

A Study on Return to Player Percentage via Form-Based Betting in Football

Bachelor's Thesis  
Konrad Kivekäs  
Aalto University School of Business  
Information and Service Management  
Spring 2024

---

<b>Author</b>	Konrad Kivekäs	
<b>Title of thesis</b>	A Study on Return to Player Percentage via Form-Based Betting in Football	
<b>Degree</b>	Bachelor's degree	
<b>Degree programme</b>	Information and Service Management	
<b>Thesis advisor(s)</b>	Tomi Seppälä	
<b>Year of approval</b>	<b>Number of pages</b>	<b>Language</b>
2024	22	English

## Abstract

Betting on sports has been done for over two millennia. Apart from the more casual gambling aspect of the sports betting markets different strategies and prediction models have been built to find profits for investors. The possibilities that sport betting markets as a ground for investments and experimentation provide has been researched thoroughly through the last half a century. Sports betting markets have offered similar opportunities for experimentation and investing as financial markets and these markets have been seen as quite like one another. Some claim to have cracked the secrets of sports betting and some do indeed earn a profit from their investment. The reality still is that the business could never have boomed if every player could be able to make profits and the business thrives upon losing players. All is tied to the house's edge or the bookies cut. Bookies determine the odds and offer these to the players. If a player would blindly choose any odd the expected return for the investment would be 93 % of the initial investment. This figure is the return to player percentage, and it showcases how the bookies make their money.

The primary objective of this thesis is to examine if momentum or a team's current form can be used as an indicator of how they will play their next match. The study will be made by examining the return to player percentage that form based betting offers in different football leagues. Effects of momentum have been researched and momentum shifts particularly have been investigated in the field of sports betting and finance. This thesis focuses on the last 5 games a football team has played and makes a verdict on how the team has performed. By looking at the last 5 games a team has played a bet will or will not be placed depending on the momentum or more simply put the form of the team. Using historical data from 2012-2023 different teams from the English Premier League, Italian Serie A, German Bundesliga and Finnish Veikkausliiga a program is used to cycle through all teams that have played in this period. Two separate studies are made on this premise, one that examines if winning teams keep on winning and another that examines if losing teams keep on losing. The results are pooled by league and return to player percentages, average odds on placed bets, standard deviation of return to player percentages and bets placed are stored for each team. The secondary objective of this thesis is to determine are there discrepancies amongst the leagues and do some leagues offer higher return to player percentages using the two strategies described.

Apart from the quantitative analysis of the historical data a literature review is provided introducing the frameworks and studies done in the field. Examinations on strategies used in sports betting, how the market fluctuates and the rationale of gamblers are provided.

The results seem to be in favor of form-based betting. The implementations of the two strategies for the four different leagues yielded results that 7 out of 8 times the return to player percentages were higher than the return to player percentage that bookies offer (93%). The strategies do not offer a chance to make profits but form as a key performance indicator could be considered.

**Keywords** Sports betting, Momentum, Football, Return to Player

## Table of Contents

Abstract

1	Introduction .....	2
1.1	Research objectives and research questions .....	3
1.2	Scope of research.....	4
1.3	Methodology.....	6
1.4	Structure of the research .....	7
2	Theoretical background .....	8
3	Methodology .....	12
4	Results.....	15
5	Discussions and conclusions.....	19
5.1	Implications to research.....	21
5.2	Implications to practice.....	22
5.3	Limitations and future research.....	22
	References.....	24

## List of Tables

Table 1.	The winning team's key figures compiled by league during seasons 2012-2023.....	15
Table 2.	The losing team's key figures compiled by league during seasons 2012-2023.....	16

# 1 Introduction

Organized sports have been around for a long time. From the ancient civilizations of Greece and Rome there are events such as Olympic games that live up to this very day. As long as organized sports have existed there is a high possibility that betting on different outcomes has already been a commonplace practise (Matheson, 2021.). As organised sports has grown as an industry so has the gambling aspect. According to a study made by Deloitte (2019) the revenue generated by the sports industry in 2018 was estimated to be around \$71.5 billion. As organized sports have risen to be a booming business, sports betting has also become more and more prevalent each passing year. According to Statista (2023) the revenue generated was 15.88 billion USD in 2018. In 2023 the same data highlighted by Statista rounds up to 43.55 billion USD. The growth has been indeed rapid, and the trend is that online sports betting will keep on growing as a market.

As the sports betting market expands the possibility for odds efficiency problems comes to question. Sports betting markets have a history of not drawing that much attention as financial markets (Hausch & Ziemba, 1995). This enables the possibility of testing new ways to generate enhanced revenues by implementing new strategies to see if there are inefficiencies in the sports betting markets (Hausch & Ziemba, 1995). Forecasting sports results is something that has been attempted for years with a variety of tools and models. Machine learning being one of the most prominent prospects that enables better data analysis and different strategies to be found in sports betting (Bunker & Thabtah, 2019). In this thesis the focus is on the results of football games. As sports betting markets emerged there have been many attempts to predict outcomes and a high focus on scorelines. Focusing on results rather than scorelines offers a simpler way of embracing the problem of prediction analysis (Goddard & Asimakopoulos, 2004).

Does a good run of form indicate that the next match is also due to be won or is this where it all ends? Questions like this surrounding momentum have been tried to answer many times using a variety of methods. Studies have been made on how some see that the momentum is something that will persist and some will say that now has to be the moment when the streak ends (Ayton & Fischer, 2004). This thesis does not take into account feelings or bias. This thesis provides 2 alternative strategies that will be used to demonstrate momentum betting. The aim is to uncover how betting on momentum in

teams may lead to higher return to player percentages than the what the average is that betting bookies offer.

