harhaa kerrallaan.

Käyttäytymistieteellisenä selitysmallina Newall ja Cortis korostavat kahta mekanismia. prospektiteorian mukainen todennäköisyyksien vääristyminen tuottaa molempien harhojen suuntaisia käyttäytymismalleja eri todennäköisyysalueilla. Pieniä todennäköisyyksiä yliarvioidaan, kun taas suuria aliarvioidaan, mikä johtaa altavastaajien ylipelaamiseen matalien todennäköisyyksien alueella ja suosikkien ylipelaamiseen korkeiden todennäköisyyksien alueella. Toiseksi urheilu tuottaa tarpeeksi vakuuttavia narratiiveja vahvistaakseen molemmat harhat, sillä poikkeukselliset altavastaajien voitot

jäävät mieleen ja houkuttelevat veikkaamaan liikaa altavastaajia, kun taas dominoivat suosikkijoukkueet voivat saada osakseen liiallista luottamusta.

Newall ja Cortisin kirjallisuuskatsaus on tämän tutkimuksen kannalta erityisen arvokas, koska se tarjoaa suoran ennusteen Liigan 1X2-markkinalle. Koska 1X2-markkina on kolmen lopputuloksen markkina, jossa yksittäisen lopputuloksen todennäköisyys on tyypillisesti alle 50 prosenttia, katsauksen kehyksen mukaan mahdollinen harha ilmeni todennäköisimmin perinteisessä muodossaan. Tämä on teoreettisesti erilainen ennuste kuin Woodlandin NHL-tutkimusten pohjalta tehtävä, jotka perustuvat kaksiulotteiseen moneyline-markkinaan ja joissa käänteinen harha on dominoiva löydös. Tässä mielessä Newall ja Cortisin kehys perustelee, miksi Liigan 1X2-markkina ansaitsee oman empiirisen tarkastelunsa eikä NHL-tulosten automaattinen soveltaminen ole perusteltua.

Whelan (2024) täydentää Newall ja Cortisin empiiristä katsausta tarjoamalla uuden teoreettisen selityksen sille, miksi suosikki–altavastaajaharha esiintyy nimenomaan kilpailullisilla kiinteiden kertoimien markkinoilla. Tämä on kirjallisuuden kannalta merkittävä selitys, sillä aiemmat teoreettiset selitykset on kehitetty pääosin hevoskilpailuiden parimutuel-markkinoita silmällä pitäen, eikä niitä voida suoraan soveltaa nykypäivän kilpailullisiin online-vedonlyöntimarkkinoihin.

Whelan rakentaa mallinsa kahdesta komponentista. Ensimmäinen on vedonlyöjien erimielisyys ottelun todennäköisyyksistä. Kun eri ihmiset arvioivat todennäköisyydet eri tavoin, eikä kukaan tiedä täysin oikeaa lopputulosta. Toinen ja teoreettisesti keskeisempi komponentti on vedonvälittäjien riskin kaihtaminen. Kun vahva suosikki kohtaa selkeän altavastaajan, suurin osa vedonlyöjistä haluaa lyödä vetonsa suosikille. Tämä tarkoittaa, että vedonvälittäjä kerää paljon vetoja suosikille, ja jos suosikki voittaa, vedonvälittäjä joutuu maksamaan paljon rahaa ulos ja kärsii siten suuret tappiot. Riskiä kaihtava vedonvälittäjä reagoi tähän heikentämällä suosikin kertoimia, mikä tarkoittaa suosikin implisiittisen todennäköisyyden nostamista yli sen todellisen arvon ja altavastaajan kertoimien parantamista vastaavasti. Tämä on juuri suosikki–altavastaajaharhan ydin.

Whelanin mallin tärkein teoreettinen johtopäätös on se, että harha voi syntyä täysin rationaalisessa ja kilpailullisessa ympäristössä ilman irrationaalisia vedonlyöjiä tai monopolistista vedonvälittäjää. Tämä asettaa kyseenalaiseksi yksinkertaisen tulkinnan, jonka mukaan harhan olemassaolo todistaa markkinoiden irrationaalisuuden. Whelanin selitys on yhteensopiva myös Shinin (1993) mallin kanssa, jossa harha syntyy vedonvälittäjän reaktiona sisäpiiritiedon uhkaan. Molemmat mallit ennustavat harhaa, mutta eri mekanismien kautta, ja todellisuudessa molemmat harhat saattavat olla läsnä samanaikaisesti.

Whelanin (2024) tutkimus on tämän tutkimuksen kannalta merkittävä erityisesti tulkinnallisesta näkökulmasta. Mikäli Liigan kertoimissa havaitaan suosikki-altavastaajaharhaa, löydös on yhdenmukainen useamman eri selityksen kanssa. Se voi kertoa vedonlyöjien prospektiteorian mukaisesta todennäköisyysvääristymisestä, vedonvälittäjän riskin kaihtamisesta tai sisäpiiritietojen vaikutuksesta. Pelkkä harhan havaitseminen ei riitä erottamaan näitä selityksiä toisistaan, mikä on tärkeä rajoite tuloksia tulkittaessa.

3.4 Metodologiset haasteet ja kotiedun hinnoittelu

Vedonlyöntimarkkinoiden tehokkuustutkimukseen liittyy useita metodologisia haasteita, jotka voivat vaikuttaa merkittävästi tulosten tulkintaan. Kaksi viimeaikaista tutkimusta käsittelee näitä haasteita erityisen selkeästi. Winkelmann, Ötting, Deutscher ja Makarewicz (2024) tarkastelevat tilastollisten testien luotettavuutta, kun taas Fischer ja Haucap (2022) analysoivat kotiedun hinnoittelua ja markkinoiden kykyä sopeutua ennakoimattomiin muutoksiin.

Winkelmann ja muut (2024) tekemässä tutkimuksessa kyseenalaistetaan suurta osaa aiemmasta vedonlyöntimarkkinoiden tehokkuutta tutkivasta kirjallisuudesta. Tutkimuksen mukaan, aiempi kirjallisuus antaa vaikutelman, että vedonlyöntimarkkinoiden tehottomuudet ovat laajalle levinneitä, mutta useimmat

tutkimukset nojaavat rajalliseen aineistoon eivätkä huomioi havaittujen harhojen ajallista pysyvyyttä. Tähän ongelmaan Winkelmann ja muut (2024) vastaavat yhdistämällä Monte Carlo -simulaatioihin perustuvan metodologisen analyysin laajaan empiiriseen aineistoon.

Simulaatioanalyysissä tutkijat generoivat keinotekoisia aineistoja, joissa markkinatehokkuus on täysin kontrolloitu. Tällä tavoin he voivat mitata, kuinka usein tavanomainen logistinen regressiotesti antaa väärän positiivisen tuloksen eli hylkää markkinatehokkuuden hypoteesin silloinkin, kun markkinat todellisuudessa ovat tehokkaat. Tulokset ovat huolestuttavia aiemman kirjallisuuden kannalta, sillä pienet otoskoot johtavat merkittävään tyyppin 1 virheen kasvuun, ja tilastollisesti merkitseviä tehottomuuksia havaitaan säännöllisesti myös täysin tehokkailla simuloituilla markkinoilla. Toisin sanoen yksittäisen kauden aineistoon perustuva tulos voi helposti olla tilastollinen harha eikä todellinen markkinatehottomuus.

Empiirisessä osuudessa Winkelmann ja muut (2024) analysoivat 14 peräkkäistä kautta viiden eurooppalaisen pääjalkapallosarjan aineistolla, yhteensä noin 25 000 ottelua. Tulos on selkeä. Vaikka tehottomuuksia esiintyy yksittäisillä kausilla, ne eivät ole pysyviä tai systemaattisia eri sarjoja tarkasteltaessa. Lisäksi simulaatioanalyysit osoittavat, että tilastollisesti merkitsevät vaikutukset yksittäisillä kausilla ovat odotettavissa myös täysin tehokkailla markkinoilla pelkästään satunnaisen vaihtelun seurauksena. Kirjallisuudessa laajasti raportoitujen tehottomuuksien taustalla saattaa siis olla osittain julkaisuvinouma. Tutkimukset, jotka löytävät merkitseviä vinoumia, päätyvät julkaistuiksi todennäköisemmin kuin tutkimukset, jotka eivät löydä mitään.

Winkelmann ja muut (2024) tekemä tutkimus on tämän tutkimuksen kannalta metodologisesti hyödyllinen. Tämän tutkimuksen Liigan aineisto kattaa kymmenen kautta ja noin 4 000–4 500 ottelua, mikä on huomattavasti laajempi kuin monissa muissa tutkimuksissa käytetyt yksittäisen kauden otokset. Tästä huolimatta Winkelmann ja muut (2024) tekemästä tutkimuksesta saatu varoitus pysyvyyden testaamisesta on

otettava vakavasti. Mikäli vinoumia havaitaan, on tärkeää testata erikseen, ovatko ne pysyviä kaikilla kausilla vai esiintyvätkö ne vain muutamalla yksittäisellä kaudella. Pelkkä tilastollinen merkitsevyys koko aineistossa ei riitä osoittamaan, että kyseessä on rakenteellinen ja hyödynnettävissä oleva tehottomuus. Tämä näkökulma ohjaa suoraan tämän tutkimuksen analyttistä lähestymistapaa. Tulokset raportoidaan sekä koko tarkastelujaksolle että kausikohtaisesti, jotta harhojen mahdollinen ajallinen vaihtelu voidaan tunnistaa.

Fischer ja Haucap (2022) lähestyvät markkinatehokkuutta eri näkökulmasta tarkastelemalla kotiedun hinnoittelua ja vedonlyöntimarkkinoiden kykyä sopeutua ennakoimattomiin shokkeihin. Tutkimus hyödyntää koronapandemiaa luonnollisena kokeena, joka mahdollisti kausaalisen päättelyn kotiedun lähteistä. Kun Saksan Bundesliiga palasi pelaamaan maaliskuussa 2020 täysin tyhjillä stadioneilla, tarjoutui ainutlaatuinen mahdollisuus erottaa yleisön vaikutus kotiedun muista osatekijöistä.

Tutkimuksen aineisto kattaa 6 120 ottelua kausien 2011/2012–2020/2021 ajalta Saksan kahdesta korkeimmasta jalkapallosarjasta. Difference-in-differences-regressiolla tutkijat havaitsivat, että kotiedun lasku oli selkeä Bundesliigassa. Noin kolmasosa aiemmasta kotiedusta katosi yleisön puuttuessa, mutta Bundesliiga 2:ssa taas muutosta ei havaittu lainkaan. Erityisen kiinnostavaa on se, että joukkueet, jotka olivat tottuneet suureen yleisömäärään, kärsivät enemmän ja tarvitsivat pidemmän ajan kotiedun palauttamiseksi, mikä viittaa osittain myös psykologiseen mekanismiin.

Markkinatehokkuuden kannalta Fischer ja Haucapin keskeisin löydös koskee vedonlyöntimarkkinoiden reagointinopeutta. Lyhyellä aikavälillä vedonlyöntikertoimet eivät heijastaneet asianmukaisesti tyhjien katsomoiden vaikutusta kotietuun, ja yksinkertaiset vedonlyöntistrategiat olisivat voineet hyödyntää tätä hinnoitteluvirhettä. Pitkällä aikavälillä markkinat kuitenkin sopeutuivat tehokkaasti. Kertoimet alkoivat heijastaa muuttunutta kotietua, ja liigakohtaiset erot hinnoiteltiin kertoimiin asianmukaisesti. Tämä lyhyen ja pitkän aikavälin ero on teoreettisesti merkittävä löydös,

joka on yhdenmukainen markkinoiden oppimisen kanssa. Sen mukaan vedonvälittäjät eivät osaa välittömästi hinnoitella täysin ennennäkemätöntä shokkia, mutta sopeutuvat nopeasti uuteen informaatioon.

Tämän tutkimuksen kannalta Fischer ja Haucap (2022) on mielenkiintoinen kahdella tavalla. Ensinnäkin kotiedun hinnoittelu on yksi Liigan tutkimuksen keskeisistä kysymyksistä, ja tähän Fischer ja Haucap tarjoavat suoran vertailukohtan jalkapallomarkkinoilta. Jääkiekossa kotiedun luonne voi erota jalkapallosta, koska yleisön rooli suljetussa hallissa saattaa olla erilainen kuin avoimella stadionilla, ja tämä tekee Liigan kotiedun hinnoittelun itsenäisestä tarkastelusta mielenkiintoisen. Toiseksi Fischer ja Haucapin löydös lyhyen aikavälin tehottomuudesta ja pitkän aikavälin sopeutumisesta tukee Winkelmann ja muut tekemää tutkimuksen argumenttia siitä, että markkinatehottomuudet voivat olla väliaikaisia ja katoavia eivätkä välttämättä rakenteellisia.

Yhdessä Winkelmann ja muut sekä Fischer ja Haucap muodostavat tärkeän metodologisen kehyksen tämän tutkimuksen tulosten tulkinnalle. Ne muistuttavat kahdesta keskeisestä varovaisuuden periaatteesta, tilastollinen merkitsevyys yksittäisellä kaudella ei todista pysyvää tehottomuutta, ja havaittu tehottomuus voi olla pikemminkin väliaikainen sopeutumisilmiö kuin rakenteellinen markkinavika. Nämä periaatteet ohjaavat suoraan sitä, miten tässä tutkimuksessa raportoitavia tuloksia tulkitaan ja mitä johtopäätöksiä niistä voidaan tehdä.

3.5 Yhteenveto

Aiemman tutkimuksen kokonaisuutta tarkasteltaessa voidaan todeta, että vedonlyöntimarkkinoiden tehokkuustutkimus on kehittynyt usealla rintamalla. Markkinatehokkuuden testaaminen on laajentunut yksinkertaisista kalibraatiotesteistä kohti monimutkaisempia ekonometrisia menetelmiä, ja tutkimuskenttä on siirtynyt tarkastelemaan yhä tarkemmin yksittäisiä vinoumia ja niiden taustalla olevia mekanismeja. Jääkiekon osalta keskeisimmän vertailukohtan tarjoavat Woodland ja

Woodlandin (1994, 2001, 2010, 2011) NHL-tutkimukset, joissa havaittiin viitteitä käänteisestä suosikki-altavastaaja-harhasta. Tämä tutkimus täydentää kyseistä tutkimusperinnettä tarkastelemalla vastaavia kysymyksiä eurooppalaisessa ja huomattavasti pienemmässä jääkiekkoliigassa.

Kotiedun hinnoittelun osalta Fischer ja Haucap (2022) ovat osoittaneet, että markkinoiden koko ja likviditeetti vaikuttavat siihen, kuinka tehokkaasti kotietu heijastuu kertoimiin. Tämä tarjoaa suoran vertailukohdan Liigan kotietuanalyysille, jossa markkinan pienemmän koon voidaan odottaa vaikuttavan hinnoittelun tarkkuuteen. Suosikki-altavastaaja-harhan osalta Snowberg ja Wolfers (2010) sekä Newall ja Cortis (2021) ovat dokumentoineet ilmiön laajuuden useilla eri markkinoilla, mutta jääkiekon osalta tulokset ovat olleet osin ristiriitaisia Woodlandien tutkimusten kanssa. Winkelmann ja muut (2024) puolestaan korostavat metodologisten valintojen merkitystä tehokkuustesteissä, mikä on ohjannut tämän tutkimuksen analyyttisiä ratkaisuja erityisesti kalibraatioanalyysin ja strategiasimulointien osalta.

Tämä tutkimus sijoittuu osaksi jatkuvasti kasvavaa kirjallisuutta, joka tarkastelee vedonlyöntimarkkinoiden tehokkuutta pienemmissä ja vähemmän tutkituissa sarjoissa. Liiga tarjoaa kiinnostavan tutkimusympäristön siksi, että se eroaa aiemmissä tutkimuksissa tarkastelluista markkinoista sekä lajin että markkinan koon puolesta. Suuri osa vedonlyöntimarkkinoiden tehokkuustutkimuksesta keskittyy jalkapalloon tai suuriin Pohjois-Amerikan ammattilaissarjoihin, joissa markkinoiden likviditeetti on huomattavasti korkeampi. Tutkimalla Liigan kertoimia voidaan arvioida, missä määrin markkinatehokkuus on universaali ilmiö vai riippuuko se markkinan koosta ja informaatioympäristöstä. Tältä osin tutkimus täydentää Moskowitzin esittämää näkemystä vedonlyöntimarkkinoista rahoitusteorian laboratoriona, jossa markkinatehokkuuden hypoteesia voidaan testata kontrolloidussa ympäristössä.

4 Vedonlyöntimarkkinoiden institutionaalinen tausta

Vedonlyöntimarkkinoiden toiminnan ymmärtäminen on ensiarvoisen tärkeää, kun lähdetään arvioimaan markkinoiden tehokkuutta ja kertoimien informaationsisältöä. Toisin kuin rahoitusmarkkinoilla, urheiluvedonlyöntimarkkinoilla hinnat eli kertoimet asetetaan vedonvälittäjien eli ns. bookkerien toimesta. Vedonvälittäjät toimivat siis markkinatakaajina ja vastapuolina vedonlyöjille. (Sauer, 1998). Tämän vuoksi kertoimet eivät ole pelkästään neutraaleja todennäköisyysestimaatteja, vaan ne heijastavat myös vedonvälittäjän riskienhallintaa, markkinakysyntää ja strategisia hinnoittelupäätöksiä (Humphreys, Rodriguez & Simmons 2017, s. 3–20).

Vedonlyöntimarkkinoiden institutionaalinen rakenne on muuttunut merkittävästi viimeisen kahden vuosikymmenen aikana erityisesti internetin ja digitaalisten vedonlyöntialustojen yleistymisen seurauksena. Online-vedonlyönti on lisännyt kilpailua vedonvälittäjien välillä, parantanut informaation saatavuutta sekä kasvattanut vedonlyöntivolyyymia monessa eri urheilulajissa. Nämä tekijät voivat vaikuttaa siihen, kuinka nopeasti uusi informaatio heijastuu kertoimiin ja kuinka tehokkaasti markkinat aggregoivat yksittäisten toimijoiden odotuksia. (Humphreys, Rodriguez & Simmons, s. 3–20).

Tässä luvussa tarkastellaan vedonlyöntimarkkinoiden institutionaalisia piirteitä, jotka voivat vaikuttaa markkinoihin erilaisin tavoin. Ensin käsitellään online-vedonlyöntimarkkinoiden kehitystä ja vedonlyöntimarkkinoita tänä päivänä sekä niiden merkitystä kilpailun ja likviditeetin kannalta. Tämän jälkeen tarkastellaan erilaisia vedonlyöntimarkkinoiden toimintatapoja, kuten vedonlyöntipörsejä sekä perinteisempiä bookmaker-markkinoita. Lisäksi analysoidaan vedonlyöntikertoimien muodostumista sekä vedonvälittäjien marginaalia. Lopuksi tarkastellaan vedonlyöjien käyttäytymistä sekä Liigan asemaa kansainvälisillä vedonlyöntimarkkinoilla.

