



**MS ritgerð**  
**Fjármálahagfræði**

**Árangur í fótbolta**  
Áhrif fólksfjölda og tekna

Þórir Gunnarsson

Leiðbeinandi: Helgi Tómasson

Hagfræðideild

Júní 2015



**HÁSKÓLI ÍSLANDS**

**Árangur í fótbolta**  
***Áhrif fólksfjölda og tekna***

Pórir Gunnarsson

Lokaverkefni til MS - gráðu í fjármálahagfræði

Leiðbeinandi: Helgi Tómasson

Hagfræðideild

Félagsvísindasvið Háskóla Íslands

Júní 2015

Árangur í fótbolta.

Ritgerð þessi er 30 eininga lokaverkefni til MS prófs við Hagfræðideild, Félagsvísindasviðs Háskóla Íslands.

© 2015 Þórir Gunnarsson

Ritgerðina má ekki afrita nema með leyfi höfundar.

Prentun: Háskólaprent

Reykjavík, 2015

## Formáli

Þetta 30 ECTS eininga verkefni er lagt fram til meistaraþrófs í fjármálahagfræði við Háskóla Íslands. Verkefnið er unnið undir handleiðslu Helga Tómassonar, prófessors við Hagfræðideild, og honum er kærlega þökkun veitt aðstoð.

Hagfræðingarnir Már Kristjánsson og Arnaldur Sölvi Kristjánsson veittu dýrmæt ráð. Þeir eiga þakkið skildar fyrir. Gunnar Jónsson, faðir minn, veitti ómetanlega aðstoð og þakka ég honum sérstaklega fyrir það.

## Útdráttur

Lagt er upp með að kanna hvort tengsl séu á milli árangurs í fótbolta, fólksfjölda og tekna. Árangur er skilgreindur á tvo ólíka vegu og viðeigandi líkön sett upp fyrir hvora skilgreiningu fyrir sig. Annars vegar eru sett fram logit líkön og hins vegar panel eða langsníðslíkön. Upplýsingar um árangur eru sóttar af vef Knattspyrnusambands Íslands. Mannfjölda- og tekjugögn koma frá Ríkisskattstjóra. Áhrif fólksfjölda á árangur eru greinileg en niðurstöður gefa ekki tilefni til ályktunar um áhrif tekna.

Þá er einnig farið yfir hagfræði íþróttá, sérstaklega fótbolta, og tæpt á rannsóknum og greinum um viðfangsefnið. Íslenskur fótolti er settur í samhengi við rannsóknir og greinar þar sem hæfir. Þá er því velt upp hvort skilgreina megi betur hlutverk íslenskra íþróttafélaga þar sem þau njóta stuðnings sveitarfélaga.

## Efnisyfirlit

Formáli .....	4
Útdráttur .....	5
Efnisyfirlit .....	6
Myndaskrá .....	8
Töfluskrá.....	8
1 Inngangur.....	9
2 Farið yfir sviðið.....	13
2.1 Fræði um íþróttir .....	13
2.1.1 Samhengi við íslenskan fótbolta .....	15
2.2 Fræði um fótbolta .....	17
2.2.1 Samhengi við íslenskan fótbolta .....	17
2.3 Uppgangur tölfræðilegrar greiningar á íþróttum.....	18
2.4 Líkön til skýringar árangurs í íþróttum .....	19
2.4.1 Samhengi við íslenskan fótbolta .....	21
2.5 Líkön af úrslitum leikja .....	22
2.6 Hagkvæmni veðmálamarkaða og fyrirsjáanleiki úrslita .....	26
2.7 Líkön til skýringar árangurs í fótbolta.....	27
3 Gögn.....	32
3.1 Umfjöllun.....	32
3.2 Úrlausnarefni.....	32
4 Tvíkostalíkön af árangri .....	34
4.1 Verklag og skilgreiningar .....	34
4.2 Tvíkostalíkön.....	36
5 Langsniðslíkön af árangri .....	45

5.1	Verklag og skilgreiningar .....	45
5.2	Langsniðslíkön .....	50
6	Niðurstöður.....	58
7	Umræður .....	59
	Heimildir.....	60

## Myndaskrá

Mynd 1. Raunárangur og mót líkana um líkur á árangri.....	42
Mynd 2. Raunárangur og mat líkans um líkur á árangri .....	44
Mynd 3. Fjöldi og tekjur á móti árangri .....	46
Mynd 4. Þróun meðaltekna yfir tíma, m.v. verðlag ársins 2012 .....	48
Mynd 5. Jaðaráhrif skýristærða á háðu breytuna .....	57

## Töfluskrá

Tafla 1. Breytur .....	36
Tafla 2. Jaðaráhrif skýribreyta .....	37
Tafla 3. Munur á líkum efri og neðri fjórðunga .....	38
Tafla 4. Mát nokkurra logit líkana.....	39
Tafla 5. Skýristærðir og háða stærðin eftir sveitarfélögum.....	47
Tafla 6. Skýristærðir og háða stærðin eftir tímabilum .....	49



## 1 Inngangur

Knattspyrna eða fótbolti er vinsælasta íþrótt í heimi og hefur verið það í talsverðan tíma (Palacios-Huerta, 2004). Miðað við stöðu fótboltans og gríðarlega umfjöllun um hann skyldi maður ætla að fjallað hafi verið ítarlega um öll svið hans. Ekki hefur skort á að fjallað sé um einstaka leiki, lið eða mót. Þar er þó gjarnan farið fremur grunnt yfir viðfangsefnið þar sem einkennandi eru deilur um ákveðin atvik eða dóma, t.a.m. hin sífellda deila um hvort bolti fari í hönd eða hönd í bolta. Sú umræða er sjálfsgöð og skiljanleg en ekki líkleg til þess að skilja mikið eftir sig. Líklegra er að raunveruleg eftirtekja verði af því að reyna að greina hvað það sé sem skiptir sköpum til að ná árangri.

Hagfræðin er í síauknum mæli notuð til þess að greina ýmiskonar vandamál og leita svara við margskonar spurningum. Hagfræðiritum skrifuðum fyrir breiðan hóp lesenda, ekki einvörðungis þá sem hafa fræðilegan bakgrunn, hefur fjölgað mjög undanfarin árin. Má nefna bækur eins og *The Undercover economist* (Harford, 2007), *Freakonomics* (Levitt og Dubner, 2006) og *Nudge* (Thaler og Sunstein, 2008). Sammerkt bókum af þessum toga er að umfjöllunarefni eru gjarnan fjölbreyttari en hefðbundið er í hagfræðinni. Auk þess er dregið úr fræðilegri umfjöllun í þessum ritum. Lesendum er þá í sjálfsvald sett hvort þeir kafi dýpra með lestri neðanmálgreina og því að rýna í heimildir. Líkön, töl og tæki hagfræðinnar nýtast vel til þess að fjalla um hin ýmsu málefni, hversu léttvæg sem þau virðast.

Fótbolti er dæmi um það sem nú er fjallað um en hefði vart talist hefðbundið viðfangsefni hagfræðinnar áður. Nefna má *Soccernomics* (Kuper og Szymanski, 2014) og *The Numbers Game* (Anderson og Sally, 2013) sem dæmi um bækur sem skoða fótbolta með augum hagfræðinnar. Þrjú af fjórum höfundum bókanna hafa verið prófessorar í hagfræði og byggja bækurnar á fræðilegum undirstöðum. Engu að síður eru efnistökin aðgengileg og bækurnar miðaðar að breiðum hópi lesenda. Kenningar og greiningar vísa í rannsóknir gerðar á fótbolta og þannig fæst aukin víska og sannfæring í því sem dregið er fram.

Vegna þess hversu umfjöllun um fótbolta er útbreidd bæði í daglegu tali og í fjölmiðlum er ljóst að af nógu er að taka í umfjöllun um íþróttina. Ef vilji er fyrir hendi má ítarskoða og rannsaka ótal atriði. Kostir til rannsókna á fótbolta eru margir. Niðurstöður leikja eru skýrar og þeim er almennt safnað saman í aðgengileg gagnasöfn. Jafnframt er mikil aukning í mælingu og söfnun upplýsinga um atvik sem eiga sér stað á meðan leik stendur. Fótbolti er einnig athyglisvert rannsóknarefni að því leitinu til að markmið þátttakenda eru ljós. Þá eru reglur leiksins fremur einfaldar, leikir taka fyrirfram ákveðinn tíma og úrslit að leik loknum eru ótvíræð, sama hve mörg deiluefni kunna að koma upp meðan á leik stendur.

Fótbolti hefur þann stóra kost sem rannsóknarefni, að gögn um rekstur eru almennt aðgengileg vegna kvaða frá viðkomandi knattspyrnusamböndum. Þau setja sér reglur sem gera þeim kleift að gæta þess að félög fari ekki í gjaldþrot eða stundi hegðun sem dregur úr samkeppni á fótboltavellinum. Liðum eru sett ýmiskonar skilyrði umfram það sem leiðir af landslögum og opinber skil á rekstrargögnum veita aðhald og koma í veg fyrir þrot. Rekstrargögnin bætast ofan á öll þau gagnasöfn sem leiðir af leiknum sjálfum með stöðutöflum, markatölum, sigurhlutföllum og fjölmörgu öðru. Aðgengileiki gagna auðveldar hverskyns rannsóknir. Sannfærandi kenningar fá aukinn trúverðugleika ef styðja má við þær með rannsóknum byggðum á viðurkenndum vísindalegum eða fræðilegum aðferðum. Þannig má sannreyna eða afsanna fjölda kenninga með því að beita fræðilegum aðferðum við greiningu gagna sem verða til um fótbolta.

Vinsældir fótbolta gera hann að áhugaverðu rannsóknarandlagi. Vinsældirnar má t.d. nýta við kennslu á efni sem talið er gagnslaust eða óspennandi. Dæmi má finna til notkunar í kennslu á hinum ýmsu fögum s.s. tölfræði, lögfræði, hagrannsóknnum eða hagfræði. Dreifing fjölda marka í fótbolta er notuð sem dæmi í kennslubók í tölfræði (Moroney, 1965). Þekkt og áhugavert dómssmál er kennt við fótboltamanninn Bosman. Ágreiningurinn í málinu snerist um atvinnufrelsi leikmannsins. Bosman dómurinn gjörbylti leikmannamarkaðnum og reglur um félagaskipti leikmanna þurftu allsherjar yfirhalningu í kjölfar hans. Um það má t.a.m. lesa í greininni „The Bosman Ruling and the Emergence of a Single Market in Soccer Talent“ (Antonioni og Cubbin, 2000).

Hagrannsóknir skipa veigamikinn sess í bókinni *Economics of Football* eftir Dobson og Goddard (2001). Bæði hagrannsóknir og hagfræði koma við sögu í nýlegri grein um

hvernig samleitni, þekkt kenning úr þjóðhagfræði um að lönd hafi til lengri tíma tilhneigingu til að vaxa saman, fellur að frammistöðu landsliða í fótbolta (Szymanski, 2014). Á sama tíma og niðurstöður þjóðhagfræðilegra rannsókna varðandi samleitni í landsframleiðlu landa hafa verið tvíræðar, eru vísbendingar um samleitni árangurs landsliða tíundaðar í greininni. Telja má upp fjölda annarra dæma en þau framantöldu nægja til þess að sýna að margvísleg áhugaverð rannsóknarefni finnast í fótboltaheiminum.

Í hagfræði eru fræðileg líkön notuð til að skýra heiminn. Sérstaklega er líkönum beitt til skoðunar á áhrifum ýmiskonar breytinga. Skellir eru sagðir eiga sér stað og líkön eru nýtt til að skoða hvað á sér stað og bera saman stöðu yfir tíma og fyrir og eftir skell. Fjármagn og vinnuafli skipa gjarnan sess í þannig líkönum. Ásamt landi og/eða náttúruauðlindum eru þessi atriði það sem kallast framleiðsluþættir og eru grunnur margra hagfræðilíkana. Afurð er sköpuð úr framleiðsluþáttum. Heil svið hagfræðinnar snúa beint að skoðun á þessum þáttum. Vinnumarkaðshagfræði snýr að vinnuafli og umhverfi þess. Þjóðhagfræði má nefna sem dæmi um svið hagfræðinnar sem skoðar fjármagn, en í raun er fjármagn gegnumgangandi skoðað í öllum sviðum hagfræðinnar. Umhverfis- og auðlindahagfræði snýr svo sérstaklega að land- og auðlindanýtingu. Skilningur á því hvaða kraftar virka á þessa þætti, og með hvaða hætti, eru ein af undirstöðum hagfræðinnar.

Fjármagn og vinnuafli eru framleiðsluþættir sem eru nauðsynlegir allri framleiðslu. Hverskonar framleiðsla sem á sér stað byggist að minnsta kosti á notkun þessara tveggja þátta. Á það jafnt við í íþróttum og á öðrum sviðum. Framleiðsla íþróttar er ekki jafn áþreifanleg vara og á við um flesta aðra framleiðslu. Íþróttir framleiða leiki til áhorfs. Árangur í íþróttum er önnur „vara“ sem keppst er um að framleiða. Mælingar á fjármagni og vinnuafli geta verið með ýmsu móti. Einföld aðferð við að tölusetja fjármagn og vinnuafli er að mæla fjölda fólks og tekjur þess. Sú aðferð er notuð í þeim líkönum sem sett eru upp hér. Líta má á fólksfjölda sem mælikvarða á stærð vinnuafli og einnig má líta á landsframleiðslu á mann sem mælikvarða á fjármagn. Sýnt hefur verið fram á að alþjóðlegur árangur í íþróttum skýrist að einhverju leyti af fólksfjölda og landsframleiðslu á mann. Forvitnilegt er að færa þannig rannsókn á smærra stig, eða

sveitarfélög, og athuga hvort skýra megi árangur í fótbolta með fjármagni og vinnuafli, eða tekjum og fjölda.

Í ritgerðinni eru sett fram líkön til þess að greina og tengja saman árangur knattspyrnuliða á Íslandi við fólksfjöldann á svæðinu sem þau telja sitt og tekjur þeirra sem þar búa. Farið er yfir hagfræði íþróttanna og fótbolta og skoðaðar ýmsar hagfræðilegar rannsóknir á íþróttum, sérstaklega fótboldatengdar. Líkön sem gerð hafa verið til að líkja eftir úrslitum leikja eru tínd til, einkum þau sem skoða úrslit út frá þjóðhagfræðilegum stærðum eins og mannfjölda og tekjum. Eftir það eru líkön sett upp í viðleitni til þess að athuga hvort þau styðji kenningu um að árangur liða skýrist af fólksfjölda og/eða tekjum. Skipta má uppsettum líkönum í tvo flokka. Árangur er skilgreindur á tvo ólíka vegu. Fyrst er árangur skilgreindur sem tvíkosta breyta og logit líkön nýtt til að greina hann. Síðan er skilgreining árangurs búin til út frá stigum í fótbolta og er árangurinn skoðaður yfir tíma. Langsniðslíkön eru notuð til að greina síðari skilgreininguna. Í lokin eru niðurstöður líkana gerðar upp og þær bornar saman við fyrri verk.

Líkönin eru valin eftir skoðun á fyrri verkum og í tengslum við þau gögn sem eru fyrir hendi. Til glöggvunar fylgir ítarleg lýsing á tæknilegri útfærslu líkananna. Því næst er ályktað um líkönin og þau vegin og metin eftir því hversu áreiðanleg þau eru. Þá er leitast við að gera úrbætur á líkönunum. Loks eru stíkar metnir og farið yfir niðurstöður.

## 2 Farið yfir sviðið

### 2.1 Fræði um íþróttir

Endurtekið er vísað í þrjár greinar sem undirstöðu hagfræðilegrar greiningar á íþróttum (Avgerinou, 2007; Bougheas og Downward, 2000; Dobson og Goddard, 2001; Procházka, 2012; Szymanski, 2003). Elst hagfræðilegra greininga á hópíþróttum er grein eftir Rottenberg (1956), Neale (1964) fylgdi í kjölfarið og síðan Sloane (1971). Rottenberg tók fyrir aðstæður á vinnumarkaði í hafnabolta í greininni „The Baseball Players' Labor Market“. Neale voru hugleikin einkenni atvinnuþróttar og skrifaði ritgerðina „The Peculiar Economics of Professional Sports“. Sloane rýndi í hegðun hjá knattspyrnufélögum og dró saman í greininni „The Economics of Professional Football“.

Vinumarkaður atvinnuhafnabolta var greindur af Rottenberg og þá sérstaklega hinar ýmsu hömlur sem valdi leikmanna til þess að semja um kaup og kjör eða félagaskipti voru settar. Hömlurnar voru tilraun til þess að stuðla að jafnari dreifingu hæfileikraríkra leikmanna sem myndi leiða af sér jafnari og meira spennandi keppni. Takmarkanir á rétti ríkustu liðanna til kaupa á bestu leikmönnum, sem gerðu leiki fyrirsjáanlegri, voru talin viðunandi réttlæting fyrir viðvarandi hömlum á leikmannamörkuðum, jafnvel í hinum annars markaðssinnuðu Bandaríkjum.

Í greininni færir Rottenberg (1956) rök fyrir því að frjálsum vinnumarkaður leikmanna leiði ekki endilega til þess að bestu leikmennirnir myndu safnast til fárra ríkra liða. Hafnabolti er hópíþrótt og vegna minnkandi jaðarframlegðar viðbótarleikmanns er virði hvers nýs stjörnuleikmanns æ minna fyrir lið með marga hæfileikaríka leikmenn. Verðmæti stjörnuleikmanns fyrir lakt lið er meira þar sem jaðarframlegð hans er meiri í slíku liði en sem enn einnar stjörnu í afburða liði. Frelsi á leikmannamarkaði þýðir því ekki endilega allsherjar uppkaup ríkra liða á bestu leikmönnum. Rottenberg færir í raun hefðbundin hagfræðileg rök fyrir samkeppni og að af frjálsum markaði leiði hagkvæma dreifingu og notkun aðfanga og vinnuafls.

Sérkennileg hagfræði atvinnuþróttar er kveikjan að skrifum og þýðing á heiti greinar Neale (1964). Sameiginleg framléiðsla keppenda er greind og því er velt upp að „varan“ sem framléidd er af íþróttum sé kappleikir til áhorfs og besta varan sé keppnir þar sem leikar eru jafnir og spennandi. Bent er á það sem kallað er Louis-Schmeling þversögnin. Nafnið vísar í tvo heimsmeistara þungavigtar í hnefaleikum, svarta Bandaríkjamanninn

Joe Louis og mótherja hans, Max Schmeling frá Þýskalandi nasismans. Bardagar þeirra þóttu sérlega spennandi hvoru tveggja vegna stöðu alþjóðastjórn mála og sem bardagi heimsmeistara við öflugan áður ósigraðan andstæðing. Reiknað er með að eins og aðrir vilji heimsmeistari í boxi hámarka hagnað. Til þess þarf hann öflugan mótherja og eftir því sem hann er öflugri vaxa tekjur af bardaganum vegna aukins áhuga. Einokun, sem almennt er talin eftirsóknarverð staða, fylgdi tekjutap einokarans þar sem fæð verðugra andstæðinga leiddi til óspennandi kappleikja. Þversögnin felst í því að samkeppni í íþróttum er af hinu góða fyrir tekjur keppenda.

Neale (1964) fer yfir sérkenni samkeppni á sviði íþróttar og færir rök fyrir því að líta eigi á íþróttabandalög, eða þau félög sem halda utan um mót og keppnir, t.d. KSÍ eða NBA, sem fyrirtæki í hagfræðilegum skilningi (e. firm). Samkeppni fer svo fram á milli íþróttabandalaga um áhorf en innan þeirra er samvinna fremur en samkeppni á milli liða og einstaklinga. Þá er fullyrt að ljóst sé að atvinnuþróttir búi við náttúrulega einokun bæði vegna hlutverks og uppbyggingar markaðanna sem þær fara fram á.