## 1.1 Research objectives and research questions

This thesis delves into the English Premier League, German Bundesliga, Italian Serie A and Finnish Veikkasuliiga and aims to find possible surpluses to returns utilizing form-based betting. The study is made by examining the return to player percentage (RTP) that two different betting strategies offer. The rationale is simple: if a team is doing well bet for their victory. If a team is doing poorly bet for their loss. In the methodology chapter I go into more depth on in what is deemed “good form” and what is deemed “bad form”. The form is calculated simply looking at the results of the last five matches a team has played. The primary research question of the thesis attempts to answer is:

- How does betting on form affect the return to player percentage?

The goal is to collect data from the last 10 years from a few different football leagues in Europe and compare the results. The aim is to see if form-based betting does or does not offer higher return to player percentages than the average return to player percentage and. As different bookies have different tools to analyse and comprise their odds there is fluctuation with how the average return to player (RTP) percentage is calculated. I'll be using an RTP % that the Finnish bookie “Veikkaus” states to be 93% and compare that to the findings the form-based betting offers. The 93% return to player percentage means that if someone were to bet on a football matches blindfolded, on the long run they would earn 93 percent of their placed bets back. The odds that will be used are from Veikkaus and Bet365. Veikkaus being the only bookie to offer odds for the Finnish Veikkkausliiga during seasons 2012-2023. Odds from Bet365 are used for the other leagues as this as a well-established bookie in the sports betting world that offers steady odds, not the best but not the worst. This naturally leaves room for optimization as each bet could be scrutinized more deeply as there is an abundance of different bookies that each offer their unique odds for a match. This optimization is too excessive for this thesis and therefore the odds offered by Veikkaus and bet365 will be used.

As comprised earlier a comparison among leagues could offer insights into possible variance and market inefficiencies. The secondary question of the thesis:

- What differences can be found among leagues in form-based betting?

As the RTP percentages, average odd and bets placed on are all calculated whilst answering the primary question this secondary question is quite a natural continuation. As the initial interest of how the return to player percentages are discussed when attempting to answer the primary research question this secondary question provides perspective into how leagues differ from one another. The differences of the leagues regarding these key numbers will be discussed.

These two questions will be the beating heart of this thesis. By investigating these from the data compiled a conclusive summary of the results can be drawn. As there are multiple different leagues used in the modelling of the strategies a more cohesive set of results may be established.

## **1.2 Scope of research**

The scope in this research is simplified quite harshly. The scope of the research is limited to scrutinizing the results on European top leagues and the Finnish topflight football league. Only the national leagues will be used in the study this excludes competitions such as the Champions League which is regarded as the most prestigious competition in Europe. This exclusion is done for the sake of simplicity as the structure of the Champions League differs greatly from the national leagues.

Bets are made on form and form only. Form is based on the results of the last 5 matches. There is no specific answer to what is seen as good or bad form. There are alternating views on how form is determined but no clearcut option. This thesis the points gathered from the last 5 matches will determine this. If the number of points gathered is 9 or higher (3 for a victory, 1 for a draw and 0 for a loss) a team is deemed to be in good form and on their next match a bet will be placed on their victory. When a team has accumulated 4 points or less in the last 5 games they will be deemed to be in bad form and a bet will be placed on their loss. I was not able to find the number of points that would be universally seen as a team playing well or a team playing poorly. Therefore, I

came up with the limits of 9 and 4 myself. These tallies are far enough from each other that there is middle ground and a clear separation. It is debatable if 9 points out of 5 games is seen as good form, but it will suffice for this thesis.

Football is a fickle sport where the emphasis whether an individual player will play affects the odds drastically. This thesis however does not take into account injuries, red cards or other missing outs of players. This exclusion has to be made even though in some cases it has drastic affects to a team winning potential if a key player is not going to play. As betting odds fluctuate and mirror these changes the need for closer player surveillance is not required in this study.

This thesis is not aiming to create a predictive betting model that is able to find returns higher than 100%, rather evaluating the possibility of form-based betting offering higher yields than the average yields. As the form-based betting is derived by one team at a time it does not consider anything other than the form that the currently selected team is in. This means that there are overlaps where teams in good form are playing against teams in good form and vice versa for the teams in bad form. This does not have any effect on the bets placed.

As mentioned earlier this thesis does not consider maximizing the returns by finding the best odd for each individual match. There is no ready-made program that shows the required data in the right format, so some concessions had to be made. The sheer number of bookies makes this such a daunting task that settling for one bookie per league will have to do. Picking out the bookie to be used in the calculations will be something of the middle of the pack. This will offer a clearer picture with lesser variance in the results.

### 1.3 Methodology

The methods used for this thesis were data-analysis and literature review. Data-analysis is highlighted in this thesis as there is an abundance of quantitative data that can be scrutinized upon. The data used in this thesis is from <https://www.football-data.co.uk/>. For some leagues this site offers football data from the season 2000/2001 onwards. As there is no data available for all the leagues from this early on, decision was made to use historical data from the last 11 years (2012-2023). This time period is long enough to showcase variance and results to draw conclusions upon as there are over 30000 matches that are looked over in the analysis per strategy.

The data collected from this site provides a CSV file that has the data already sorted by: country, league, year, date, time, home team, away team, home goals, away goals and odds for a home win, draw or an away win. This data has been formatted beforehand. In this thesis the wanted data is sorted out using the programming language Python.