4.1 Online-vedonlyöntimarkkinoiden kehitys

Urheiluvedonlyönti on pitkään ollut osa rahapelimarkkinoita, mutta markkinoiden rakenne on muuttunut merkittävästi erityisesti internetin yleistymisen myötä. Aiemmin vedonlyönti tapahtui pääasiassa fyysisissä vedonlyöntipisteissä tai kansallisten monopoliyhtiöiden kautta, mikä rajoitti kilpailua ja vedonlyöntitarjontaa. Internetti ja digitalisaatio mahdollisti vedonlyöntialustojen kehittymisen 1990-luvun lopulta lähtien. Tämä mahdollisti vedonlyöntipalveluiden tarjoamisen kansainvälisesti ja laajensi huomattavasti markkinoiden kokoa. Online-vedonlyönnin myötä markkinoille on tullut suuri määrä uusia toimijoita, mikä on lisännyt kilpailua ja parantanut vedonlyöjien mahdollisuuksia vertailla eri vedonvälittäjien tarjoamia kertoimia (Humphreys, Rodríguez & Simmons 2017, s. 3–20).

Online-vedonlyönnin yleistyminen on lisännyt vedonvälittäjien välistä kilpailua sekä tarjonnut vedonlyöjille pääsyn useiden eri toimijoiden kertoimiin samanaikaisesti. Vedonlyöjät voivat vertailla kertoimia eri vedonvälittäjien välillä ja valita itselleen parhaimman vaihtoehdon. Tämä on lisännyt markkinoiden läpinäkyvyyttä ja kaventanut vedonvälittäjien mahdollisuuksia ylläpitää suuria marginaaleja ilman, että vedonlyöjät siirtyvät kilpailijan palveluun. (Sauer 1998). Aiemmassa tutkimuksessa on esitetty, että vedonlyöntimarkkinoiden digitalisaatio on lisännyt hintakilpailua sekä nopeuttanut informaation leviämistä markkinoilla.

Digitalisaatio on myös lisännyt markkinoiden likviditeettiä. Suurempi vedonlyöjien määrä tarkoittaa, että markkinoille virtaa enemmän informaatiota ja panoksia, mikä voi parantaa kertoimien informaationsisältöä. Taloustieteellisessä kirjallisuudessa vedonlyöntimarkkinoita on usein tarkasteltu informaatiota aggregoivina markkinoina, joissa yksittäisten toimijoiden arviot tapahtumien todennäköisyyksistä yhdistyvät markkinahintaan eli vedonlyöntikertoimeen (Sauer 1998). Teoriassa suurempi likviditeetti ja kilpailu voivat lisätä markkinoiden tehokkuutta, koska hinnoitteluvirheet havaitaan ja korjataan nopeammin.

Samalla vedonlyöntiyhtiöt ovat kehittäneet entistä kehittyneempiä analyysimalleja ja tilastollisia työkaluja kertoimien asettamiseen. Vedonvälittäjät hyödyntävät usein tilastollisia ennustemalleja, historiallista dataa sekä markkinainformaatiota arvioidessaan otteluiden todennäköisyyksiä. Lisäksi kertoimia voidaan päivittää markkinoilta saadun informaation perusteella, kun vedonlyöjien panokset paljastavat uusia näkemyksiä ottelun todennäköisyyksistä. Vedonlyöntiyhtiöt ovat myös ruvenneet profiloimaan vedonlyöjiä, ja seuraavat etenkin voitollisten, hyvin menestyvien pelaajien pelikäyttäytymistä, ja mahdollisesti muokkaavat kertoimia heidän vetojensa perusteella. (Vaughan Williams, 2005. s. 11–30). Tämän vuoksi vedonlyöntimarkkinat muistuttavat joiltakin osin rahoitusmarkkinoita, joissa hinnat muuttuvat jatkuvasti uuden informaation perusteella.

Toisaalta online-vedonlyöntimarkkinoiden kasvu on myös muuttanut vedonvälittäjien toimintaympäristöä. Vedonvälittäjät seuraavat aktiivisesti kilpailijoiden kertoimia ja reagoivat nopeasti markkinamuutoksiin. Kertoimet voivat muuttua nopeasti uuden informaation, kuten loukkaantumisten tai kokoonpanomuutosten seurauksena. Lisäksi vedonvälittäjät voivat muuttaa kertoimia reagoidakseen vedonlyöjien panostuskäyttäytymiseen ja tasapainottaakseen riskejään eri lopputulosten välillä. (Humphreys, Rodríguez & Simmons 2017, s. 3–20).

Vedonlyöntimarkkinoiden globalisoituminen tarkoittaa myös sitä, että monet urheilusarjat, mukaan lukien pienemmät kansalliset liigat, ovat osa laajempaa kansainvälistä vedonlyöntimarkkinaa. Vaikka yksittäisen sarjan paikallinen kiinnostus voi olla rajallinen, vedonlyöntikohteet ovat usein tarjolla useiden kansainvälisten vedonvälittäjien kautta. Tämä voi lisätä markkinoiden informaationsisältöä ja parantaa hinnoittelun tehokkuutta myös pienemmissä sarjoissa. Toisaalta pienempien sarjojen kohdalla markkinoiden likviditeetti voi olla alhaisempi kuin suurissa kansainvälisissä liigoissa, mikä saattaa lisätä hinnoitteluvirheiden mahdollisuutta. (Humphreys, Rodríguez & Simmons 2017, s. 3–20).

4.2 Bookmaker-markkinat ja vedonlyöntipörssit

Urheiluviedonlyöntimarkkinoilla voidaan erottaa kaksi keskeistä toimintamallia, jotka ovat perinteiset bookmaker-markkinat ja vedonlyöntipörssit. Näiden kahden markkinan toimintalogiikka ja hinnan muodostus eroavat toisistaan merkittävästi, mikä voi vaikuttaa siihen, kuinka hyvin tapahtumat heijastavat tapahtumien todellisia todennäköisyyksiä. Tämän vuoksi on tärkeää ymmärtää millaisessa institutionaalisessa ympäristössä hinnat muodostuvat, kun lähdetään arvioimaan vedonlyöntimarkkinoiden tehokkuutta.

Perinteisessä bookmaker-mallissa vedonvälittäjä toimii vedonlyöjän vastapuolena. Vedonvälittäjä asettaa kertoimet eri lopputuloksille ja ottaa vastaan vedonlyöjien panoksia. Tavoitteena ei ole ainoastaan ennustaa tapahtuman todennäköisyyksiä mahdollisimman tarkasti, vaan myös hallita vedonvälittäjän omaa riskiä. Käytännössä tämä tarkoittaa sitä, että kertoimet sisältävät vedonvälittäjän marginaalin, jonka avulla yritys varmistaa positiivisen odotusarvon pitkällä aikavälillä (Sauer 1998). Marginaali syntyy siitä, että kaikkien mahdollisten lopputulosten implisiittisten todennäköisyyksien summa ylittää yhden.

Vedonvälittäjät käyttävät kertoimien asettamisessa tyypillisesti yhdistelmää tilastollisia malleja, asiantuntija-arvioita sekä markkinainformaatiota. Alkuperäiset kertoimet muodostetaan usein analysoimalla historiallista dataa, joukkueiden suorituskykyä sekä muita ottelun lopputulokseen vaikuttavia tekijöitä. Tämän jälkeen kertoimia voidaan muuttaa markkinoilta saadun informaation perusteella. Jos tiettyyn lopputulokseen kohdistuu paljon panoksia, vedonvälittäjä voi muuttaa kertoimia tasapainottaakseen riskinsä eri lopputulosten välillä (Vaughan Williams, 2005. s. 11-30). Tämän vuoksi kertoimet eivät ole pelkästään todennäköisysestimaatteja, vaan ne voivat heijastaa myös vedonlyöjien käyttäytymistä.

Vedonlyöntipörssit muodostavat vaihtoehdoisen markkinarakenteen. Vedonlyöntipörssissä vedonvälittäjä ei toimi vastapuolena vedoille, vaan alusta yhdistää

vedonlyöjät, jotka ovat valmiita ottamaan vastakkaiset näkemykset tietyn tapahtuman lopputuloksesta. Tällöin yksi vedonlyöjä voi esimerkiksi lyödä vetoa tietyn joukkueen voiton puolesta, kun taas toinen vedonlyöjä ottaa vedon vastaan ja panostaa kyseistä lopputulosta vastaan. Vedonlyöntipörssin ylläpitäjä kerää tyypillisesti komission toteutuneista voitoista, mutta ei kannaa suoraa riskiä vedonlyöntituloksista. (Smith, Paton & Vaughan Williams, 2006).

Vedonlyöntipörssien toimintaa on usein verrattu rahoitusmarkkinoihin, koska hinnat muodostuvat suoraan kysynnän ja tarjonnan vuorovaikutuksessa. Tämän vuoksi vedonlyöntipörssijä on pidetty potentiaalisesti tehokkaampina informaation aggregoijina kuin perinteisiä bookmaker-markkinoita. Kun markkinaosapuolet voivat vapaasti asettaa tarjouksia eri kertoimille, markkinahinta voi heijastaa tarkemmin osallistujien kollektiivista näkemystä tapahtuman todennäköisyydestä (Vaughan Williams, 2005).

Empiirinen tutkimus on tarkastellut myös sitä, miten bookmaker-markkinoiden ja vedonlyöntipörssien hinnat eroavat toisistaan. Joissakin tutkimuksissa on havaittu, että vedonlyöntipörssien hinnat voivat olla hieman paremmin kalibroituja suhteessa toteutuneisiin tuloksiin. Toisaalta erot markkinoiden välillä eivät aina ole suuria, mikä viittaa siihen, että myös perinteiset bookmaker-markkinat voivat aggregoida informaatiota suhteellisen tehokkaasti (Humphreys, Rodríguez & Simmons, 2017).

Vedonlyöntimarkkinoiden rakenteen kannalta on myös tärkeää huomioida vedonlyöntiyhtiöiden erilaiset liiketoimintamallit. Osa vedonvälittäjistä pyrkii aktiivisesti tasapainottamaan panoksia eri lopputulosten välillä, jolloin heidän riskinsä pienenee. Toiset vedonvälittäjät puolestaan hyväksyvät suurempaa riskiä ja keskittyvät tarjoamaan markkinoille kilpailukykyisiä kertoimia houkutellessaan ammattimaisia vedonlyöjiä. Useissa tutkimuksissa esimerkiksi Pinnaclen tai B365:n kaltaisia vedonvälittäjiä on pidetty erityisen informatiivisina alhaisten marginaalien ja korkeiden panosrajojen

vuoksi. Tässä tutkimuksessa käytetään B365:n kertoimia, koska ne ovat saatavilla kattavasti koko tarkastelujaksolta.

Yhteenvetona voidaan todeta, että vedonlyöntimarkkinoiden institutionaalinen rakenne vaikuttaa keskeisesti siihen, miten kertoimet muodostuvat ja kuinka hyvin ne heijastavat tapahtumien todellisia todennäköisyyksiä. Perinteisillä bookmaker-markkinoilla hinnanmuodostukseen vaikuttavat sekä vedonvälittäjän riskinhallinta että vedonlyöjien kysyntä, kun taas vedonlyöntipörssissä hinnat määräytyvät suuremmin kysynnän ja tarjonnan perusteella. Näiden erojen ymmärtäminen on tärkeää, kun arvioidaan vedonlyöntikertoimien informaation sisältöä ja markkinoiden tehokkuutta.

4.3 Kertoimien muodostuminen ja vedonvälittäjän marginaali

Vedonlyöntikertoimet ovat vedonlyöntimarkkinoiden hintoja. Ne ilmaisevat kuinka suuren palautuksen vedonlyöjä saa panokselleen, jos valittu lopputulos toteutuu. Taloustieteellisestä näkökulmasta kertoimet ovat kiinnostavia siksi, että ne voidaan tulkita tapahtumien implisiittisiksi todennäköisyyksiksi. Kyseinen tulkinta ei ole kuitenkaan täysin suoraviivainen, sillä lähtökohtaisesti aina vedonvälittäjän tarjoama kerroin sisältää marginaalin, jonka avulla vedonvälittäjä pyrkii varmistamaan itselleen positiivisen odotetun tuoton pitkällä aikavälillä. Tämän vuoksi vedonlyöntikertoimet eivät sellaisenaan vastaa puhtaita todennäköisyysarvioita, vaan ne sisältävät sekä informaatiota ottelun lopputuloksen odotetusta todennäköisyydestä että vedonvälittäjän kaupallisesta hinnoittelusta (Sauer, 1998). Vedonvälittäjän marginaalia voidaan havainnollistaa kolikonheitolla, jossa kummankin lopputuloksen todennäköisyys on 50 %. Todennäköisyyden käänteisarvon mukaan kerroin olisi tasan 2, mutta vedonvälittäjä tarjoaa esimerkiksi kertoimen 1.98 molemmille vaihtoehdoille. Tällöin implisiittiset todennäköisyydet ovat 0,505 ja niiden summa ylittää yhden. Tämä ylimenevä osa kuvastaa vedonvälittäjän marginaalia, tässä tapauksessa yhtä prosenttia.

Kiinteä kertoimisessa vedonlyönnissä vedonvälittäjä asettaa kertoimet ennen vedonlyöntikohteen sulkeutumista. Kertoimien lähtökohtana on arvio eri lopputulosten todennäköisyyksistä, mutta käytännössä hinnoittelu perustuu useiden tekijöiden yhdistelmään. Vedonvälittäjät hyödyntävät historiallista dataa, tilastollisia malleja, joukkueita ja pelaajia koskevaa informaatiota sekä markkinasignaaleja. Tämän lisäksi kertoimien asettamisessa voidaan ottaa huomioon vedonlyöjien odotettu käyttäytyminen. Kertoimien tehtävänä ei siis ole ainoastaan heijastaa "oikeaa" todennäköisyyttä, vaan myös tuottaa vedonvälittäjälle hinnoittelurakenne, joka on kilpailukykyinen mutta samalla riittävän kannattava. Vedonlyöntimarkkinoiden taloustieteellisessä kirjallisuudessa tätä on pidetty keskeisenä erona suhteessa sellaisiin markkinoihin, joissa hinnat muodostuvat puhtaasti kysynnän ja tarjonnan välisessä kaupankäynnissä (Humphreys, Rodríguez & Simmons, 2017; Smith, Paton & Vaughan Williams, 2006).

Implisiittinen todennäköisyys saadaan yksinkertaisimmillaan käänteislukuna kertoimesta. Jos kotivoiton kerroin on esimerkiksi 2,00, sitä vastaava raaka implisiittinen todennäköisyys on $1/2,00$ eli 0,50. Useamman lopputuloksen markkinoilla, kuten 1X2-vedonlyönnissä, näiden käänteislukujen summa ylittää kuitenkin tavallisesti yhden. Tämä ylitys kuvastaa vedonvälittäjän marginaalia. Jos esimerkiksi kotivoiton, tasapelin ja vierasvoiton kertoimista lasketut implisiittiset todennäköisyydet summautuvat arvoon 1,06, markkina sisältää kuuden prosentin marginaalin. Tutkimuskirjallisuudessa juuri tämän marginaalin poistaminen on välttämätön vaihe silloin, kun kertoimia käytetään otteluiden todennäköisyysestimaattien arviointiin tai markkinatehokkuuden testaamiseen (Sauer, 1998).

Vedonvälittäjän marginaali on keskeinen osa vedonlyöntiyhtiön liiketoimintamallia. Se toimii suojana sekä ennustevirheitä että epätasapainoista panosjakaumaa vastaan. Mitä suurempi marginaali on, sitä heikompi palautusaste vedonlyöjälle jää keskimäärin. Siksi alhaisen marginaalin vedonvälittäjiä pidetään tutkimuksissa usein parempina markkinan

kollektiivisen näkemyksen mittareina kuin korkeamman marginaalin toimijoita. (Humphreys, Rodríguez & Simmons, 2017).

Kertoimien muodostuminen on myös dynaaminen prosessi. Markkinaaan avattavat alkuperäiset kertoimet eivät yleensä säily muuttumattomina, vaan niitä päivitetään ennen ottelun alkua uuden informaation ja panosvirtojen perusteella. Jos markkina saa tietoa esimerkiksi avainpelaajan loukkaantumisesta, maalivahdin vaihtumisesta tai muusta ottelun voimasuhteita muuttavasta tekijästä, vedonvälittäjä voi muuttaa kertoimia nopeasti. Kertoimet voivat muuttua myös ilman yksittäistä uutistapahtumaa, jos vedonlyöjien panostus käyttäytyy tavalla, joka paljastaa markkinoille uutta informaatiota. Tällöin markkina toimii ainakin osittain informaatiota kokoavana mekanismina. Vaughan Williams. L (2005) korostaa, että vedonvälittäjien kertoimet sisältävät huomattavan määrän ennusteinformaatiota, mikä osaltaan selittää niiden laajan käytön urheilutulosten ennustamisessa.

Vedonvälittäjät avaavat kertoimia markkinoille hyvin vaihtelevasti, joskus useita viikkoja ja jopa kuukausia aiemmin, kun taas johonkin pienemmän lajin otteluun saatetaan avata vedonlyöntikohteet vasta muutamia tunteja ennen ottelun alkua. Kertoimet voidaan jaotella aukeamis- ja sulkeutumiskertoimiin. Kohteen avauskerroin viittaa niihin kertoimiin, joilla vedonlyöntikohde avataan markkinoille. Sulkeutumiskerroin puolestaan tarkoittaa juuri ennen kohteen sulkeutumista vallinneita kertoimia. Sulkeutumiskertoimia on usein pidetty informatiivisempina kuin avauskertoimia koska ne ehtivät sisällyttää enemmän hajautunutta markkinainformaatiota. Mitä lähemmäs ottelun alkua tullaan, sitä enemmän markkina ehtii reagoida uusiin tietoihin ja panosvirtoihin. Tämän vuoksi sulkeutumiskertoimet soveltuvat yleensä paremmin markkinatehokkuuden testaamiseen kuin varhaiset avauskertoimet, jos tutkimuksen tavoitteena on arvioida markkinan lopullista todennäköisyysarviota (Smith, Paton & Vaughan Williams, 2006; Forrest, McHale & McAuley, 2007). Tässä tutkimuksessa käytetään nimenomaan Liigan otteluiden sulkeutumiskertoimia.

Kertoimien muodostumista ei siis pidä ymmärtää puhtaasti mekaanisena todennäköisyysslaskentana. Useat tutkimukset korostavat, että vedonvälittäjät voivat hinnoitella markkinaa myös strategisesti. Jos vedonvälittäjä tietää tietyn tyyppisten kohteiden houkuttelevan erityisen paljon panoksia, se voi hinnoitella nämä kohteet tavalla, joka hyödyntää vedonlyöjien käyttäytymistäipumuksia. Tämä näkökulma on erityisen tärkeä suosikki–altavastaajajarhan tulkinnassa. Jos altavastaajat ovat systemaattisesti vedonlyöjien suosiossa suhteessa niiden todelliseen voittotodennäköisyyteen, vedonvälittäjällä voi olla kannustin pitää niiden kertoimet hieman ”liian matalina” kysynnän vuoksi. Tällöin havaittu vinouma ei välttämättä johdu yksinomaan informaation puutteesta, vaan myös strategisesta hinnanmuodostuksesta (Sauer 1998; Humphreys, Rodríguez & Simmons ,2017).