Sloane (1971) skoðar enskan atvinnumannafótbolta m.t.t. forsendu sem áður hafði verið gengið út frá í hagfræðilegum verkum um íþróttir. Miðað var við að íþróttafélög leituðust fyrst og fremst við að hámarka hagnað, eins og almennt er reiknað með að fyrirtæki geri. Athugun Sloane á þessari forsendu leiðir í ljós að hún á ekki við um enskan fótbolta. Spá um hegðun félaga byggð á forsendunni passar ekki við hegðun þeirra í reynd. Nokkrar mögulegar háværkunarforsendur eru skoðaðar og hve vel þær falla að raunverulegri hegðun íþróttafélaganna. Bornar eru upp forsendur um háværkun hagnaðar, öryggi liðs, aðsókn leikja og nytjar. Sloane telur þá nálgun að lið leitist við að hámarka nytjar best falla að hegðun þeirra. Imprað er á því vandamáli við forsendu um háværkun nytja að skilgreina megri nytjar á þann veg að nánast hvaða hegðun sem er sé viðleitni til nytjahávrkunar. Forsenda um háværkun nytja fellur þó vel að fótbolta þar sem almenn samstaða er um að árangur í íþróttinni sé einkanlega frammistaða í leikjum og nytjar af fótbolta hljóttist af árangri.

Margar hugmynda Neale og Sloane stangast á. Sloane lítur á íþróttafélög sem sjálfstæðar einingar eða fyrirtæki í samkeppni þar sem ákveðið samráð á sér stað. Félag í samráði taka gjarnan sameiginlegar ákvarðanir en Sloane lítur svo á, ólíkt Neale, að þrátt fyrir sameiginlegar ákvarðanir eigi að líta á íþróttafélög fremur en íþróttabandalög

sem fyrirtæki í hagfræðilegum skilningi. Sameiginleg ákvarðanatataka er hvorki nauðsynlegt né nægjanlegt skilyrði fyrir því að líta á bandalög sem fyrirtæki í hagfræðilegum skilningi. Forsenda Neale um að íþróttafélög leitist helst við að hámarka hagnað er einnig í andstöðu við niðurstöðu Sloane. Forsenda Neale er þó nær lagi um bandarísk íþróttafélög en evrópsk. Íþróttafélög í Bandaríkjunum hafa fremur sýnt tilhneigingu til þess að leitast við að hámarka hagnað með því að skila reglulegum hagnaði. Sloane sýnir fram á að forsendan sé ekki altæk með því að sýna fram á að hún standist ekki séu ensk knattspyrnufélög skoðuð.

Sloane (1971) gerir grein fyrir ólíkum áhrifum forsendna um hámarks hagnaðar fremur en nytja á niðurstöðu um hvernig dreifingu hæfileikaríkra leikmanna yrði háttáð á frjálsum leikmannamarkaði. Líkt og fyrr segir yrði fallandi jaðarframlegð til þess að leikmenn dreifðust sé miðað við hámarks hagnaðar (Rottenberg, 1956). Auk þess myndu miklir yfirburðir leiða til óspennandi leikja sem drægi úr aðsókn og tekjum þannig að leikmenn dreifðust nokkuð miðað við forsendu um hámarks hagnaðar. Sé hins vegar miðað við hámarks hagnaðar nytja eða árangurs eru engar skorður við því hve lið geta safnað til sín mörgum stjörnuleikmönnum til þess að auka árangur á vellinum. Rök fyrir hömlum á frjálsum markaði með leikmenn eru þannig til staðar ef leitast á við að hafa leiki spennandi. Með kenningu Neale (1964) í huga, um að líta á íþróttabandalög sem fyrirtæki (e. firm), má einnig velja því fyrir sér hvort takmarkanir á rétti leikmanna til þess að semja um félagaskipti, kaup og kjör, séu til þess fallin að hámarka hagnað heildarinnar. Takmarkanirnar verða til þess að liðin stunda hálfgerða rentusókn, þar sem gengið er á hlut leikmanna, til að auka hagnað hjá liðunum.

### **2.1.1 Samhengi við íslenskan fótbolta**

Markaður knattspyrnumanna á Íslandi er frjálur, sérstaklega í samanburði við markað hafnaboltaleikmanna þess tíma sem Rottenberg (1956) fjallaði um. Kenning Rottenberg um að á frjálsum markaði leikmanna dreifist þeir hæfileikaríku milli liða vegna hagkvæmrar ráðstöfunar leikmanna milli liða fellur ágætlega að íslenskum fótbolta. Því til stuðnings má nefna að á síðustu 20 tímabilum úrvalsdeildarinnar hafa sjö mismunandi lið unnið deildina. Á sama tímabili hafa einungis fjögur lið unnið ensku úrvalsdeildina.

Forvitnilegt er að velta fyrir sér hlutverki og markmiðum íslenskra íþróttafélaga í samhengi við kenningu Sloane (1971) um misjöfn áhrif ólíkra hámarkunarforsendna. Hlutverk íslenskra íþróttafélaga er víðtækt, þau sjá um margskyns íþróttir og þjálfa fólk á öllum aldri. Félögin eru ekki í eigu fjárfesta heldur eru byggð upp af félagsmönnum með stuðningi sveitarfélaga flest hver. Samfélagslegt hlutverk þeirra er boðað m.a. í reglum KSÍ þar sem kvaðir eru á liðum sem taka þátt í efri deildum um að vera með yngri flokka starf. Því verður ekki haldið fram að helsta markmið félaganna sé að hámarka hagnað. Fremur er hlutverk þeirra að hámarka nytjar. Af hegðun liða og orðum þeirra sem stýra þeim má ætla að árangur á vellinum vegi þungt í því víðfeðma sviði sem fallið getur undir hugtakið nytjar.

Eignarhaldið á íslenskum íþróttafélögum er athyglisvert og má reyndar velta því fyrir sér hvort eignarhaldi á þeim sé yfirleitt til að dreifa. Þau eru um margt líkust samvinnufélögum í eigu félagsmannanna, sem eru þá þeir sem stunda íþróttir undir merkjum félagsins og jafnvel stuðningsmenn. Sveitarfélögin leika almennt stórt hlutverk í starfi íþróttafélaga og styðja þau til þess að gegna samfélagsþjónustuhlutverki í starfi yngri flokka, ásamt því að sjá um rekstur íþróttamannvirkja sem sveitarfélögin hafa meira og minna reist. Á sama tíma og þessu hlutverki er sinnt keppa íþróttafélögin í meistaraflokkum og fylgja því gjarnan veruleg fjárútlát. Í öllu falli felst talsverð óvissa í tekjum meistaraflokka, sem tengist m.a. árangri á vellinum í gegnum miðasölu og verðlaunafé. Þessi ólíku hlutverk og markmið stangast á. Sjálfstæði íþróttafélaga frá sveitarfélögum, sem iðulega standa undir stórum hluta rekstrarkostnaðar, minnir um sumt á fræga vísun þingmanns í biblíuna í umfjöllun um vörslu og ráðstöfun vandalausra á peningum annarra sem fé án hirðis (Pétur Blöndal, 1994).

Fótbolti er gífurlega vinsæl íþrótt sem keppt er í um allan heim og mikil samkeppni er um að næla í góða leikmenn. Vegna þess að íslenski leikmannamarkaðurinn er opinn fyrir erlend lið má líta svo á að efri mörk séu á hve góðir leikmenn spili í íslensku deildinni. Bestu bitarnir eru tíndir út til erlendra liða sem eru efnaðri en þau íslensku. Með auknum peningum í fótbolta eykst vandinn þar sem íslensk lið hafa engan grundvöll til þess að keppa um að halda í bestu leikmennina. Þess vegna og fyrir það að leikmenn hafa metnað til að komast í erlend lið er erfiðara fyrir íslensk lið að kaupa árangur. Góðir og metnaðarfullir leikmenn horfa út fyrir landsteinana og upphæðir sem



viðgangast eru þannig að íslensk félög þurfa að sætta sig við að hafa engin áhrif á þessum stóra markaði og mega sín lítils í að halda í leikmenn sem vilja og geta farið út.

## **2.2 Fræði um fótbolta**

Fótbolti hefur orðið innblástur ótal fræðigreina og athyglisvert er hve fjölbreytt efni þeirra er. Þannig hefur til dæmis verið farið yfir það hvort knattspyrnulið njóti góðs af því að heimavöllur sé á gervigrasi (Barnett og Hilditch, 1993). Hlutlægni og samkvæmni knattspyrnudómara hefur verið rannsökuð (Dawson, Dobson, Goddard og Wilson, 2007). Kenningu um samkvæmni milli dómara er hafnað þar sem gerla má greina mun á milli þeirra. Auk þess er munur á fjölda brota heimaliðs og aðkomuliðs. Munurinn virðist stafa af tilhneigingu dómara til þess að dæma fremur heimaliðum í hag. Leikjafræði og vítaspyrnur atvinnufótboltamanna hafa verið bornar saman og samræmist hegðun leikmannanna bestu hegðun skv. kenningum leikjafræði (Chiappori, Levitt og Groseclose, 2002; Palacios-Huerta, 2003). Aðsókn á fótboltaleiki er meðal rannsóknæfnis. Líkt og við má búast er meiri aðsókn á leiki sem eru taldir líklegir til þess að vera jafnir. Einnig er meiri aðsókn á leiki liða sem eru góð í þeim skilningi að standa sig betur en flest önnur (Forrest og Simmons, 2002; J. García og Rodríguez, 2002).

Talsvert hefur verið skoðað hvernig hægt sé að gera knattspyrnu meira spennandi, svo sem með því að gera deildir jafnari (Szymanski, 2007; Vrooman, 2007). Allmikið hefur verið um rannsóknir á áhrifum þjálfara á árangur liða. Nefna má rannsóknir á knattspyrnudeildum Þýskalands (Frick, Barros og Prinz, 2010; Frick og Simmons, 2008), Hollands (Bruinshoofd og ter Weel, 2003), Ítalíu (Paola og Scoppa, 2011) og Englands (Audas, Dobson og Goddard, 1997). Áhrif þess að skipta um þjálfara eru sérstaklega skoðuð og er nokkur samhljómur rannsókna um að þjálfaraskipti skili takmörkuðum árangri.

### **2.2.1 Samhengi við íslenskan fótbolta**

Yfirfæra má margar þessara erlendu rannsókna á íslenskan fótbolta og skoða hvernig þær eigi við. Vinsæl rannsókn er að skoða á áhrif þess að skipta um þjálfara á árangur og væri einfalt og skemmtilegt að færa það yfir á íslenskan fótbolta. Talsverð umræða hefur verið um notkun gervigrasvalla á Íslandi og mætti athuga hvort lið með þannig heimavöll njóti góðs af því. Aðsóknartölur á leiki íslenskra liða liggja fyrir og mætti nota í

rannsókn á áhrifum áhorfenda á niðurstöður leikja. Sérstaklega væri áhugavert að komast yfir gögn um rekstur íslenskra fótboltaliða og yfirfæra erlendar rannsóknir um rekstur fótboltaliða á íslensk lið, enda félagsformið afar ólíkt því sem annarstaðar tíðkast.

### 2.3 Uppgangur tölfræðilegrar greiningar á íþróttum

Hagfræðin er í auknum mæli notuð til umfjöllunar um álitamál tengd fótbolta og íþróttum almennt. Bækur, tímarits- og fræðigreinar eru nú skrifaðar um spurningar sem kvikna varðandi fótbolta og umgjörðina í kringum hann.

Sérstaklega hefur notkun tölfræði í almennri umfjöllun og greiningu á íþróttum aukist. Það á jafnt við hvort sem verið er að skoða rekstur liða, ársreikninga, launakostnað, leikmannakaup o.s.frv. eða kappleiki og þau atvik sem eiga sér stað á keppnisvællinum. Þróunina má að miklu leyti rekja til útgáfu bókarinnar *Moneyball* eftir Michael Lewis (2003). Í henni segir Lewis sögu undirmáls hafnaboltaliðs sem snýst til betri vegar eftir að gömlum tuggum og klisjum er hafnað en þess í stað treyst á frumlega hugsun og fræðilegar aðferðir. Í þessu ferli eru reyndir ráðgjafar reknir og ungir hagfræðingar leysa þá af hólmi. Árangur liðsins, og síðar frásögn Lewis, vakti verðskuldaða athygli. Vinsæl kvikmynd var gerð á grunni sögu bókarinnar og margskonar lærdómur hefur verið dreginn af frásögninni. Þar má ekki síst nefna að líta eigi hefðir og settar venjur gagnrýnum augum. Vísindalegum aðferðum er beitt til þess að sannreyna það sem talinn hefur verið sannleikur og styðji tölurnar hann ekki, er honum og hugmyndum byggðum á honum hiklaust hafnað. Miskunnarlaust er farið yfir hvað virkar og hvað megi fara betur og því fleygt sem ekki stenst skoðun. Þar sem tölurnar eru óvægnar mæta nýjar aðferðir andstöðu og þeir sem starfað hafa eftir eldri aðferðum eru tortryggir. Með því að líta hafnabolta gagnrýnum augum og skoða hvað það væri sem raunverulega skóp sigra voru fundnir möguleikar til að auka sigurlíkur. Til að mynda var venja að skoða vel hve vel hafnaboltaleikmenn hittu boltann að jafnaði. Leikmenn sem náðu að halda góðu hlutfalli hitra kasta kostuðu meira á markaði leikmanna. Með því að einblína á þessa stærð var litið hjá því að hægt er að komast í fyrstu höfn, og þ.a.l. auka líkur á að skora stig, með því að fá að ganga í fyrstu höfn, sem er leyft ef kastarinn hittir ekki tiltekinn ramma oftast en þrisvar eða vegna þess að kastað er í leikmanninn. Með því að bæta við fjölda skipta sem leikmaður fær að ganga í fyrstu höfn við það hve

hittinn hann er fæst betri mynd af raunverulegu virði hans fyrir sigurlíkur liðs. Með notkun þessarar nýju tölfræði mátti um tíma finna „hagstæða“ leikmenn.

Gleggsta dæmið um mýtu sem ekki stóðst var hve yfirborðskenndar ráðningardeildir voru. Dómar voru felldir út frá útliti án þess að frekari rökstuðning þyrfti til. Leikmenn voru afskrifaðir sem féllu ekki að staðalímynd og þeir sem gerðu það á yfirborðinu áttu svo órökrétt forskot á aðra um framgang. Viðteknir fordómar, gjarnan byggðir á útliti og almennri grunnhyggni, leiddu af sér möguleika á að spila á þá. Segja má að til staðar hafi verið tækifæri til að hagnast með því að spila rétt á mýtur leikmannamarkaðarins.

Kenning um að samband launa og framlegðar leikmanns til liðs hafi verið óskilvirk er færð fram í *Moneyball*. Kenningin hefur verið sannreynd fræðilega (Hakes og Sauer, 2006) en í þeirri athugun er þó gerð grein fyrir því að stuttu eftir útgáfu bókar Lewis hætti „högnunartækifærið“ að vera til staðar á leikmannamörkuðum. Sambærileg rannsókn hefur verið gerð þar sem árangur fótboldaliða er skoðaður í tengslum við hlutfall svartra leikmanna og launakostnaðar (Szymanski, 2000). Þjögun var til staðar þar sem laun svartra fótboldamanna voru ekki í réttu hlutfalli við getu þeirra. Ekki þurfti þó bók til þess að þjögunin leystist heldur greiddi markaðurinn úr málinu þegar tekið var eftir misræminu og hæfileikar svartra leikmanna fóru að endurspeglast rétt í launum þeirra.

## 2.4 Líkön til skýringar árangurs í íþróttum

Ýmsir hafa ætlað sér að skýra hvað sker úr um árangur í íþróttum á milli þjóða. Í þeirri viðleitni hefur þáttum að baki árangri verið skipt í þrennt og er árangur skilgreindur sem fjöldi verðlaunapeninga á Ólympíuleikum (De Bosscher, De Knop, Van Bottenburg og Shibli, 2006). Skipt er í stærsta, mið og smæsta flokkinn (e. macro-, meso- and micro-level) og litið yfir mikilvæga þætti sem geta leitt til árangurs á alþjóðavettvangi í íþróttum í hverjum flokki. Undir stóra flokkinn falla þættir eins og fólksfjöldi, veðurfar, menning og hagsæld. Sameiginlegt þáttum efsta flokksins er að erfitt er að hafa áhrif á þá, þeir eru tregbreytanlegir í besta lagi, ef yfir höfuð má hafa áhrif á þá. Í miðflokk falla atriði eins og straumar og stefnur í stjórnámálum og hjá þeim sem fara með stjórn íþróttamála. Á þessu stigi eru tækifæri til þess að hafa jákvæð áhrif á árangur með góðri umgjörð í kringum íþróttir. Smæsti flokkurinn nær yfir íþróttamennina sjálfa og þeirra næsta umhverfi svo sem þjálfara, fjölskyldu og vini. Á smæsta stiginu má hafa áhrif á

einhverja þætti, t.a.m. þjálfunaraðferðir, meðan öðrum, t.d. erfðum, verður ekki haggað. Þessi flokkun er að sjálfsögðu einföldun og áhrifaþættir geta verið við mörk eða fallið á milli stiga. Að flokkun lokinni er litið yfir rannsóknir í hverjum flokki fyrir sig.

Flestar rannsóknir gerðar á uppsprettu árangurs í íþróttum falla undir stærsta flokkinn (De Bosscher o.fl., 2006). Atriði stærsta flokksins eru flest mælanleg og almennt til um þau gögn. Aðgengi að gögnum gerir skiljanlegt að sá flokkur hafi helst verið rannsakaður. Þær rannsóknir hafa margar verið aðhvarfsgreining eða skoðun á fylgni á þáttum sem skýra eiga árangur. Þannig hafa þættir eins og fólksfjöldi og lands- eða þjóðarframleiðsla á mann oft verið skoðaðir og reglubundið verið sýnt fram á að þeir skýri um helming breytileika (e. total variance) á alþjóðlegum árangri í íþróttum.

De Bosscher og meðhöfundar (2006) furða sig á því hve fáar rannsóknar hafa verið gerðar á árangursþáttum sem falla undir miðflokkinn. Fjárútlát íþróttabandalaga eru sett í samhengi við það hve miklu sé ósvarað þegar kemur að því hvað raunverulega skili árangri í íþróttum. Auk þess eru mestu möguleikarnir til þess að hafa áhrif í þessum flokki. Höfundar telja að betur megi gera í að skoða það sem fellur undir miðflokk til að markvissari árangur náist sem réttlæti fjáraustur.

Þættir miðflokksins mælast ekki svo glögg. Fæð rannsókna í flokknum stafar eflaust af erfiðleikum við gefa einstökum þáttum flokksins einkunnir eða setja á þá tölugildi. Því hafa ýmsar leiðir verið farnar í rannsóknum sem falla undir hann. Má nefna lýsingar og samanburð á skipulagi og stofnunum tengdum íþróttum á milli landa. Athuganir á forsendum þess að íþróttarárangur náist yfir höfuð hafa einnig verið gerðar, ásamt skoðunum á því sem þjóðir sem ná árangri eiga samþætt. Af þessum verkum er dreginn ýmiskonar lærdómur um hvað skipti máli til að ná alþjóðlegum árangri. Dæmi þar um er mikilvægi skýrra markmiða og markvissra hlutverka stofnana sem viðkoma íþróttum, að snemma sé tekið eftir efnilegum einstaklingum og rétt hlúð að þeim, möguleikar á fjárhagslegum stuðningi og vel skipulögð mót með fjölþjóðlegri þátttöku. De Bosscher og meðhöfundar (2006) gera þó grein fyrir því að engum hefur tekist að búa til fræðilegt líkan með skýrum mælanlegum þáttum svo hægt sé að sannreyna kenningar þeirra með tölfræðilegum aðferðum.

Í smæsta flokknum er hægt að nálgast og ræða beint við þá sem falla undir hann, íþróttamenn og þjálfara, öfugt við stofnanir og lönd í miðflokknum. Skýrari niðurstöður

hafa því fengist í flokknum af áhrifaþáttum bæði til góðs og ills á árangur í íþróttum. Svör íþróttamanna bera með sér að hvatning, áræðni og einlægur áhugi er þeim efst í huga varðandi það hvað vegi mest í árangri í hjá þeim sjálfum (De Bosscher o.fl., 2006). Varðandi smæsta flokkinn er einkunnagjöf þátta flokksins einnig vandamál. Svörin eru þó skýrari en í miðflokknum þar sem þau stafa beint frá íþróttamönnum.

#### **2.4.1 Samhengi við íslenskan fótbolta**

Tilfni er til að huga vandlega að því hvernig staðið er að baki íþróttum hér á landi, hvert markmið íþróttafélaga sé og hvaða leið eigi að fara að því marki. Íþróttafélög á Íslandi eru sérstök vegna blandaðs hlutverks nokkurs konar samfélagsþjónustu og þess að keppa um árangur meðal þeirra bestu á landinu, jafnvel á mörgum sviðum í einu. Skýrar má setja fram hverju lið stefni að og íþróttasambönd og sveitarfélög þurfa að velta fyrir sér hverju þau leitist við að ná fram.