There was a variety of different bet providers that make the optimization of betting more efficient and complex. A decision was made to use only the odds that bet365 provided as this company is an esteemed bet provider. For the Finnish Veikkausliiga the Finnish bookie Veikkaus is used for the odds as other bookies don't offer odds for the Veikkausliiga in the aforementioned database.

There are not any readymade programs that provide the means to sort the data. A program was created in Python to help in the data-analysis. The results from the Python program are stored into an Excel file where the final results are calculated.

The literature reviewed for this thesis offers insights into what has been researched on this field and what conclusions have been made. The literature used in this was amassed from different articles and research papers found using Google Scholar and Scopus. The literature used provides findings that have been made in the field and a structure to the thesis. Such as introducing known concepts, data and phenomenon that are known to happen in sports betting. The sports betting markets are constantly evolving and adapting as data analytics have become increasingly more effective. Development of the



market has of course outdated some of the strategies viewed in the literature review section, but this is to be expected. The main goal of the literature review is to introduce studied concepts and provide to the reader. These concepts include betting strategies, how the market has been exploited and has been attempted to find hidden profits for bettors. A brief examination on the rationale of people who bet on the sports betting markets will be provided also.

#### **1.4 Structure of the research**

The thesis will continue in the following structure. Chapter 2 guides the reader in by introducing the theoretical background and established concepts for the thesis. Historical views on what has been done in the sports betting markets in regards of strategies and findings will be introduced. What have been the motivations behind examining the sports betting markets as a viable investment option. Overview of these topics will provide the backbone to why sports betting markets have been and are utilized from the point of view of investors/players. The developments of the market will be examined. Examining these will provide the grounds for the rest of the research and help in formulating answers to future implications.

Chapter 3 delves into the methodology used in the thesis. Brief examination of the methodologies used in similar studies will be provided. The main point of this chapter will be explaining thoroughly how the data used for data-analysis was collected, formatted and analysed. An explanation on how the program created for data-analysis was used and the limitations of the program.

In chapter 4 the results derived from the data-analysis will be showcased. Presentation will offer a means to compare the findings in the data with ease. Findings and comparisons between strategies and will be noted.

Chapter 5 will finish of the thesis. The results drawn from the data-analysis will be given a verdict and conclusions will be made in this chapter. Future implications will be discussed, and the thesis concluded.

## **2 Theoretical background**

This chapter introduces and examines the theoretical background and frameworks used in this thesis. A brief overview of sports betting and what kind of betting strategies have been used in the past will be examined. Different betting strategies offer insights into what have been already tried and what findings have been made in the field. Efficiency on the sports betting market is examined and findings that have been made are presented. The studies on efficiency provide the reasoning as in to why there are profits to be made in the sports betting market. As sports betting is gambling and the baseline is that the house always wins an examination of the ethics of the sports betting industry offers perspective into the business.

Due to the rapid growth of the sports betting industry, it has drawn a lot of attention into itself. The accessibility of sports betting and the overall scale of the market has attracted a lot of players to the market (Parke & Parke, 2019). Hence the market has grown so rapidly Still the most prominent issue in the examination of sports betting markets remains the predictability of match results (Goddard & Asimakopoulos, 2004) . The market would be a lot nicer to players if the match results were not so ambiguous and unpredictable. For decades now sports betting markets have evolved and morphed into what they are today. A huge emergence on the market was the at first the era of online betting and then later online betting globally (Franck et al., 2009). The sports betting market thanks to its large volatility and flexibility is at times seen as an alternative to the stock market (Thaler & Ziemba, 1988). Sports betting strategies have been implemented and are constantly still being developed.

### **2.1 The efficiency of sports betting markets**

The efficiency or rather the lack of it in the sports betting market has been seen as an opportunity to grab onto (Goddard & Asimakopoulos, 2004). Numerous bookies using their own methods to derive the odds have proven to offer an industrious bettor to find inefficiencies in the odds. This is where profit is made and what has arduously been researched for decades.

Arbitrages being the ultimate form of discrepancies and inefficiencies with odds enabling the player to earn profit risk-free. Described as “the simultaneous purchase and sale of the same, or essentially similar, security in two different markets for advantageously different prices” (Sharpe & Alexander, 1990; cited in Shleifer & Vishny, 1997, p.35) arbitrages are the dream of sports bettors. Arbitrages have a history of being a bit under

the radar in sports betting before the introduction of online betting (Vlastakis et al., 2009) The rise of online sports betting and the growing number of bookies competing for customers have enabled arbitrages to happen. Arbitrages have been scrutinized and modelled quite thoroughly even in academic literature as of the writing of this thesis and there probably won't be similar opportunities to utilize arbitrages as thoroughly in the days to come.

## **2.2 The effect of momentum and rationale of bettors**

Momentum is something that has been researched for decades in the financial markets. Momentum and value have been determined to go hand in hand in the equity markets offering a premium for investors (Asness et al., 2013). How momentum is measured is not that simply concluded but simply put momentum is an ongoing sustained level of performance (Lehman & Hahn, 2013). Momentum is not as closely examined in European football as it is in the national basketball association (NBA) and national football league (NFL). There is an abundance of academic literature written of the NBA and NFL and among these papers are studies made on momentum. Papers on momentum can be applied to football as it shares the same elements as these other two sports.