Tämän tutkimuksen kannalta kertoimien muodostumisen ja marginaalin ymmärtäminen on olennaista kahdesta syystä. Ensinnäkin tutkimuksessa käytettävät B365:n 1X2-kertoimet on muunnettava implisiittisiksi todennäköisyyksiksi tavalla, joka ottaa vedonvälittäjän marginaalin huomioon. Toiseksi mahdollisten systemaattisten vinoumien tulkinta edellyttää ymmärrystä siitä, että kertoimet voivat heijastaa samanaikaisesti sekä informaatiota ottelun lopputuloksesta että vedonvälittäjän kaupallisia ja riskienhallinnallisia tavoitteita. Tästä seuraa, että markkinoiden tehokkuutta ei tule arvioida ainoastaan tarkastelemalla kertoimia ennusteina, vaan myös huomioimalla institutionaalinen ympäristö, jossa nämä hinnat muodostuvat.

4.4 Vedonlyöjien motiivit ja käyttäytyminen

Vedonlyöntimarkkinoiden toimintaa ei voida ymmärtää pelkästään tarkastelemalla vedonvälittäjien hinnoittelua tai markkinarakenteita, sillä myös vedonlyöjien käyttäytymisellä on keskeinen rooli markkinoiden toiminnassa. Vedonlyöjien päätöksentekoon vaikuttavat sekä rationaaliset että käyttäytymistaloustieteessä tunnistetut psykologiset tekijät. Näiden tekijöiden seurauksena vedonlyöjien panostuskäyttäytyminen ei välttämättä perustu pelkästään odotusarvon maksimointiin,

vaan siihen voivat vaikuttaa esimerkiksi riskimieltymykset, viihdemotiivit ja erilaiset kognitiiviset vinoumat.

Taloustieteellisessä kirjallisuudessa vedonlyöntiä on tarkasteltu niin sijoitustoiminnan kuin viihdepelaamisen näkökulmasta. Osa vedonlyöjistä pyrkii systemaattisesti tunnistamaan kohteita, joissa vedonvälittäjän kertoimet eivät vastaa heidän omaa arviotansa tapahtuman todennäköisyydestä. Tällöin vedonlyönti vaikuttaa sijoitustoiminnalta, jossa vedonlyöjä pyrkii löytämään positiivisen odotusarvon kohteita. Samanaikaisesti suuri osa vedonlyöjistä osallistuu markkinoille kuitenkin ensisijaisesti viihdemotiivien vuoksi. Tällöin päätökset eivät välttämättä perustu tarkkaan todennäköisyysanalyysiin, vaan vedonlyönti voi liittyä esimerkiksi urheilun seuraamisen lisäämiseen tai jännityksen kokemiseen. (Humphreys, Rodríguez & Simmons, 2017).

Käyttäytymistaloustieteellinen tutkimus on osoittanut, että yksilöt arvioivat todennäköisyyksiä usein systemaattisesti vääristyneellä tavalla. Erityisesti pienten todennäköisyyksien tapahtumia voidaan yliarvioida, kun taas suurten todennäköisyyksien tapahtumia aliarvioidaan. Tämä ilmiö on keskeinen osa prospektiteoriaa, jonka mukaan yksilöt eivät arvioi riskejä lineaarisesti suhteessa niiden todennäköisyyksiin (Kahneman, 2012). Vedonlyöntimarkkinoilla tämä voi näkyä esimerkiksi siinä, että altavastaajat houkuttelevat suhteettoman paljon panoksia, koska niiden tarjoamat korkeat kertoimet tekevät mahdollisesta voitosta houkuttelevan.

Yksi tunnetuimmista vedonlyöntimarkkinoilla havaituista ilmiöistä on niin sanottu suosikki–altavastaajaharha. Useissa tutkimuksissa on havaittu, että altavastaajien kertoimet voivat olla suhteessa liian matalia niiden todelliseen voittotodennäköisyyteen nähden. Toisin sanoen vedonlyöjät ovat valmiita panostamaan altavastaajiin suuremmalla todennäköisyydellä kuin mitä puhtaasti rationaalinen odotusarvo edellyttäisi. Tämä ilmiö on dokumentoitu useilla eri vedonlyöntimarkkinoilla ja urheilulajeissa (Sauer, 1998). Suosikki–altavastaajaharhaa on selitetty sekä vedonlyöjien riskihakuisuudella että vedonvälittäjien strategisella hinnoittelulla.

Vedonlyöjien käyttäytymiseen voivat vaikuttaa myös informaation käsittelyyn liittyvät tekijät. Esimerkiksi tuoreisiin tapahtumiin liittyvä informaatio voi saada suhteettoman suuren painoarvon päätöksenteossa. Joukkueen viimeaikaiset voitot tai tappiot voivat vaikuttaa vedonlyöjien käsityksiin joukkueen todellisesta tasosta enemmän kuin mitä objektiivinen analyysi edellyttäisi. Tällaisia ilmiöitä on kuvattu muun muassa ylireagointina ja alireagointina informaatioon, ja ne ovat tuttuja myös rahoitusmarkkinoiden tutkimuksesta. (Barberis, Shleifer & Vishny, 1998).

Lisäksi vedonlyöjien päätöksentekoon voivat vaikuttaa sosiaaliset tekijät sekä joukkueisiin liittyvät tunnesiteet. Vedonlyöjät voivat esimerkiksi suosia kotimaansa tai suosikkijoukkueensa voittoa riippumatta siitä, millaisia kertoimia markkina tarjoaa. Tällainen käyttäytyminen voi aiheuttaa markkinoille kysyntävetoisia vääristymiä, mikäli tiettyihin kohteisiin kohdistuu jatkuvasti enemmän panoksia kuin niiden todennäköisyydet oikeuttaisivat.

Vedonlyöntimarkkinoiden tehokkuuden kannalta keskeinen kysymys on se, missä määrin tämän kaltaiset käyttäytymisperäiset vinoumat heijastuvat markkinahintoihin. Jos markkinoilla toimii riittävästi informoituja ja rationaalisia vedonlyöjiä, he voivat hyödyntää mahdollisia hinnoitteluvirheitä ja korjata ne nopeasti markkinahintaan. Tällöin markkinat voivat pysyä tehokkaina huolimatta yksittäisten vedonlyöjien irrationaalisesta käyttäytymisestä. Jos taas käyttäytymisperäiset vinoumat ovat laajalle levinneitä ja pysyviä, ne voivat johtaa systemaattisiin poikkeamiin kertoimien ja toteutuneiden todennäköisyyksien välillä.

Tässä tutkimuksessa tarkastellaan erityisesti vedonlyöjien käyttäytymistä mahdollisten vinoumien näkökulmasta. Jos Liigan vedonlyöntimarkkinoilla havaitaan esimerkiksi suosikki-altavastaja-harhaa tai muita systemaattisia tuottokuvioita, voivat ne viitata siihen, että vedonlyöjien käyttäytymisellä on vaikutusta markkinahintoihin. Toisaalta jos kertoimista johdetut todennäköisyydet vastaavat hyvin toteutuneita lopputuloksia eikä

yksinkertaisilla strategioilla saavuteta ylituottoa, tämä tukisi näkemystä siitä, että markkina pystyy aggregoimaan yksittäisten toimijoiden informaation tehokkaasti.

4.5 Liiga vedonlyöntikohteena

Suomalainen jääkiekkoliiga Liiga muodostaa kiinnostavan tutkimuskohteen vedonlyöntimarkkinoiden tarkasteluun. Kansainvälisessä vedonlyöntitutkimuksessa on analysoitu laajasti suuria urheilusarjoja, kuten Englannin Valioliigaa tai Pohjois-Amerikan ammattilaisliigoja, mutta pienempiä sarjoja koskeva tutkimus on ollut paljon vähäisempää. Tämän vuoksi Liiga tarjoaa mahdollisuuden tarkastella, miten vedonlyöntimarkkinat toimivat ympäristössä, jossa markkinan koko ja kansainvälinen kiinnostus ovat suhteellisen rajallisia verrattuna maailman suurimpiin urheilusarjoihin.

Liiga on Suomen korkein jääkiekon sarjataso ja yksi Euroopan merkittävimmistä jääkiekkosarjoista. Sarjan joukkue määrä vaihtelee tutkimuksen ajanjakson 2014–2024 14–16 joukkueen välillä. Joukkueet kohtaavat toisensa useamman kerran runkosarjassa ennen pudotuspelejä. Tässä tutkimuksessa käytetään vain runkosarjaotteluita eikä pudotuspeliotteluita ole huomioitu. Jääkiekko on Suomessa yksi suosituimmista ja seuratuimmista urheilulajeista, joten kotimaisilla vedonlyöjillä on usein hyvä tietopohja sarjan joukkueista, pelaajista ja ottelutilanteista.

Vedonlyöntimarkkinoiden näkökulmasta Liiga on kuitenkin huomattavasti pienempi kohde kuin esimerkiksi suurimmat jalkapalloliigat. Vedonlyöntivolyyymi on tyypillisesti alhaisempi, ja kansainvälisten vedonlyöjien kiinnostus sarjaa kohtaan voi olla rajallisempaa. Taloustieteellisessä kirjallisuudessa markkinoiden koko ja likviditeetti on usein liitetty markkinoiden tehokkuuteen. Suuremmilla markkinoilla toimii enemmän analytikoita ja ammattimaisia vedonlyöjiä, mikä voi lisätä kilpailua ja nopeuttaa hinnoitteluvirheiden korjaantumista. Pienemmillä markkinoilla vastaava informaatio ei välttämättä ole yhtä laajasti saatavilla, mikä voi teoriassa lisätä mahdollisuutta systemaattisiin hinnoitteluvirheisiin (Sauer, 1998).

Toisaalta vedonlyöntimarkkinoiden globalisoituminen on muuttanut tilannetta myös pienempien sarjojen osalta. Useat kansainväliset vedonvälittäjät tarjoavat nykyisin kertoimia myös sellaisiin sarjoihin, jotka ovat maantieteellisesti tai kaupallisesti rajatumpia. Tämä tarkoittaa, että myös Liigan ottelut ovat osa laajempaa kansainvälistä vedonlyöntimarkkinaa. Vedonlyöjät eri maista voivat osallistua markkinoille, mikä voi lisätä likviditeettiä ja parantaa tai heikentää kertoimien informaation sisältöä. Samalla vedonvälittäjät voivat hyödyntää kansainvälistä dataa ja tilastollisia malleja kertoimien asettamisessa.

Urheiluna jääkiekko tuo myös hieman erityispiirteitä vedonlyöntimarkkinoihin. Jääkiekko-otteluissa yksittäisten tapahtumien, kuten maalien määrä on suhteellisen pieni verrattuna esimerkiksi koripalloon. Tämä lisää lopputulosten satunnaisuutta ja voi vaikeuttaa otteluiden todennäköisyyksien tarkkaa arviointia. Tämän lisäksi esimerkiksi maalivahtien suoritustaso, pelaajien loukkaantumiset ja joukkueiden otteluruuhkat voivat aiheuttaa merkittävästi yksittäisten otteluiden voimasuhteisiin. Tällaiset tekijät voivat lisätä epävarmuutta kertoimien muodostumisessa ja mahdollisesti luoda tilanteita, joissa markkina ei täysin heijasta todellisia todennäköisyyksiä.

Liigan vedonlyöntimarkkinoita koskeva empiirinen tutkimus on kuitenkin suhteellisen rajallista. Suuri osa vedonlyöntimarkkinoita käsittelevästä kirjallisuudesta keskittyy jalkapalloon tai suuriin Pohjois-Amerikan sarjoihin, kuten jääkiekon NHL:ään tai amerikkalaisen jalkapallon NFL:ään, joissa vedonlyöntivolyymit ovat huomattavasti suurempia. Tästä syystä suomalaisen jääkiekkoliigan vedonlyöntikertoimien systemaattinen analyysi voi tuottaa uutta empiiristä tutkimusta siitä, miten vedonlyöntimarkkinat toimivat pienemmissä urheilusarjoissa.

Tässä tutkimuksessa Liigan vedonlyöntimarkkinoiden tehokkuutta tarkastellaan kausien 2014–2024 runkosarjaotteluiden aineiston avulla. Vedonlyöntikertoimina käytetään B365-vedonvälittäjän tarjoamia 1X2-kertoimia, jotka on kerätty Oddsportal-

verkkosivustolta. B365 on yksi maailman suurimmista urheiluedonlyöntiyhtiöistä ja sen kertoimet ovat laajasti saatavilla useille eri urheilusarjoille. Näin ollen niitä voidaan pitää edustavana mittarina vedonlyöntimarkkinoiden kollektiiviselle näkemykselle otteluiden todennäköisyyksistä.

Liigan tarkastelu tarjoaa siten mahdollisuuden arvioida, missä määrin vedonlyöntimarkkinoiden tehokkuus riippuu markkinan koosta, likviditeetistä ja informaatioympäristöstä. Mikäli kertoimista johdetut implisiittiset todennäköisyydet vastaavat hyvin toteutuneita tuloksia eikä yksinkertaisilla vedonlyöntistrategioilla saavuteta ylituottoa, tämä viittaisi siihen, että myös pienemmän urheilusarjan vedonlyöntimarkkinat voivat toimia suhteellisen tehokkaasti. Toisaalta systemaattiset vinoumat voisivat viitata siihen, että markkinan koko tai vedonlyöjien käyttäytyminen vaikuttaa hinnoitteluun.

5 Aineisto

Tässä tutkimuksessa analysoidaan suomalaisen jääkiekkoliiga Liigan ottelukertoimia ja ottelutuloksia usean kauden ajalta. Tutkimusaineisto koostuu ottelutuloksista sekä vedonlyöntikertoimista. Ottelutulokset kuvaavat pelattujen otteluiden lopputuloksia, kun taas kertoimet vedonlyöntimarkkinoiden ennakkokäsityksiä otteluiden todennäköisistä lopputuloksista.

Vedonlyöntimarkkinoiden tehokkuuden arviointi edellyttää näiden kahden tietolähteen yhdistämistä. Kertoimista voidaan johtaa implisiittiset todennäköisyydet eri lopputuloksille, minkä jälkeen näitä todennäköisyyksiä voidaan verrata toteutuneisiin ottelutuloksiin. Mikäli kertoimet heijastavat tehokkaasti markkinoilla saatavilla olevaa informaatiota, implisiittisten todennäköisyyksien tulisi vastata pitkällä aikavälillä toteutuneita lopputulosten suhteellisia osuuksia.

Tutkimusaineisto kattaa useita Liigan runkosarjakausia, mikä mahdollistaa vedonlyöntimarkkinoiden tehokkuuden tarkastelun myös ajallisesti. Usean kauden aineisto lisää analyysin luotettavuutta, koska yksittäisten otteluiden satunnaisvaihtelun vaikutus pienenee suuremmassa aineistossa.

5.1 Otteluaineisto

Tutkimuksessa käytetty otteluaineisto koostuu Liigan runkosarjaotteluista kymmeneltä kaudelta, kausien 2014–2015 ja 2024–2025 väliseltä ajalta. Aineisto sisältää kaikkien näiden kausien runkosarjaottelut, mikä tarkoittaa yhteensä 4495 ottelua. Ottelumäärä kausittain vaihtelee hieman, sillä kyseisen ajanjakson aikana Liiga on tehnyt erilaisia muutoksia sarjamalliin, joukkueiden määrä on vaihdellut sekä koronapandemian aikana vuosina 2020–2022 joitakin runkosarjaotteluita on peruttu ja jäänyt pelaamatta.

Otteluaineisto sisältää kunkin ottelun osalta keskeiset perustiedot, kuten ottelupäivän, kotijoukkueen, vierasjoukkueen sekä ottelun lopputuloksen. Lopputulos on luokiteltu kolmen mahdollisen lopputuloksen mukaan: kotivoitto, tasapeli tai vierasvoitto. Tasapeli tarkoittaa tässä yhteydessä tilannetta varsinaisen peliajan jälkeen, koska 1X2-vedonlyöntimarkkinoilla jatkoaika ja voittolaukauskilpailu eivät vaikuta vedon lopputulokseen.

Otteluaineisto on kerätty Liigan virallisista tilastoista. Viralliset ottelutiedot tarjoavat luotettavan lähteen toteutuneille lopputuloksille ja varmistavat, että aineisto kattaa kaikki sarjassa pelatut ottelut. Otteluaineisto muodostaa tutkimuksen empiirisen analyysin perustan, koska sen avulla voidaan arvioida, kuinka hyvin vedonlyöntikertoimet ennustavat otteluiden lopputuloksia.

Usean kauden aineiston käyttäminen on tärkeää myös siksi, että yksittäisiin otteluihin liittyy merkittävää satunnaisvaihtelua. Esimerkiksi jääkiekossa yksittäisten otteluiden lopputuloksiin voivat vaikuttaa monet tekijät, kuten loukkaantumiset, maalivahtien suoritukset tai ottelun aikana tapahtuvat sattumat. Suurempi otoskoko mahdollistaa sen, että tällainen satunnaisvaihtelu tasoittuu analyysissä.

Taulukossa 1 esitetään ottelumäärät sekä lopputulosten jakautuminen kausittain. Koko tarkastelujakson aikana pelattiin yhteensä 4495 runkosarjaottelua. Aineiston esikäsitelyssä havaittiin yksi selkeä datavirhe, jossa kotivoiton kerroin oli ilmeisen virheellinen. Kyseinen ottelutulos poistettiin analyysistä, joten lopullinen analysoitava aineisto sisältää 4494 ottelua. Ottelumäärä pysyi useimmilla kausilla 450 ottelussa, mutta kauden 2019–2020 ottelumäärä jäi 443:een ja kauden 2020–2021 ottelumäärä 424:ään. Nämä poikkeamat selittyvät koronapandemian aiheuttamilla otteluiden peruutuksilla. Tuoreimman kauden eli kauden 2024–2025 ottelumäärä on 480, joka on sarjamallin muutoksen takia hieman suurempi.

Taulukko 1. Liigan runkosarjaotteluiden jakautuminen kausittain

Kausi	Ottelut	Kotivoitot	Koti-%	Tasapelit	Tasapeli-%	Vierasvoitot	Vieras-%
2015–2016	450	200	44,4	103	22,9	147	32,7
2016–2017	450	188	41,8	113	25,1	149	33,1
2017–2018	450	204	45,3	110	24,4	136	30,2
2018–2019	450	202	44,9	107	23,8	141	31,3
2019–2020	443	178	40,2	96	21,7	169	38,1
2020–2021	424	179	42,2	93	21,9	152	35,8
2021–2022	447	183	40,9	100	22,4	164	36,7
2022–2023	450	205	45,6	104	23,1	141	31,3
2023–2024	450	193	42,9	111	24,7	146	32,4
2024–2025	480	208	43,3	111	23,1	161	33,5
Yhteensä	4 494	1 940	43,2	1 048	23,3	1 506	33,5

Taulukko 1 kertoo, että lopputulosten jakautuminen on pysynyt suhteellisen vakaana tarkastelujakson aikana. Kotivoitto on yleisin lopputulos kaikilla kausilla, ja koko aineistossa kotivoittojen osuus on 43,2 prosenttia otteluista. Vierasvoittoja on 33,5 prosenttia ja tasapelejä 23,3 prosenttia otteluista. Kausittaisessa tarkastelussa havaittavin poikkeus on kausi 2019–2020, jolloin kotivoittoprosentti laski 40,2 prosenttiin ja vierasvoittoprosentti nousi 38,1 prosenttiin. Tämä poikkeama ajoittuu koronapandemian alkamiseen, jonka seurauksena osa kauden loppuotteluista pelattiin ilman yleisöä. Kotiedun heikkeneminen yleisöttömissä otteluissa on havaittu myös aiemmassa tutkimuksessa (Fischer & Haucap, 2022).