Hætta er á umboðs- og freistnivanda vegna þess að hlutverk, markmið og stefna eru ekki nægilega skýr hjá íþróttafélögum studdum af sveitarfélögum. Freistnivandinn (e. moral hazard) kemur fram þegar meiri áhætta er tekin en ef sá sem tekur hana myndi bera áhættuna sjálfur. Umboðsvandi (e. principal-agent problem) er sambærilegt vandamál sem kemur upp þegar upplýsingar eru ósamhverfar, ekki þær sömu milli aðilanna, og hagsmunir þess sem umboð veitir og þess sem með umboð fer eru ekki alveg þeir sömu. Vegna ósamhverfra upplýsinga, eða takmarkaðrar þekkingar þess sem veitir umboðið, er dregið úr getu hans til að finna að því sem betur má fara. Aðgát í rekstri félags getur þá misfarist. Má nefna hættu á að viðleitni sveitarfélags til að sinna almenntri lýðheilsu verði til þess að óafvitandi sé borgað með verkefni íþróttafélags ótengdu lýðheilsu, eins og samkeppni um árangur í afreksíþróttum. Dæmi sem kom upp fyrir nokkrum árum er að Hafnarfjarðarbær þurfti að hlaupa undir bagga með Knattspyrnufélaginu Haukum sem höfðu skuldsett sig um of (Garðar Örn Úlfarsson, 2013). Skilgreina þarf vel hlutverk íþróttafélaga, sérstaklega þegar rekstur þeirra er fjármagnaður með skattheimtu. Þörf er á að huga vel að þessum málum og getur þá verið gagnlegt að skoða það verk sem unnið hefur verið fræðilega tengt íþróttum. Af nægu er að taka.

## 2.5 Líkön af úrslitum leikja

Fjöldi líkana hefur verið smíðaður í þeim tilgangi að leitast við að sjá fyrir úrslit einstakra fótboltaleikja eða móta. Þessi mikli áhugi á að spá rétt fyrir um ákveðin úrslit skýrist, sem margt annað, að verulegu leyti af mögulegum fjárhagslegum ávinningi. Búi einhver yfir aðferð til þess að geta sagt fyrir um úrslit öðrum fremur væri auðvelt að auðgast á veðmálamörkuðum tengdum fótbolta, sem er drifkraftur slíkra rannsókna.

Talsvert er liðið síðan byrjað var að líkja eftir fótbolta með tölfræði og líkönum. Árið 1951 kom fyrsta útgáfa tölfræðibókar Moroney (1965) ætluð almenningi út. Höfðað er til fjöldans og notagildi tölfræði boðuð í kafla um Poisson dreifingu með dæmi úr fótbolta. Að því marki er líkt ágætlega eftir dreifingu marka í fótboltaleikjum með Poisson dreifingu. Reep og Benjamin (1968) líkja eftir þáttum fótbolta, s.s. dreifingu markaskorunar líkt og Moroney, með neikvæðu tvíkostalíkani (e. negative binomial) árið 1968 og klykkja út með skrifum um hve ríkjandi hendingar virðast í fótbolta. Sérstaklega er tekið fram að sé litið á staka leiki virðist lítið samband á milli fjölda skota að marki og markaskorunar. Séu hins vegar margir leikir skoðaðir í einu er samband markskota og marka ljóst. Reep og Benjamin telja því slembinn þátt (e. stochastic element) spila inn í staka leiki. Reep og Benjamin, með aðstoð Pollard (1971), betrumbæta rannsókn sína, m.a. vegna mistaka í gagnasafni, og skoða þá einnig fleiri knattleiki. Hill (1974) telur Reed og Benjamin gera of mikið úr hlut hendinga í fótbolta. Fylgni spáa sérfræðinga fótboltatímarits við lokastöðu knattspyrnudeilda fær Hill til að álykta að hæfileikar hafi eitthvað með úrslit að gera. Þessar rannsóknir eru unnar út frá samtölum leikja og lengri tímabil skoðuð. Rannsóknir miðaðar við einstaka leiki fylgdu.

Maher (1982) líkir eftir markaskorun í einstökum þegar spiluðum leikjum. Stíkar fyrir styrk sóknar og varnar, bæði heimaliðs og útiliðs, eru metnir með aðferð mesta sennileika (e. maximum likelihood estimation). Möt á stikum styrkleika eru notuð til að búa til meðaltöl sem stungið er inn í líkan þar sem mörk heima- og útiliðs fylgja óháðum Poisson dreifingum. Viðurkennt er að betri endurspeglun á raunveruleikanum tæki tillit til fylgni í fjölda marka heima- og útiliðs. Líkan Maher er fremur hermílikan en spálíkan, þar sem úrslit liggja fyrir. Líkön til að spá úrslitum leikja byggjast mörg á vinnu Maher.

Dixon og Coles (1997) búa til spálíkan ætlað til að hagnast á fótboltaveðmálum. Þeir rekja atriði sem leitast á við að taka inn í líkön fótboltaúrslita. Atriðin eru að reikna með

misjöfnum styrk liða, forskoti af að spila á heimavelli, að forspárgildi sé í nýlegum úrslitum, að ólíkir mælikvarðar séu fyrir styrk sóknar og varnar og að inn í mælingu nýlegra úrslita spili misjafn styrkur liða. Eftir fremsta megni þróa Dixon og Coles (1997) líkan Maher (1982) á grundvelli þessara atriða. Nýjungar eru að líkanið leyfir notkun ófullkominna gagnasafna, úrslita margra deilda, auk þess sem tillit er tekið til sveiflna í frammistöðu liða í líkaninu ólíkt því sem var í eldra líkani. Styrkleikamöt í líkani Maher eru fastar á meðan Dixon og Coles taka sérstakt tillit til nýjustu úrslita og leyfa styrkleikamötum að þróast og líkja betur eftir því að styrkur liða er breytilegur. Veðmálaregla ásamt spálíkani gefur af sér arðbæra stefnu miðað við tímabilið sem þeir skoða. Þó er ólíklegt að það sé enn raunin því veðmögurum fer fram í að spá úrslitum og stilla stuðla af eftir því, eins og sýnt hefur verið fram á (Forrest, Goddard og Simmons, 2005).

Rue og Salvesen (2000) smíðuðu líkan til þess að spá fyrir um úrslit leikja og heilla tímabila. Skýristærðir eru búnar til úr því sem þeir telja mestu skipta, styrkleika sóknar liðs og styrkleika varnar og er hvorutveggja metið út frá markaskorun. Bayesískar aðferðir eru nýttar við smíði líkans sem leitt er af nýjustu úrslitum og í keyrslu Markov keðju Monte Carlo aðgerða sem spá fyrir um einstök úrslit og heil tímabil úr samanlögðum spám úrslita. Rue og Salvesen telja sitt líkan nýta vel gögnin sem það er fódrað á. Einnig er talið að líkanið skili viðunandi spám réttlættum með samanburði spáa við raunveruleg úrslit úr enskri knattspyrnu.

Koning (2000) rannsakaði hollenska knattspyrnu m.t.t. þess hvort dregið hefði úr samkeppni. Hann bjó til raðað probit (e. ordered probit) líkan sem spáði úrslitum beint fremur en markaskorun. Líkanið var þróun á eldra líkani (Neumann og Tamura, 1996), sem notað var til þess að skoða jafnvægi liða í NFL-deildinni eða ameríska fótboltanum. Skýribreytur eru mælikvarðar á misjöfnum styrkleika liða en leiðrétt er fyrir heimavallarforskot, sem er viðbót Koning við eldra líkan. Koning kaus notkun þessa líkans sökum einfaldleika þess. Styrkur liða er mældur í einum stika og er hann auðveldlega aðgreindur frá forskoti af því að spila á heimavelli. Jafnvægi liða sveiflast á tímabilinu frá 1956-1997 en engin sérstök leitni er í sveiflunum og ályktað að ekki hafi dregið úr jafnvægi á tímabilinu.

Crowder, Dixon, Ledford og Robinson (2002) reyndu að endurbæta líkan Dixon og Coles (1997) með slembipætti ásamt því að hafa það kvikara (e. dynamic) en áður. Mat á styrk sóknar og varnar í eldra líkani voru fastar sem tóku þó sérstakt tillit til nýafstaðinna leikja en með slembipáttum er tekið inn í nýrra líkan að hendingar spili gjarnan þátt í fótbolta. Samanburður líkananna sýnir að þau spá á móta vel fyrir um úrslit en endurbætta líkanið er almennara og á því rétt á sér.

Karlis og Ntzoufras (2003) settu upp nokkur tvívíð Poisson (e. bivariate Poisson) líkön sem líkja eftir markaskorun í fótbolta. Dæmi um líkan notað til að spá mörkum sem þeir setja fram á almennu formi er:

$$(X_i, Y_i) \sim BP(\lambda_{1i}, \lambda_{2i}, \lambda_{3i})$$

þar sem

$$\log(\lambda_{1i}) = \mu + sókn_{h_i} + vörn_{g_i} + heimavallaráhrif$$

$$\log(\lambda_{2i}) = \mu + sókn_{g_i} + vörn_{h_i}$$

Áætluð eru mörk heimaliðs í leik  $i$  með  $X_i$  og útiliðs með  $Y_i$ . Líkön leiða af sér  $\lambda_1$  og  $\lambda_2$  sem eru mót á meðaltölum notuð í tvívíða Poisson dreifingu til að spá mörkum. Liðurinn  $\lambda_3$  inniheldur mat á samdreifni (e. covariance)  $X$  og  $Y$  og hefur áhrif á markaspá. Meðalfjöldi marka er táknaður með  $\mu$ . Stikar fyrir styrk sóknar og varnar eru táknaðir með  $sókn$  og  $vörn$ . Áhrif þeirra eru metin með aðhvarfi sem frávik frá styrk meðalliðs. Lágvísar  $h_i$  og  $g_i$  vísa til heima,  $h$ , og útiliðs,  $g$ , og leiks  $i$ . Áhrif þess að spila á heimavelli koma inn í reikninginn sem *heimavallaráhrif*. Aðhvarfslíkön eru notuð til þess að meta stika notaða í ofangreindu spálíkani. Spálíkanið byggir því óbeint á gögnum, gögn eru ekki notuð í líkaninu en stikar sem það byggir á eru fengnir með aðhvarfslíkönunum sem nýta gögn. Fallið er frá forsendu um að mörk liða í leikjum séu óháð og eldri líkön þannig endurbætt. Kynnt voru líkön sem standast þrátt fyrir að reiknað sé með að fylgni sé í markaskorun liða í leikjum. Annar vandi eldri líkana er að hlutfall jafntefla í spám stemmir ekki við hve tíð þau eru í raun. Tilraun til að leiðrétta dreifingu úrslita líkans að veruleika er gerð með þætti sem eykur líkur á jafnteflum. Þá er leiðrétt vegna þess sem á ensku heitir overdispersion og felst í því að dreifni er meiri en meðaltal Poisson dreifingar. Síðar notuðu Karlis og Ntzoufras (2009) aðra nálgun en áður þar sem líkan spáir markamun í leikjum fremur en mörkum hvors liðs fyrir sig. Með því miða við markamun er komist hjá vandamáli vegna fylgni í markafjölda innan leikja.



Þetta líkan markamunar gefur líka af sér fleiri jafnteflaspár en fyrri líkön og því þarf ekki að laga líkanið að því að jafnteflaspár séu of fáar.

Goddard og Asimakopoulos (2004) hönnuðu líkan, sem spáir beint fyrir um úrslit, með raðaðri probit aðhvarfsgreiningu. Kostur við að spá um úrslit fremur en mörk er að komist er hjá vandamáli sem fylgir því að mörk heima- og útiliðs í leik eru ekki óháðar stærðir. Hvern leik spila tvö lið og ljóst er fjöldi marka mótherjans skiptir verulegu máli varðandi hvatann til að skora fleiri. Ef mótherjanum hefur ekki tekist að skora mark og lið hefur skorað þrjú færast áherslur gjarnan í að verja það sem komið er. Í stöðunni 3-3 er talsvert meiri hvati fyrir lið að skora næsta mark sem getur orðið sigurmarkið. Líkön sem spá fyrir um mörk heima- og útiliðs þurfa því að fela í sér leiðréttingu fyrir fylgni markaskorunar en ekki úrslitalíkön. Fjöldi nýjunga var tekinn inn í líkan Goddard og Asimakopoulos (2004) sem byggist ekki einungis á fyrri úrslitum eins og eldri líkön. Mikilvægi leiks, fallhætta, barátta um að komast upp um deild, þátttaka í bikarkeppnum og vegalengdir milli leikvanga liða eru nýjungar. Allir þessir þættir hafa áhrif á úrslit og nýjungarnar eru felldar inn í líkanið.

Tvær algengar nálganir að því að spá um úrslit leikja voru bornar saman af Goddard (2005). Hann bar saman spár um markaskorun liða, og þannig óbeinar spár um útkomu leikja, eða beinar spár um útkomu leiks, sigur, jafntefli eða tap. Fyrri nálgunin var með tvívíðri Poisson aðhvarfsgreiningu (e. bivariate Poisson regression), sambærilegri þeirri frá Karlis og Ntzoufras (2003), sem gefur af sér líkön sem notuð eru til að spá fyrir um mörk beggja liða í leik. Sú síðari er gerð með raðaðri probit aðhvarfsgreiningu þar sem afleidd líkön gefa af sér beina spá um úrslit. Líkön Goddard standa sig betur í öllum tilfellum en einföld spáregla byggð á fyrri hlutfallslegri skiptingu úrslita, sem gefur úrslit í sömu hlutföllum og reynslan hefur verið. Líkönin reynast öll ámóta spávís. Innan þeirra er þó líkan með blandaðri aðferð sem gefur bestu niðurstöðu. Þá eru hvoru tveggja háð breyta fyrir eldri úrslit og tafirín skýribreyta byggð á fjölda marka notuð í mati líkans með röðuðu probit aðhvarfi.

Graham og Stott (2008) útfærðu probit líkan byggt á líkani Koning (2000) og spá úrslitum beint. Kvikir stikar fyrir styrk liða eru metnir með sennileikafalli (e. likelihood function) og dregst úr vægi eldri leikja í stikum með veldishrönnun (e. exponential decay). Hraði hrönnunar er stilltur með hámrökun á spágetu líkansins á úrtak utan þess

sem notað er í mati líkansins (e. holdout sample). Líkanið notuðu þeir til að sýna fram á að kerfisbundnar bjaganir eru í stuðlum veðmangara. Bjaganirnar verða þó ekki skýrðar af því að veðmangarar sleppi mikilvægum upplýsingum við að stilla stuðla eða því að óviðkomandi upplýsingar séu notaðar.

Spann og Skiera (2009) báru saman forspárgildi spáa af þrennskona mismunandi uppruna. Árangur spámarkaða (e. prediction markets), stuðla veðmangara og spáa íþróttasérfræðinga (e. tipsters) við að spá réttum úrslitum í þýskum fótbolta er tekinn fyrir. Sérfræðingarnir standa sig verst en spámarkaðir og stuðlar veðmangara standa sig nokkuð vel. Reynslan af ólíku aðferðunum er svo miðuð við hver ávöxtun væri af beitingu þeirra á veðmálamarkaði. Ef ekki væri fyrir háa þóknun, 25%, komast Spann og Skiera að því að af spám spámarkaða og þeim sem leiða af stuðlum veðmangara hefði mátt búa til arðbær kerfi veðmála.

## **2.6 Hagkvæmni veðmálamarkaða og fyrirsjáanleiki úrslita**

Náin tengsl hafa löngum verið á milli íþróttu og veðmála. Því er ekki að furða að talsverð vinna hafi verið lögð í að búa til líkön til að spá um úrslit í því skyni að auka hagnaðarvon í veðmálum. Fjöldi líkana hefur verið smíðaður og má ætla að þau verði æ fullkomnari í að spá réttum úrslitum og spádómsgeta þeirra meiri. Þó hefur einnig verið tekið skref til baka frá tilraunum til að spá fyrir um ákveðin úrslit. Þannig hafa vangaveltur vaknað um hvort veðmálamarkaðir séu hagkvæmir. Þá er stærri mynd skoðuð fremur en ákveðin leikir og því velt upp hvort yfir höfuð sé hægt að hagnast skipulega á veðmálamörkuðum. Hagkvæmni veðmálamarkaða tengdum íþróttum í Bandaríkjunum hefur verið skoðuð í nokkuð langan tíma. Fyrir tæpri hálfri öld skoðaði Pankoff (1968) skilvirkni veðmálamarkaða sem tengdust amerískum fótbolta. Á þessum tíma var Nóbelsverðlaunahafinn Eugene Fama að þróa hugmyndir sínar um skilvirkni markaða og birti síðar í greininni „Efficient Capital Markets“ (1970). Pankoff naut aðstoðar Fama og bar saman skilvirkni fjármálamarkaða og veðmálamarkaða. Markaður er sagður skilvirkur eða hagkvæmur þegar fjöldi rökhugsandi þátttakenda er á markaðnum sem leitast við að hámarka hagnað út frá sömu aðgengilegu opinberu upplýsingunum. Pankoff færir rök fyrir því að margt sé sameiginlegt með veðmálamörkuðum og fjármálamörkuðum og að veðmálamarkaðir virðist skilvirkir. Í ítarlegri grein Sauer (1998) um hagfræði veðmálamarkaða tekur hann í meginatriðum undir hugmyndir Pankoff um

að veðmálamarkaðir séu fremur hagkvæmir. Sauer tekur þó fram að undantekningar séu á því og telur hann að betur megi ná utan um undantekningarnar með kenningum og líkönum.

Án þess að fara frekar yfir líkön sem gerð hafa verið má líta betur til kenningar um skilvirka markaði í samhengi við veðmálamarkaði. Fyrirfram mætti ætla að tilraunir til að hagnast á veðmálamörkuðum með notkun fullkomins spálíkans hafi ekki tekist. Kenningin um skilvirka markaði gengur í grófum dráttum út á að þar sem fjárfestar búa allir við sömu upplýsingar getur einstakur fjárfestir ekki gert ráð fyrir að regluleg ávöxtun hjá honum verði umfram meðalávöxtun markaðar án þess að taka aukna áhættu. Ætla má að kenning um skilvirka markaði nái yfir veðmálamarkaði. Veðmálafyrirtæki hafa sömu upplýsingar og viðskiptavinir þeirra og stilla stuðla eftir því. Félögin myndu vart halda áfram starfsemi ef hagnaður væri ekki af henni. Ljóst er að veðbankar stilla stuðla sér í hag og nýta sér spágetu sína til að hagnast á þeim sem veðja (Levitt, 2004). Álykta má að fullkomið líkan til þess að hagnast á veðmálastarfsemi hafi ekki fundist. Veðmálastarfsemi myndi þá leggjast af eða, það sem er líklegra er, að hún myndi laga sig að nýjum veruleika og taka tillit til þeirra nýju upplýsinga sem komnar eru fram hafi tekist að búa til fullkomið líkan. Kannski er ekki hægt að útiloka að fullkomið spálíkan hafi verið búið til. Hafi svo verið er sá sem það gerði hins vegar væntanlega líka nægilega skynsamur til þess að halda því út af fyrir sig, sem væri forsenda þess að geta viðhaldið því sem peningamyllu.

## **2.7 Líkön til skýringar árangurs í fótbolta**

Fá líkön er að finna til útskýringar á árangri í fótbolta yfir lengri tímabil. Fræðigreinum þar um hefur þó fjölgað liðinn áratug eða svo. Þær byggjast á eldri greiningum á árangri þjóða á Ólympíuleikum. Houston og Wilson (2002) bjuggu ekki að eldri hagrannsóknnum sem almennt skýra árangur í fótbolta. Þeir skoðuðu árangur í fótbolta, eftir styrkleikalista landsliða frá Alþjóðaknattspyrnusambandinu (FIFA), og tekjur, miðað við verga landsframleiðslu á mann og fólksfjölda. Byggðu þeir á kenningu Gary Becker (1965) um ráðstöfun tíma og eftirspurn í tómsundur. Fótbolti er þá flokkaður sem tómsundur, sem aftur teljast venjuleg vara, í þeim skilningi að eftirspurn eftir tómsundum eykst með tekjum. Þá er vísað í kenningu um minnkandi jaðarafköst og ráð gert fyrir að áhrif viðbótar tómsunda á árangur fari minnkandi. Sett er upp líkan með

aðferð minnstu kvaðrata og eru jákvæð áhrif tekna, nátturulegra logra fólksfjölda og fyrri þátttöku í heimsmeistaramóti á styrkleikaröðun þjóða skv. lista FIFA. Neikvæð áhrif eru af tekjum í öðru veldi sem talið er styðja kenningu um minnkandi jaðaráhrif tekna á árangur.