Momentum is a double-edged knife as there are two kinds of momentum: positive momentum (increases performance) and negative momentum (decreases performance) (Lehman & Hahn, 2013). It would seem that the constant shift of momentum having an effect on player behaviour makes it at times a hard object to research thoroughly but attempts have been made. It is suggested that momentum and risk-taking go hand in hand (Lehman & Hahn, 2013). Teams and players with a "hot hand" (performing well) will attempt riskier plays that possibly lead to more victories and a gain in positive momentum. On the other hand, players, and teams with "cold feet" (performing poorly) will avoid risky plays due to a fear of a negative outcome feeding themselves into the loop of bad momentum (Lehman & Hahn, 2013).

The effect of momentum can be seen in the manner of how the original odds for a match are published as. As stated in *The Economic Journal*, Volume 114, Issue 495, April 2004, Pages 223–246, "In sports betting, bookmakers announce a price, after which adjustments are small and infrequent. Bookmakers do not play the traditional role of market makers matching buyers and sellers but, rather, take large positions with respect to the outcome of game". The bookies derive odds that have their own margin included in it and then depending on the bets placed the odds fluctuate. More bets on a home team win means that the odds will lower and the odds for a draw or an away team win will rise.

I was not able to find academic papers written on the subject of how bettors place bets and how it affects the odds. Bettors are a varied group of people and something that for over a half a century have not been seen as rational (Rosett, 1965). Understanding the odds and how they are contrived has always been somewhat unclear for the average bettor (Rosett, 1965). There have been studies made on how different people see “hot hands” and the gambler’s fallacy. Some people do tend to believe that a winning streak predicts future wins, and some people are more prone to believing that the end of the winning streak is just around the corner (Ayton & Fischer, 2004). These are big principles in gambling and offer insight into the rationale behind gambler’s decision-making. Unfortunately, due to the fluctuating of the odds, being dependent on the number of bets placed complicates the studying of how bettors place bets in online sports betting. From smaller articles that are not academically renowned the consensus is that people tend to like to bet on winners (Nitrobet 2023). This would imply that odds for winning teams tend to become lower and lower as more bets are placed making the odds less and less viable for betting.

### **2.3 Betting strategies and findings**

This chapter offers a glimpse into what kind of strategies have been used and what do they revolve around. Prediction models are at the forefront of betting strategies and that is not surprising. Prediction strategies have been studied and are studied constantly. The idea of being able to create a winning strategy based on historical data has been attempted in many forms. Machine learning offers ample opportunity to expand the betting strategies (Hubáček et al., 2019). As the betting markets grow so do the capabilities of software used in analysis creating new ways to find inefficiencies and values. Machine learning is a key element that can be used to create prediction models (Bunker & Thabtah, 2019). Machine learning is utilized in sports betting analysis and is a great asset for it (Bunker & Thabtah, 2019).

There have been made studies on betting strategies as attempts to scale out inefficiencies in the sports betting markets. As mentioned earlier sports betting markets have a likeness to financial markets and therefore offer ample opportunities for implementing different betting strategies (Moskowitz, 2021). The study made by Tobias J Moskowitz (Asset Pricing and Sports Betting, 2021) provides insight into how strategies used in trading in the more traditional financial markets can be implemented into the sports betting markets. The findings implied that the strategies implemented cannot turn a profit thanks to the bookies cut (Moskowitz, 2021). There is still much to be explored and

the idea of the sports betting markets being a laboratory for testing strategies persists (Moskowitz, 2021).

## **3 Methodology**

### **3.1 Data collection**

The data collected for this thesis was found at a sports betting analysis website <https://www.football-data.co.uk/>. This site offers an abundance of historical data for European topflight football leagues and smaller leagues such as the Finnish league that is also scrutinized in this thesis. The historical data at the website is compiled at some leagues from 2001 onwards and some leagues have data only from 2012 to this day. This is the major reason as to why this thesis only focuses on the seasons from 2012 to 2023. The websites data seems to be valid and trustworthy as there has been data collection ongoing for over two decades and there is no inherent reason to falsify the data. The data collected there is firsthand data and there are no exterior sources mentioned.

The data per league is comprised in different excel CSV files. For the Finnish league the collection of the data was simple as there is only one file that holds all the data from 2012 to 2023. The more major leagues used in this thesis are arranged seasonally so each season's data is in its own file and needs to be combined manually. The data is already formatted in a way that it is almost organized universally among the leagues. Some minor changes had to be made such as deleting a few columns from here and there but the baseline for all the files for the different leagues remain the same. The data is organized chronologically starting from 2012 and ending in 2023. The data follows the format of the header line which is: Country, League, Season, Date, Time, Home, Away, Home goals (HG), Away goals (AG), Result, Odd home win, Odd draw and Odd away win. There is a lot more data in the file such as more bookies, referees etc. but these are not essential for

After the data was collected and formatted it was filtered using the programming language Python. Python is a valid tool to be utilized in this thesis as it offers a variety of possibilities for data-analysis and presenting the results via Python is very straightforward. There is no ready-made function or tool in Python that works as a filter for the data that was needed for this thesis, so a unique program had to be self-made. There were 2 programs created, one for filtering the data for winning teams and one where the focus is on losing teams. This helps in making the programs simpler and the results easier to examine. The programs work in a very similar fashion and are essentially the same apart from the other program presenting results for the bets placed for winning teams and the other for bets placed on losing teams.