5.2 Vedonlyöntikertoimet

Vedonlyöntikertoimina tutkimuksessa käytetään B365-vedonvälittäjän tarjoamia 1X2-kertoimia. B365 on yksi maailman suurimmista urheiluedonlyöntiyhtiöistä ja sen kertoimet ovat laajasti saatavilla useille eri urheilusarjoille. Tästä syystä B365:n kertoimia voidaan pitää hyvänä indikaattorina vedonlyöntimarkkinoiden kollektiiviselle näkemykselle otteluiden lopputulosten todennäköisyyksistä. Useissa vedonlyöntimarkkinoiden tehokkuutta käsittelevissä tutkimuksissa on käytetty juuri B365:n kertoimia niiden laajan saatavuuden ja markkinaa hyvin edustavan luonteen vuoksi.

Vedonlyöntikertoimet ovat kerätty Oddsportal-verkkosivustolta (Oddsportal, 2025), joka toimii vedonlyöntikertoimien vertailu sekä datasivustona. Oddsportal kerää ja arkistoi useiden vedonvälittäjien tarjoamia kertoimia, jonka vuoksi sitä käytetään usein historiallisten vedonlyöntikertoimien lähteenä tutkimuksissa. Sivusto mahdollistaa myös yksittäisten vedonvälittäjien kertoimien tarkastelun pitkällä aikavälillä, mikä on tämän tutkimuksen kannalta keskeistä.

Tutkimuksessa käytetään 1X2-kertoimia. 1X2-markkinassa pelaaja voi panostaa kolmeen mahdolliseen lopputulokseen: kotivoittoon (1), tasapeliin (X) tai vierasvoittoon (2). Jääkiekon tapauksessa tasapeli tarkoittaa tilannetta, jossa joukkueet ovat tasatilanteessa varsinaisen peliajan jälkeen, eikä jatkoaikaa tai voittolaukauskilpailua huomioida vedon lopputuloksessa. Kerätty aineisto sisältää jokaiseen 4494 otteluun kertoimet kotivoitolle, vierasvoitolle sekä tasapelille. Näin ollen aineisto sisältää yli 13 000 vedonlyöntikerrointa. Aineistossa olevat kertoimet ovat sulkeutumiskertoimia eli ne kuvastavat kertoimia juuri ennen kunkin ottelun alkua. Koska aineisto kattaa yhteensä kymmenen kauden runkosarjaottelut, aineisto muodostaa laajan kokonaisuuden vedonlyöntimarkkinoiden arvioimista ottelutodennäköisyyksistä. Jokaisesta ottelusta saatavat kolme kerrointa mahdollistavat myös eri lopputulosten erillisen tarkastelun sekä vedonlyöntistrategioiden simuloinnin analyysivaiheessa. Tämän rakenteen ansiosta

aineisto soveltuu hyvin vedonlyöntimarkkinoiden kalibraation, mahdollisten vinoumien sekä markkinatehokkuuden empiiriseen arviointiin.

Taulukossa 2 esitetään B365:n 1X2-kertoimien keskeiset tunnusluvut kotivoiton, tasapelin ja vierasvoiton osalta koko tarkastelujakson ajalta. Kotivoiton keskimääräinen kerroin on 2,34, mikä on implisiittisenä todennäköisyytenä noin 43 prosenttia. Tasapelin keskimääräinen kerroin on 4,28 ja vierasvoiton 3,27. Kertoimien mediaanit ovat kaikissa lopputulosluokissa lähellä keskiarvoja, mikä viittaa siihen, että kertoimien jakauma on kohtuullisen symmetrinen. Kotivoiton osalta mediaani 2,11 on kuitenkin selvästi pienempi kuin keskiarvo 2,34, mikä viittaa oikealle vinoutuneeseen jakaumaan. Aineistossa on siis jonkin verran poikkeuksellisen korkeita kotivoittokertoimia, jotka nostavat keskiarvoa mediaanin yläpuolelle. Sama ilmiö on havaittavissa vierasvoiton kertoimissa, joissa mediaani 2,90 on selvästi pienempi kuin keskiarvo 3,27.

Tasapelin kertoimet eroavat selvästi kotivoiton ja vierasvoiton kertoimista. Tasapelin keskimääräinen kerroin 4,28 on huomattavasti korkeampi kuin muiden lopputulosten kertoimet, mikä heijastaa tasapelin alhaisempaa todennäköisyyttä jääkiekko-otteluissa. Tasapelin kertoimien keskihajonta on lisäksi selvästi pienempi kuin muiden lopputulosten, vain 0,53, kun kotivoiton vastaava luku on 1,00 ja vierasvoiton 1,54. Tämä viittaa siihen, että tasapelin todennäköisyys vaihtelee otteluiden välillä selvästi kapeammalla välillä kuin kotivoiton tai vierasvoiton todennäköisyydet. Tämä tarkoittaa käytännössä sitä, että vedonvälittäjän näkemys tasapelin todennäköisyydestä on ottelusta toiseen huomattavasti vakaampi kuin näkemys siitä, kumpi joukkue voittaa.

Taulukko 2. B365:n 1X2-kertoimien tunnusluvut koko tarkastelujakson aineistossa

Lopputulos	Keskiarvo	Mediaani	Keskihajonta	Minimi	Maksimi
Kotivoitto	2,34	2,11	1,00	1,10	18,00
Tasapeli	4,28	4,20	0,53	3,30	12,00
Vierasvoitto	3,27	2,90	1,54	1,11	20,00

Kertoimien vaihteluvälit ovat laajoja erityisesti kotivoiton ja vierasvoiton osalta. Kotivoiton kerroin vaihtelee 1,10:stä 18,00:aan ja vierasvoiton kerroin 1,11:stä 20,00:aan. Nämä ääriarvot kuvastavat aineiston sisältävän sekä hyvin tasaväkisiä otteluita että otteluita, joissa toinen joukkue on ollut selkeä suosikki. Tasapelin vaihteluväli on huomattavasti kapeampi, 3,30:stä 12,00:aan, mikä on johdonmukainen tasapelin kertoimien alhaisen hajonnan kanssa.

Vedonlyöntikertoimien ja ottelutulosten yhdistäminen mahdollistaa vedonlyöntimarkkinoiden tehokkuuden empiirisen tarkastelun. Jos markkinat ovat tehokkaat, kertoimista johdettujen implisiittisten todennäköisyyksien tulisi vastata pitkällä aikavälillä toteutuneita lopputulosten osuuksia. Mahdolliset systemaattiset poikkeamat näiden välillä voivat viitata markkinoilla esiintyviin vinoumiin tai tehottomuuksiin.

5.3 Aineiston käsittely

Ennen varsinaista analyysia tutkimusaineisto käsiteltiin vaiheittain. Käsittelyn tavoitteena oli muuntaa raakamuotoiset vedonlyöntikertoimet analyysiin soveltuviksi implisiittisiksi todennäköisyyksiksi sekä rakentaa tarvittavat muuttujat kalibraatioanalyysia, suosikki-altavastaaja-analyysia ja eri vedonlyöntistrategioiden simulointia varten. Lisäksi aineisto tarkistettiin mahdollisten datavirheiden varalta.

Aineiston esikäsittelyssä havaittiin yksi selkeä datavirhe, jossa kotivoiton kerroin oli ilmeisen virheellinen poiketen huomattavasti kaikista muista aineiston kotivoittokertoimista. Kyseinen ottelu poistettiin analyysistä, jolloin lopullinen analyysiaineisto sisältää 4 494 ottelua. Datavirheiden tunnistaminen ja poistaminen on standardikäytäntö empiirisessä tutkimuksessa, ja yksittäisen havainnon poistaminen näin suuresta aineistosta ei vaikuta tulosten luotettavuuteen.

Aineiston käsittelyn ensimmäisessä vaiheessa kustakin ottelusta laskettiin raakamuotoiset implisiittiset todennäköisyydet kotivoitolle, tasapelille ja vierasvoitolle. Implisiittinen todennäköisyys saadaan laskemalla kertoimen käänteisluku. Lopputuloksen i raakamuotoinen implisiittinen todennäköisyys lasketaan kaavan (1) mukaisesti:

$$p_i = \frac{1}{o_i} \quad (1)$$

jossa o_i on lopputulosta i vastaava vedonlyöntikerroin. Jos kotivoiton kerroin on esimerkiksi 2,00, sitä vastaava raaka implisiittinen todennäköisyys on $1/2,00$ eli 0,50. Vastaavasti tasapelin kerrointa 4,50 vastaava implisiittinen todennäköisyys on $1/4,50 \approx 0,222$ ja vierasvoiton kerrointa 3,60 vastaava implisiittinen todennäköisyys on $1/3,60 \approx 0,278$.

Kuten aiemmin luvussa 4.3 todettiin, vedonvälittäjän kertoimet sisältävät marginaalin, jonka vuoksi kaikkien kolmen lopputuloksen raakamuotoisten implisiittisten todennäköisyyksien summa ylittää yhden. Vedonvälittäjän marginaali M lasketaan kaavan (2) mukaisesti:

$$M = \left(\frac{1}{o_1} + \frac{1}{o_x} + \frac{1}{o_2} \right) - 1 \quad (2)$$

jossa o_1 , o_x ja o_2 ovat kotivoiton, vierasvoiton ja tasapelin kertoimet. Jos implisiittisten todennäköisyyksien summa on esimerkiksi 1,067, marginaali on 6,7 prosenttia. Tässä tutkimuksessa marginaali on laskettu jokaiselle ottelulle erikseen, jolloin voidaan tarkastella marginaalin vaihtelua sekä otteluiden että kausien välillä.

Jotta kertoimista johdettuja todennäköisyyksiä voidaan käyttää markkinatehokkuuden arvioinnissa, raakamuotoiset implisiittiset todennäköisyydet on normalisoitava siten, että ne summautuvat yhteen. Normalisoitu implisiittinen todennäköisyys lasketaan kaavan (3) mukaisesti:

$$p_i = \frac{p^*i}{(p_1 + p_x + p_2)} \quad (3)$$

jossa p_i on lopputuloksen i raaka implisiittinen todennäköisyys ja nimittäjä on kaikkien kolmen lopputuloksen raakamuotoisten implisiittisten todennäköisyyksien summa. Normalisoinnin jälkeen kotivoiton, vierasvoiton ja tasapelin implisiittiset todennäköisyydet summautuvat yhteen, ja niitä voidaan käyttää vedonvälittäjän marginaalista poistettuina todennäköisysestimaatteina. Kyseinen normalisointimenetelmä on vakiintunut käytäntö vedonlyöntimarkkinoilla tehokkuustutkimuksien osalta (Sauer, 1998; Woodland & Woodland, 1994).

Toisessa vaiheessa kullekin ottelulle muodostettiin dummy-muuttujat toteutuneille lopputuloksille. Jokaiselle lopputulokselle luotiin erilliset binäärimuuttujat, jotka saavat arvon yksi, jos kyseinen lopputulos toteutui, ja muussa tapauksessa arvon nolla. Nämä muuttujat mahdollistavat normalisoitujen implisiittisten todennäköisyyksien vertailun vertaamisen toteutuneisiin lopputuloksiin, joka taas on kalibraatioanalyysin perusta.

Seuraavaksi aineisto jaettiin todennäköisyysluokkiin kalibraatioanalyysia varten. Kotivoiton ja vierasvoiton todennäköisyydet luokiteltiin 0,05:n levyisiin luokkiin, kun taas tasapelin todennäköisyydet luokiteltiin 0,02:n levyisiin luokkiin. Tasapeleihin valittiin kapeampi luokkaväli, koska tasapelien implisiittiset todennäköisyydet vaihtelevat selvästi kapeammalla välillä kuin kotivoiton tai vierasvoiton todennäköisyydet, jolloin 0,05:n luokkaväli tuottaisi liian harvoja luokkia. Kalibraatioanalyysissä huomioidaan vain ne todennäköisyysluokat, joissa havaintoja on vähintään 30. Tällä rajauksella pyritään vähentämään satunnaisvaihtelun vaikutusta yksittäisten luokkien tuloksiin.

Neljännessä vaiheessa muodostettiin suosikki-altavastaaja-analyysia varten tarvittavat muuttujat. Kussakin ottelussa suosikiksi määriteltiin se lopputulos, jolla on korkein normalisoitu implisiittinen todennäköisyys, ja altavastaajaksi se lopputulos, jolla on alhaisin normalisoitu implisiittinen todennäköisyys. Tämä luokittelu tehtiin sekä koko 1X2-markkinalle että erikseen joukkuepohjaista analyysia varten, jossa tasapeliä ei oteta huomioon. Jälkimmäisessä tarkastelussa suosikki ja altavastaaja määritellään pelkästään

kotivoiton ja vierasvoiton implisiittisten todennäköisyyksien perusteella. Tämä jako on perusteltu siksi, että suuri osa aineiston altavastaaajatilanteista on sellaisia, joissa korkein kerroin on tasapelillä. Tällöin pelkästään 1X2-markkinaan perustuva suosikki-altavastaaaja-jako ei välttämättä kuvaa joukkueiden välistä voimasuhdetta yhtä puhtaasti kuin joukkuepohjainen tarkastelu.

Seuraavaksi laskettiin eri vedonlyöntistrategioiden simulointia varten tarvittavat tuottomuuttujat. Kullekin ottelulle laskettiin nettotuotto kaavan (4) mukaisesti:

$$R_i = o_i - 1, \text{ jos veto osuu} \\ \text{tai } R_i = -1, \text{ jos veto ei osu} \quad (4)$$

jossa o_i on pelatun lopputuloksen kerroin. Tämä laskutapa vastaa tilannetta, jossa vedonlyöjä panostaa yhden yksikön kutakin ottelua kohti. Käytännössä siis, jos veto osuu, niin tuotto on kerroin miinus yksi, kun taas vedon hävitessä tuotto on miinus yksi. Strategian kokonaistuottoa mitataan sijoitetun pääoman tuottoasteella eli ROI:lla, joka on laskettu kaavan (5) mukaisesti:

$$ROI = \left(\frac{1}{N}\right) \sum R_i \quad (5)$$

jossa N on pelattujen vetojen kokonaismäärä ja R_i on yksittäisen vedon nettotuotto. ROI ilmaisee keskimääräisen tuoton tai tappion suhteessa panokseen. Strategia on keskimäärin tappiollinen, jos ROI on negatiivinen, kun taas positiivinen ROI viittaa siihen, että strategia tuottaa voittoa. Näitä tuottomuuttujia hyödynnetään myöhemmin sekä yksittäisten strategioiden kannattavuuden arvioinnissa että strategioiden välisten erojen tilastollisessa testauksessa.

Aineiston käsittely toteutettiin Excelissä ja Python ohjelmointikielellä hyödyntäen pandas-kirjastoa datan käsittelyyn sekä scipy- ja statsmodels-kirjastoja tilastolliseen analyysiin. Analyysissä käytetyt menetelmät ja niiden tulokset esitetään tarkemmin seuraavissa luvuissa.

6 Empiirinen analyysi

Seuraavaksi tässä luvussa esitellään tutkimuksen empiirinen analyysi sekä siitä saadut tulokset. Luku etenee yksi analyysi kerrallaan esittelemällä käytetyt menetelmät sekä niillä saadut tulokset. Analyysi alkaa vedonvälittäjän marginaalin tarkastelusta, minkä jälkeen arvioidaan kertoimien kalibraatiota ja ennustetarkkuutta. Tämän jälkeen tarkastellaan suosikki-altavastaaaja-harhaa, kotiedun hinnoittelua sekä kertoimien ennustevoimaa logistisen regression avulla. Lopuksi simuloidaan useita yksinkertaisia vedonlyöntistrategioita ja arvioidaan niiden tuottoja suhteessa markkinatehokkuuden viitekehukseen.

6.1 Vedonvälittäjän marginaali

Vedonvälittäjän marginaali kuvastaa sitä osuutta, jonka vedonvälittäjä pidättää itsellään eli osuutta, jolla vedonvälittäjä tekee omasta toiminnastaan kannattavaa. Marginaali lasketaan aiemmin esitetyn kaavan (2) mukaisesti raakamuotoisten implisiittisten todennäköisyyksien summana vähennettynä yhdellä. Mitä suurempi marginaali on, sitä heikompi odotusarvo vedonlyöjälle jää keskimäärin. Marginaalianalyysi on markkinatehokkuuden arvioinnin kannalta tärkeä lähtökohta, sillä se määrittää sen kynnyksen, jonka yläpuolelle vedonlyöjän on ylitettävä saavuttaakseen positiivisen odotusarvon.

Taulukossa 3 esitetään vedonvälittäjän marginaalin tunnusluvut kausittain. Koko tarkastelujakson aineistossa B365:n marginaalin keskiarvo on 6,72 prosenttia ja mediaani 6,64 prosenttia. Keskiarvon ja mediaanin läheisyys toisiinsa viittaa siihen, että marginaalin jakauma on kohtuullisen symmetrinen eikä yksittäisiä poikkeuksellisia havaintoja ole paljon. Marginaalin vaihteluväli on kuitenkin huomattava, sillä pienimmillään marginaali on 4,07 prosenttia ja suurimmillaan 12,49 prosenttia. Tämä vaihteluväli kuvastaa sitä, että yksittäisten otteluiden hinnoittelussa on merkittäviä eroja.

Kun marginaalin tunnuslukuja tarkastellaan kausittain, huomataan marginaalin keskiarvon vaihtelevan 6,33 prosentin ja 7,59 prosentin välillä. Alhaisin keskimääräinen marginaali on kaudella 2016–2017 ja korkein kaudella 2020–2021. Koronapandemian aikana kauden 2020–2021 marginaali nousi selvästi muita kausia korkeammaksi, mikä voisi viitata esimerkiksi siihen, että vedonvälittäjä kompensoi pandemian aiheuttamaa lisääntynyttä epävarmuutta kasvattamalla marginaaliaan. Myös yksittäisten otteluiden maksimimarginaali on korkeimmillaan juuri koronakausilla 2020–2021 ja 2021–2022, joiden maksimit ovat 12,38 ja 12,49 prosenttia. Kauden 2015–2016 maksimimarginaali 11,23 prosenttia on myös poikkeuksellisen korkea, mutta tämä saattaa heijastaa tarkastelujakson alkuvaiheen hinnoittelukäytäntöjä.