Eins og nefnt var að framan hafa rannsóknir við skýringar á dreifingu verðlauna á milli þjóða á Ólympíuleikum verið notaðar sem grunnur að sambærilegum rannsóknum á fótbolta. Þar um má nefna grein Hoffmann, Lee Chew Ging og Ramasamy (2002a). Þeir nýta eigin rannsókn á árangri á Ólympíuleikum (2002b) í rannsókn á árangri í fótbolta. Þar sem félags- og efnahagslegar aðstæður þjóða gefa fyrirheit um ólympískan árangur telja þeir að vert sé að skoða fótbolta með sömu aðferðum. Líkt Houston og Wilson (2002) er styrkleikalisti landsliða settur saman af FIFA skoðaður m.t.t. fólksfjölda, þróunar fótbolta, vergrar þjóðarframléiðslu, meðalhita og menningar. Hoffmann, Ging og Ramasamy telja sig sýna fram á að framannefndir þættir skýri a.m.k. að hluta til hvernig landslið raðist á styrkleikalista FIFA. Grein Hoffmann og félaga er fremur ósannfærandi og vekur fremur upp spurningar en að svara þeim. Dæmi um það er hvernig tilraun er gerð til þess að meta áhrif menningar með því að sameina gervibreytu (e. dummy variable) fyrir rómönsk lönd og fólksfjöldabreytu.<sup>1</sup> Skoðuð hvor fyrir sig tekst ekki að sýna fram á áhrif þeirra í þeim líkönum sem búin voru til en saman mynda þær breytu sem hefur mælanleg áhrif sem sagt er að réttlæti notkun breytunnar. Þá hefur verið bent á að þjögung er einnig til staðar í úrtaki þjóða sem miðað er við í rannsókninni. Úrtakið er valið út frá þeim þjóðum sem unnið hafa til verðlauna á Ólympíuleikum (Macmillan & Smith, 2007). Úrtakið felur í sér að þjóðir sem líklegar eru til árangurs veljast í það og það gefur því ekki rétta mynd af þýði. Auk þess eru leifaliðir líkansins ekki normaldreifðir sem bendir til þess að skilgreina þurfi það betur.

Rannsókn Hoffmann, Ging, Matheson og Ramasamy (2006) er endurtekning á fyrra verki þriggja þeirra (2002a) en þar er árangur kvennalandsliða skoðaður. Í meginatriðum er komist að sömu niðurstöðum og áður. Umgetin rómönsk gervibreyta, sem er tengd fólksfjölda, hefur þó ekki áhrif hjá konum. Athugun á áhrifum pólitískra atriða bendir til

---

<sup>1</sup> Rómönsk lönd, eða Latin eins og þau eru nefnd í greininni, eru ríki Mið- og Suður-Ameríku ásamt Portúgal og Spáni.

Þess að þau geti að einhverju marki skýrt röðun á styrkleikalista FIFA. Það tekur annars vegar til hlutfallslegs munar meðallauna kynja, sem er notaður sem mælikvarði á jafnrétti, og hins vegar þess hvort konur komi frá landi sem hefur verið undir kommúnískri stjórn. Hvorutveggja jafnrétti og kommúnísk stjórn hafa jákvæð áhrif á árangur landsliðskvemma. Niðurstöðum verður þó að taka með sama fyrirvara og verkið sem byggt er á því úrtakið hefði mátt velja betur.

Macmillan og Smith (2007) fá afnot af gögnum Hoffmann, Ging og Ramasamy (2002a) og bæta við gagnasafnið 100 löndum og fjölga breytum til að laga það sem miður fór. Aftur er rannsakað hvað skýrir styrkleikaröðun landsliða í fótbolta til lengri tíma. Ólíkt fyrri niðurstöðu, sem miðuð var við þrengra úrtak, eru áhrif fólksfjölda á árangur landsliða miðað við nýtt, stærra, og að þeir segja, óbjagað úrtak, ljós. Einnig er sýnt fram á minnkandi jaðaráhrif fólksfjölda á árangur. Mannfjöldi virðist styrkja landslið að ákveðnu marki en fer svo að draga úr áhrifunum eftir því sem fjöldinn eykst. Í viðleitni til að mæla áhrif fótboltamenningar eru skoðuð tengsl árangurs við hve langt er síðan fyrsti landsleikur þjóðar var leikinn. Sú breyta reynist skýra hluta árangurs þar sem „eldri“ landslið njóta góðs af aldri sínum á styrkleikalista FIFA.

Torgler leitar skýringa á árangri landsliða á heimsmeistaramótum með hagrannsókn í grein (2006). Könnuð eru áhrif hefðar fyrir fótbolta, ásamt efnahagslegum, lýðfræðilegum og veðurfarslegum þáttum á árangur landsliða. Ný nálgun að árangri landsliða er viðhöfð. Nálgunin felst í því að árangur er mældur sem samanlögð stigasöfnun þjóða úr heimsmeistaramótunum 1930-2002. Komist er að því að mælikvarðar á hefð fyrir fótbolta og efnahagslegir þættir skýri árangur en fólksfjöldi og veður hafi ekki áhrif. Niðurstöðurnar koma Torgler ekki á óvart og bendir hann því til stuðnings á að fámennar þjóðir eins og Holland, Belgía og Svíþjóð hafi náð talsverðum árangri á vellinum en fjölmennar þjóðir eins og Kína og Indland ekki. Velta má fyrir sér hvort aðferðir Torgler séu viðunandi þar sem meðaltöl breyta yfir mislangan tíma eru notuð til að skýra summu annarrar breytu yfir enn annað tímabil.

Hagfræðingar frá La Coruña á Spáni rannsökuðu árangur í fótbolta út frá borgum í Evrópu fremur en löndum (P. C. García, Castro og Santos, 2007). Þeir telja sig vera þá fyrstu til að skoða árangur þar sem mælingar eru miðaðar við borgir fremur en lönd. Árangur í fótbolta kalla þeir að hafa á liði að skipa sem keppir í fremstu röð í fótbolta í

Evrópu. Liðum og borgum er stillt upp í úrvalsflokk og ekki úrvalsflokk. Skipun í flokka fer eftir því hvort lið hefur keppt í riðlakeppni meistaradeildar evrópska knattspyrnusambandsins á tímabilinu frá 1992 til 2007. Úrtakið eru alls 186 borgir og skipa 41 þeirra úrvalsflokkinn. Líkan er sett upp til að skýra tvíkosta dreifingu borganna. Skýristærðir í logit líkani eru fólksfjöldi og verg landframléiðsla (VLF) á mann í borgunum. Niðurstöðurnar styðja kenningu höfundar um að aukinn fólksfjöldi og aukin velmegun, m.v. VLF á mann, skili betri árangri í fótbolta. Að auki er komist að því að fólksfjöldi vegur hlutfallslega meira en VLF á mann.

Leeds og Leeds (2009) greina alþjóðlegan árangur í fótbolta út frá sambandi við stofnanir landa. Líkön eru sett fram sem skilgreina eiga árangur með fjölda breyta. Má nefna fjármagn og vinnuafli, pólitískar- og stofnanabreytur, t.a.m. hvort löndum sé stjórnað af kommúnistum eða hvort þar sé olía. Gervibreytur eru fyrir nýlendur, heimsálfur og árangur félagsliða landa í álfukeppnum félagsliða. Mikill fjöldi breyta lýsir sér í því að fleiri en 30 stuðlar eru metnir í hverju líkani. Niðurstöður áhrifa þess að vera fjölmenn og með mikla landsframléiðslu á mann samræmast fyrri rannsóknum þar sem bæði atriðin skipta verulegu máli. Áhrifa gætir af tæpum helmingi atriðanna sem metin eru í líkönunum. Notað er hefðbundið líkan aðferð minnstu kvaðrata og svo neikvætt tvíkostalíkan. Sýnt er fram á áhrif fólksfjölda og vergrar landsframléiðslu á mann í öllum líkönunum. Sama á við um gervibreytu fyrrverandi nýlenduhera, gervibreytu landa sem eru í asíska knattspyrnusambandinu og gervibreytu árangurs landa í Suður-Ameríkukeppni félagsliða. Allar þessar breytur hafa jákvæð formerki að undanskildri þeirri fyrir asíska sambandið. Jákvæðu formerkin gefa til kynna að þegar atriðin metin með stikunum eigi við, eða aukist, sé árangur líklegri. Aðild að asíska sambandinu dró því úr líkum á árangri en annað bætti líkurnar á honum. Sé sýnt fram á áhrif atriðis á það almennt við í báðum tegundum líkana, sem höfundar telja treysta betur niðurstöður þeirra. Líkönin eru þó svo umfangsmikil að það er líkt og höfundar séu á nótaveiðum í eldiskví. Höfundar koma sjálfir inn á að verk þeirra sé það víðfeðmasta sem kynnt hefur verið. Að því sögðu er ályktað að stofnanir þjóða móti alþjóðlegan árangur.

Af framansögðu ágripi má sjá að enn er verk fyrir höndum við að finna rót árangurs í fótbolta. Reynt hefur verið að sýna fram á áhrif fólksfjölda og hagsældar sem hefur

tekist að einhverju leiti. Á sama tíma hefur reynst erfiðara að meta áhrif annarra atriða svo sem menningar, stjórnmála og veðurs.

## 3 Gögn

### 3.1 Umfjöllun

Úrslit leikja og stöðutöflur eru sóttar af heimasíðu Knattspyrnusambands Íslands („Mótalisti | Knattspyrnusamband Íslands“, e.d.). Gott safn úrslita er að finna á síðunni úr leikjum hvort sem um ræðir félagslið eða landslið, konur eða karla, meistaraflokka eða þá yngri. Þó hefur misfarist að huga nægilega að því til hvers verið er að safna gögnunum nema þá einungis til að fletta upp einstökum úrslitum eða ákveðinni stöðutöflu.

Viðmótið sem úrslit og stöðutöflur eru kölluð fram úr kallar eftir því að söfnun sé þrengd um of. Vandí er að einungis er hægt að kalla fram úrslit og stöðutöflu eins móts í einu og því er seinvirkt að safna miklu magni gagna. Vandkvæði fylgja framsetningu úrslita og stöðutafna. Úrslit og töflur birtast í vafra og ekki er boðið uppá möguleika til þess að hlaða þeim niður á neinu formi, hvorki sem textaskjali né í töflureikni og hvað þá á öðru formi. Aftur leiðir þetta til þess að seinvirkt er að safna gögnum þar sem annað hvort þarf að klippa og líma gögn fram og til baka eða sækja inn í töflureikni í gegnum vefviðmót sem hann býður uppá. Hvernig sem gögn KSÍ eru sótt verður að eiga við framsetningu þeirra handvirkt svo mögulegt sé að vinna frekar með þau. Svör sambandsins við fyrirspurn um hvort gögnin væru ekki til á þægilegra og aðgengilegra formi voru að svo væri ekki en að varla geti tekið svo langan tíma að ná í gögnin (skrifstofa KSÍ, símtal, sumarið 2014).

Skiljanlega er aðra sögu að segja af þeim gögnum sem Hagstofa Íslands býður uppá. Einfalt er að sækja gögn hennar og þau eru í aðgengilegu formi eins og búast má við af stofnuninni. Vandinn felst þá fremur í því hvernig eigi að skilgreina nákvæmlega hvers sé leitað.

Gögn um skýristærðir líkana eru sótt af vef Ríkisskattstjóra, sem var vandræðalaust og einfalt.

### 3.2 Úrlausnarefni

Upphaflega var lagt upp með að skoða árangur knattspyrnuliða á Íslandi fremur en sveitarfélaga í sambandi við fjölda fólks að baki þeim og tekna þess. Mannfjöldatölur miðaðar við öll póstnúmer og einnig fyrir hverfi Reykjavíkur eru til frá Hagstofunni.



Tímabil mannfjöldagagnanna er þó skemmra en þeirra sem notast er við frá Ríkisskattstjóra. Ekki er vitað um gögn um tekjur ákveðinna hverfa og póstnúmera eða viðunandi nálgun á tekjum með öðrum sambærilegum gögnum. Ekki þótti verjandi að ein skýristærð miðaðist við hverfi eða póstnúmer liða og hin við sveitarfélagið sem lið kæmi úr. Talsverð vandkvæði hefðu einnig fylgt því að ákveða hvernig festa ætti lið niður á póstnúmer eða skipta þeim upp á milli liða. Aðdáendur liða einskorðast ekki við ákveðin póstnúmer, lið koma úr sömu póstnúmerum og fylgi margra þeirra nær yfir mörg póstnúmer eða hverfi. Vegna þessara vandkvæða var leitað annarra leiða til athugunar á áhrifum tekna og mannfjölda á árangur. Þannig hefur spilað inn í val á líkönnum aðengi að gögnum og á hvaða formi þau eru.

## 4 Tvíkostalíkön af árangri

### 4.1 Verklag og skilgreiningar

Skilgreining árangurs hlýtur að vera forsenda fyrir smíði líkans af árangri í fótbolta. Hér er árangri stillt upp sem tvíkosta breytu, annað hvort hefur honum verið náð eða ekki. Árangur er skilgreindur á þann veg að hafa spilað í röð fremstu liða á Íslandi, ef leikið hefur verið í úrvalsdeild eða efstu deild knattspyrnu á Íslandi, á tímabilinu 1992-2014. Með því að nota þetta tímabil er árangur skoðaður frá sama upphafspunkti og í panel líkönum sem fylgja í næsta kafla. Tímabil panel líkana takmarkast af framboði gagna en þau eru einungis til fyrir tímabilið 1992-2012. Ákveðið var að miða við sama upphafspunkt í tvíkosta logit líkani. Árangurinn er miðaður við sveitarfélög. Hafi sveitarfélag átt lið í efstu deild knattspyrnu á Íslandi á árunum 1992-2014 hefur það náð árangri í fótbolta, annars ekki. Háða breytan í gagnasafninu er því annað hvort núll, ef þessari skilgreiningu árangurs hefur ekki verið náð, eða einn, ef árangur hefur náðst. Með þessari nálgun verður til líkan þar sem útkoman er mat á líkum á því að sveitarfélag hýsi úrvalsdeildarlið. Í næsta kafla er árangur skilgreindur á annan máta sem farið er betur yfir þar.

Óháðu breyturnar sem kannað er hvort skýri árangur eru annars vegar mannfjöldi í sveitarfélögum og hins vegar tekjur íbúa þeirra. Gögn skýristærða eru fengin af vef ríkisskattstjóra (RSK), gagnasafnið heitir *heildartekjur eftir sveitarfélögum tekjuárin 1992-2012* („Álagning einstaklinga“, e.d.). Gagnasafnið er stórt og fjöldi breyta tekinn saman yfir margra ára tímabil fyrir öll sveitarfélög. Í rannsókn með sambærilegum líkönum, þar sem árangri sveitarfélaga í fótbolta er stillt upp sem tvíkosta breytu, er miðað við nýjustu mælingar skýristærða og sú aðferð að miða við nýjustu mælingar er tekin upp (P. C. García o.fl., 2007). Fyrir uppsetta líkanið eru tíndar út nýjustu mælingar breytanna tveggja fyrir öll sveitarfélög, en þær eru frá 2012. Heildartekjur eru notaðar sem tekjubreyta en undir hana falla allar tekjur sem fólk telur til skatts, laun, opinberar bætur, ellilífeyrir, fjármagnstekjur o.s.frv. Í gögnum RSK er búið að leggja saman tekjur allra einstaklinga í viðkomandi sveitarfélagi. Til að mælt sé á sama grunni er deilt í heildartöluna með fjölda skattgreiðenda. Mannfjöldabreytan sem notuð er vísar til fjölda svokallaðra skattskyldra eininga eða skattgreiðenda, sem eru einstaklingar, hjón eða sambúðarfólk. Til stóð að nota fólksfjöldataölur sveitarfélaga frá Hagstofunni en í

stað þess er notast við fjölda skattskyldra eininga úr sama gagnasafni og tekjugögn koma úr. Athugun leiddi í ljós að notkun fjölda skattskyldra eininga væri algjörlega sambærileg því að nota fjölda skattgreiðenda, þar sem fylgni fjölda skattskyldra eininga og fólksfjölda í sveitarfélögum skv. Hagstofunni reyndist vera fullkomin fyrir árið 2012 (0,999 skv. útreikningi) og því talin vera vel viðundandi nálgun á mannfjölda. Hvoru tveggja mælingum á mannfjölda og tekna er umbreytt á nýjan kvarða fyrir aðhvarfsgreiningar svo stíkar líkana séu ekki óþarflega smáir og illskiljanlegir. Í tekjumælingar er deilt með milljón og þúsund í mælingar mannfjölda. Fylgni á milli óháðu breytanna var athuguð til að forðast hættu á marglínuleika. Fylgni tekna skattgreiðenda í sveitarfélagi og fjölda þeirra er takmörkuð eða 0,127 og því ekki hætta á marglínuleika vegna fylgninnar.

Árangur er mældur í þrepum, sveitarfélög hafa annað hvort náð árangri eða ekki og er skipt í flokkana *úrval* og *ekki úrval* eftir því. Hefðbundin línuleg líkön eru ekki viðunandi til þess að líkja eftir þeim eiginleika háðu breytunnar að hún sé tvíkosta. Hefðbundnar forsendur um normal dreifingu og einsdreifni (e. homoscedasticity) bresta og líkan venjulegrar aðferðar minnstu kvaðrata og sambærileg líkön því ekki hentugt. Logit líkön henta vel til þess að líkja eftir áhrifum skýribreyta á tvíkosta háðar breytu. Almennt logit líkan er:

$$P(Y_i = 1) = \frac{\exp(\mathbf{x}_i \cdot \boldsymbol{\beta})}{1 + \exp(\mathbf{x}_i \cdot \boldsymbol{\beta})}$$

Í þessari jöfnu er  $Y$  háða breytan,  $P(\cdot)$  líkur á að háða breytan sé í stöðu ( $\cdot$ ),  $\mathbf{x}_i$  er vektor skýristærða og  $\boldsymbol{\beta}$  er vektor stíka. Venja er að áhugaverða atburðinum eða þættinum sé stillt upp sem einum og þá er  $P(Y = 1)$  líkur á að hann eigi sér stað. Líkurnar velta á stikum skýristærða  $\boldsymbol{\beta}$  fengnum úr logit aðhvarfsgreiningu og skýribreytunum  $\mathbf{x}_i$ . Stíkar aðhvarfsgreiningar,  $\boldsymbol{\beta}$ , eru fastar stærðir. Þær eru annars vegar fasti líkansins og hins vegar stíkar skýribreyta, en skýribreytur,  $\mathbf{x}_i$ , taka mismunandi gildi. Stíkin  $\boldsymbol{\beta}$  segir til um áhrif  $\mathbf{x}_i$ , einstakrar skýristærðar, á logra líkinda þess að  $Y_i = 1$  stýrt fyrir áhrifum annarra skýristærða ef fleiri en ein skýristærð er í líkani. Þannig er  $\exp(\boldsymbol{\beta})$  margföldunaráhrif á líkindin (e. odds) af breytingu á  $\mathbf{x}_i$  upp á eina einingu þegar öðrum skýristærðum er haldið föstum. Formerki  $\boldsymbol{\beta}$  segir til um hvort líkur á að  $Y_i = 1$  aukist eða úr þeim dragi við stærri  $\mathbf{x}_i$ . Hraði breytingar á líkindin  $P(Y_i = 1)$  eykst eftir því sem  $|\boldsymbol{\beta}|$ , tölugildi  $\boldsymbol{\beta}$ , er meira. Ef stíkin  $\boldsymbol{\beta} = 0$  þá er  $Y_i$  óháð

$x_i$ . Líkindahlutfall (e. odds ratio) er  $\exp(\beta)$  eða líkindi þegar  $X = x + 1$  á móti líkindi við  $X = x$ . Fastinn er yfirleitt ekki skoðaður sérstaklega og engin sérstök ástæða yfirleitt talin til þess. Logit líkanið er tvíkosta þannig að líkur á engum atburði eða þætti er einfalt að fá vegna þess að  $P(Y_i = 1) = 1 - P(Y_i = 0)$ .