The baseline program asks for a file and the user types the filename. The program opens the found file and asks for a name for a team. Typing the name of the team makes the program read every line of the file where said team is either the home team or the away team. The program records each result as a win, draw or a loss and stores the result in a list. A function was created that enables the program to cycle through the results list and determine if the team is in “good form” or “bad form” by examining total points gained from the 5 matches for the team. This also means that the first 5 matches per season are not valid for bets. After receiving a result from the function, the program determines whether there should be a bet placed on the next match. If a bet is placed the result of the match determines whether the bet yielded in a loss or a win. Cumulative counters for bets placed, profits, losses, rounds won, rounds lost and the bet odd store the data from a betting round. After this the results list apprehends the result. The earlier mentioned form function determines whether the next match is a betting round after it determines the form of the team in question by examining the results of the teams last 5 games. This loop repeats itself until there are no more lines in the file that is being read. At the end of the program displays the team’s name, return to player percentage, total profit/loss, average odd, bet rounds won and bet rounds lost.

This is not the perfect way to read results and determine should bets be placed but it is an effective one. There were some empty lines in the files that were skipped. These lines were empty due to cancellations of matches but as they could be skipped and were so rare they were quite an insignificant problem. The program starts placing bets only after the 5<sup>th</sup> match as it needs to establish a form for the team. After the 5<sup>th</sup> match the program continuously cycles through matches and seasons without a stop.

This program required the user to type a team name each time a new team was to be run through the program. This means that each team are viewed as individuals and don’t take anything else into account other than their own results. This at times leads to conflicts of interest if teams that played against each other were deemed to be in the same form, bets were placed for both the home team winning and the away team winning.

After a team is run through the “good form” and “bad form” programs the results are written into Excel. The key figures RTP, Profit/loss, Average odd, bets won, bets lost and bets total are written for each team in a league individually. After the individual data is collected the total RTP of the league was calculated by taking a weighted average of the individual RTPs in a league. Total bets were calculated simply by adding the totals together. A leagues average odd was calculated by taking a weighted average of the average odds per team. Total profit/loss was calculated by summing individual results

together. After the calculations have been done for each team the leagues averages are calculated from each team in their respective league. This was done for each league for the 2 different strategies used. The results were compiled to the same sheet and exhibited in the results chapter. The results from each league are used in comparison in both of the strategies.

### **3.2 Literature review and the validity of the data**

There is a focus on the collected quantitative data used in this thesis. The literature review was done using Google Scholar and Scopus to scope out articles, books and other findings that offer frameworks for the thesis. There are also more unofficial sources that were found using the regular Google search engine as there are gaps in the literature review that have not been accounted for in scientific research. The raw data that was used in this thesis was collected from a website found on Google. The validity of the data is noted to be precise as the website used is a firsthand source for data collection. The website has amassed data from all kinds of different leagues and sports for over 20 years and is updated regularly. This gives the data creditability as there is no reason to offer and keep collecting and offering data in such a vast amount if the numbers would not even be correct.

The materials used in the literature review are a mix of highly cited research papers all the way to papers that are not as well known. The highly cited papers have been published in creditable and prestigious websites such as “The Journal of Finance” and these papers can be deemed as credible sources. Papers written on betting strategies are a bit ambiguous as there has not been that many highly esteemed research papers done on specifically sports betting strategies. Some of the strategies overviewed are a bit inspired in the nature but as sports betting markets attract a variety of players the more inspired and unknown research offers ideas on the rationale of the player base in sports betting.



## 4 Results

In this chapter the results and findings from the analysed data will be showcased. The key figures are presented in tables. This chapter first and foremost focuses on the results of the two betting strategies that have been used to generate the data. Comparisons between leagues and the different strategies will be looked more in-depth. The results will be compared to earlier findings in the field that were explored in chapter 2 “Theoretical frameworks and background”.

### 4.1 RTP percentage comparisons among leagues, strategies and bookies.

The tables where the results are pooled will be presented. The highest figure in each column is bolded. Primarily, the most important indicator of the results the RTP (%) is presented. Then the RTP % that the Finnish bookie Veikkaus offers is presented. Total bets placed is also presented and the average odd for a placed bet. Finally, the standard deviation of the RTP percentages of all teams in a league is calculated and shown.

#### 4.1.1 Winning teams’ key figures results

Table 1. **The winning** team’s key figures compiled by league during seasons 2012-2023. RTP(%), Veikkaus RTP (%), total bets placed, average odd on placed bets and standard deviation of RTP percentages in the league.

League	RTP (%)	Veikkaus RTP (%)	Total Bets	Average Odd	RTP Stdev
Premier League	0.957	0.93	<b>2658</b>	2.714	<b>0.306</b>
Bundesliga	0.926	0.93	2018	<b>2.753</b>	0.193
Serie A	0.944	0.93	2556	2.551	0.303
Veikkausliiga	<b>0.964</b>	0.93	1371	2.708	0.229

As can be seen from table 1 almost all of the league RTP percentages were higher than what the bookie Veikkaus claims to offer. For winning strategies displayed in the table above Veikkausliiga seems to offer the highest returns 0.964% when placing bets for

winning teams. The Serie A and the English Premier League also offer high returns of over 93% that is deemed as the baseline of return to player at least according to the Finnish bookie Veikkaus. Bundesliga is the only one that comes short of this 93% mark by a close call of being within one percent of the threshold. Total bets placed are naturally higher in the Premier League and Serie A due to there being more teams in these leagues leading into more games. Average odd when a bet has been placed is quite close among the Premier League, Serie A and Bundesliga teams. Highest odds have been offered in the Bundesliga. Veikkausliiga and the Premier League are close to on one another with the difference being only 0.006. Serie A falls from the other 3 in this regard as the average odd is 2.551 which is quite a bit lower than the second lowest Veikkausliiga 2.708. The standard deviation (sample) derived from the RTP percentages is very close with the Premier League and the Serie A. Veikkausliiga has offered a smaller standard deviation for return to player percentages and Bundesliga is in its own right the league with the lowest standard deviation regarding RTP percentages.