Taulukko 3. Vedonvälittäjän marginaalin tunnusluvut kausittain (%)

Kausi	Keskiarvo	Mediaani	Minimi	Maksimi
2015–2016	7,04	6,97	6,15	11,23
2016–2017	6,33	6,26	4,24	7,48
2017–2018	6,48	6,49	5,34	8,01
2018–2019	6,64	6,66	5,34	7,85
2019–2020	6,58	6,62	5,09	8,60
2020–2021	7,59	6,85	5,40	12,38
2021–2022	6,95	6,64	5,34	12,49
2022–2023	6,56	6,53	5,31	7,74
2023–2024	6,47	6,47	5,16	7,72

Kausi	Keskiarvo	Mediaani	Minimi	Maksimi
2024–2025	6,59	6,62	4,07	8,54
Koko aineisto	6,72	6,64	4,07	12,49

Liigassa vedonvälittäjän marginaalin taso on selvästi korkeampi kuin suurissa kansainvälisissä urheilusarjoissa, joissa marginaalit ovat tyypillisesti 4–6 prosentin tasolla. Tämä on linjassa aiemman kirjallisuuden kanssa, jonka mukaan pienempien ja vähemmän likvidien markkinoiden kertoimet sisältävät suuremman vedonvälittäjän marginaalin (Humphreys, Rodriguez & Simmons, 2017). Käytännössä korkeampi vedonvälittäjän marginaali tarkoittaa sitä, että vedonlyöntimarkkinoilla vedonlyöjän on vaikeampaa saavuttaa positiivista odotusarvoa. Tämä on tärkeä lähtökohta tietää myöhempää eri strategioiden simulaatiota varten.

6.2 Kalibraatioanalyysi

Kalibraatioanalyysin tavoitteena on arvioida, kuinka hyvin vedonlyöntikertoimista johdetut implisiittiset todennäköisyydet vastaavat toteutuneita lopputulosten osuuksia. Täysin kalibroituineilla markkinoilla ottelut, joissa kotivoiton implisiittinen todennäköisyys on esimerkiksi 40 prosenttia, pitäisi päättyä kotivoittoon juuri 40 prosentin todennäköisyydellä pitkällä aikavälillä. Systemaattiset poikkeamat tästä suhteesta viittaavat markkinoiden virheelliseen hinnoitteluun.

Kalibraatioanalyysi toteutettiin ryhmittelemällä ottelut normalisoitujen implisiittisten todennäköisyyksien perusteella todennäköisyysluokkiin. Kotivoiton ja vierasvoiton osalta käytettiin 0,05:n levyisiä luokkia, mikä tuottaa riittävän määrän havaintoja kutakin luokkaa kohti samalla kuin luokkien määrä pysyy helposti tulkittavana. Tasapelin osalta käytettiin kapeampaa 0,02:n luokkaväliä, koska tasapelin implisiittiset todennäköisyydet

vaihtelevat huomattavasti kapeammalla välillä kuin kotivoiton tai vierasvoiton todennäköisyydet. Leveämpi 0,05:n luokkaväli tuottaisi tasapelille vain muutaman luokan, mikä heikentäisi analyysin tarkkuutta. Molemmissa tapauksissa luokkavälin valinnassa pyrittiin tasapainottamaan luokkien riittävä havaintomäärä ja analyysin tarkkuus. Kustakin luokasta laskettiin sekä keskimääräinen implisiittinen todennäköisyys että toteutunut voitto-osuus. Tilastollinen merkitsevyys testattiin kaksipuoleisella z-testillä, joka testaa eroaako toteutunut osuus tilastollisesti merkitsevästi implisiittisestä todennäköisyydestä. Analyysissa huomioitiin vain ne luokat, joissa havaintoja on vähintään 30, jotta satunnaisvaihtelun vaikutus tuloksiin olisi mahdollisimman pieni.

Taulukossa 4 esitetään kotivoiton kalibraatioanalyysin tulokset. Koko aineiston koossa kotivoiton keskimääräinen implisiittinen todennäköisyys on 44,5 prosenttia ja toteutunut kotivoitto-osuus on 43,2 prosenttia. Kertoimet siis yliarvioivat kotivoiton todennäköisyyden, mutta ero on todella pieni. Yksittäisiä todennäköisyysluokkia tarkastelemalla havaitaan, että p-arvot ovat kaikissa luokissa selvästi yli 0,05, mikä tarkoittaa, että yhdessäkään luokassa ero implisiittisen todennäköisyyden ja toteutuneen osuuden välillä ei ole tilastollisesti merkitsevä. Kotivoiton kertoimet ovat siten kohtuullisen hyvin kalibroituineet koko todennäköisyyskaalalla.

Taulukko 4. Kotivoiton todennäköisyyksien kalibraatio

Todennäköisyysluokka	Ennustettu todennäköisyys	Toteutunut osuus	Havaintojen määrä	p-arvo
0,10	13,3 %	19,4 %	31	0,323
0,15	17,8 %	9,1 %	88	0,033
0,20	22,8 %	19,3 %	150	0,310
0,25	27,7 %	24,5 %	278	0,228
0,30	32,7 %	31,6 %	462	0,613

Todennäköisyysluokka	Ennustettu todennäköisyys	Toteutunut osuus	Havaintojen määrä	p-arvo
0,35	37,4 %	35,5 %	586	0,345
0,40	42,6 %	43,2 %	755	0,748
0,45	47,6 %	46,0 %	652	0,431
0,50	52,4 %	51,5 %	582	0,667
0,55	57,3 %	54,5 %	376	0,269
0,60	62,2 %	62,4 %	242	0,951
0,65	67,0 %	64,2 %	173	0,426
0,70	72,1 %	71,2 %	66	0,868
0,75	77,2 %	80,6 %	31	0,644
Yhteensä	44,5 %	43,2 %	4 472	

Taulukossa 5 esitetään tasapelin kalibraatioanalyysin tulokset. Taulukosta on poistettu luokat, joissa havaintojen määrä on alle 30. P-arvo on laskettu kaksipuoleisella z-testillä. Tasapelin keskimääräinen implisiittinen todennäköisyys on 22,2 prosenttia ja toteutunut tasapeliosuus on 23,3 prosenttia. Kertoimet siis aliarvioivat tasapelin todennäköisyyden hieman, eli tasapelejä tapahtuu käytännössä hieman useammin kuin kertoimet ennakoivat. Myöskään tasapelin osalta yksikään todennäköisyysluokka ei tuota tilastollisesti merkitsevää eroa, joskin luokka 0,20 on lähellä merkitsevyysrajaa p-arvolla 0,116. Tasapelin kertoimet ovat siis myös melko hyvin kalibroituineet, vaikka pieni systemaattinen alihinnoittelu on havaittavissa.

Taulukko 5. Tasapelien todennäköisyyksien kalibraatio

Todennäköisyysluokka	Ennustettu todennäköisyys	Toteutunut osuus	Havaintojen määrä	p-arvo
0,14	15,1 %	12,2 %	49	0,578
0,16	17,2 %	14,6 %	192	0,328
0,18	19,2 %	23,5 %	324	0,055
0,20	21,2 %	23,3 %	974	0,116
0,22	22,9 %	23,5 %	2319	0,494
0,24	24,6 %	26,2 %	587	0,348
Yhteensä	22,2 %	23,3 %	4 445	

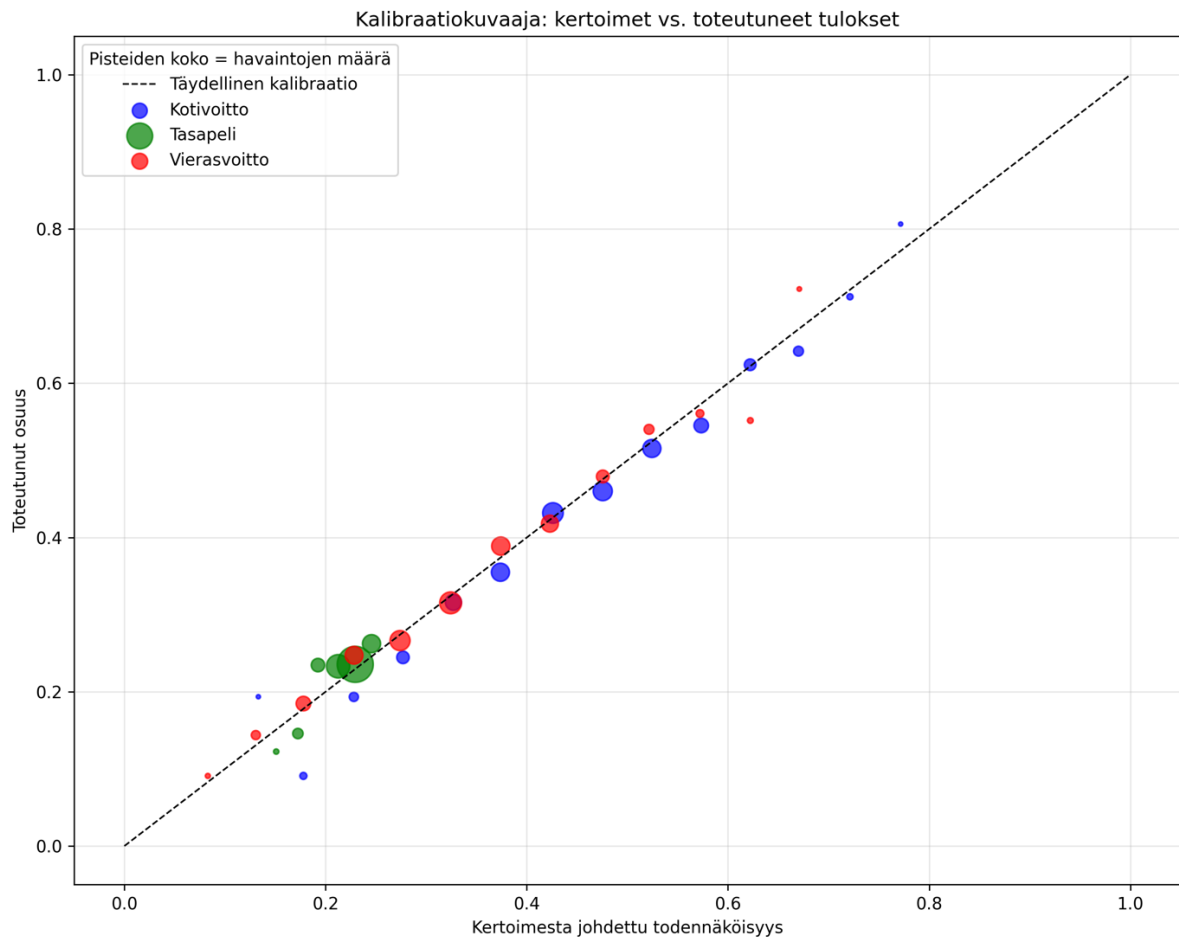
Taulukossa 6 esitetään vierasvoiton kalibraatioanalyysin tulokset. Taulukosta on poistettu luokat, joissa havaintojen määrä on alle 30. P-arvo on laskettu kaksipuoleisella z-testillä. Vierasvoiton keskimääräinen implisiittinen todennäköisyys on 33,1 prosenttia ja toteutunut vierasvoittojen osuus 33,3 prosenttia, eli ero on todella pieni. Vierasvoiton kertoimet ovat kaikista kolmesta lopputuloksesta parhaiten kalibroittuneet, ja p-arvot ovat niin ikään kaikissa luokissa selvästi yli 0,05. Tulos viittaa siihen, että B365 hinnoittelee vierasvoiton todennäköisyyden erittäin tarkasti suhteessa toteutuneisiin lopputuloksiin.

Taulukko 6. Vierasvoiton todennäköisyyksien kalibraatio

Todennäköisyysluokka	Ennustettu todennäköisyys	Toteutunut osuus	Havaintojen määrä	p-arvo
0,05	8,3 %	9,1 %	44	0,848
0,10	13,1 %	14,4 %	146	0,636
0,15	17,8 %	18,5 %	379	0,732

Todennäköisyysluokka	Ennustettu todennäköisyys	Toteutunut osuus	Havaintojen määrä	p-arvo
0,20	22,8 %	24,7 %	562	0,283
0,25	27,4 %	26,6 %	717	0,646
0,30	32,4 %	31,5 %	850	0,575
0,35	37,4 %	38,9 %	594	0,458
0,40	42,3 %	41,8 %	512	0,818
0,45	47,6 %	47,9 %	288	0,903
0,50	52,2 %	54,0 %	174	0,621
0,55	57,2 %	56,1 %	107	0,811
0,60	62,2 %	55,2 %	58	0,268
0,65	67,1 %	72,2 %	36	0,512
Yhteensä	33,1 %	33,3 %	4 467	

Kalibraatioanalyysin tuloksia havainnollistetaan kuviossa 1, jossa esitellään kaikkien kolmen lopputuloksen implisiittiset todennäköisyydet suhteessa toteutuneisiin osuuksiin. Täydellinen kalibraatio vastaa mustalla katkoviivalla merkittyä $y=x$ -suoraa, jolloin kertoimista johdettu todennäköisyys vastaisi täsmälleen toteutunutta osuutta. Pisteiden koko kuvaa havaintojen määrää kussakin todennäköisyysluokassa. Kuvioista havaitaan, että suurin osa pisteistä sijoittuu lähelle täydellisen kalibraation suoraa. Erityisesti suurimmat pisteet, joissa havaintoja on eniten, osuvat lähimmäksi suoraa, mikä on tilastollisesti odotettua. Pienet poikkeamat näkyvät lähinnä matalissa todennäköisyysluokissa, joissa havaintojen määrä on pieni ja satunnaisvaihtelu suurta.



Kuvio 1. Kalibraatioanalyysi

Kertoimien ennustetarkkuutta mitattiin myös brier scoren avulla. Brier score on ennustetarkkuuden kokonaismittari, joka lasketaan ennustettujen todennäköisyyksien ja toteutuneiden lopputulosten välisten neliöityjen virheiden keskiarvona. Pienempi brier score viittaa parempaan ennustetarkkuuteen. Koko aineiston kokonais- brier score on 0,6153, kun taas naiivilla mallilla, joka käyttää pelkästään toteutuneita keskimääräisiä osuuksia ennusteena, vastaava arvo on 0,6470. Kertoimet parantavat siis ennustetarkkuutta 0,0316 yksiköllä suhteessa naiiviin malliin, mikä osoittaa sen, että kertoimet sisältävät aitoa ennustevoimaa pelkkään historialliseen keskiarvoon verrattuna. Kotivoiton brier score on 0,2276, tasapelin 0,1786 ja vierasvoiton 0,2091. Tasapelin alhaisin brier score viittaa siihen, että tasapelin todennäköisyys on kertoimissa parhaiten kalibroitu, mikä on johdonmukainen kalibraatiotaulukoiden tulosten kanssa.

Kokonaisuutena tekemämme kalibraatioanalyysi osoittaa, että B365:n kertoimet ovat Liigan otteluissa kohtuullisen hyvin kalibroitu. Yksikään todennäköisyysluokka ei tuota tilastollisesti merkitsevää eroa implisiittisen todennäköisyyden ja toteutuneen osuuden välillä kotivoiton tai vierasvoiton osalta. Tasapelin osalta havaittiin pieni, mutta ei tilastollisesti merkitsevä, taipumus alihinnoitella tasapelin todennäköisyyttä. Nämä tulokset tukevat näkemystä siitä, että Liigan vedonlyöntimarkkinat ovat kohtuullisen tehokkaat, vaikka pieniä systemaattisia vinoumia on havaittavissa.

6.3 Suosikki-altavastaaja-analyysi

Suosikki-altavastaaja-analyysin tavoitteena on selittää, esiintyykö Liigan osalta vedonlyöntimarkkinoilla systemaattista vinoumaa suosikkien ja altavastaajien hinnoittelussa. Suosikki-altavastaaja-harha tarkoittaa siis tilannetta, jossa suosikkien kertoimet ovat systemaattisesti hinnoiteltu liian mataliksi niiden todelliseen voittotodennäköisyyteen nähden, kun taas altavastaajien kertoimet ovat liian korkeat. Tällöin suosikkeihin pelaaminen tuottaa keskimäärin heikomman tuoton kuin altavastaajien pelaaminen.

Jokaisessa ottelussa suosikiksi määriteltiin se lopputulos, jolla on matalin normalisoitu implisiittinen todennäköisyys ja vastaavasti altavastaajaksi se lopputulos, jolla on korkein normalisoitu implisiittinen todennäköisyys. Analyysi toteutettiin lopulta kahdella tavalla. Ensimmäisessä tarkastelussa huomioitiin koko 1X2-markkina, jolloin suosikki tai altavastaaja voi olla myös tasapeli. Toisessa tarkastelussa jätettiin tasapeli pois, ja rajattiin tarkastelu pelkästään joukkueiden väliseen vastakkainasetteluun, koska suurin osa aineiston altavastaajakertoimista oli sellaisia, joissa korkein kerroin on tasapelillä. Tällöin pelkästään 1X2-markkinaan perustuva jako ei välttämättä kuvaa joukkueiden välistä voimasuhdetta yhtä puhtaasti. Tässä tarkastelussa tasapeliin päättyvä ottelu tulkitaan tappioksi molemmille joukkueille, koska tutkimusaineisto ei sisällä moneyline-kertoimia. Tämä lähestymistapa on metodologisesti hyvin yksinkertaistettu, mutta se tarjoaa lisänäkökulman joukkueiden välisen voimasuhteen ja suosikki-altavastaaja-

asetelman tarkasteluun puhtaammin ilman tasapelin vaikutusta. Strategioiden tilastollinen merkitsevyys testattiin yksisuuntaisella t-testillä, jossa nollahypoteesina on, että strategian keskimääräinen tuotto on nolla. Strategioiden välinen ero testattiin kahden riippumattoman otoksen t-testillä.

Taulukossa 7 esitetään suosikki-altavastaaja-strategioiden tuotot sekä 1X2-markkinoilla, että ilman tasapeliä. Koko 1X2-markkinalla suosikin ROI on -7,07 prosenttia, ja se on erittäin merkitsevästi negatiivinen ($p < 0,001$). Käytännössä tämä tarkoittaa, että pelkästään suosikkeja pelaamalla toiminta on tappiollista. Altavastaajan ROI on sen sijaan vain -1,97 prosenttia, ja sen p-arvo on 0,472. Tutkimuksen kannalta tämä on tärkeä havainto, sillä voidaan todeta, että altavastaajia pelaaminen ei ole yhtä selvästi tappiollista kuin suosikkeihin pelaaminen. Strategioiden välinen ero on lähes tilastollisesti merkitsevä ($p = 0,099$), mikä antaa viitteitä suosikki-altavastaaja-harhan olemassaolosta Liigan vedonlyöntimarkkinoilla.

Taulukko 7. Suosikki- ja altavastaajastrategioiden tuotot

Strategia	Voitto-%	ROI	t-arvo	p-arvo
1X2-markkina				
Suosikki	49,4 %	-7,07 %	-4,904	< 0,001
Altavastaaja	22,8 %	-1,97 %	-0,719	0,472
Strategioiden ero			-1,649	0,099
Ilman tasapeliä				
Suosikki	49,4 %	-7,07 %	-4,904	< 0,001
Altavastaaja	27,3 %	-8,48 %	-3,595	0,0003
Strategioiden ero			0,507	0,612

Kun tasapeli jätetään analyysistä pois, ja tarkastellaan pelkästään joukkueiden välistä vastakkainasettelua, tulokset muuttuvat. Suosikin ROI pysyy tismalleen samana -7,07 prosenttia, mutta altavastaajan ROI laskee merkittävästi -8,48 prosenttiin. Tämän lisäksi myös se on nyt tilastollisesti merkitsevä ($p = 0,0003$). Strategioiden välinen ero puolestaan ei ole enää tilastollisesti merkitsevä ($p = 0,612$). Tämä tulos vahvistaa aiemman havainnon siitä, että suurin osa aineiston altavastaaja asemista on tasapelejä, joiden poistaminen muuttaa analyysin tuloksia merkittävästi.