**Tafla 1. Breytur**

Breyta	Skilgreining	Uppruni	Gildi
Úrvalsbærir	Sveitarfélög sem átt hafa lið í úrvalsdeild '92-'14	KSÍ	0 eða 1
Fjöldi	Skattskyldar einingar í sveitarfélagi	RSK	32 - 73.143
Tekjur	Heildartekjur í svfl. deilt með fjölda	RSK	3,6-7,9 Milljónir

Öll sveitarfélög eru í gagnasafninu, en þau voru 75 árið 2012. Þrettán þeirra falla í úrvalsflokkinn og 62 utan hans. Sveitarfélög sem skipa úrvalsflokkinn eru að jafnaði með hærri tekjur og meiri mannfjölda. Mjög mikill munur er á meðalfjöldanum, sem er rúm 11 þúsund í úrvalsflokknum en einungis um 650 í hinum flokknum. Mikill munur á fjölda sést einnig af því að fjölmennasta sveitarfélag sem ekki nær úrvalsflokki telur 4595 skattgreiðendur, eða innan við helming meðalfjöldans í úrvalsflokknum. Ekki er eins gríðarlegur munur á meðaltekjum flokkanna, sem eru um 6 milljónir kr. á ári í úrvalsflokknum og rúmar 5 milljónir kr. í hinum. Dreifni fjöldans er margfalt meiri í úrvalsflokknum en dreifni tekna er litlu meiri í honum. Skýristærðir eru báðar jákvætt skekktar (e. skewed) í átt að lægri gildum, sérstaklega fjöldamælingar.

## 4.2 Tvíkostalíkon

Línuleg líkon venjulegrar aðferðar minnstu kvaðrata (VAMK) passa illa þegar skýristærð er tvíkosta breyta, þar sem erfitt er að líkja eftir raunverulegri dreifingu skýristærðar með þeim. Línuleg VAMK líkon henta auk þess ekki þar sem mót þeirra geta leitt til niðurstaðna utan núlls eða eins, sem eru þau gildi sem háða úrvalsbreytan getur tekið. Ekkert er því til fyrirstöðu að línulegt líkan leiði til spáa sem eru minni en núll prósent eða meiri en hundrað prósent, sem augljóslega geta ekki verið réttar. Á þessi gögn er því hentugra að keyra logit aðhvarfsgreiningu þar sem niðurstöður geta einungis verið mót á líkum á milli núll og hundrað prósent. Fyrsta tilraun til þess að líkja eftir gögnunum er logit líkan sem tekur einungis mið af tekjum og fjölda og er á forminu:

$$P(\text{úrval}) = \frac{\exp(\beta_0 + \beta_1 Fjöldi + \beta_2 tekjur)}{1 + \exp(\beta_0 + \beta_1 Fjöldi + \beta_2 tekjur)}$$

Samkvæmt þessu líkani fara líkur á því að sveitarfélag sé í úrvalsflokknum eftir fjölda og tekjum íbúa. Tölfræðiforrit er notað til þess að meta líkanið og með ítrun eru stíkar líkansins metnir með aðferð mesta sennileika (e. maximum likelihood estimation). Niðurstöður líkansins er að finna í töflu um helstu niðurstöður logit líkans tekna og fjölda.

Logit líkan tekna og fjölda  
Tvíkosta háð breyta: úrvalsflokkur eða ekki

	Stíki	Staðalskekkja	z	p-gildi
Fasti	-7,304	3,077	-2,37	0,018
Fjöldi	0,993	0,366	2,72	0,007
Tekjur	0,679	0,530	1,28	0,200

Stíkar fjölda og fasta eru með hóflegar staðalskekkjur borið saman við mót stíkana en staðalskekkja tekjustika er fremur mikil. Formerki stíka eru eins og búist var við bæði jákvæð fyrir fjöldabreytu og tekjur. Sveitarfélag eykur líkur á því að falla í úrvalsflokk með auknum mannfjölda.

Auknar tekjur íbúa sveitarfélags virðast auka líkur á sveitarfélagið falli í úrvalsflokk, en ekki er hægt að fullyrða um áhrifin vegna stærðar staðalskekkju í mati stíka.

Fyrir meðalsveitarfélag eru líkur á að sveitarfélagið falli í úrvalsflokk 22,7%.

Ætlaðar líkur á sveitarfélag falli í úrvalsflokk margfaldast, að öðru óbreyttu, um líkindahlutfallið (e. odds ratio)  $\exp(0,993) = 2,699$  við hverja þúsund skattgreiðendur eða skattskyldar einingar, en að framan var gerð grein fyrir sem nálgun á mannfjölda sveitarfélaga. Með sama hætti margfaldast líkur á sveitarfélag falli í úrvalsflokk um líkindahlutfallið (e. odds ratio)  $\exp(0,679) = 1,972$  við hverja milljón sem meðaltekjur aukast um að öðru óbreyttu.

Tafla 2. Jaðaráhrif skýribreyta við meðaltöl

Breyta	Jaðaráhrif
Fjöldi	0,174
Tekjur	0,119

Jaðaráhrif við meðaltöl skýribreyta segir til um breytingu á líkum á að falla í úrvalsflokk við smávægilega aukningu á skýristærð við meðaltal hennar. Til að mynda ef

fjöldi er aukinn um hundrað er  $0,1 \cdot 0,174 = 0,0174\%$  meiri líkur á að falla í úrvalsflokk, hundrað skattgreiðendur er  $0,1$  í líkaninu vegna skölunar sem miðuð er við að 1000 séu 1 til að unnið sé með auðskiljanlegri stærðir. Ef meðaltekjur í sveitarfélagi aukast um 100 þúsund,  $0,1$  í líkani þar sem skali miðar við milljónir, eru jaðaráhrifin við meðaltöl  $0,1 \cdot 0,119 = 0,0119\%$ .

**Tafla 3. Munur á líkum efri og neðri fjórðunga**

Breyta	Mat	Staðalskekkja	Samanburður	Breyting á líkum
Fjöldi	0,993	0,366	efra/neðri fjórðungar m.v. meðaltekjur	8,5% -3,1% = 5,4%
Tekjur	0,679	0,530	efra/neðri fjórðungar m.v. meðalfjölda	28,0% -16,2% = 11,8%

Það eru 5,4% meiri líkur á sveitarfélag falli í úrvalsflokk ef borinn er saman mannfjöldi miðaður við fyrsta fjórðungsmark og yfir í þriðja fjórðungsmark þegar tekjum er haldið föstum í meðaltali. Breytingin á líkunum er meiri, 11,8%, ef fjölda er haldið föstum og tekjur eru skoðaðar miðaðar við sömu fjórðunga.

Þegar háð stærð er tvíkosta, eða skiptist í fleiri flokka, getur það gerst þegar skýristærð er yfir tiltekinni stærð að hún tengist einungis einu ákveðnu gildi eða flokki háðu stærðarinnar. Þá á sér stað það sem er kallað aðskilnaður (e. separation). Vísbendingar um að breyta í líkaninu leiði til fullkomins- eða hálffullkomins aðskilnaðar (e. complete or quasi-complete separation) komu upp. Fullkominn aðskilnaður þýðir að af gildi skýribreytu megi fullkomlega ætla hvert gildi háðu breytunnar sé. Hálffullkominn aðskilnaður er eins, skýribreyta segir fullkomlega til um gildi háðu breytunnar, en munurinn er sá að í einhverjum tilvikum tekst það ekki og aðskilnaðurinn því hálffullkominn. Fyrir sum gildi skýristærðar mælist háða breytan alltaf í sama flokknum í fullkomnum aðskilnaði. Skýribreyta veldur aðskilnaði. Það kann að hljóma undarlega að það sé vandamál að skýribreyta nái algjörlega að segja fyrir um ástand háðrar breytu. Vandamál sem gjarnan stafar af aðskilnaði er að stíkar fá mjög ýkt gildi þegar sennileikamöt eru notuð til að meta þá. Gildi þeirra geta þá orðið annað hvort mjög há eða svo lág að þau nálgist núll. Möt á á frávikum verða einnig iðulega mjög stór. Oft er reynt að losna við vanda sem stafar af aðskilnaði með stækkun gagnasafns. Þá er vonast eftir fjölbreyttari gildum á skýristærð þannig að losnað sé við aðskilnaðinn. Einnig má reyna að komast fyrir vandann með því að skilgreina líkan uppá nýtt. Aðrar aðferðir eru

einnig notaðar svo sem endurskilgreining skýristærða eða jafnvel að varpa fyrir róða skýristærð sem aðskilnaður stafar af.

Engri þeirra aðferða sem nota má til þess að losna við aðskilnað er beitt hér. Ekki er talin þörf á því. Aðskilnaður getur komið upp þegar raunverulega er mikið samband milli skýribreyta og háðrar breytu. Gagnasafnið sem miðað er við verður ekki stækkað þar sem gögnin ná til allra sveitarfélaga ársins 2012. Gildi stika eru ekki ýkt og staðalfrávik þeirra eru fremur hófleg miðað við það sem getur gerst þegar aðskilnaður kemur upp. Farið verður yfir fleiri líkön og þau prófuð en það er ekki gert í tilraun til að leysa hugsanlegan vanda af aðskilnaði. Fremur er athugað hvort önnur líkön reynist betur skv. hefðbundnum mælikvörðum, eins og farið verður yfir.

Grunurinn um fullkominn eða hálffullkominn aðskilnað virðist stafa af því hve vel breyta mannfjölda segir til um árangur. Hún kemur einnig upp í líkani þar sem einungis fjöldabreyta er notuð til að skýra árangur en ekki ef líkan er keyrt þar sem árangur skýrist bara af tekjum. Athugun leiddi einnig í ljós að losnað er við meldinguna ef öfgagildinu (e. outlier) Reykjavík er hent úr gagnasafni áður en líkanið er keyrt, en þar er langmesti fjöldinn. Óráðlegt er talið að kasta til hliðar eins mikilvægum þætti líkansins og mannfjöldabreytan er þar sem samband virðist vera til staðar skv. líkönum. Það er stutt með kenningum og fjölda fyrri rannsókna. Ekki verður heldur átt við gagnasafnið með því að henda út öfgagildum. Líkan leitt af athugun af áhrifum þess gaf sömu mót á stikum þannig að engin þörf virðist á slíkri valkvæðri og huglægri „lagfæringu“.

**Tafla 4. Mát nokkurra logit líkana**

Líkan	Skýribreytur	-2log Likelih.	Frígráður	AIC	Samanborin líkön	Munur á Deviance	corr(y, $\hat{\mu}$ )
1	F+T+F <sup>2</sup> +T <sup>2</sup>	31,75	70	41,75	-	-	0,73
2	F+T+F <sup>2</sup>	31,75	71	39,75	2-1	0,00	0,73
3	F+T+T <sup>2</sup>	31,86	71	39,86	3-1	0,11	0,73
4	F+T	31,86	72	37,86	4-3	0,00	0,73
5	F+F <sup>2</sup>	33,29	72	39,29	5-2	1,53	0,73
6	T+T <sup>2</sup>	57,61	72	63,61	6-3	25,75	0,04
7	F	33,49	73	37,49	7-4	1,63	0,73
8	T	59,00	73	63,00	8-4	27,14	0,36
9	Fasti	69,17	74	71,17	9-7	35,68	-

Í töflu með mátum nokkurra logit líkana er borið saman hve vel líkön með ólíkum samsetningum skýristærða passa við árangursgögn. Smærra líkan er borið saman við stærra líkan sem inniheldur það smærra og viðbót sem felld hefur verið út og má sjá í dálknum samanborin líkön hvaða líkan eru borin saman í hverju tilfalli fyrir sig. Stærsta líkanið er skoðað fyrst en í því eru skýristærðir fjöldi og tekjur ásamt báðum breytum í öðru veldi. Líkön með breytunum í öðru veldi eru prófuð til að athuga hvort vaxandi eða dvínandi jaðaráhrif tekna eða fjölda séu til staðar. Gjarnan er reiknað með minnkandi jaðaráhrifum hvors tveggja og skoðun á jaðaráhrifum liggur því beint við. „Deviance“ eða munur á log likelihood er notaður til að bera saman stærri líkön og minni. Minni líkönin eru innifalin í þeim stærri þannig að þau eru það sem eftir stendur af stærra líkani þegar hluti þess hefur verið felldur út.

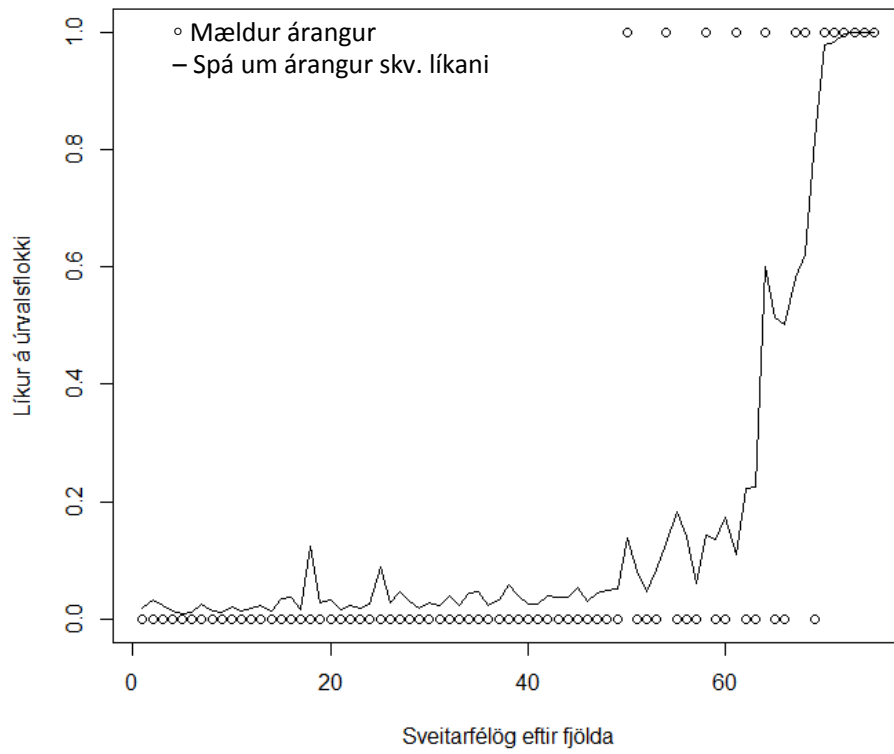
Frá stærsta líkani og niður úr eru áhrif úrfellinganna ekki þannig að álykta megi að smærri líkön séu verri fyrr en komið er að líkani 6. Líkan 6 er borið saman við líkan 3 og skoðað hvernig áhrif það hefur að fella algjörlega út fjöldaskýristærð og notast einungis við tekjuskýristærð og veldi hennar. „Deviance“ munur líkananna er talsverður og komist er að því að líkanið versni við að fella út fólksfjöldaskýristærð. Líkan 3 er því eftir þessum mælikvarða betra en líkan 6. Í líkönum 7 og 8 standa eftir stakar skýristærðir fjölda annars vegar og tekna hins vegar. Þau eru bæði borin saman við líkan 4, sem samanstendur af þeim báðum. Ekki er hægt að fullyrða út frá stærð „deviance“ munar líkana að líkan sem einungis inniheldur fjöldaskýristærð sé verra en stærra líkanið. Áhrif þess að standa einungis eftir með tekjuskýristærð eru þannig að ekki er hægt að réttlæta úrfellingu fjöldaskýristærðar úr líkani 4, líkanið versnar við að skýristærðin sé felld út. Líkan 9 er eingöngu byggt á fasta og ekki skýristærð. Borið saman við líkan 7 versnar líkanið við að fella út fjöldaskýristærðina. Því má ætla að árangur skýrist a.m.k. að einhverju leyti af fjölda.

Gæði líkana eru einnig borin saman með því að skoða upplýsingagildi Akaike (e. Akaike information criterion, AIC). Betra gildi AIC er í líkönum sem falla vel að gögnum en þó er refsað fyrir fjölda stika í líkani til að draga úr hættu á ofmáti (e. overfit). Lægri AIC gildi eru betri. Líkan 7, sem einungis inniheldur fjöldaskýristærð, er með lægsta AIC allra líkana eða 37,5, sem er litlu hærra en AIC líkans 4 (37,9) þar sem skýristærðir eru fjöldi og tekjur. Miðað við skoðun á gildum AIC er líkan 7 best og líkan 4 kemur næst.

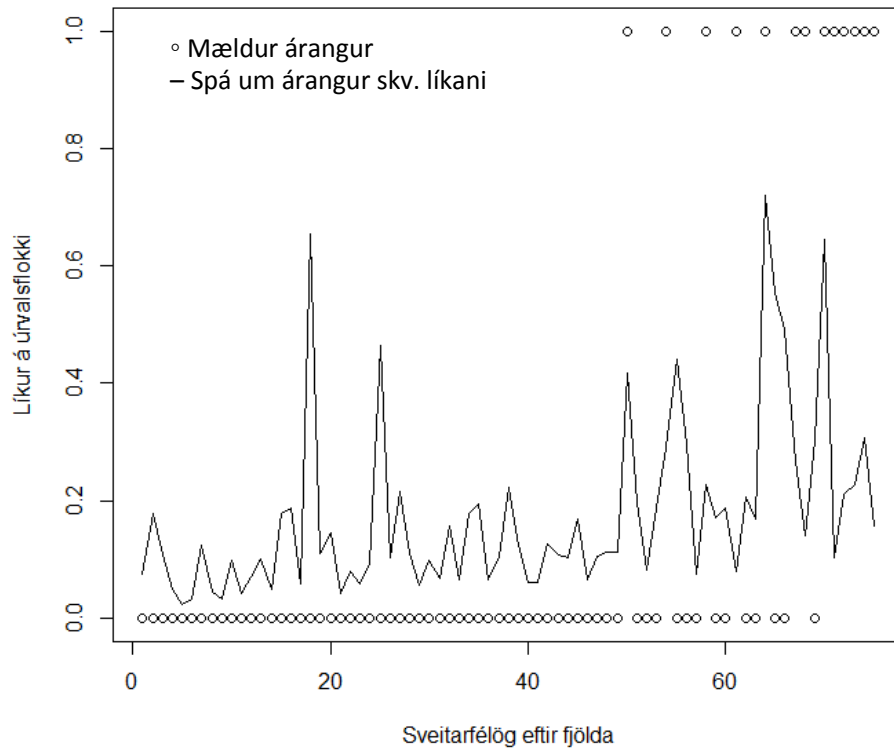


Til að mæla spágetu líkana og bera þau saman er fylgni háðu breytunnar um árangur og ætlaðra gilda hennar úr líkönum (e. fitted values), eða líkum líkana á því að falla í úrvalsflokk, reiknuð. Sést að fylgnin er minnst þegar fjöldi er ekki hafður með í líkani. Önnur líkön en nr. 6 og 8 hafa öll nokkurn veginn sömu fylgnina á milli háðu breytunnar og líkum líkana. Mikil fylgni mata annarra líkana við háðu breytuna getur þó bent til þess að líkönin séu nokkuð góð ef einungis þessi mælikvarði er skoðaður. Til glöggvunar á því hvað verið er að bera saman má skoða myndir þar sem líkum á að falla í úrvalsflokk úr líkönum 4 og 8 er stillt upp á mynd með raunmælingunum. Mælikvarði á fylgni er hversu vel línur, sem eru áætlanir líkana, fylgja punktunum, sem eru raunverulegu mælingar árangurs sveitarfélaga. Ljóst er af myndum að líkan 4 stendur sig mun betur og fellur betur að raunverulegu gögnunum en líkan 8. Myndunum er svo þannig stillt upp að frá vinstri til hægri eftir lárétta ásnum eykst fjöldi skattgreiðeinda og af myndum virðast því sjáanleg tengsl á milli fjölda og árangurs þar sem aukinn fjöldi tryggir betri árangur.