#### 4.1.2 Losing teams key figures results.

League	RTP (%)	Veikkaus RTP (%)	Total Bets	Average Odd	RTP Stdev
Premier League	<b>0.964</b>	0.93	<b>2152</b>	2.514	0.204
Bundesliga	0.945	0.93	1991	2.410	<b>0.376</b>
Serie A	0.936	0.93	2115	2.582	0.187
Veikkausliiga	0.960	0.93	1102	<b>2.629</b>	0.158

Table 2. **The losing** team's key figures compiled by league during seasons 2012-2023. RTP(%), Veikkaus RTP (%), total bets placed, average odd on placed bets and standard deviation of RTP percentages in the league.

As can be seen from table 2, the highest RTP (%) in the losing teams strategy is found in the English Premier League **0.964**. Finnish Veikkausliiga is a close second with **0.960**. The other remaining leagues are not quite so high but both have offered returns that are higher than the bookies **0.93**. Total bets placed once again showcase disparity among leagues due to the different sizes of leagues. The highest average odd on bets placed is in found in the Veikkausliiga **2.629**. Serie A is close to **2.6**, Premier League drops off closer to **2.5** and the Bundesliga has offered on average odds that are 0.219 lower than that of Veikkausliiga. Standard deviation (sample) regarding return to player percentage is by

far highest in the Bundesliga at **0.376**. Serie A and Premier League are close to 0.2 and Veikkausliiga is standing the lowest at **0.158**.

#### **4.1.3 RTP percentage comparisons and rationale**

The average RTP (%) offered by bookies is quite ambiguous as there are so many bookies offering the chance to place bets. The Finnish gambling restriction service “Peluri” has surmised that the RTP (%) in sports betting is around 88–92.5 % in games that have fixed odds. Games with fluctuating odds offer an RTP that is surmised to be at maximum 85%. The column “Veikkaus RTP %” are the return to player percentages that the Finnish bookie Veikkaus has stated that their odds offer to players in each game in the corresponding league. In comparison to the 93% RTP only one instance is found where the form-based betting is underachieving of the 93% RTP. Bundesliga winning teams achieved a close 0.926 % just short of the 93%. The other 7 instances all surpassed the 93 % mark. This could imply that form-based betting does indeed offer higher yields to players.

Based purely on RTP % it is hard to declare which strategy offers the better results. Different leagues offer different results, but the scope of the results is close and similar in all the leagues. Other than the Bundesliga having a difference of 0.19 in the RTP percents the other leagues in terms of RTP are within 0.1 when comparing the 2 different strategies.

#### **4.2 Comparison of total bets placed, average odds and standard deviations**

The total bets placed on average is ~300 higher for the leagues in figure 1. The total amount of bets was a positive surprise as there being thousands of bets across leagues makes the results more valid. The Bundesliga is a league with only 18 teams whilst the Serie A and the Premier League are leagues with 20 teams per league. Finland's Veikkausliiga is the smallest league with twelve teams. This is the most vital reason as to why there are less bets placed in the Finnish Veikkausliiga and German Bundesliga.

Overall, the total amount of bets placed is higher in table 1 and that may be due to the more lenient view of a team being in-form and therefore eligible for a bet. As the “good form” used in this thesis is 9 or more points out of the last 5 games it is a bit more forgiving than the established “bad form” that is 4 points out of the last 5 games. What is

meant by this is that staying eligible for a bet for a team in good form is easier than for a team in bad form. A team in good form can at maximum lose 2 games in a row and still be eligible for a bet. The losing teams can lose eligibility for betting in 2 games as if they win 2 in a row, they won't be a valid betting target for at least the next 4 games. This is probably why there are so many more bets placed in figure 1 when compared to figure 2. Alas the difference is still not anything groundbreaking but of notice.

The difference in the average odds is something that is quite interesting to examine as there is a clear implication of the odds being higher for teams to keep on winning rather than losing. Only the Serie A offers on average higher odds for losing teams to keep on losing and even this is a very close call as the difference is only **0.031**. All the other leagues offered higher average odds for winning teams to keep on winning with the most drastic difference being in the Bundesliga where the difference between the 2 strategies was **0.343**. The difference between strategies regarding the average odd was very notable for the Premier League and for Veikkausliiga. Even though there are noticeable differences between the 2 strategies the yields in return to player are strikingly similar and the impact of the average odd is something that could be further investigated.

Standard deviation of return to player percentages offer perspective on the distribution of the return to player percentages among leagues. Losing teams in Bundesliga have accumulated the most volatile standard deviation by far with a value of **0.376**. No other value in the losing table comes close to this value. The winning table offers values that surpass **0.3** but still are shadowed by the volatility of the Bundesliga losing teams RTPs. Other leagues were more volatile in the winning table than the losing table with the exception of Bundesliga where the volatility drops below **0.2**. This is the most noticeable shift in volatility between strategies and leagues.

## 5 Discussion and conclusions

### 5.1 Conclusions on the betting strategies

As a reminder the research questions that were aimed to be answered in this thesis were:

- How does betting on form affect the return to player percentage?
- What differences can be found among leagues in form-based betting?