Taulukossa 8 esitetään suosikki- ja altavastaajastrategioiden ROI kausittain. Suosikin ROI on negatiivinen jokaisella yksittäisellä kaudella. Aineiston alkupään kausilla suosikin ROI oli noin -3,5 ja -6,1 prosentin tasolla, mutta viimeisille kausille se on laskenut jo -9,2 ja -13,1 prosentin tuntumaan. Tämä viittaa siihen, että suosikkien ylihinnointelu on pahentunut tarkastelujakson aikana. Altavastaajan ROI vaihtelee huomattavasti enemmän kausittain, ja se on joillakin kausilla jopa selvästi positiivinen. Erityisesti kaudet 2024 – 2025 (+10,15%), 2021-2022 (9,16%) ja 2018-2019 (+5,50%) pistävät silmään positiivisella ROI:llaan. Toisaalta taas joillakin kausilla altavastaajan ROI on hyvin selvästi negatiivinen, kuten kausilla 2022–2023 (-24,16 %) ja 2017–2018 (-8,77 %). Tämä melko suuri kausittainen vaihtelu viittaa siihen, että altavastaajaharha ei ole täysin systemaattinen ja johdonmukainen Liigan vedonlyöntimarkkinoilla.

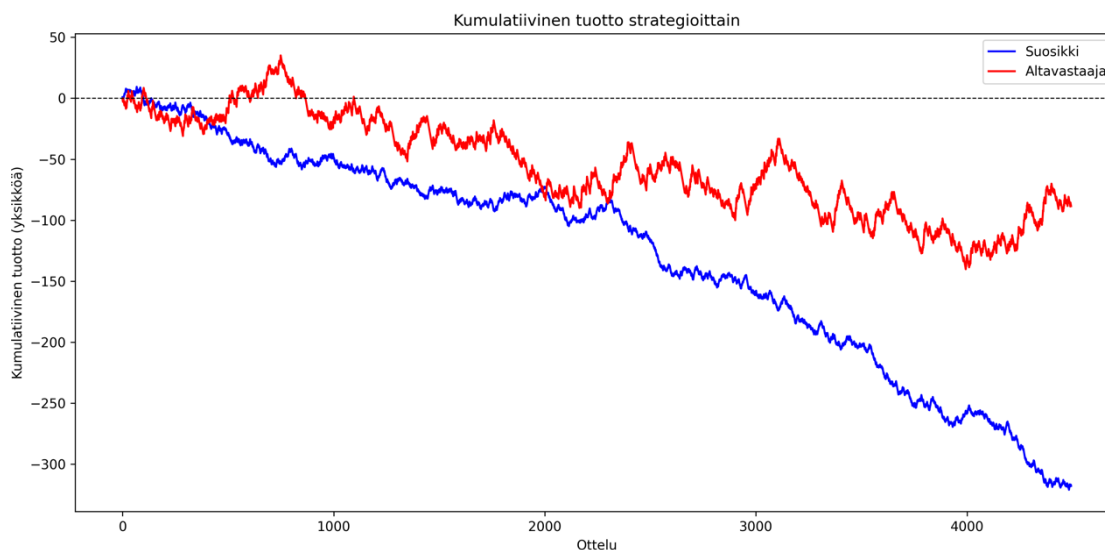
Taulukko 8. Suosikki- ja altavastaajastrategioiden ROI kausittain (%)

Kausi	Suosikki ROI	Altavastaaja ROI	Ottelut
2015–2016	-6,06	-3,99	450
2016–2017	-5,10	1,23	450
2017–2018	-3,52	-8,77	450
2018–2019	-4,20	5,50	450

Kausi	Suosikki ROI	Altavastaaja ROI	Ottelut
2019–2020	-2,64	-7,39	443
2020–2021	-10,03	-3,84	424
2021–2022	-7,12	9,16	447
2022–2023	-9,42	-24,16	450
2023–2024	-9,23	-6,42	450
2024–2025	-13,14	10,15	480
Yhteensä	-7,07	-1,97	4 494

Kokonaisuutena suosikki-altavastaaja-analyysi antaa viitteitä siitä, että Liigan vedonlyöntimarkkinoilla esiintyy lievää suosikki-altavastaaja-harhaa. Altavastaajan ROI ei poikkea tilastollisesti merkitsevästi nolasta koko 1X2-markkinalla, kun taas suosikin ROI on erittäin merkitsevästi negatiivinen. Strategioiden välinen ero on lähes tilastollisesti merkitsevä, mikä tukee näkemystä siitä, että vedonlyöjän näkökulmasta suosikkeja hinnoitellaan markkinoilla systemaattisesti liian epäsuotuisasti.

Kumulatiivinen tuottokuvaaja havainnollistaa strategioiden pitkän aikavälin kehitystä. Kuviossa 2 esitetään suosikki- ja altavastaajastrategioiden kumulatiivinen tuotto koko tarkastelujakson ajalta. Suosikkistrategia laskee tasaisesti ja jyrkästi koko aineiston läpi päätyen noin -320 yksikköön 4 494 ottelun jälkeen. Altavastaaja-strategia on huomattavasti vaihtelevampi ja pysyttelee lähellä nolaa tarkastelujakson alkupuolella, mutta päättyy lopulta noin -110 yksikköön. Kuvaaja havainnollistaa selkeästi sen, että vaikka altavastaajastrategia on suosikkistrategiaa parempi, kumpikaan ei tuota positiivista odotusarvoa pitkällä aikavälillä.



Kuvio 2. Suosikki- ja altavastaajastrategioiden kumulatiivinen tuotto

Suhteessa aiempaan jääkiekkotutkimukseen havainto on mielenkiintoinen. Woodland ja Woodland (1994, 2001) havaitsivat jääkiekon NHL:ssä käänteisen suosikki-altavastaaja-harhan, jossa suosikit oli alihinnoiteltu ja altavastaajat ylihinnoiteltuja suhteessa toteutuneisiin voittotodennäköisyyksiin. Liigan aineistossa tilanne näyttää olevan päinvastainen eli tulokset ovat lähempänä klassista suosikki-altavastaaja-harhaa, jossa altavastaajiin panostaminen on suhteellisesti kannattavampaa kuin suosikkeihin panostaminen. Tämä ero voi heijastaa markkinoiden kokoeroa, sillä NHL on huomattavasti suurempi ja likvidimpi markkina kuin Liiga. Se voi myös kertoa vedonlyöjien erilaisista käyttäytymistäipumuksista pienemmillä markkinoilla. Kausittainen tarkastelu osoittaa kuitenkin, että ilmiö ei ole täysin johdonmukainen, vaan altavastaajan ROI vaihtelee merkittävästi eri kausien välillä. Tämä viittaa siihen, että mahdollinen harha ei ole pysyvä markkinatehottomuus vaan pikemminkin satunnainen ilmiö, joka vaihtelee markkinaolosuhteiden mukaan.

6.4 Multinomiaalinen logistinen regressio

Multinomiaalisen logistisen regression avulla arvioimme, sisältävätkö vedonlyöntikertoimet tilastollisesti merkitsevää ennustevoimaa otteluiden

lopputulosten suhteen. Tutkimuksessa käytettiin multinomiaalista logistista regressiota (multinomial log), sillä se soveltuu tavallista logistista regressiota paremmin, kun pyritään selittämään kolmea selitettävää arvoa, kotivoittoa, vierasvoittoa ja tasapeliä. Binäärinen logistinen regressio yhdistäisi tasapelin ja vierasvoiton samaan luokkaan, jolloin käytännössä selitettävä muuttuja olisi vain ”kotivoitto tai ei”.

Multinomiaalisessa logistisessa regressiossa yksi lopputulos valitaan vertailuluokaksi, johon muita lopputuloksia verrataan. Tässä tutkimuksessa vertailuluokaksi valittiin kotivoitto, jolloin malli tuottaa kaksi vertailua, jotka ovat tasapeli suhteessa kotivoittoon ja vierasvoitto suhteessa kotivoittoon. Selittävinä muuttujina käytetään kotivoiton ja tasapelin normalisoituja implisiittisiä todennäköisyyksiä. Vierasvoiton implisiittinen todennäköisyys jätettiin mallista pois, koska kotivoiton, tasapelin ja vierasvoiton normalisoidut implisiittiset todennäköisyydet summautuvat aina yhteen, mikä taas aiheuttaa täydellisen multikollineaarisuuden kaikkien kolmen muuttujan ollessa mallissa samanaikaisesti.

Taulukossa 9 esitetään regression tulokset. Taulukon yläosa kuvaa tasapelin ja kotivoiton välistä vertailua. Kotivoiton implisiittinen todennäköisyys on tässä vertailussa erittäin merkitsevä ($\beta = -3,404$, $p < 0,001$) eli mitä suurempi kotivoiton implisiittinen todennäköisyys on, sitä epätodennäköisemmin ottelu päättyy tasapeliin suhteessa kotivoittoon. Tasapelin implisiittinen todennäköisyys ei sen sijaan ole merkitsevä tässä vertailussa ($p = 0,710$), mikä viittaa siihen, että tasapelin kerroin ei tuo lisäinformaatiota kotivoiton ja tasapelin välisen eron selittämiseen, kun kotivoiton todennäköisyys on jo mallissa mukana.

Toinen osa yläosan taulukosta kuvaa vierasvoiton ja kotivoiton välistä vertailua. Kotivoiton implisiittinen todennäköisyys on tässäkin vertailussa erittäin merkitsevä ($\beta = -5,775$, $p < 0,001$) eli mitä suurempi kotivoiton implisiittinen todennäköisyys on, sitä epätodennäköisemmin ottelu päättyy vierasvoittoon suhteessa kotivoittoon. Kertoimen itseisarvo on suurempi kuin tasapelin vertailussa, mikä on loogista, tällöin kotivoiton

todennäköisyyden kasvaessa vierasvoitto on tasapeliä epätodennäköisempi lopputulos. Tasapelin implisiittinen todennäköisyys ei ole merkitsevä myöskään tässä vertailussa ($p = 0,255$).

Taulukko 9. Multinomiaalinen logistinen regressio

Muuttuja	Kerroin	Keskivirhe	z-arvo	p-arvo
Tasapeli vs. kotivoitto				
Vakio	1,137	0,613	1,855	0,064
Kotivoiton implisiittinen todennäköisyys	-3,404	0,384	-8,860	< 0,001
Tasapelin implisiittinen todennäköisyys	-0,840	2,256	-0,372	0,710
Vierasvoitto vs. kotivoitto				
Vakio	2,839	0,555	5,115	< 0,001
Kotivoiton implisiittinen todennäköisyys	-5,775	0,345	-16,718	< 0,001
Tasapelin implisiittinen todennäköisyys	-2,391	2,101	-1,138	0,255
Tunnusluku		Arvo		
Havaintoja		4 494		
Pseudo R ²		0,0408		
Oikein luokiteltu		49,3 %		
LLR p-arvo		< 0,001		

Mallin kokonaisselitysvoima on pseudo R²:lla mitattuna 0,0408, mikä on hieman pienempi kuin vastaava arvo binäärisessä logistisessa regressiossa. Tämä on odotettua, koska kolmen luokan ennustaminen on tilastollisesti haastavampaa kuin kahden luokan. Malli luokittelee 49,3 prosenttia otteluista oikein, mikä on selvästi satunnaista arvailua parempi tulos kolmen tasaveroisen luokan tilanteessa, jossa satunnainen arvaus tuottaisi noin 33 prosentin tarkkuuden. LLR p-arvo on alle 0,001, mikä osoittaa, että malli on kokonaisuutena tilastollisesti merkitsevä verrattuna malliin ilman selittäviä muuttujia.

Multinomiaalisen logistisen regression tulokset osoittavat, että kotivoiton implisiittinen todennäköisyys on vahva ja tilastollisesti erittäin merkitsevä selittäjä sekä tasapelin että

vierasvoiton todennäköisyydelle suhteessa kotivoittoon. Tasapelin implisiittinen todennäköisyys ei sen sijaan ole kummassakaan vertailussa merkitsevä, mikä voi viitata siihen, että tasapelin kerroin ei tarjoa sellaista lisäinformaatiota, jota kotivoiton kerroin ei jo sisällä. Kokonaisuutena tulokset vahvistavat, että B365:n kertoimet sisältävät merkittävää ennustevoimaa Liigan otteluiden lopputulosten suhteen, mikä on linjassa markkinatehokkuuden kanssa.

6.5 Kotietu

Kotiedulla tarkoitetaan ilmiötä, jossa kotijoukkue pärjää vierasjoukkuetta paremmin omalla kotikentällään. Kotiedun on havaittu olevan merkittävä tekijä useissa eri urheilulajeissa. Sitä on selitetty muun muassa kotiyleisön tuesta, matkustamisesta aiheutuvasta rasituksesta vierasjoukkueelle sekä tuomareiden mahdollisesta myönteisyydestä kotijoukkuetta kohtaan (Fischer & Haucap, 2022). Vedonlyöntimarkkinoiden tehokkuuden kannalta on tärkeää arvioida, heijastuuko kotietu oikein kertoimiin. Jos kotietu on kertoimissa ylihinnoiteltu, kotijoukkueeseen pelaaminen tuottaa systemaattisesti heikompaa tuottoa kuin vierailevaan joukkueeseen pelaaminen, ja päinvastoin.

Koko tarkastelujakson aineistossa kotivoittoja on 43,2 prosenttia, tasapelejä 23,3 prosenttia ja vierasvoittoja 33,5 prosenttia otteluista. Kotivoitto on selvästi yleisin lopputulos, mikä osoittaa kotiedun olevan todellinen ilmiö Liigassa. Kotiedun hinnoittelua tarkasteltaessa havaitaan, että kotivoiton keskimääräinen implisiittinen todennäköisyys on 44,5 prosenttia, kun toteutunut kotivoittojen osuus on 43,2 prosenttia. Vastaavasti vierasvoiton implisiittinen todennäköisyys on 33,3 prosenttia ja toteutunut osuus 33,5 prosenttia. Kertoimet siis yliarvioivat kotivoiton todennäköisyyden ja aliarvioivat vierasvoiton todennäköisyyden hieman, mikä viittaa siihen, että kotietu on kertoimissa lievästi ylihinnoiteltu. Erot ovat kuitenkin pieniä, alle 1,5 prosenttiyksikköä, mikä viittaa siihen, että vedonvälittäjä hinnoittelee kotiedun pääosin oikein. Ero on kuitenkin todella pieni, joten sen perusteella ei voi tehdä vahvoja

johtopäätöksiä kotiedun systemaattisesta virhehinnoittelusta. Ero voi heijastua esimerkiksi satunnaisvaihtelusta tai marginaalin epätasaisesta jakautumisesta lopputulosten välille.

Taulukossa 10 esitellään kotivoitto- ja vierasvoittoprosentit sekä kotiedun hinnoitteluero kausittain. Kotivoittoprosentti vaihtelee 40,2 ja 45,6 prosentin välillä eri kausilla. Selkein poikkeus on kausi 2019–2020, jolloin kotivoittoprosentti laski 40,2 prosenttiin ja vierasvoittoprosentti nousi 38,1 prosenttiin. Tämä poikkeama ajoittuu koronapandemian alkamiseen, jonka seurauksena osa kauden loppuotteluista pelattiin ilman yleisöä ennen kuin kausi keskeytettiin kokonaan. Kotiedun hinnoitteluero on negatiivinen useimmilla kausilla, eli kertoimet arvioivat kotivoiton toteutunutta osuutta korkeammaksi. Poikkeus on kausi 2022–2023, jolloin toteutunut kotivoittoprosentti ylitti implisiittisen todennäköisyyden.

Taulukko 10. Kotivoitto- ja vierasvoitto% sekä kotiedun hinnoitteluero kausittain

Kausi	Kotivoitto%	Vierasvoitto%	Impl. tod. koti%	Ero (tot.-impl.)
2015–2016	44,4	32,7	45,3	-0,9
2016–2017	41,8	33,1	44,6	-2,8
2017–2018	45,3	30,2	45,2	0,1
2018–2019	44,9	31,3	45,5	-0,6
2019–2020	40,2	38,1	45,0	-4,8
2020–2021	42,2	35,8	43,0	-0,8
2021–2022	40,9	36,7	43,4	-2,5
2022–2023	45,6	31,3	43,9	1,7
2023–2024	42,9	32,4	44,4	-1,5

Kausi	Kotivoitto%	Vierasvoitto%	Impl. tod. koti%	Ero (tot.-impl.)
2024–2025	43,3	33,5	44,8	-1,5
Yhteensä	43,2	33,5	44,5	-1,3

Kotiedun merkitystä tarkasteltiin myös koronapandemian näkökulmasta jakamalla aineisto kolmeen osaan: ennen koronaa, koronan aikana ja koronan jälkeen. Jako perustuu siihen, että koronapandemia vaikutti merkittävästi otteluiden peliolosuhteisiin, erityisesti yleisön läsnäoloon. Koronan aikaiseksi jaksoksi on määritelty maaliskuusta 2020 elokuuhun 2022, jolloin otteluita pelattiin osittain tai kokonaan ilman yleisöä tai rajoitetulla katsojakapasiteetilla. Taulukossa 11 esitetään lopputulosten jakautuminen ja kotiedun hinnoittelueron näiden kolmen ajanjakson mukaan.

Taulukko 11. Kotiedun muutos koronapandemian aikana

Ajanjakso	Ottelut	Kotivoitto%	Vierasvoitto%	Tasapeli%	Impl. tod. koti%	Ero (tot.-impl.)
Ennen koronaa (2015–2020)	2 209	43,4	33,2	23,4	45,3	-1,8
Koronan aikana (2020–2022)	905	41,4	35,9	22,7	43,0	-1,8
Koronan jälkeen (2022–2025)	1 380	43,9	32,5	23,6	44,8	-0,9

Taulukon 11 tulokset osoittavat, että kotivoittoprosentti laski koronan aikana 43,4 prosentista 41,4 prosenttiin, kun vastaavasti vierasvoittoprosentti nousi 33,2 prosentista 35,9 prosenttiin. Tämä on johdonmukainen havainto sen kirjallisuuden kanssa, jonka mukaan yleisön läsnäolo vaikuttaa merkittävästi kotiedun suuruuteen (Fischer & Haucap, 2022), joskin kausaalisuhteen täydellinen vahvistaminen edellyttäisi tarkempaa analyysiä. Ilman yleisöä pelatuissa otteluissa kotijoukkueen etu pienenee, koska yleisön tuki ja sitä kautta myös kotikenttä menettävät merkitystään. Koronan jälkeen kotivoittoprosentti on palannut lähelle ennen pandemiaa vallinnutta tasoa, mikä viittaa siihen, että kotiedun palautuminen normaaliksi on tapahtunut suhteellisen nopeasti rajoitusten poistuttua. Mielenkiintoista on, että kotiedun hinnoitteluerot pysyi lähes samana ennen koronaa ja koronan aikana, mikä tarkoittaa, että vedonvälittäjä ei täysin sopeutunut kotiedun heikkenemiseen koronajakson aikana. Tämä voidaan tulkita merkinä markkinoiden hitaasta reagoinnista poikkeukselliseen tilanteeseen.