### Líkan 4



### Líkan 8



Mynd 1. Raunárangur og mót líkana um líkur á árangri

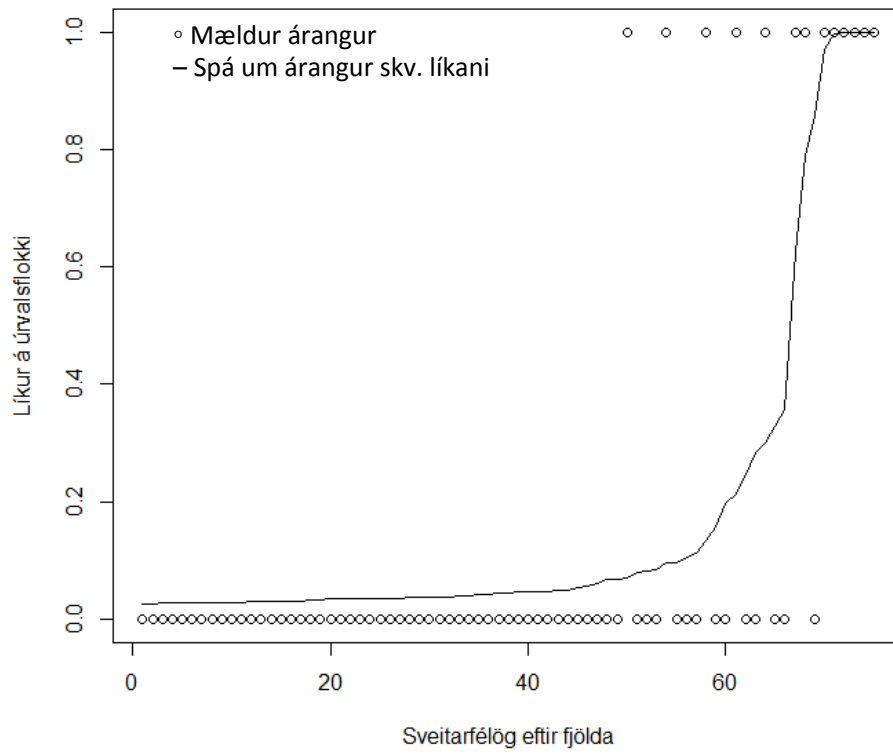
Með samanburði líkana er komist að því að logit líkönin sem innihalda skýristærð eða stærðir í öðru veldi bæta litlu við smærri líkön. Þau líkja ekki það markvert betur eftir gögnum að réttlæti að stækka líkönin með veldisskýristærðum. Ekki tekst að sýna fram á að veldisskýristærðir bæti nokkru við líkönin, einungis í einu líkananna er lítil vísbending um það. Það er í líkani 5,  $f + f^2$ , þar sem stiki veldisskýristærðar er lág neikvæð tala sem gæti bent til minnkandi jaðaráhrifa fjölda. Líkön sem reynast best eru líkan 7, sem einungis inniheldur fjöldaskýristærð, og líkan 4, sem samanstendur af fjölda- og tekjuskýristærðum og hefur þegar verið farið yfir. Skoða má niðurstöður líkans 7 í töflu.

Logit líkan fjölda  
Tvíkosta háð breyta: úrvalsflokkur eða ekki

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>z</i>	<i>p-gildi</i>
Fasti	-3,661	0,745	-4,91	<0,001
Fjöldi	1,191	0,370	3,22	0,001

Niðurstöður líkansins gefa til kynna að fjöldi hafi áhrif á árangurslíkur. Formerki stika fjöldans er jákvætt sem búið var við, og því meiri líkur á að falla í árangursflokk með auknum fjölda. Stiki fjöldans er metinn sem stærri tala en áður eða 1,191 og ætlaðar líkur á að sveitarfélag falli í úrvalsflokk margfaldast, að öðru óbreyttu, um  $\exp(1,191) = 3,29$  við hverja þúsund skattgreiðendur, áður  $\exp(0,993) = 2,699$ . Í sveitarfélagi með meðalfjölda eru líkur á að sveitarfélagið falli í úrvalsflokk skv. líkani 7 22,8% eða 0,1% hærra en matið úr líkani 4 sem telst sambærilegt. Eins og áður hefur aukinn mannfjöldi greinileg áhrif á líkur á að sveitarfélag falli í úrvalsflokk. Mynd af mati líkans 7 á líkum sveitarfélaga á árangri sýnir að líkanið fangar nokkuð raunverulega dreifingu árangurs.

### Líkan 7



Mynd 2. Raunárangur og mat líkans 7 um líkur á árangri

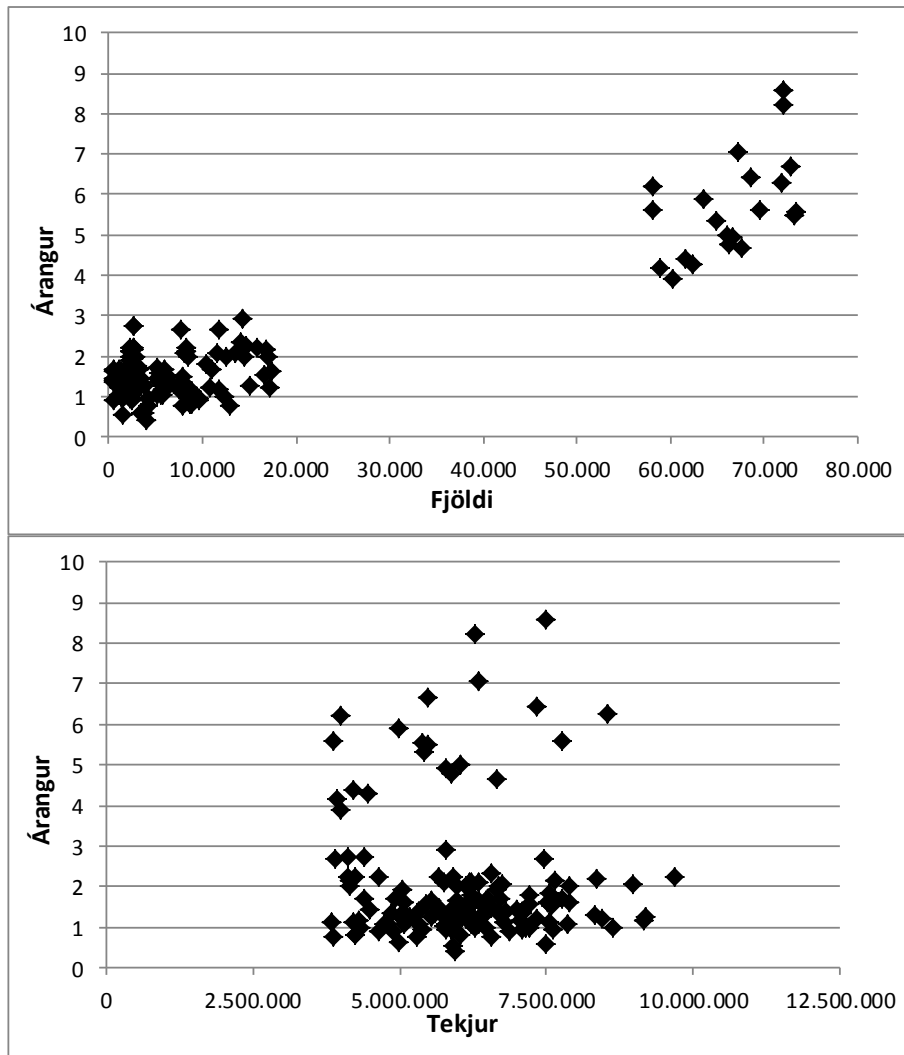
## 5 Langsniðslíkön af árangri

### 5.1 Verklag og skilgreiningar

Fyrst ber að útskýra annars konar skilgreiningu árangurs. Árangri er stillt upp sem samanlögðum meðalstigum liða sveitarfélaga á hverju tímabili úrvalsdeildarinnar, efstu deildar íslenskrar knattspyrnu, frá 1992-2012. Fyrir hvert tímabil eru meðalstig liða reiknuð. Ef fleiri en eitt lið koma úr sveitarfélagi eru meðalstig liðanna lögð saman. Í hverju sveitarfélagi fyrir sig eru lögð saman meðalstig liðanna sem þaðan koma. Þannig er mynduð ný háð breyta, samanlögð meðalstig liða sveitarfélags. Miðað er við meðalstig liða á leik yfir heil tímabil fremur en heildarstigafjölda vegna þess að liðum í úrvalsdeildinni fjölgaði á tímabilinu. Með því að deila fjölda leikja í stigin fást meðalstig á leik. Vissulega eykst summa meðalstiga líka þar sem liðum fjölgaði um tvö árið 2008 en sú aukning er minni hlutfallslega en aukning heildarstiga. Heildarstigum fjölgaði verulega með fjölgun liða enda fjölgaði leikjum liðanna úr 18 í 22 á tímabili og liðunum um tvö. Með notkun summu meðalstiga fremur en summu heildarstiga er komist fyrir vandann af fjölgun leikja, en eftir stendur þó aukning vegna þess að liðum fjölgaði um tvö. Tímabilið sem er til skoðunar takmarkast við gagnasafn skýristærða frá ríkisskattstjóra og nær því til áranna 1992-2012.

Ástæða þess að lögð eru saman meðalstig liða innan sveitarfélags er að gögn skýristærða eru miðuð við sveitarfélög. Ef árangursbreytan væri miðuð við lið fremur en sveitarfélög væri ólíkt gengi liða innan sveitarfélags skoðað með sömu skýristærðum. Þannig gæti árangur Íslandsmeistara og annars liðs sem féll á sama tímabili verið skoðaður með nákvæmlega sömu skýristærðum, komi þau úr sama sveitarfélagi.

Sömu skýristærðir eru notaðar og í logit líkönum. Skýristærðirnar takmarkast þó ekki við nýjustu mælingar heldur eru skýristærðir skoðaðar yfir allt tímabilið. Þar sem bæði sveitarfélög og tímabil eru mörg eru gögnin á langsniðs- eða panel formi og því viðeigandi að nota panel líkön. Meðalheildartekjur skattgreiðenda innan sveitarfélags er aftur tekjubreyta og fjöldi skattskyldra eininga eða skattgreiðenda í sveitarfélagi er einnig notaður aftur sem nálgun á íbúafjölda. Þar sem tekjur eru skoðaðar yfir langan tíma er leiðrétt fyrir þróun verðlags þannig að raunveruleg þróun tekna er skoðuð. Leiðrétt er miðað við meðalstöðu vísitölu neysluverðs á hverju ári. Verðlag ársins 2012, nýjasta árs úrtaksins, er viðmiðunarár verðlags.



**Mynd 3. Fjöldi og tekjur á móti árangri**

Sveitarfélög sem falla undir úrtakstímabilið hafa misoft átt lið í úrvalsdeildinni. Reykjavík hefur t.a.m. átt lið í úrvalsdeildinni allt tímabilið, önnur sveitarfélög sjaldnar svo sem Borgarbyggð þaðan sem Skallagrímur keppti í eitt skipti í úrvalsdeildinni, árið 1997.

Vandi sem kemur upp við skoðun á svo löngu tímabili er að talsvert hefur verið um sameiningar sveitarfélaga. Við sameiningu tekur íbúafjöldi nýja stærra sveitarfélagsins eðlilega stökk. Sama stökk verður ekki á meðalheildartekjum þar sem stærðin er alltaf á sama grunni en þó má ætla einhverjar breytingar á meðalheildartekjum vegna sameininga. Miðað er við meðaltekjur skattgreiðenda og sá grunnur breytist ekki þótt skattgreiðendum fjölgi. Reyndin er þó sú að einungis lið Keflavíkur spilaði í úrvalsdeild fyrir og eftir sameiningu sveitarfélags á tímabilinu sem er skoðað. Við sameininguna fjölgaði skattgreiðendum í sveitarfélagi liðsins úr tæpum 4 þúsundum í rúm 5 þúsund.

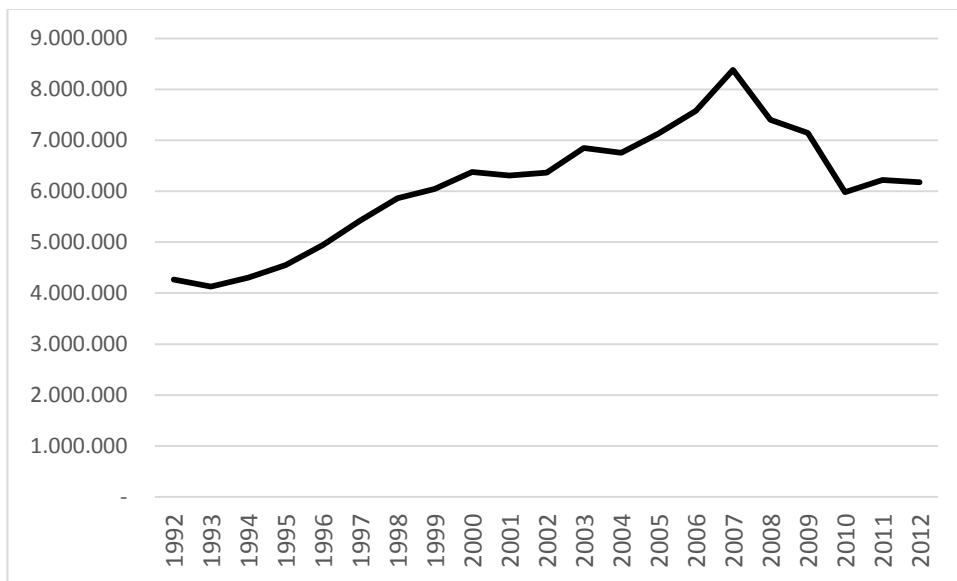
Knattspyrnuliðið Keflavík keppti einu sinni fyrir sveitarfélagið Keflavík á viðmiðunartímabilinu en frá árinu 1994 hefur liðið tilheyrt Reykjanesbæ. Áhrif sameininga koma því óverulega inn í úrtakið. Tekið er á vanda við sameiningu sveitarfélaga í Reykjanesbæ með því að hafa mælingar sveitarfélaganna í sitt hvoru lagi, þótt taki til sama liðsins. Fyrir sameiningu er ein mæling árangurs í Keflavíkurkaupstað og fjöldi og tekjur ársins látnar fylgja, eftir sameiningu er fjöldi mælinga á árangri fyrir Reykjanesbæ. Allar eru mælingarnar á árangri ÍBK. Þar sem mælingar skýristærða fyrir Keflavíkurkaupstað og Reykjanesbæ eru ekki byggðar á sama grunni eru sveitarfélögin höfð hvort fyrir sig í gagnasafninu.

**Tafla 5. Skýristærðir og háða stærðin eftir sveitarfélögum**

	Meðaltal samanlagðra meðalstiga	Staðalfrávik samanlagðra meðalstiga	Meðaltal Fjölda	Staðalfrávik Fjölda	Meðaltal meðalheildartekna	Staðalfrávik meðalheildartekna	Tíðni í úrvalsdeild
Akranes	1,77	0,53	2.898	296	5.814.450	1.186.102	18
Akureyri	1,38	0,72	8.378	680	5.014.165	1.072.759	7
Garðabær	1,10	0,44	4.926	1.038	7.019.330	1.390.209	8
Grindavík	1,17	0,27	1.244	192	6.662.729	1.101.509	17
Hafnafjörður	1,96	0,54	11.671	2.411	6.158.426	1.440.080	16
Kópavogur	1,35	0,49	13.515	3.396	6.335.065	1.663.977	14
Ólafsfjörður	1,39	0,28	546	26	6.097.512	952.180	6
Reykjanesbær	1,35	0,27	6.362	1.131	5.633.084	762.851	18
Reykjavík	5,67	1,25	66.455	5.156	5.682.839	1.358.525	21
Árborg	0,86	0,13	4.167	21	5.337.707	53.459	2
Vestmannaeyjar	1,51	0,45	2.309	66	6.375.578	1.107.669	19

Keflavíkurkaupstaður og Borgarbyggð eru einu sveitarfélögin á úrtakstímabilinu sem hafa einungis einu sinni átt lið í úrvalsdeildinni. Gögnin um liðin eru ekki á panel formi þar sem sveitarfélögin falla aðeins eitt árið inn í úrtakið og vantar því tímavidd til að uppfylla skilyrði þess að vera panel gögn. Þau eru höfð utan við líkönin þar sem mót á fixed eða random effect fyrir sveitarfélögin væru byggð á of litlum upplýsingum og keyrsla panel líkana á mælingar sem hafa enga tímavidd eru ekki réttlætanager. Sveitarfélögin sem skipa gagnasafnið eru 11 talsins. Þau hafa misoft verið í úrvalsdeild á tímabilinu 1992-2012, allt frá því að vera alltaf, eins og við á um Reykjavík, niður í lágmarkið til að vera notað í líkanasmíðum, tvisvar, sem Sveitarfélagið Árborg hefur náð. Töflur með skýristærðum og óháðu stærðinni hafa verið settar upp hvort tveggja

miðaðar við sveitarfélög og svo eftir tímabilum. Í töflu með árlegum tölum er í útreikningi miðað við sveitarfélög sem falla inní gagnasafnið á hverju ári. Breytilegt er því hvaða sveitarfélög eru í útreikningi hvers árs. Árborg er t.a.m. með árin 2010 og 2012, árin sem sveitarfélagið átti lið í úrvalsdeild, önnur ár eru tölur sveitarfélagsins ekki með í útreikningi. Í töflu sveitarfélaganna er útreikningur sambærilegur þar sem aðeins er miðað við þau ár sem sveitarfélögin áttu lið í úrvalsdeild. Árborg er aftur gott dæmi þar sem meðaltöl og frávík eru reiknuð eftir þeim tveimur tímabilum sem lið sveitarfélagsins keppti í úrvalsdeild.



**Mynd 4. Þróun meðaltekna yfir tíma, m.v. verðlag ársins 2012**



Tafla 6. Skýristærðir og háða stærðin eftir tímabilum

	Meðaltal samanlagðra meðalstiga	Staðalfrávik samanlagðra meðalstiga	Meðaltal Fjölda	Staðalfrávik Fjölda	Meðaltal meðal-heildartekna	Staðalfrávik meðal-heildartekna
1992	2,33	2,05	14.604	21.450	4.263.824	348.426
1993	2,54	1,86	15.856	23.789	4.129.916	355.618
1994	1,70	1,15	12.270	19.011	4.306.489	395.397
1995	1,78	1,03	11.227	20.100	4.550.998	392.700
1996	1,74	1,15	10.867	20.701	4.943.120	449.791
1997	1,87	1,22	11.154	22.649	5.421.613	576.957
1998	2,28	1,80	12.575	25.065	5.865.024	660.819
1999	1,93	1,54	12.693	23.351	6.044.511	530.796
2000	1,69	1,35	11.946	22.396	6.375.685	619.853
2001	2,02	1,29	14.535	23.257	6.305.390	300.330
2002	1,92	1,40	13.902	23.255	6.367.127	625.033
2003	2,37	2,31	15.513	25.609	6.848.201	1.214.447
2004	1,94	1,27	14.340	23.814	6.753.578	792.663
2005	2,39	2,08	15.523	26.286	7.129.078	346.663
2006	1,93	1,66	15.768	24.308	7.577.130	550.193
2007	2,68	2,05	22.381	28.102	8.383.187	1.047.270
2008	2,83	2,89	19.343	26.540	7.402.943	658.695
2009	2,40	2,61	17.182	24.887	7.144.392	1.288.530
2010	2,07	1,56	15.801	23.961	5.983.848	810.433
2011	2,06	1,91	16.458	23.396	6.220.781	963.163
2012	1,85	1,45	14.533	22.648	6.174.903	965.836

Einkenni panel gagna er að einstaklingum eða hópum er fylgt eftir í tíma. Það sem gerir panel líkön sérstök er auka vídd vegna þess að gögn ná hvoru tveggja yfir tíma og á milli hópa sem miðað er við. Almennt panel líkan er á forminu:

$$y_{it} = \mathbf{X}'_{it}\boldsymbol{\beta} + \varepsilon_{it}, i = 1, \dots, N; t = 1, \dots, T$$

þar sem  $i$  stendur fyrir þversniðsvídd gagna, sem geta t.d. verið fjölskyldur, einstaklingar, heimili, fyrirtæki eða lönd, og  $t$  stendur fyrir tímavídd. Stærðirnar  $\mathbf{X}_{it}$  og  $\boldsymbol{\beta}$  eru vektorar af stærð  $k + 1$  og  $k$  er fjöldi breyta. Úrtaksstærð þversniðsvíddar er  $N$  og tímavídd er af lengdinni  $T$ .