In quite a surprising manner the results were overwhelmingly positive. There seems to be a clear indication that betting on form does offer a higher return to player percentage than what the bookies market as the average return to player percentage of 93%. Of course, to be noted is that even though the highest return to player percentage was 96.4% this still means that the strategy even in the best-case scenario on average loses money each bet. This does not mean that the thesis concludes on a sore note but instead as the research question dictates conclusion of form-based betting affecting the return to player percentage can be concluded. Both strategies used in this thesis, betting on winners to keep on winning and for losers to keep on losing on average provided higher return to player percentages than the average bookies offer. Only 1 instance was discovered where the return to player percentage lower than the fabled 93% and that was betting for winning teams to keep on winning in the Bundesliga (92.6%) This is not such a dramatic drop after all, 2018 bets were placed and still the strategy almost managed to be on par with the average bookies RTP percentage.

The differences between the two different strategies were strikingly similar. This makes it quite hard to conclude that which strategy is the superior one. If the return to player percentages were just summed together and divided the results would very narrowly shine light to betting on losers to keep on losing. The difference is quite minimal, and each league has a better and worse alternative regarding these two strategies. Both strategies had the same highest return to player percentage of 96.4 % which is remarkably high. Remarkably high in the sense that each of the bets placed were not optimized in any way. The careful selection of the most optimal bookie for each match was not utilized at all and this would seem at least to my eye like a possibility for bettors to pounce on as there could be significant profits found in this sort of optimization. Also, what is of high interest is that there was no subjective element in each bet placed and not even the form of the other team was taken to account. This means that biases, line-ups, historical data and whatever a bettor might use to gain an advantage for the next bet was not utilized at all in this thesis.

When comparing the leagues one of the first things to note is how high return to player percentages Veikkasuliiga offered as the strategies offered 0.964 and 0.96 for return to player percentages. These are remarkably high and if a similar finding could be made then the RTPs could bounce up to the fabled 100% that every player dreams of. The rather small Finnish league may offer higher numbers due to there being less games played making the data not as valid as the other leagues. Might also be that as the league is much more unknown in the mainstream as these other leagues that the odds being offered are simply not as accurately calculated as in the bigger more known leagues.

Bundesliga offers the least bang for your buck. This could be due to the great powershifts in the Bundesliga where teams tend to rise and fall more frequently than in other leagues. A study made by McKinsey & Company in 2020 showed that Bundesliga teams have a unique structure from other teams due to fan ownership of teams. This fan ownership could be why Bundesliga is a bit unique and more unpredictable than teams in other leagues but that is up to speculation.

Serie A offers steady returns to the player with not much of a difference between strategies. It would be interesting to combine these strategies into a coherent strategy and compare the results to these findings. Serie A is also famous for tactical football where the score margins are not so radical as in other leagues.

The English Premier league offers the betting possibilities with the highest number of total bets placed in both strategies. As mentioned earlier due to the number of teams in leagues it was only between the Serie A and the Premier League. What could be viewed as a positive surprise was the high return to player percentages that the English Premier league offered in both strategies. The powershifts in the league are rather high and this can be seen in the number of bets placed. Losers tend to keep losing more and winners tend to win more. Thanks to the huge differences between the clubs there are at times very large odds in matches, maximum being 41 in odds offered by Bet365.

A more carefully built program that considers the opponents form at a given time would offer an interesting comparison for this strategy.

## 5.1 Implications to research

As stated earlier the indication is that form-based betting can offer higher return to player percentages. The strategy is not fool proof as all instances did not yield results that were greater than what bookies offer on average. The form-based betting results using 4 different leagues at an interval of 20 years would indicate that betting on losing teams to keep on losing does offer extra yields to the return to player percentage. Winners tend to keep on winning can also be surmised but this strategy fell just slightly short in one instance and therefore the indication is that the strategy does not always work.

What these findings could reinforce is that the betting markets are a great possibility in trying out strategies. The sheer amount of historical data that has been collected and the data collected at all times offers many an opportunity for bettors to experiment. As for the sports betting markets offering inefficiencies for keen investors the indication would be that there are no outright inefficiencies in the market that can be found using this kind of a strategy. But even as the results never surpassed the return to player percentage of 100% there seems to be some value in utilizing form-based betting as the yields were higher than what the bookies stated to be the maximum RTP per match (93%). There are some inefficiencies that could be utilized and combined to other strategies, but the scope to which these may be researched is something that cannot be thoroughly scrutinized in this thesis.

The efficiency of the sports betting markets is something to view as an essential thing to examine whenever there is idea of a new betting strategy to be implemented. The efficiency of the markets is key to finding value where it was deemed that there isn't any and I'd say that the efficiency aspect is something that should be focused upon. The implication of this thesis suggests that there are inefficiencies that can be utilized, and the next step would be to find inefficiencies that are there to be exploited in order to create value instead of losing a little less than others. Other than the inefficiencies of the market and arbitrages to an extent the literature written on sports betting is in a sense in an odd place as there are brilliant papers written on larger concepts and phenomena but then an abundance of papers written on a lot more trivial things that offer very little to topics that just slightly deviate from these papers.

All in all I'd say that the implication that form-based betting can provide additional yields and could be added to these prior implications and findings.

## **5.2 Implications to practice**

What do the possibilities of these results entail? As the results were almost unanimously positive with the one slight deviation the form of a team could be taken into consideration before placing a bet. As the form is only one factor of a team bettors don't have to blindly always bet when the form of a team so dictates but it could be used to influence decision. As a key performance indicator combined with other factors that bettors view as important such as: injuries, red cards, and line-ups.