Kotiedun taloudellista merkitystä arvioitiin lisäksi tarkastelemalla kotijoukkue- ja vierasjoukkuestrategioiden tuottoja. Kotijoukkuestrategiassa pelataan aina kotijoukkueen voiton puolesta ja vierasjoukkuestrategiassa aina vierailevaa joukkuetta. Kotijoukkuestrategian ROI on -10,18 prosenttia ja se on erittäin merkitsevästi negatiivinen ($t = -6,110$, $p < 0,001$). Vierasjoukkuestrategian ROI on -5,53 prosenttia ja sekin on tilastollisesti merkitsevästi negatiivinen ($t = -2,510$, $p = 0,012$). Vierasjoukkueeseen pelaaminen on siis selvästi kannattavampaa kuin kotijoukkueeseen pelaaminen, vaikka molemmat strategiat ovat tappiollisia. Tämä tukee havaintoa kotiedun ylihinnittelusta kertoimissa eli käytännössä kotivoiton todennäköisyyttä arvioidaan systemaattisesti liian korkeaksi, minkä seurauksena kotijoukkueeseen panostavan vedonlyöjän odotusarvo on heikompi kuin vierasjoukkueeseen panostavan.

Kokonaisuutena kotiedun analyysi osoittaa, että kotietu on todellinen ja merkittävä ilmiö Liigassa, mutta vedonvälittäjä yli hinnoittelee sen systemaattisesti kertoimissa. Kotietu on siis kertoimissa läsnä, mutta hieman liian voimakkaana, minkä seurauksena kotijoukkueeseen pelaaminen on selvästi tappiollisempaa kuin vierailevaan

joukkueeseen pelaaminen. Koronapandemian aikainen kotiedun heikkeneminen oli selvästi havaittavissa aineistossa, mutta vedonvälittäjä ei täysin reagoanut tähän muutokseen kertoimissaan, mikä voidaan tulkita merkinä markkinoiden rajoittuneesta sopeutumiskyvystä poikkeuksellisissa olosuhteissa. Kotiedun analyysin tuloksia tulkittaessa on otettava myös huomioon, että käytetyt menetelmät eivät mahdollista kotiedun täydellistä eristämistä joukkueiden välisistä vahvuuseroista. Kotivoiton implisiittinen todennäköisyys siis sisältää väistämättä molemmat tekijät, eli sen, että peli pelataan kotona, että sen kuinka vahva kotijoukkue on suhteessa vierasjoukkueeseen. Puhtaan kotiedun mittaaminen edellyttäisi joukkueiden absoluuttisen vahvuuden kontrolloimista, mikä ei tämän tutkimuksen aineiston puitteissa ole mahdollista. Havaitut erot kotivoitto- ja vierasvoittostrategioiden tuotoissa voivat siten heijastaa myös muita tekijöitä kuin pelkästään kotiedun hinnoittelua.

6.6 Vedonlyöntistrategioiden simulointi

Vedonlyöntistrategioiden simulointi on keskeinen osa markkinatehokkuuden empiiristä testausta. Jos vedonlyöntimarkkinat ovat tehokkaat, yksinkertaisilla ja julkisesti saatavilla olevaan informaatioon perustuvilla strategioilla ei pitäisi pystyä saavuttamaan positiivista odotusarvoa pitkällä aikavälillä. Perusstrategioiden, kuten suosikki-, altavastaaaja-, kotijoukkue- ja vierasjoukkuestrategioiden tulokset on raportoitu luvuissa 6.3 ja 6.5 osana suosikki-altavastaaaja- ja kotiedun analyyssejä. Tässä luvussa tarkastellaan monimutkaisempia strategioita, jotka pyrkivät hyödyntämään markkinoilla mahdollisesti esiintyviä vinoumia tarkemmin rajattujen kriteerien avulla. Kaikissa simuloinneissa oletetaan, että vedonlyöjä panostaa yhden yksikön kutakin ottelua kohti, ja tuottoa mitataan kaavan (5) mukaisella ROI-luvulla. Tilastollinen merkitsevyys testataan yksisuuntaisella t-testillä, jossa nollahypoteesina on, että strategian keskimääräinen tuotto on nolla.

6.6.1 Perusstrategiat

Raja-arvostrategioissa vedonlyöntikohteet rajataan tietyn kertoimen ylä- tai alarajan perusteella. Tavoitteena on selvittää, esiintyykö tietyissä kerroinluokissa systemaattisia vinoumia, joita laajemmat perusstrategiat eivät paljasta. Suosikkistrategioissa testataan eri alarajoilla rajattuja strategioita, joissa pelataan vain selkeitä suosikkeja. Altavastajastrategioissa puolestaan pelataan vain korkeakertoimisia kohteita, joissa altavastajan asema on erityisen selvä. Tällä tavoin voidaan arvioida, onko suosikki-altavastaja-harha voimakkaampaa tietyllä kerroinalueella kuin toisella.

Taulukossa 12 esitetään suosikkistrategioiden tulokset eri raja-arvoilla. Kaikki suosikkistrategiat ovat tappiollisia raja-arvosta riippumatta. ROI vaihtelee -7,76 prosentin ja -9,11 prosentin välillä, eikä yhdenkään strategian ROI poikkea merkittävästi perussuosikkistrategian -7,07 prosentin tasosta. Tulos osoittaa, että suosikkien ylihinnoittelu ei ole erityisen voimakasta millään tietyllä kerroinalueella, vaan se on tasainen ilmiö kaikkien suosikkikertoimien osalta. Kaikki strategiat ovat tilastollisesti merkitsevästi negatiivisia, mikä vahvistaa suosikkeihin pelaamisen systemaattisen tappiollisuuden.

Taulukko 12. Suosikkistrategiat eri raja-arvoilla

Strategia	N	Voitto-%	ROI	t-arvo	p-arvo
Suosikki kerroin < 1,50	358	67,3 %	-8,65 %	-2,550	0,011
Suosikki kerroin < 1,60	605	64,6 %	-7,76 %	-2,759	0,006
Suosikki kerroin < 1,80	1 162	59,7 %	-7,82 %	-3,295	0,001
Suosikki kerroin < 2,00	1 820	55,2 %	-9,11 %	-4,936	< 0,001

Taulukossa 13 esitetään puolestaan altavastaajastrategioiden tulokset eri raja-arvoilla. Altavastaaja-strategiat poikkeavat selvästi suosikkistrategioista. Mitä korkeampi kertoimen alaraja on, sitä heikompi ROI on. Kerroin yli 3,00 tuottaa ROI:n -19,94 prosenttia, kerroin yli 4,00 ROI on jo -25,78 prosenttia ja kerroin yli 5,00 ROI peräti -35,28 prosenttia. Liigan aineistossa erittäin suurikertoimiset altavastaajat ovat selvästi ylihinnoiteltuja vedonlyöjän näkökulmasta, eli niiden voittotodennäköisyys on pienempi kuin kertoimet antavat ymmärtää. Kerroin yli 5,00 strategian p-arvo 0,093 ylittää tavanomaisesti käytetyn 0,05 merkitsevyysrajan, mikä johtuu osittain pienestä otoskoosta (N = 90).

Taulukko 13. Altavastaajastrategiat eri raja-arvoilla

Strategia	N	Voitto-%	ROI	t-arvo	p-arvo
Altavastaaja kerroin > 3,00	630	21,3 %	-19,94 %	-3,130	0,002
Altavastaaja kerroin > 4,00	206	14,6 %	-25,78 %	-2,005	0,046
Altavastaaja kerroin > 5,00	90	10,0 %	-35,28 %	-1,692	0,093

Raja-arvostrategioiden tulokset osoittavat, että systemaattiset vinoumat eivät rajoitu tiettyyn kerroinalueeseen suosikkien osalta, vaan suosikkeihin pelaaminen on tasaisen tappiollista kertoimista riippumatta. Altavastaajien osalta tulokset ovat yllättäviä. Erittäin korkeakertoimisiin altavastaajiin pelaaminen on selvästi tappiollisempaa kuin matalamman kertoimen altavastaajiin pelaaminen. Tämä viittaa siihen, että Liigan vedonlyöntimarkkinoilla suuria altavastaajia hinnoitellaan pikemminkin liian edullisesti vedonvälittäjälle kuin vedonlyöjälle, mikä on vastoin klassista suosikki-altavastaaja-harhaa.

6.6.2 Kehittyneemmät strategiat

Perus- ja raja-arvostrategioiden lisäksi testattiin kaksi kehittyneempää vedonlyöntistrategiaa, jotka pyrkivät hyödyntämään tilastollista mallia yksinkertaisten sääntöjen sijaan. Nämä strategiat ovat arvovetomalli ja Kelly-kriteeri. Molempien strategioiden lähtökohtana on logistisen regression tuottama ennuste kotivoiton todennäköisyydelle, jonka perusteella arvioidaan, onko yksittäinen vedonlyöntikohde alihinnoiteltu suhteessa kertoimeen. Strategioiden arviointi toteutettiin sekä in-sample-että out-of-sample-testauksena. In-sample-testauksessa strategia muodostetaan ja testataan samalla aineistolla, kun taas out-of-sample-testauksessa strategia muodostetaan historiallisella datalla ja testataan sen jälkeen datalla, jota malli ei ole ”nähty”. Out-of-sample-testaus on tärkeää, jotta vältytään in-sample ylisovittumiselta, joka voi tuottaa liian optimistisen arvion strategian todellisesta suorituskyvystä. Taulukossa 14 esitellään molempien strategioiden tulokset.

Tässä tutkimuksessa aineisto jaettiin kahteen osaan. In-sample-aineistoon, joka kattaa kaudet 2015–2021 (2 667 ottelua), ja out-of-sample-aineistoon, joka kattaa kaudet 2021–2025 (1 827 ottelua). Strategiat muodostettiin in-sample-aineiston perusteella ja testattiin out-of-sample-aineistolla. Vertailun vuoksi raportoidaan myös in-sample-tulokset, jotka kuvaavat strategian teoreettista suorituskyyä tilanteessa, jossa koko aineisto on käytössä. Taulukossa 14 esitetään kehittyneempien strategioiden in-sample- ja out-of-sample-tulokset rinnakkain.

Arvovetomallissa hyödynnetään logistisen regression ennustamaa kotivoiton todennäköisyyttä vertailuarvona. Kustakin todennäköisyysluokasta lasketaan keskimääräinen kotivoiton kerroin, ja arvoveto tunnistetaan tilanteessa, jossa yksittäisen ottelun kerroin ylittää luokansa keskimääräisen kertoimen vähintään kahdella prosentilla. Idea on, että tällaisessa tilanteessa kerroin on poikkeuksellisen korkea suhteessa vastaavanlaisiin otteluihin, mikä voisi viitata ylikertoimeen. In-sample-aineistossa arvovetomalli tunnisti yhteensä 1 441 arvovetoa, mikä vastaa 32,1 prosenttia

kaikista otteluista. Arvovetomalli tuotti ROI:n -12,51 prosenttia, ja tulos on erittäin merkitsevästi negatiivinen ($t = -3,991$, $p < 0,001$). Out-of-sample-aineistossa malli tunnisti 591 arvovetoa, mikä vastaa 32,3 prosenttia out-of-sample-otteluista. Out-of-sample ROI oli -5,58 prosenttia, eikä se poikennut tilastollisesti merkitsevästi nolasta ($t = -1,110$, $p = 0,267$). Silmiinpistävää on se, että in-sample ROI oli selvästi heikompi kuin out-of-sample ROI, mikä on tyypillistä ylisovittumisen seurauksena. Out-of-sample-tulos on siten informatiivisempi arvio strategian todellisesta suorituskyvystä. Vaikka out-of-sample ROI ei ole tilastollisesti merkitsevästi negatiivinen, se on silti negatiivinen, eikä strategia tuota positiivista odotusarvoa.

Kelly-kriteeri on matemaattinen menetelmä, joka laskee kullekin vedolle optimaalisen panoksen koon suhteessa sen odotettuun arvoon. Kelly-kriteerin mukainen optimaalinen panostusosuus lasketaan kaavan (6) mukaisesti:

$$f = \frac{(p \times b - (1 - p))}{b} \quad (6)$$

jossa p on ennustettu voittotodennäköisyys ja b on nettokertoimet eli kerroin miinus yksi. Jos Kelly-kriteerin arvo on positiivinen, veto on teoriassa arvokas ja kannattaa pelata. Jos arvo on negatiivinen tai nolla, vetoa ei tule pelata lainkaan. Tässä tutkimuksessa käytetään puoli-Kellyä, jossa suositeltu panostusosuus puolitetaan. Puoli-Kelly on käytännössä vakiintunut konservatiivisempi lähestymistapa, joka vähentää vedonlyöntikassaan kohdistuvaa riskiä säilyttäen silti suurimman osan strategian teoreettisesta hyödystä.

In-sample havaintoja Kelly-kriteeri tunnisti yhteensä 42 positiivisen odotusarvon vetoa, mikä vastaa vain 0,9 prosenttia kaikista otteluista. Kelly-strategia tuotti ROI:n -55,71 prosenttia, joskin tulos ei ole tilastollisesti merkitsevää pienen otoskoon vuoksi ($t = -1,983$, $p = 0,054$). Out-of-sample-aineistossa malli tunnisti vain 7 vetoa, mikä vastaa 0,4 prosenttia out-of-sample-otteluista. Out-of-sample ROI oli -100,00 prosenttia ($t = -3,488$, $p = 0,013$), eli strategia hävisi jokaisen vedon. On kuitenkin tärkeää huomata, että seitsemän sekä neljäkymmenen kahden ottelun otoskoko on liian pieni luotettaviin

johtopäätöksiin. Jo pelkkä vetomäärän pienuus kertoo jotain oleellista, sillä malli löytää äärimmäisen harvoin tilanteita, joissa se tunnistaa positiivisen odotetun arvon, mikä on itsessään johdonmukainen tulos markkinatehokkuuden näkökulmasta. Positiivisen Kelly-arvon saavat kohteet ovat todennäköisesti satunnaisia poikkeamia eikä aitoja alihinnoiteltuja kohteita.

Taulukko 14. Kehittyneempien strategioiden tuotot

Strategia	In-sample N	In-sample ROI	Out-of- sample N	Out-of-sample ROI	OOS p- arvo
Arvovetomalli	1 441	-12,51 %	591	-5,58 %	0,267
Kelly-kriteeri (puoli-Kelly)	42	-55,71 %	7	-100,00 %	0,013

Kokonaisuutena kehittyneemmät strategiat eivät tuota positiivista odotusarvoa kummallakaan tavalla mitattuna. Arvovetomalli tuottaa out-of-sample-aineistossa paremman tuloksen kuin in-sample-aineistossa, mikä on tyypillistä in-sample-ylisovittumisen seurauksena, mutta pysyy silti selvästi negatiivisena. Tulos on linjassa markkinatehokkuuden kanssa. Vaikka strategiat hyödyntävät tilastollista mallinnusta yksinkertaisten sääntöjen sijaan, ne eivät pysty tunnistamaan sellaisia alihinnoiteltuja kohteita, joita vedonvälittäjä ei olisi jo hinnoitellut oikein.

6.6.3 Yhteenveto strategioista

Taulukossa 15 esitetään kaikkien tässä tutkimuksessa simuloitujen vedonlyöntistrategioiden tulokset koottuna yhteen. Mukana ovat sekä luvuissa 6.3 ja 6.5 raportoidut perusstrategiat että tässä luvussa esitetyt raja-arvo- ja kehittyneemmät strategiat.

Taulukko 15. Kaikkien strategioiden yhteenveto

Strategia	N	ROI	p-arvo
Perusstrategiat			
Suosikki (1X2)	4 494	-7,07 %	< 0,001
Altavastaja (1X2)	4 494	-1,97 %	0,472
Kotijoukkue	4 494	-10,18 %	< 0,001
Vierasjoukkue	4 494	-5,53 %	0,012
Raja-arvostrategiat: suosikit			
Suosikki kerroin < 1,50	358	-8,65 %	0,011
Suosikki kerroin < 1,60	605	-7,76 %	0,006
Suosikki kerroin < 1,80	1 162	-7,82 %	0,001
Suosikki kerroin < 2,00	1 820	-9,11 %	< 0,001
Raja-arvostrategiat: altavastajat			
Altavastaja kerroin > 3,00	630	-19,94 %	0,002
Altavastaja kerroin > 4,00	206	-25,78 %	0,046
Altavastaja kerroin > 5,00	90	-35,28 %	0,093
Kehittyneemmät strategiat			
Arvovetomalli	1 441	-12,51 %	< 0,001
Kelly-kriteeri (puoli-Kelly)	42	-55,71 %	0,054

Yhteenvetotaulukon tulokset osoittavat johdonmukaisesti, että yksikään testatuista strategioista ei tuota positiivista odotusarvoa. Kaikki strategiat ovat tappiollisia, ja suurin osa niistä on tilastollisesti merkitsevästi negatiivisia. Ainoa poikkeus on altavastajastrategia 1X2-markkinalla, jonka ROI -1,97 prosenttia ei poikkea tilastollisesti merkitsevästi nolasta. Tämä ei kuitenkaan tarkoita, että kyseinen strategia olisi

kannattava, vaan ainoastaan että sen tappiollisuudesta ei ole riittävää tilastollista näyttöä näin suurellakin aineistolla.

Tulokset ovat selvästi linjassa tehokkaiden markkinoiden hypoteesin kanssa. Yksinkertaisimmatkaan strategiat, kuten aina suosikkiin tai aina vierasjoukkueeseen pelaaminen, eivät tuota positiivista odotusarvoa, eikä kehittyneempien tilastollisten menetelmien hyödyntäminen paranna tilannetta. Vedonvälittäjän marginaali muodostaa käytännöllisen esteen positiivisen odotusarvon saavuttamiselle, ja marginaalin ylittäminen edellyttäisi sellaista informaatioetua tai ennustustarkkuutta, jota pelkällä julkisesti saatavilla olevalla datalla ei pystytä saavuttamaan ainakaan näillä kyseisillä strategioilla.

6.7 Yhteenveto

Tässä luvussa kootaan yhteen empiirisen analyysin keskeisimmät tulokset ja arvioidaan niiden merkitystä tutkimuskysymysten kannalta. Analyysi kattoi vedonvälittäjän marginaalin tarkastelun, kertoimien kalibraation, suosikki-altavastaja-harhan, logistisen regression, kotiedun hinnoittelun sekä useiden vedonlyöntistrategioiden simuloinnin.

Ensimmäisessä tutkimuskysymyksessä pohdittiin, esiintyykö Liigan ottelukertoimissa systemaattisia vinoumia. Analyysin perusteella voidaan todeta, että vinoumia esiintyy, mutta ne ovat pääosin pieniä eivätkä johdonmukaisesti tilastollisesti merkitseviä. Kalibraatioanalyysi osoitti, että kertoimet ovat kohtuullisen hyvin kalibroituineet kaikissa kolmessa lopputuloksessa. Yksikään todennäköisyysluokka ei tuottanut tilastollisesti merkitsevää eroa kotivoiton tai vierasvoiton osalta, ja tasapelin osalta havaittiin vain pieni mutta ei tilastollisesti merkitsevä alihinnoittelu. Suosikki-altavastaja-analyysi antoi viitteitä lievistä klassisesta suosikki-altavastaja-harhasta, sillä altavastajan ROI ei poikennut merkittävästi nollassa, kun taas suosikin ROI oli erittäin merkittävästi negatiivinen. Kotiedun osalta havaittiin systemaattinen mutta pieni ylihinnottelu, sillä kertoimet arvioivat kotivoiton todennäköisyyden keskimäärin 1,3 prosenttiyksikköä toteutunutta korkeammaksi.