Þrjú líkön eru gjarnan notuð til að kryfja panel gagnasöfn. Einfaldasta form aðhvarfsgreiningar á panel gögn, þar sem gögnunum er slegið saman (e. pooled regression), lítur fram hjá mögulegum mun viðmiðunarhópa og reiknar ekki með misjöfnum áhrifum skýristærða yfir tíma. Almennt pooled líkan er:

$$y_j = \mathbf{X}'_j \boldsymbol{\beta} + \varepsilon_j, j = 1, \dots, N \cdot T$$

Hvorki er tillit tekið til tíma né þversniðsvídda í þessu almenna líkani. Oft er of mikil einföldun að taka ekki tillit til sérstakrar uppbyggingar panel gagna. Með því að reyna að líkja betur eftir misjöfnum einkennum einstaklinga, eða hópa, og/eða ólíkum áhrifum yfir tíma má búa til betri líkön. Þá er hægt að taka tillit til þess með líkönum þar sem áhrif einstaklinga eru fixed effects eða random effects, sem gæti útlagst á íslensku sem líkön með föstum eða slembnum áhrifum. Almenn fixed eða random effect líkön eru:

$$y_{it} = u_i + \mathbf{X}'_{it} \boldsymbol{\beta} + \varepsilon_{it}, i = 1, \dots, N; t = 1, \dots, T$$

Þar sem  $u_i$  er einstaklingsbundinn liður sem getur verið fixed eða random. Í fixed effect líkönum er gert ráð fyrir að einstaklingsáhrifin séu hugsanlega háð skýristærð sem tekið er inn í líkanasmíð í gegnum fasta fyrir hvern einstakling og/eða tímabil. Í random effect líkönum er gert ráð fyrir að einstaklingsáhrif séu ekki tengd neinum skýristærðum heldur stafi af ólíkri dreifingu leifaliða á milli hópa eða tímabila. Einstaklingsáhrifin, eða tímaáhrif, eru því ekki reiknuð með fastanum eða stikum skýristærða heldur eru tekin inn í útreikning leifaliðs fyrir hvern einstakling eða tímabil og líkön þannig aðlöguð að gögnum.

Líkön sem farið er yfir eru með skýristærðunum línulegum og svo skýristærðunum í öðru veldi. Líkön með fallform af öðru tagi voru prófuð s.s. með kvaðratrót skýristærða, lógaritma og veldisvísisföllum (e. exponential). Niðurstöður þeirra voru ekki þannig að tilefni sé talið til að bæta við umfjöllun um þau líkön þar sem þau bæta litlu við.

## 5.2 Langsniðslíkön

Áður en lagt er í flóknari líkön eru skoðuð panel líkön sem einungis eru skipuð skýristærðunum tveimur línulega. Fyrst ber að skoða pooled VAMK líkan. Í líkaninu er gert ráð fyrir að fasti og stikar séu fastar stærðir óháð hópum eða tíma. Líkanið er:

$$samanlögð\ meðalstig_j = \beta_0 + \beta_1 fjöldi_j + \beta_2 tekjur_j + \varepsilon_j$$

Þar sem ekkert tillit er tekið til sveitarfélaga- eða tímavídda í líkaninu eru ekki lágvísar (e. subscript) fyrir sveitarfélög,  $i$ , og tímabil,  $t$ , í jöfnunni. Í staðinn kemur  $j$  sem stendur fyrir mælingu óháða tíma eða sveitarfélögum, öllu gagnasafninu er skeytt saman í pooled líkan sem lítur fram hjá hvoru tveggja. Stendur  $j$  þá fyrir tiltekið sveitarfélag á tilteknum tíma og mælingar  $j$  eru af fjöldanum  $N \cdot T$ . Jafnan segir að meðalstig í

sveitarfélagi skýrist af fasta  $\beta_0$  að viðbættum *fjölda* í sama sveitarfélagi margfölduðu með stika fjöldans,  $\beta_1$ , auk (meðal)tekna íbúa sveitarfélagsins margfölduðum með stika tekna,  $\beta_2$ . Þá er gert ráð fyrir að einhver skekkja geti verið í matinu fyrir hvert sveitarfélag og fellur hún undir leifaliðinn,  $\varepsilon_j$ .

Líkan 1: Pooled VAMK líkan  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	0,907	0,265	3,42	<0,001
Fjöldi	0,068	0,003	27,83	<0,001
Tekjur	0,029	0,042	0,69	0,490

Aðhvarfslíkanið er:

$$\text{samanlögð meðalstig} = 0,91 + 0,07\text{fjöldi} + 0,03\text{tekjur}$$

Báðir stikar skýristærða hafa jákvæð formerki, eins og búist var við. Meðalstig í sveitarfélagi aukast því þegar fjöldi eða tekjur aukast. Aukist fjöldaskýristærð í sveitarfélagi um eina einingu, sem er 1000 skattgreiðendur, eykst háða breytan í sveitarfélaginu, samanlögð meðalstig, um 0,07 stig. Þegar meðaltekjur íbúa sveitarfélags hækka um eina milljón, eininguna sem miðað er við, eykst háða breytan í sveitarfélaginu um 0,03 stig. Mat á stikanum er þó þannig að varhugavert er að álykta of mikið um hann og má sjá það t.a.m. af því að staðalskekkjan er stærri tala en stikinn.

Áhrif fjöldans eru veruleg en áhrif tekna ekki vel greinanleg. F-próf með tilgátu um að stikar líkans séu jafnir núlli gefur tilefni til að álykta að þeir séu ekki allir jafnir núlli. Tekjustiki er þó talsvert frá því að gefa afdráttarlausu niðurstöðu og því tilefni til að skoða hvort gera megji betur. Því er rétt að skoða frekari líkön og fara betur yfir þau próf sem eiga við.

Sambærilegt fixed effect líkan gerir ráð fyrir að áhrif skýristærða séu fastar stærðir óháðar hópum eða tíma en fasti hvers sveitarfélags, eða tímabils ef líkani er þannig stillt upp, sé ólíkur. Líkanið er:

$$\text{samanlögð meðalstig}_{it} = (\beta_0 + u_i) + \beta_1\text{fjöldi}_{it} + \beta_2\text{tekjur}_{it} + \varepsilon_{it}$$

Munur fixed effects líkans og pooled líkans er  $u_i$  sem er fixed effect stiki hvers sveitarfélags eða föst áhrif í hverju sveitarfélagi önnur en sameiginlegi fastinn  $\beta_0$ . Áhrif fixed effect stikans  $u_i$  eru gjarnan mæld með gervibreytum og stendur  $u_i$  þá fyrir gervibreytu sveitarfélags  $i$ .

Líkan 2: Fixed-effects  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	0,353	0,311	1,14	0,258
Fjöldi	0,127	0,023	5,39	<0,001
Tekjur	-0,019	0,047	-0,40	0,687

Með F-prófi er metið hvort fixed effect líkan eigi við með því að skoða hvort hóparnir hafi sameiginlegan fasta. Niðurstöður prófsins eru að hægt sé að hafna núlltilgátu um sameiginlegan fasta. Niðurstaða prófsins felur með öðrum orðum í sér að fixed effects hluti líkansins bæti það umfram pooled VAMK líkanið. Þá er summa leifaliða í öðru veldi í fixed effect líkaninu lægri en í pooled líkaninu, 42,61 á móti 59,37, og bendir mælikvarðinn til þess að fixed effect líkanið passi betur við gögnin. Samanburður á mælikvarða Akaike á upplýsingagildi (e. Akaike information criteria, AIC) á pooled og fixed effects líkönum bendir einnig til þess að fixed effects líkanið sé betra. Gildi AIC eru annars vegar 288,96 fyrir pooled líkanið en 260,29 fyrir fixed effect líkanið og lægra AIC gildi er betra. Ályktað er að nota eigi fixed effect líkan fremur en pooled líkan.

Niðurstöður fixed effects aðhvarflíkansins eru um margt sambærilegar við þær úr pooled líkaninu. Aftur eru áhrif fjöldabreytu afgerandi í máti líkansins. Stiki fjöldaskýribreytu er enn jákvæður en stiki tekjuskýribreytu verður reyndar neikvæður. Stiki tekjuskýribreytu er þó eins og áður þannig að staðalskekkja stikans veldur óvissu um hvoru megin við núll áhrif skýristærðarinnar falla í raun og því hæpið að álykta of mikið um áhrif hennar. Fasti fixed effects líkansins er heldur lægri en stafar það að verulegu leiti til af því að nú hefur hvert sveitarfélag aukalega sinn eigin fixed effect stika sem dregur úr mikilvægi sameiginlega stikans. Fjölgun skattgreiðenda um þúsund leiðir skv. fixed effect aðhvarflíkaninu til 0,023 aukningar í samanlögðum meðalstigum liða sveitarfélags. Aukning meðaltekna íbúa um milljón leiðir til þess að dregur úr háðu breytunni um 0,019 stig en eins og áður segir er varhugavert að fullyrða um áhrif skýribreytu út frá stikanum.

Líkan 3: Random-effects  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	0,765	0,270	2,84	0,005
Fjöldi	0,072	0,006	12,06	<0,001
Tekjur	0,038	0,040	0,96	0,337

Með líkani eftir random effects aðferð er leystur vandi vegna munar á venjulegu pooled aðhvarfslíkani og gögnum með því að bæta liðum við leifalið þar sem áhrif einstaklinga eða hópa koma fram sem sérstök mæliskekkja hvers hóps, einstaklings eða tíma, til viðbótar við þá sameiginlegu. Líkanið er:

$$samanlögð\ meðalstigi_{it} = \beta_0 + \beta_1 fjöldi_{it} + \beta_2 tekjur_{it} + (\varepsilon_{it} + u_i)$$

Viðbótin við pooled líkanið,  $u_i$ , er liður notaður til að grípa misleitni (e. heterogeneity) á milli einstaklinga eða hópa. Í fixed effects líkani fer aðlögunin fram með sérstökum fasta fyrir einstaklinga eða hópa en í random effects líkani gerist það í sérstakri viðbót við leifaliðinn.

Með Breusch-Pagan prófinu (BP) er metið hvort slembin áhrif (e. random effect) séu til staðar með því að bera saman random effect líkan og hliðstætt pooled VAMK líkan. Núlltilgátan er að dreifniliður (e. variance component) hópanna sé núll og að ekki séu til staðar random effect áhrif hjá einstaklingum eða hópum. Núlltilgátu BP-prófs er hafnað í Random effects líkaninu og komist að því að líkanið passi betur við gögn en pooled líkan.

Hausmann próf er notað til samanburðar á fixed effects líkani og random effects líkani af sama tagi. Núlltilgáta prófsins er að engin fylgni sé milli einstaklings eða hópa og gildi á skýristærð. Prófið er notað til aðstoðar við að skera úr þegar bæði Breusch Pagan próf fyrir random effects líkan og F-próf fyrir fixed effects líkan gefa til kynna að líkönin séu betri en pooled líkan. Núlltilgátu Hausmann prófs er hafnað og ályktað að fixed effect líkanið henti betur. Komast má að sömu niðurstöðu með samanburði á mælikvarða Akaike á upplýsingagildi AIC fixed effect líkansins en það er nokkru lægra (260,5), sem er betra, en random effect líkansins (291,4). Summa leifaliða í öðru veldi úr random effect líkaninu er hærri (60,38) en hvoru tveggja í pooled líkaninu (59,37) og fixed effect líkaninu (42,61). Líkanið passar því síður við gögnin.

Þessar niðurstöður ólíkra mælikvarða á mátgæðum leiða til þess að ekki þörf á að túlka random effect líkanið. Fixed effect líkanið reynist betur. Bæði líkönin geta ekki verið rétt, og því einungis þörf á túlkun á því sem er réttara út frá mælikvörðum sem eru skoðaðir. Þar sem fixed effect líkanið er betra má álykta að sömu kraftar og ákvarða hæfni í fótbolta ákvarði einnig tekjur eða fjölda.

Athugun á jaðaráhrifum skýristærða er forvitnileg. Pooled líkan sem metur jaðaráhrif skýristærða er:

$$\text{samanlögð meðalstig}_j = \beta_0 + \beta_1 \text{fjöldi}_j + \beta_2 \text{tekjur}_j + \beta_3 \text{fjöldi}_j^2 + \beta_4 \text{tekjur}_j^2 + \varepsilon_j$$

Við fyrra pooled líkan bætast skýristærðir *fjöldi* og *tekna* í öðru veldi. Líkt og felst í pooled líkönum er aftur hvorki tekið tillit til tíma- né þversniðsvíddar og lágvísir *j* því í jöfnu, ekki *i* eða *t*.

Líkan 4: Pooled VAMK  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	2,114	0,982	2,15	0,033
Fjöldi	0,006	0,014	0,43	0,670
Tekjur	-0,263	0,312	-0,84	0,401
Fjöldi <sup>2</sup>	0,001	0,0002	4,49	<0,001
Tekjur <sup>2</sup>	0,022	0,025	0,91	0,364

Því verður ekki haldið fram að líkanið sé gott. Ekki tekst að sýna að skýristærðir hafi áhrif. Túlkun á líkani þar sem niðurstöður eru eins óafgerandi og raunin er með pooled líkanið er hæpin. Ekki verður talið rétt að fara yfir niðurstöður bæði vegna þess og hins að engin sannfæring og mikil óvissa fylgdi allri túlkun á líkaninu.

Líkan 5: Fixed-effects  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	2,130	0,932	2,29	0,024
Fjöldi	0,108	0,048	2,24	0,027
Tekjur	-0,592	0,321	-1,84	0,068
Fjöldi <sup>2</sup>	0,0003	0,0004	0,83	0,405
Tekjur <sup>2</sup>	0,045	0,025	1,84	0,068

Athugað er hvort fixed effect líkan eigi við fremur en pooled líkan með F-prófi. Af prófinu leiðir að núlltilgátu um sameiginlegan fasta er hafnað og komist að því að fixed effect líkan passi betur en samsvarandi pooled líkan. Summa leifaliða fixed effect líkansins í öðru veldi er minni en í pooled líkaninu, 41,28 á móti 51,95, sem er önnur

vísbending um að fixed effect líkanið eigi frekar við. Upplýsingagildi Akaike fyrir fixed effect líkanið er 259,91 en gildið fyrir pooled líkan er hærra, og verra, eða 273,46.

Líkan 6: Random-effects  
Háða breytan: Samanlögð meðalstig

	<i>Stiki</i>	<i>Staðalskekkja</i>	<i>t-gildi</i>	<i>p-gildi</i>
Fasti	2,083	0,932	2,23	0,027
Fjöldi	0,014	0,021	0,64	0,522
Tekjur	-0,300	0,294	-1,02	0,309
Fjöldi <sup>2</sup>	0,0008	0,0003	2,85	0,005
Tekjur <sup>2</sup>	0,027	0,023	1,15	0,253

Random effect líkan með veldisskýrisstærðum er borið saman við pooled líkan sömu tegundar með Breusch-Pagan prófi. Núlltilgátu BP-prófs er hafnað í random effect líkaninu. Niðurstöður prófsins gefa til kynna að líkja megi betur eftir gögnum með því að taka tillit til slembinna áhrifa á milli sveitarfélaganna og nota random effect líkan fremur en pooled líkan.

Við ákvörðun á milli fixed- og random effect líkana er Hausman próf framkvæmt. Núlltilgáta prófsins er að random effect áhrif séu óháð skýristærðum. Ef núlltilgátu prófsins er hafnað er ekki rétt að nota random effect líkan þar sem tengsl eru milli random effect áhrifa og skýristærða. Við þær aðstæður er fixed effect líkan því notað. Hið gagnstæða á þó ekki endilega við ef núlltilgátunni er ekki hafnað, þ.e.a.s. að rétt sé að nota random effect. Ekki er hægt að hafna núlltilgátu Hausman prófsins. Ef núlltilgátu Hausman prófs er hafnað er rangt að nota random effect líkan.

Random effect líkanið gefur ekki tilefni til túlkunar. Hæpið er að túlka nokkuð úr stika *fjöldans* í öðru, sem er eini stikinn með hóflega staðalskekkju m.v. stærð stika, og einnig varla hægt að segja mikið um hann þar sem áhrif línulega stika *fjöldans* eru óljós. Harla lítið ótvírætt eða ápreifanlegt er að fá úr túlkun þessa líkans.

Upplýsingagildi Akaike fixed effect líkansins er betra (259,91) en random effect líkansins (275,62). Summa leifaliða í öðru veldi úr random effect líkaninu er hærri (52,72) en bæði í pooled líkaninu (51,95) og fixed effect líkaninu (41,28). Random effect líkanið fellur því verr að gögnunum skv. mælikvarðanum. Mælikvarðar á gæðum líkana benda til þess að fixed effect líkanið sé betra. Vegna þess og þar sem loku er ekki fyrir það skotið að nota fixed effect líkan þó núlltilgátu Hausman prófs sé ekki hafnað er random effect líkanið ekki túlkað heldur komist að því að fixed effect líkanið sé réttara.

Vegna þess að fixed effect líkanið reynist betur er rétt að skoða það nánar. Sameiginlegi fastinn segir lítið í fixed effect líkani. Að honum frátöldum er einungis stiki *fjöldans* í fixed effect líkaninu með staðalskekkju sem er minni en helmingur af mati stika. Stiki *fjöldans* er 0,108 og staðalskekkja 0,048, formerki stikans er jákvætt líkt og búist var við. Stiki *fjöldans* í öðru veldi gefur óljósa niðurstöðu. Stikar *tekju*- breytanna eru með staðalskekkjur minni en mót stikanna en þó fremur miklar. Formerki tekjustikanna koma á óvart. Samkvæmt niðurstöðum fixed effect aðhvarfslíkansins eru neikvæð áhrif af tekjum en dregur úr viðbótar neikvæðum áhrifum af auknum tekjum þar sem jaðaráhrif tekna eru jákvæð. Varlega verður þó að fara í túlkun stikanna þar sem slaka þarf á kröfum um stærð staðalskekkja þeirra, eigi að notast við þá.