## **5.3 Limitations and future research**

The limitations of this study are not hard to find. The simplistic approach to betting, the number of leagues used, and the scope of the examination is something that limits the validity of the results. The simplistic approach of having 2 different strategies instead of one that combines both offers room for comparisons but leaves a bit to be desired. The possibility of including matches that end in a draw in the strategies was also pondered upon but not implemented. A more comprehensive study that could determine for each individual what should be bet on would be an interesting prospect but hard to create. Form as I see it is usually good or bad and the in between is not discussed.

Even though there is an abundance of matches used in this study more leagues could have been implemented. This would make the results more valid and offer opportunity for more cross examination between leagues and strategies. Adding a couple of more prestigious football leagues and not so well-known leagues could really show the full scope of what form-based betting offers.

The validity of some papers in the field is questionable. There have been comprehensive studies made on the different aspects such as strategies, gamblers rationale and how the bookies operate. The bookies have not been researched that thoroughly and that may be due to a lack of transparency. To offer a more versatile perspective on the world of gambling these more unknown and less valid papers and websites work fantastically, but as sources to be used in a comprehensive study the value these add is speculative.

Form as a basis for future research offers interesting prospects for study. Momentum indeed has been examined earlier in the fields of finance and sports betting but there is always room for more. As the results yielded such positive results, building a prediction model that draws emphasis on the form of a team would be interesting to see. Thankfully



sports betting markets do offer an abundance of possibilities for new ideas and findings to be made and momentum as a key performance indicator could be a vital element in studies made later on!

## References

Ayton P & Fischer I. 2004. The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory and Cognition*, Vol 32, P. 1369-1378

Blaine H. 2023. How to Bet on Losing Teams. Retrieved May 4, 2024, From:

<https://nitrobetting.eu/guides/how-to-bet-on-losing-teams/>

Bunker R & Thabtah F. 2019. A machine learning framework for sport result prediction. *Applied Computing and Informatics*, Vol 15, Issue 1, P. 27-33

Deloitte (Giorgio P, Kelley A & Borok G). 2019. The Big Gamble

Football-data used for the data-analytics: Historical data -> Leagues. Retrieved March 7, 2024, From:

<https://www.football-data.co.uk/>

Franck E, Verbeek E & Nuesch S. 2009. Inter-market Arbitrage in Sports Betting. NCER Working Paper Series Working Paper #48 from National Centre for Econometric Research

Goddard J & Asimakopoulos I. 2004. Forecasting Football Results and the Efficiency of Fixed-odds Betting. *Journal of Forecasting*, Vol 23, issue 1, P. 51-66

Gray P & Gray S. 1997. Testing Market Efficiency: Evidence From the NFL Sports Betting Market. *The Journal of Finance*, Vol 52, issue 4, P. 1725-1737

Hausch & Ziemba. 1995. Efficiency of Sports and Lottery Betting Markets. *Handbooks in Operations Research and Management Science*, Vol 9, P. 545-580

Hubáček O, Šourek, G & Železný, F. 2019. Exploiting sports-betting market using machine learning. *International Journal of Forecasting*, Vol 35, Issue 2, P. 783-796

Lehman D & Hahn J. 2013. Momentum and organizational risk taking: Evidence from the National Football League. *Management Science*, Vol 59, Issue 4, P 852-868

Levitt S. 2004. Why are Gambling Markets Organised so Differently from Financial Markets? *The Economic Journal*, Vol 114, issue 495 P. 223-246

Matheson, V. 2021. An Overview of the Economics of Sports Gambling and an Introduction to the Symposium, Vol 47, P. 1-8

McKinsey & Company (Behrenbeck K, Tacke T, Schreiber N, Schlappa M, Krüger H, Frevel N & Beiderbeck D). 2020. Bundesliga - an economic powerhouse

Moskowitz T. 2021. Asset Pricing and Sports Betting. The Journal of Finance, Vol 76, Issue 6, P. 3153-3209

Moskowitz T, Asness C & Pedersen L. 2013. Value and Momentum Everywhere. The Journal of Finance, Vol, 68 Issue 3, P. 929-985

Palautusprosentti ja kerroin vedonlyönnissä (Return to Player Percentage and Odds in Betting). 2024. Retrieved May 4, 2024, From:

<https://www.peluuri.fi/tietopankki/palautusprosentti-ja-kerroin-vedonlyonnissa>

Parke A & Parke J. 2019. Transformation of Sports Betting into a Rapid and Continuous Gambling Activity: A Grounded Theoretical Investigation of Problem Sports Betting in Online Settings. International Journal of Mental Health and Addiction Vol 17, P. 1340-1359

Return to Player Percentage Veikkaus. 2019. Retrieved May 4, 2024, From:

<https://www.veikkaus.fi/fi/yrittys#!/article/tiedotteet/pelit/2019/03-maaliskuu/22-veikkaus-nosti-pitkavedon-palautusprosentteja-suosikkisarjoissa>

Rosett R. 1965. Gambling and Rationality. Journal of Political Economy, Vol 73, Number 6

Statista: Online Sports Betting - Worldwide. (n.d.). Retrieved May 05, 2024, From: <https://www.statista.com/outlook/dmo/eservices/online-gambling/online-sports-betting/worldwide>

Thaler & Ziemba. 1988. Anomalies Parimutuel Betting Markets: Racetracks and Lotteries. Journal of Economic Perspectives Vol 2 no 2 P. 161-174

Vlastakis N, Dotsis G, Markellos N. 2009. How efficient is the European football betting market? Evidence from arbitrage and trading strategies. Journal of Forecasting, Vol 28 issue 5 P. 426-444