Toinen tutkimuskysymys koski sitä, vastaavatko kertoimista johdetut implisiittiset todennäköisyydet toteutuneita voitto-osuuksia. Kalibraatioanalyysin ja brier scoren perusteella voidaan todeta, että kertoimet vastaavat toteutuneita osuuksia kohtuullisen hyvin. Kokonais-brier score 0,6153 on selvästi parempi kuin naiivilla mallilla saatu 0,6470, mikä osoittaa sen, että kertoimet sisältävät aitoa ennustevoimaa. Logistinen regressio vahvisti tämän havainnon, sillä kotivoiton implisiittinen todennäköisyys on erittäin merkitsevä selittäjä kotivoiton toteutumiselle. Kausittainen kalibraatioanalyysi osoitti lisäksi, että kertoimien tarkkuus on pysynyt vakaana tarkastelujakson aikana, joskin koronapandemian aikana kotivoiton alihinnoittelu oli hieman tavanomaista voimakkaampaa, mahdollisesti yleisöttömien otteluiden seurauksena.

Kolmas tutkimuskysymys koski sitä, voiko yksinkertainen vedonlyöntistrategia tuottaa positiivisen odotusarvon. Strategiasimuloinnin tulokset ovat tässä suhteessa yksiselitteiset, sillä yksikään testatuista 13 strategiasta ei tuottanut positiivista odotusarvoa. Kaikki strategiat olivat tappiollisia, ja suurin osa niistä oli tilastollisesti merkitsevästi negatiivisia. Ainoa poikkeus oli altavastajastrategia 1X2-markkinalla, jonka ROI oli -1,97 prosenttia, ei poikennut tilastollisesti merkitsevästi nolasta, mutta tämäkään ei tee strategiasta vielä kannattavaa. Kehittyneemmät strategiat, kuten arvovetomalli ja Kelly-kriteeri, eivät myöskään pystyneet tuottamaan positiivista odotusarvoa, joskin näiden tulosten tulkintaa rajoittaa tulevaisuuden näkemisen harha. Vedonvälittäjän keskimääräinen marginaali 6,72 prosenttia muodostaa käytännöllisen esteen positiivisen odotusarvon saavuttamiselle, ja tulokset viittaavat siihen, että julkisesti saatavilla olevaan dataan perustuvilla strategioilla ei pystytä ylittämään tätä kynnystä.

Neljäs tutkimuskysymys koski sitä, muuttuvatko vinoumat ja tehokkuus eri kausien välillä. Kausittainen tarkastelu osoitti, että markkinat ovat pysyneet pääosin tehokkaina koko tarkastelujakson ajan, mutta joitakin ajallisia vaihteluita on havaittavissa. Suosikin ROI on laskenut tarkastelujakson aikana, mikä viittaa siihen, että suosikkien ylihinnottelu on

pahentunut. Altavastaajan ROI vaihtelee merkittävästi kausittain, eikä se ole johdonmukaisesti positiivinen tai negatiivinen millään kaudella. Selkein ajallinen poikkeus on koronapandemian aikainen jakso, jolloin kotivoittoprosentti laski ja vierasvoittoprosentti nousi merkittävästi. Vedonvälittäjä ei täysin sopeutunut tähän muutokseen kertoimissaan, minkä voidaan ajatella olevan merkki markkinoiden rajoittuneesta sopeutumiskyvystä poikkeuksellisissa olosuhteissa.

Kokonaisuutena empiirisen analyysin tulokset tukevat näkemystä siitä, että Liigan vedonlyöntimarkkinat ovat melko tehokkaat. Kertoimet sisältävät merkittävää ennustevoimaa, ne ovat pääosin hyvin kalibroituineet ja yksinkertaisilla strategioilla ei pystytä saavuttamaan positiivista odotusarvoa. Pieniä ja osin tilastollisesti merkitseviä vinoumia esiintyy erityisesti kotiedun hinnoittelussa ja suosikki-altavastaaja-asetelmassa, mutta nämä vinoumat eivät ole riittävän suuria ylittääkseen vedonvälittäjän marginaalin. Tulokset ovat linjassa aiemman vedonlyöntimarkkinoiden tehokkuustutkimuksen kanssa, jonka mukaan erityisesti pienemmillä markkinoilla kertoimet voivat olla kohtuullisen tehokkaita, vaikka ne eivät saavutakaan suurten kansainvälisten sarjojen tehokkuuden tasoa.

7 Johtopäätökset

Tämän tutkimuksen tavoitteena oli arvioida vedonlyöntimarkkinoiden tehokkuutta suomalaisen jääkiekkoliiga Liigan ottelukertoimista. Tutkimuksessa tarkasteltiin, esiintyykö B365:n 1X2-kertoimissa systemaattisia vinoumia, vastaavatko kertoimista johdetut implisiittiset todennäköisyydet toteutuneita voitto-osuuksia, voiko yksinkertainen vedonlyöntistrategia tuottaa positiivisen odotusarvon ja muuttuvatko mahdolliset vinoumat kausien välillä. Aineisto kattoi kymmenen Liigan runkosarjakautta vuodesta 2014 vuoteen 2025, yhteensä 4 494 ottelua.

Tutkimuksen keskeinen tulos on, että Liigan vedonlyöntimarkkinat ovat melko tehokkaat. Kertoimet sisältävät merkittävää ennustevoimaa ja ne ovat pääosin hyvin kalibroituineet. Myöskään yksinkertaisilla strategioilla ei pystytä saavuttamaan positiivista odotusarvoa. Tulos on linjassa tehokkaiden markkinoiden hypoteesin kanssa, jonka mukaan markkinahintojen tulisi heijastaa kaikkea saatavilla olevaa informaatiota eikä systemaattista ylituottoa pitäisi olla saavutettavissa julkisesti saatavilla olevan datan perusteella (Fama, 1970).

Kalibraatioanalyysi osoitti, että kertoimet vastaavat toteutuneita lopputuloksia kohtuullisen tarkasti kaikissa kolmessa lopputulosluokassa. Yksikään todennäköisyysluokka ei tuottanut tilastollisesti merkitsevää eroa kotivoiton tai vierasvoiton osalta. Tasapelin osalta havaittiin pieni mutta ei tilastollisesti merkitsevä alihinnoittelu, eli tasapelejä tapahtui hieman useammin kuin kertoimet ennakoivat. Brier score -analyysi vahvisti, että kertoimet parantavat ennustetarkkuutta merkittävästi suhteessa naiiviin malliin, mikä osoittaa niiden sisältävän aitoa informaatiota otteluiden todennäköisyyksistä. Kotivoiton implisiittinen todennäköisyys on erittäin merkitsevä selittäjä kotivoiton toteutumiselle, joten logistinen regressio vahvisti tämän havainnon

Systemaattisia vinoumia havaittiin erityisesti kahdessa yhteydessä. Ensinnäkin kotiedun hinnoittelussa havaittiin pieni mutta johdonmukainen ylihinnoittelu, sillä kertoimet arvioivat kotivoiton todennäköisyyden keskimäärin 1,3 prosenttiyksikköä toteutunutta

korkeammaksi. Tämä näkyi myös strategia-analyysissä, jossa vierasjoukkueeseen pelaaminen tuotti selvästi paremman tuloksen kuin kotijoukkueeseen pelaaminen, vaikka molemmat strategiat olivat tappiollisia. Toiseksi suosikki-altavastaja-analyysi antoi viitteitä lievistä klassisesta suosikki-altavastaja-harhasta. Altavastajan ROI ei poikennut tilastollisesti merkitsevästi nolasta 1X2-markkinalla, kun taas suosikin ROI oli erittäin merkitsevästi negatiivinen. Strategioiden välinen ero oli lähes tilastollisesti merkitsevä, mikä antaa viitteitä siitä, että suosikkien todennäköisyys on kertoimissa hieman yliarvioitu. Tulos poikkeaa Woodlandin ja Woodlandin (1994, 2001) NHL:ssä havaitsemasta ilmiöstä, jossa altavastajiin pelaaminen osoittautui suhteellisesti kannattavaksi.

Vedonlyöntistrategioiden simulointi osoitti yksiselitteisesti, että yksikään testatuista 13 strategiasta ei tuottanut positiivista odotusarvoa. Vedonvälittäjän keskimääräinen marginaali 6,72 prosenttia muodostaa käytännöllisen haasteen positiivisen odotusarvon saavuttamiselle. Sekä yksinkertaiset että kehittyneemmät strategiat jäivät selvästi tappiolle simulaatiossa. Kehittyneempien strategioiden, kuten arvovetomallien ja Kelly-kriteerin heikko suoriutuminen viittaa siihen, että pelkästään julkisesti saatavilla olevaan dataan perustuvalla analyysillä ei pystytä tunnistamaan sellaisia alihinnoiteltuja kohteita, joita vedonvälittäjä ei olisi jo hinnoitellut oikein.

Kausittainen tarkastelu osoitti, että markkinat ovat pysyneet pääosin tehokkaina koko tarkastelujakson ajan, mutta joitakin ajallisia vaihteluita on havaittavissa. Selkein poikkeus oli koronapandemian aikainen jakso, jolloin kotivoittoprosentti laski ja vierasvoittoprosentti nousi merkittävästi yleisöttömien otteluiden seurauksena. Vedonvälittäjä ei täysin sopeutunut tähän muutokseen kertoimissaan, mikä on linjassa Fischer ja Haucapin (2022) havaintojen kanssa kotiedun hinnoittelun herkkyydestä markkinaolosuhteille. Suosikin ROI on lisäksi laskenut tarkastelujakson aikana, mikä viittaa siihen, että suosikkien todennäköisyys on kertoimissa yhä enemmän yliarvioitu suhteessa toteutuneisiin voitto-osuuksiin.

Lisäksi on syytä huomata, että havaitut vinoumat ovat pääosin pieniä ja osin tilastollisesti heikkoja, eikä niiden perusteella voida tehdä vahvoja johtopäätöksiä systemaattisesta virhehinnoittelusta. Esimerkiksi strategioiden väliset ROI-erot voivat selittyä useilla tekijöillä, jotka eivät liity suoraan markkinoiden tehottomuuteen. Ensinnäkin vedonvälittäjän marginaali ei välttämättä jakaudu tasaisesti eri lopputulosten välille, vaan se voi olla rakenteellisesti suurempi tietyissä kerroinluokissa. Toiseksi ottelukohtaiset tekijät, kuten avainpelaajien loukkaantumiset, maalivahdin vaihtuminen tai joukkueen otteluruuhka, voivat vaikuttaa ottelun tosiasiallisiin voimasuhteisiin tavalla, jota kertoimet eivät täysin heijasta mutta jota tämä tutkimus ei myöskään pysty kontrolloimaan. Kolmanneksi niin jääkiekko-otteluihin liittyy huomattava satunnaisvaihtelu, sillä maalimäärät ovat hyvin vaihtelevia ja yksittäiset tapahtumat kuten ylivoimatilanteet/pelaajan ulosajotilanteet tai maalivahtien suoritukset voivat ratkaista ottelun lopputuloksen tavalla, joka ei ole ennustettavissa kertoimista. Neljänneksi Liigan vedonlyöntimarkkinoiden likviditeetti on rajallinen, mikä voi aiheuttaa kertoimiin satunnaisia poikkeamia, jotka eivät heijasta systemaattista virhehinnoittelua vaan pikemminkin markkinan suhteellisen pientä kaupankäyntiä. Kaikkien näiden tekijöiden vuoksi tutkimuksen tulokset tulee ymmärtää viitteellisinä havaintoina markkinoiden toiminnasta, ja mahdollisten vinoumien tulkinnassa on syytä säilyttää varovaisuus.

Tutkimuksen tuloksia tulkittaessa on otettava huomioon muutamia heikkouksia. Ensinnäkin tutkimuksessa käytetään vain yhden vedonvälittäjän kertoimia, mikä voi vaikuttaa tulosten yleistettävyyteen. B365:n kertoimet ovat laajasti käytettyjä ja markkinaa hyvin edustavia, mutta eri vedonvälittäjien kertoimien vertailu voisi tuottaa lisäinformaatiota markkinoiden tehokkuudesta. Toiseksi kehittyneemmät strategiat kärsivät tulevaisuuden näkemisen harhasta, koska regressiomalli on estimoitu koko aineistolla. Oikeassa vedonlyöntitilanteessa malli tulisi estimoida vain menneellä datalla, mikä saattaisi muuttaa tuloksia. Lisäksi tutkimus kattaa vain runkosarjaottelut eikä pudotuspelejä, jotka voisivat tarjota erilaisen markkinaympäristön tehokkuuden testaamiseen. Täytyy myös huomioida, että Liiga on suhteellisen pieni markkina, joten

tulokset eivät välttämättä ole suoraan yleistettävissä suurempiin ja likvidimpiin vedonlyöntimarkkinoihin.

Jatkotutkimuksen kannalta olisi mielenkiintoista tarkastella useampien vedonvälittäjien kertoimia ja arvioida, eroavatko niiden hinnoittelukäytännöt toisistaan. Tutkimuksen aineistoon olisi hyvä myös lisätä moneyline-kertoimet. Lisäksi live-vedonlyönnin eli ottelun aikana tapahtuvan vedonlyönnin tehokkuuden tutkiminen voisi tuottaa uutta tietoa siitä, kuinka nopeasti markkinat reagoivat ottelun aikaisiin tapahtumiin. Myös kehittyneempien koneoppimismenetelmien hyödyntäminen ennustemallinnuksessa voisi tarjota paremman vertailukohdan markkinoiden tehokkuuden arvioimiseen. Liigan laajentuminen ja sarjan kehittyminen tarkastelujakson aikana tarjoavat myös mielenkiintoisen alustan tutkia, miten markkinoiden tehokkuus muuttuu sarjan ja sarjamallin muuttuessa.

Lähteet

Barberis, N., Shleifer, A., & Vishny, R. (1998). A Model of Investor Sentiment. *Journal of Financial Economics*. [https://doi.org/10.1016/S0304-405X\(98\)00027-0](https://doi.org/10.1016/S0304-405X(98)00027-0)

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383-417. <https://doi.org/10.2307/2325486>

Fama, E.F. (1991). Efficient Capital Markets: II. *The Journal of Finance*, 46(5), 1575-1617. <https://doi.org/10.2307/2328565>

Fischer, K., & Haucap, J. (2022). Home advantage in professional soccer and betting market efficiency: The role of spectator crowds. *Kyklos*, 75(2), 294-316. <https://doi.org/10.1111%2Fkykl.12291>

Gandar, J., Zuber, R., & Johnson, R.S. (2004). A reexamination of the efficiency of the betting market on national hockey league games. *Journal of Sports Economics*, 5(2), 152-168. <https://doi.org/10.1177/1527002503257208>

Humphreys, B., Rodriguez, P., & Simmons, R. (2017). *The economics of sports betting*. Edward Elgar.

Kahneman, D. (2012). Ajattelu nopeasti ja hitaasti. s. 319–332. *Terra Cognita*.

Levitt, S.D. (2004). Why Are Gambling Markets Organised So Differently from Financial Markets?, *Economic Journal*, 114(495), 223-246. Noudettu 28.01.2025 osoitteesta https://www.jstor.org/stable/3590093?read-now=1#page_scan_tab_contents

Moskowitz, T.J. (2021). Asset pricing and sports betting. *The Journal of Finance*, 76(6), 3153-3209. <https://doi.org/10.1111/jofi.13082>

Newall.P.W.S., & Cortis, D. (2021). Are sports bettors biased toward longshots, favorites, or both? A literature review. *Risks*, 9(1), 22. <https://doi.org/10.3390/risks9010022>

Oddsportal.com. 2026. Liiga Results, Scores & Historical Odds. <https://www.oddsportal.com/hockey/finland/liiga-2024-2025/results/>

Paul, R. J., & Weinbach, A. P. (2012). Sportsbook pricing and the behavioral biases of bettors in the NHL. *Journal of Economics and Finance*, 36(1), 123–135. <https://doi.org/10.1007/s12197-009-9112-4>

- Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4), 2021-2064. Noudettu 28.01.2025 osoitteesta <https://www.jstor.org/stable/2565046>
- Shin, H. S. (1993). Measuring the Incidence of Insider Trading in a Market for State-Contingent Claims. *Economic Journal*, 103(420), 1141-1153. <https://doi.org/10.2307/2234240>
- Smith, M.A., Paton, D., & Vaughan Williams, L., (2006). Market efficiency in Person-to-Person Betting. *Economica*, 73(292), 673-689. <https://www.jstor.org/stable/3874062>
- Snowberg, E. & Wolfers, J. (2010). Explaining the Favorite-Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy*, 118(4), 723-746. <https://doi.org/10.1086/655844>
- Vaughan Williams, L. (2005). Information efficiency in financial and betting markets. Cambridge University Press.
- Whelan, K. (2024). Risk aversion and favourite–longshot bias in a competitive fixed-odds betting market. *Economica*, 91(361), 188–209. <https://doi.org/10.1111/ecca.12500>
- Winkelmann, D., Ötting, M., Deutscher, C., & Makarewicz, T. (2024). Are betting markets inefficient? Evidence from simulations and real data. *Journal of Sports Economics*, 25(1), 54–97. <https://doi.org/10.1177/15270025231204997>
- Woodland, B. M., & Woodland, L. M. (2010). Market efficiency and the NHL totals betting market: Is there an under bias? *Economics Bulletin*, 30(4), 3122–3127. https://www.researchgate.net/publication/227410468_Market_Efficiency_and_the_NHL_totals_betting_market_Is_there_an_under_bias
- Woodland, L. M., & Woodland, B. M. (1994). Market efficiency and the favorite-longshot bias: The baseball betting market. *Journal of Finance*, 49(1), 269–279. <https://doi.org/10.1111/j.1540-6261.1994.tb04429.x>
- Woodland, L. M., & Woodland, B. M. (2001). Market efficiency and profitable wagering in the national hockey league: Can bettors score on longshots? *Southern Economic Journal*, 67(4), 983–995. <https://doi.org/10.2307/1061582>
- Woodland, L. M., & Woodland, B. M. (2011). The reverse favorite-longshot bias in the national hockey league: Do bettors still score on longshots? *Journal of Sports Economics*, 12(1), 106–117. <https://doi.org/10.1177/1527002510368792>

8 Liitteet

Liite 1. Tekoälyn käyttö tutkielmassa

Tämän pro gradu -tutkielman tekemisessä on hyödynnetty tekoälytyökalua (ChatGPT) tutkimus- ja kirjoitusprosessin tukena. Tekoälyä on käytetty muun muassa tiedonhaun apuvälineenä, suunnittelutyökaluna sekä rakenteen hahmottamisessa ja jäsentelyssä.

Tutkielman varsinainen sisältö, analyysi, tulkinat ja johtopäätökset ovat tekijän omia, eikä tekoäly ole vaikuttanut tutkimuksen tuloksiin. Tekoälyn käytössä on noudatettu Vaasan yliopiston ohjeistusta sekä hyvän tieteellisen käytännön periaatteita.