Túlkun á líkani með skýristærðum í veldi liggur ekki eins beint við og í línulegu líkani. Áhrif skýristærðar eru breytileg eftir því hvaða stig skýristærðar er skoðuð. Í fixed effects líkaninu með veldisskýristærðum fara áhrif fjölda eftir því við hvaða fjölda er miðað. Eins fara áhrif tekna eftir því hvaða tekjur er miðað. Líkanið fengið úr fixed effect aðhvarfinu er:

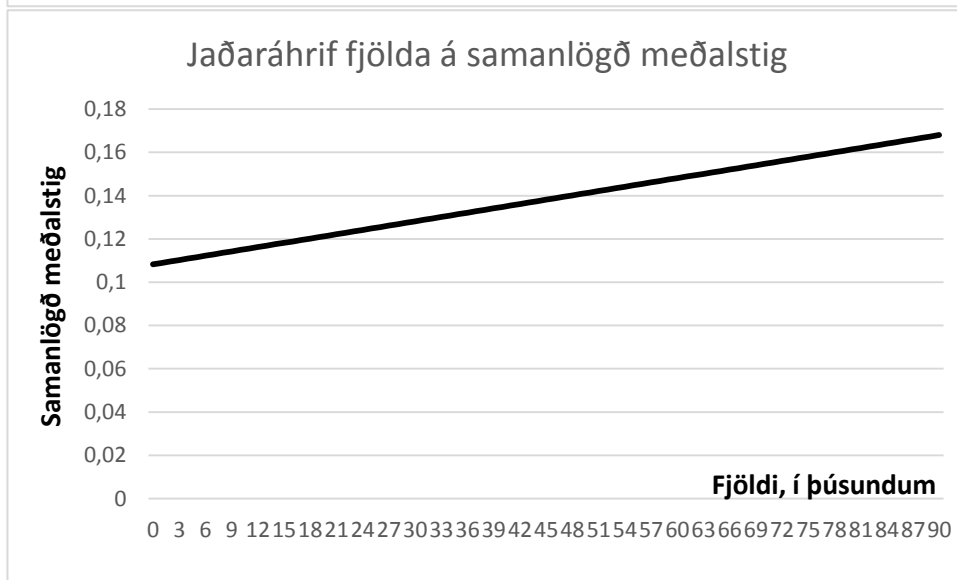
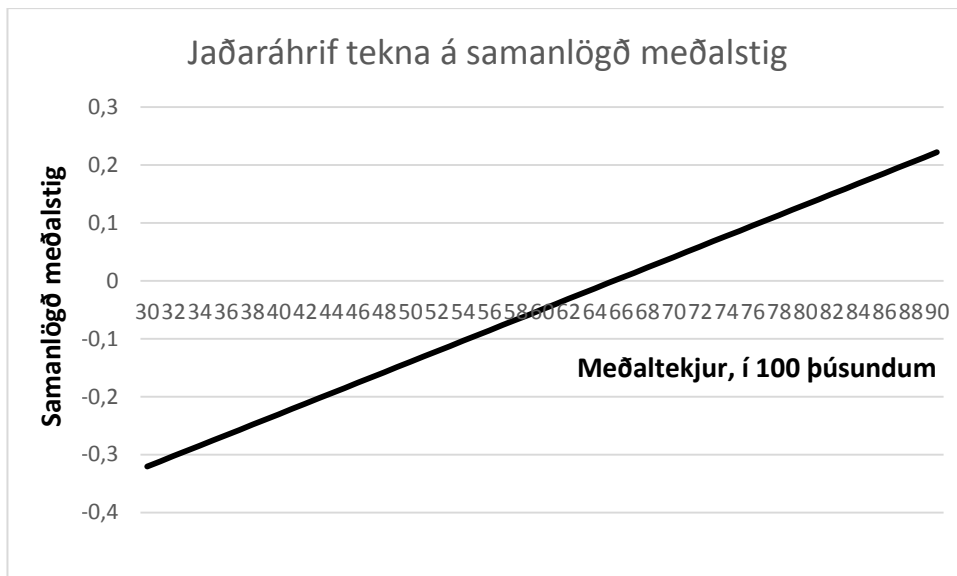
*samanlögð meðalstig*

$$= (2,13 + u_i) + ,108fjöldi - ,592tekjur + ,0003fjöldi^2 + ,045tekjur^2$$

Neikvætt formerki línulegs stika tekna hefur þau áhrif að við meðaltekjur skattskyldra eininga innan þess bils sem þær raunverulega dreifast yfir eru áhrif tekna neikvæð. Það er ekki fyrr en vel er komið út fyrir raunverulega bilið sem tekjurnar dreifast yfir sem áhrif aukinna tekna orðin jákvæð skv. fixed effect líkaninu með veldisskýristærðum. Er það við meðaltekjur meiri en 13 milljónir króna. Neikvæð áhrif tekna eru breytileg. Við lægri gildi tekna aukast neikvæðu áhrifin þangað til komið er í lágpunkt sem er þegar tekjur eru um 6,5 milljónir króna að meðaltali. Síðan dregur úr neikvæðu áhrifunum eftir lágpunkt og áhrifin verða síðan jákvæð við enn auknar tekjur. Stiki veldisskýristærðar fjöldans er lág tala og ólínuleg áhrif hans á hver áhrif fjöldans eru á háðu breytuna því lítil.

Með því að deilda fixed effect aðhvarfsjöfnuna má skoða beint jaðaráhrif skýristærða. Gagnlegt er að skoða áhrifin myndrænt.





**Mynd 5. Jaðaráhrif skýristærða á háðu breytuna samanlögð meðalstig**

Glögg má greina að jaðaráhrif tekna eru breytileg eftir því við hvaða tekjur er miðað. Jaðaráhrifin eru neikvæð þar til botni er náð í um 6,5 milljónum. Eftir það eru jaðaráhrif tekna jákvæð. Jaðaráhrif fjöldans eru jákvæð en lítil sem sést af því lína jaðaráhrifa fjöldans er fremur flöt.

## 6 Niðurstöður

Því er velt upp hvort árangur sveitarfélaga í fótbolta skýrist af tekjum íbúa og fjölda þeirra. Tvær ólíkar skilgreiningar á árangri eru því settar fram og líkön skipuð tekjum og fjölda sem skýristærðum borin saman. Önnur skilgreiningin skiptir sveitarfélögum í tvo hópa eftir því hvort lið úr sveitarfélaginu hefur tekið þátt í úrvalsdeild á síðustu rúmlega tuttugu árum eða ekki. Hin skilgreining árangurs er stig úr úrvaldsdeildinni. Ýmis líkön eru sett fram til að fanga hvað skýri hvora árangursskilgreiningu fyrir sig og skipta má þeim í tvennt. Tvíkosta logit líkön eru notuð til að líkja eftir annarri skilgreiningu árangurs og panel eða langsniðslíkön fyrir hina.

Líkönin gáfu mistraustar niðurstöður. Túlkun margra þeirra er hæpin vegna vafa um áreiðanleika. Þau líkön sem nýta má til að álykta gefa til kynna að tengsl séu á milli fjölda íbúa sveitarfélags og árangurs í fótbolta. Samræmist það því sem fyrirfram mátti reikna með út frá kenningum hagfræði og fyrri rannsóknum (P. C. García o.fl., 2007; Hoffmann o.fl., 2002a; Houston og Wilson, 2002; Macmillan og Smith, 2007). Jaðaráhrif fjölmennis á árangur eru skv. langsniðslíkönunum jákvæð og vaxandi, en lítil, og árangur því enn líklegri með auknum fjölda íbúa sveitarfélaga. Er það andstætt fyrri rannsókn þar sem jaðaráhrif fólksfjölda á árangur reyndust vera minnkandi eins og fremur er búist við (Macmillan og Smith, 2007). Vegna þess hve jaðaráhrifin eru lítil og hve misvel tekst að sýna fram á þau í líkönunum er vart rétt að lesa of mikið í þá niðurstöðu. Torveldara reynist að tengja tekjur íbúa við árangur en fjölda og gefa líkön misvísandi niðurstöður um áhrif tekna. Fyrri rannsóknir hafa sýnt að tekjur hafi jákvæð áhrif á árangur í fótbolta (P. C. García o.fl., 2007; Hoffmann o.fl., 2002a; Houston og Wilson, 2002; Macmillan og Smith, 2007). Jaðaráhrif tekna á árangur eru einnig skoðuð með líkönunum. Það bætir litlu við getu til að álykta um áhrif tekna, rannsóknir benda þó til þess að jaðaráhrif tekna á árangur í fótbolta séu minnkandi (Houston og Wilson, 2002).

## 7 Umræður

Niðurstöður líkana takmarkast af fjölda punkta (e. data points) í hvorri tegund gagnasafna. Í gagnasafni með tvíkosta háðu breytuna eru 75 punktar, sem ákvarðast af fjölda sveitarfélaga. Í 146 skipti hafa sveitarfélög sem átt hafa lið a.m.k. tvisvar í úrvalsdeild á tímabilinu 1992 – 2012 verið í úrvalsdeildinni, punktar langsniðsgagna eru jafnmargir. Fjöldi punkta eða athugna er því ekki mikill í gagnasöfnunum en þau eru þó vel nothæf fyrir líkönin sem eru sett upp.

Skilgreining árangurs spilar vafalaust inn í útkomu. Vera má að öðruvísi niðurstöður fengjust væri árangur skilgreindur á annan hátt en hér er gert. Fróðlegt væri að beita sömu aðferðum við gögn um fleiri deildir en einungis úrvalsdeildina. Árangur þyrfti þá að skilgreina á annan máta og líkja eftir árangri með nýjum nálgunum í líkönum.

Mikið hefur verið gagnrýnt hve litla umfjöllun kvennafótbolti, og almennt íþróttir kvenna, fái. Ekkert er gert til að rétta af kynjahallann hér sem er miður. Telja má það augljósan kost fyrir frekari rannsóknir.

Eðlilegt er að hugleiða hvernig nýta megi niðurstöður rannsóknarinnar. Mikil óvissa fylgir niðurstöðum í þessu verki, að frátöldu því að árangur skýrist að einhverju leyti af fólksfjölda. Því er spurning hvernig nýta eigi niðurstöður og einnig hverjum þær nýtast. Vart getur það talist sérstakt markmið sveitarfélaga að leitast við að ná árangri í fótbolta. Í öllu falli væri mjög sérstakt að leitast við að fjölga íbúum sérstaklega í þeim tilgangi.

Notagildi verksins er fremur fræðilegs eðlis þar sem farið hefur verið yfir margar fræðigreinar og rannsóknir. Þá er tæpt á álitamálum varðandi hlutverk íþróttafélaga og er verðugt verkefni að skoða gagnrýnum augum hvert það eigi að vera.

## Heimildir

- Anderson, C. og Sally, D. (2013). *The Numbers Game: Why Everything You Know About Soccer Is Wrong*. New York: Penguin Books.
- Antonioni, P. og Cubbin, J. (2000). The Bosman Ruling and the Emergence of a Single Market in Soccer Talent. *European Journal of Law and Economics*, 9(2), 157–173. doi:10.1023/A:1018778718514
- Audas, R., Dobson, S. og Goddard, J. (1997). Team performance and managerial change in the English Football League. *Economic Affairs*, 17(3), 30–36. doi:10.1111/1468-0270.00039
- Avgerinou, V. (2007). The Economics of Professional Team Sports: Content, Trends and Future Developments. *Sport Management International Journal*, 3(1), 5–17.
- Álagning einstaklinga. (e.d.). *Ríkisskattstjóri*. Sótt 5. janúar 2015 af <https://www.rsk.is/fagadilar/stadtolur-skatta/alagning-einstaklinga/nr/267>
- Barnett, V. og Hilditch, S. (1993). The Effect of an Artificial Pitch Surface on Home Team Performance in Football (Soccer). *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, 156(1), 39–50. doi:10.2307/2982859
- Becker, G. S. (1965). A Theory of the Allocation of Time. *The Economic Journal*, 75(299), 493–517. doi:10.2307/2228949
- Bosscher, V. De, Knop, P. De, Bottenburg, M. Van og Shibli, S. (2006). A Conceptual Framework for Analysing Sports Policy Factors Leading to International Sporting Success. *European Sport Management Quarterly*, 6(2), 185–215. doi:10.1080/16184740600955087
- Bougheas, S. og Downward, P. (2000). The economics of professional sports leagues: a bargaining approach. *DISCUSSION PAPERS IN ECONOMICS-UNIVERSITY OF NOTTINGHAM*, (6).
- Bruinshoofd, A. og Weel, B. ter. (2003). Manager to go? Performance dips reconsidered with evidence from Dutch football. *European Journal of Operational Research*, 148(2), 233–246. doi:10.1016/S0377-2217(02)00680-X
- Chiappori, P.-A., Levitt, S. og Groseclose, T. (2002). Testing Mixed-Strategy Equilibria When Players Are Heterogeneous: The Case of Penalty Kicks in Soccer. *The American Economic Review*, 92(4), 1138–1151.
- Crowder, M., Dixon, M., Ledford, A. og Robinson, M. (2002). Dynamic Modelling and Prediction of English Football League Matches for Betting. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 51(2), 157–168.
- Dawson, P., Dobson, S., Goddard, J. og Wilson, J. (2007). Are football referees really biased and inconsistent?: evidence on the incidence of disciplinary sanction in the

- English Premier League. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(1), 231–250. doi:10.1111/j.1467-985X.2006.00451.x
- Dixon, M. J. og Coles, S. G. (1997). Modelling Association Football Scores and Inefficiencies in the Football Betting Market. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 46(2), 265–280.
- Dobson, S. og Goddard, J. A. (2001). *The economics of football*. New York: Cambridge University Press.
- Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *The Journal of Finance*, 25(2), 383–417. doi:10.2307/2325486
- Forrest, D., Goddard, J. og Simmons, R. (2005). Odds-setters as forecasters: The case of English football. *International Journal of Forecasting*, 21(3), 551–564. doi:10.1016/j.ijforecast.2005.03.003
- Forrest, D. og Simmons, R. (2002). Outcome uncertainty and attendance demand in sport: the case of English soccer. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 51(2), 229–241. doi:10.1111/1467-9884.00314
- Frick, B., Barros, C. P. og Prinz, J. (2010). Analysing head coach dismissals in the German “Bundesliga” with a mixed logit approach. *European Journal of Operational Research*, 200(1), 151–159. doi:10.1016/j.ejor.2008.11.048
- Frick, B. og Simmons, R. (2008). The impact of managerial quality on organizational performance: evidence from German soccer. *Managerial and Decision Economics*, 29(7), 593–600. doi:10.1002/mde.1431
- García, J. og Rodríguez, P. (2002). The Determinants of Football Match Attendance Revisited Empirical Evidence From the Spanish Football League. *Journal of Sports Economics*, 3(1), 18–38. doi:10.1177/152700250200300103
- García, P. C., Castro, J. A. D. og Santos, J. M. S. (2007). The economic geography of football success: empirical evidence from european cities. *Rivista di Diritto ed Economia dello Sport*, 3(2), 67–88.
- Garðar Örn Úlfarsson. (2013, 1. ágúst). Leystir út og settir í 25 ára lántökubann. *visir.is*. Sótt 6. maí 2015 af <http://www.visir.is/leystir-ut-og-settir-i-25-ara-lantokubann/article/2013701089955>
- Goddard, J. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting*, 21(2), 331–340. doi:10.1016/j.ijforecast.2004.08.002
- Goddard, J. og Asimakopoulos, I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, 23(1), 51–66. doi:10.1002/for.877
- Graham, I. og Stott, H. (2008). Predicting bookmaker odds and efficiency for UK football. *Applied Economics*, 40(1), 99–109. doi:10.1080/00036840701728799

- Hakes, J. K. og Sauer, R. D. (2006). An economic evaluation of the Moneyball hypothesis. *The Journal of Economic Perspectives*, 20(3), 173–185.
- Harford, T. (2007). *The Undercover Economist* (Reprint edition.). New York: Random House Trade Paperbacks.
- Hill, I. D. (1974). Association Football and Statistical Inference. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 23(2), 203–208.  
doi:10.2307/2347001
- Hoffmann, R., Ging, L. C., Matheson, V. og Ramasamy, B. (2006). International women's football and gender inequality. *Applied Economics Letters*, 13(15), 999–1001.  
doi:10.1080/13504850500425774
- Hoffmann, R., Ging, L. C. og Ramasamy, B. (2002a). The Socio-Economic Determinants of International Soccer Performance. *Journal of Applied Economics*, 5(2), 253.
- Hoffmann, R., Ging, L. C. og Ramasamy, B. (2002b). Public policy and olympic success. *Applied Economics Letters*, 9(8), 545–548. doi:10.1080/13504850110102784
- Houston, R. G. og Wilson, D. P. (2002). Income, leisure and proficiency: an economic study of football performance. *Applied Economics Letters*, 9(14), 939–943.  
doi:10.1080/13504850210140150
- Karlis, D. og Ntzoufras, I. (2003). Analysis of Sports Data by Using Bivariate Poisson Models. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 52(3), 381–393.
- Karlis, D. og Ntzoufras, I. (2009). Bayesian modelling of football outcomes: using the Skellam's distribution for the goal difference. *IMA Journal of Management Mathematics*, 20(2), 133–145. doi:10.1093/imaman/dpn026
- Koning, R. H. (2000). Balance in Competition in Dutch Soccer. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 49(3), 419–431.
- Kuper, S. og Szymanski, S. (2014). *Soccernomics: Why England Loses, Why Spain, Germany, and Brazil Win, and Why the U.S., Japan, Australia--And Even Iraq--Are Destined* (3 edition.). New York, NY: Nation Books.
- Leeds, M. A. og Leeds, E. M. (2009). International Soccer Success and National Institutions. *Journal of Sports Economics*, 10(4), 369–390.  
doi:10.1177/1527002508329864
- Levitt, S. D. (2004). Why Are Gambling Markets Organised so Differently from Financial Markets? *The Economic Journal*, 114(495), 223–246.
- Levitt, S. D. og Dubner, S. J. (2006). *Freakonomics [Revised and Expanded]: A Rogue Economist Explores the Hidden Side of Everything* (Revised & Expand, Roughcut edition.). William Morrow.

- Lewis, M. (2003). *Moneyball: The Art of Winning an Unfair Game* (1st edition.). W. W. Norton & Company.
- Macmillan, P. og Smith, I. (2007). Explaining International Soccer Rankings. *Journal of Sports Economics*, 8(2), 202–213. doi:10.1177/1527002505279344
- Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica*, 36(3), 109–118. doi:10.1111/j.1467-9574.1982.tb00782.x
- Moroney, M. J. (1965). *Facts from figures* (3. útg.). Harmondsworth: Penguin.
- Mótalisti | Knattspyrnusamband Íslands. (e.d.). Sótt af <http://www.ksi.is/mot/motalisti/>
- Neale, W. C. (1964). The Peculiar Economics of Professional Sports: A Contribution to the Theory of the Firm in Sporting Competition and in Market Competition. *The Quarterly Journal of Economics*, 78(1), 1–14. doi:10.2307/1880543
- Neumann, G. R. og Tamura, R. F. (1996). *Managing Competition: The Case of the National Football League* (Working Paper No. 96-15). University of Iowa, Department of Economics. Sótt 14. janúar 2015 af <https://ideas.repec.org/p/uia/iowaec/96-15.html>
- Palacios-Huerta, I. (2003). Professionals Play Minimax. *The Review of Economic Studies*, 70(2), 395–415.
- Palacios-Huerta, I. (2004). Structural changes during a century of the world's most popular sport. *Statistical Methods and Applications*, 13(2), 241–258. doi:10.1007/s10260-004-0093-3
- Pankoff, L. D. (1968). Market Efficiency and Football Betting. *The Journal of Business*, 41(2), 203–214.
- Paola, M. D. og Scoppa, V. (2011). The Effects of Managerial Turnover: Evidence From Coach Dismissals in Italian Soccer Teams. *Journal of Sports Economics*, 1527002511402155. doi:10.1177/1527002511402155
- Pétur Blöndal. (1994, 25. maí). Fé án hirðis. *Morgunblaðið*, 32. Reykjavík.
- Procházka, D. (2012). Financial Conditions and Transparency of the Czech Professional Football Clubs. *Prague Economic Papers*, 21(4), 504–521.
- Reep, C. og Benjamin, B. (1968). Skill and Chance in Association Football. *Journal of the Royal Statistical Society. Series A (General)*, 131(4), 581–585. doi:10.2307/2343726
- Reep, C., Pollard, R. og Benjamin, B. (1971). Skill and Chance in Ball Games. *Journal of the Royal Statistical Society. Series A (General)*, 134(4), 623–629. doi:10.2307/2343657
- Rottenberg, S. (1956). The Baseball Players' Labor Market. *Journal of Political Economy*, 64(3), 242–258.

- Rue, H. og Salvesen, Ø. (2000). Prediction and Retrospective Analysis of Soccer Matches in a League. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 49(3), 399–418.
- Sauer, R. D. (1998). The Economics of Wagering Markets. *Journal of Economic Literature*, 36(4), 2021–2064.
- Sloane, P. J. (1971). SCOTTISH JOURNAL OF POLITICAL ECONOMY: THE ECONOMICS OF PROFESSIONAL FOOTBALL: THE FOOTBALL CLUB AS A UTILITY MAXIMISER\*. *Scottish Journal of Political Economy*, 18(2), 121–146.
- Spann, M. og Skiera, B. (2009). Sports forecasting: a comparison of the forecast accuracy of prediction markets, betting odds and tipsters. *Journal of Forecasting*, 28(1), 55–72. doi:10.1002/for.1091
- Szymanski, S. (2000). A Market Test for Discrimination in the English Professional Soccer Leagues. *Journal of Political Economy*, 108(3), 590–603. doi:10.1086/262130
- Szymanski, S. (2003). The Economic Design of Sporting Contests. *Journal of Economic Literature*, 41(4), 1137–1187.
- Szymanski, S. (2007). The Champions League and the Coase Theorem. *Scottish Journal of Political Economy*, 54(3), 355–373. doi:10.1111/j.1467-9485.2007.00419.x
- Szymanski, S. (2014). Convergence and Soccer. *Harvard International Review*, 36(1), 41.
- Thaler, R. H. og Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness* (1 edition.). New Haven: Yale University Press.
- Torgler, B. (2006). Historical Excellence' in Soccer World Cup Tournaments: Empirical Evidence with Data from 1930 to 2002. *Rivista di Diritto ed Economia dello Sport*, 2(1), 101–117.
- Vrooman, J. (2007). Theory of the Beautiful Game: The Unification of European Football. *Scottish Journal of Political Economy*, 54(3), 314–354. doi:10.1111/j.1467-9485.2007.00418.x