NHH

Norwegian School of Economics
Bergen, Autumn 2025

Small Leagues, Complex Betting, Big Opportunities?

Evaluating Efficiency in Football Betting Formats Before, During and After Covid-19

Eirik Kroka and Magnus Nygaard Sivesind

Supervisor: Jonas Andersson

Master thesis, Economics and Business Administration
Major: Business Analytics

NORWEGIAN SCHOOL OF ECONOMICS

This thesis was written as a part of the Master of Science in Economics and Business Administration at NHH. Please note that neither the institution nor the examiners are responsible – through the approval of this thesis – for the theories and methods used, or results and conclusions drawn in this work.

Acknowledgments

This thesis is written as part of our Master of Science degree in Business Analytics at the Norwegian School of Economics.

We would like to thank our supervisor, Jonas Andersson from the Department of Business and Management Science at NHH for his advice, insights and feedback during the writing of this thesis. Through meetings we have had great discussion in the development of this thesis and other intriguing topics.

We would also like to extend our deepest thanks to our friends and families for their constant support and encouragement. Their input and support have meant a lot to us and motivated us through this semester!

And lastly, a very special thank you to our amazing partners for their patience and understanding.

Thank you!

Norwegian School of Economics

Bergen, December 2025

Eirik Kroka

Magnus Nygaard Sivesind

Abstract

This thesis examines market efficiency in the European football betting markets by comparing the Home/Draw/Away markets in the top five leagues with the markets in Nordic leagues, and by analysing the Asian Handicap markets in the top five leagues. Market efficiency is evaluated within the framework of Thaler and Ziemba (1988), in which a betting market is considered weak efficient if no betting strategy can generate positive expected returns after accounting for the bookmaker's margin.

This analysis uses a dataset of more than 32,000 matches from the 2013/2014 to the 2024/2025 seasons and is divided into three subperiods: before, during, and after the Covid-19 pandemic. A logistic regression based on Winkelmann et al. (2024) is employed and extended with additional explanatory variables, to identify systematic biases. The results reveal several statistically significant biases across leagues and periods. However, these biases are largely temporary and vary substantially across markets, with a higher prevalence observed during and after the Covid-19 period.

To assess whether identified biases translate into exploitable inefficiencies, the study evaluates betting strategies using return on investment and applies a bootstrap procedure to account for sampling uncertainty. The results demonstrate that although several strategies yield positive average ROI estimates, most are not statistically robust once uncertainty is considered. In total, one strategy that produces returns that are consistently positive and statistically distinguishable from zero is identified. This opportunity occurs in the English Asian Handicap market during Covid-19.

Overall, this thesis has in general not found any evidence against weak efficiency based on the potential biases examined.

Contents

Acknowledgments ... i

Abstract ... ii

Contents ... iii

List of Figures ... v

List of Tables ... vi

1 Introduction ... 1

2 Theory ... 4

2.1 Background ... 4

2.2 Betting in Football ... 5
2.2.1 Betting Odds ... 6
2.2.2 Margin ... 6
2.2.3 Biases in Betting ... 8
2.2.4 Implied Probability ... 10
2.2.5 Formats of Betting ... 11

2.3 Market Efficiency ... 16
2.3.1 Efficient Market Hypothesis ... 16
2.3.2 Efficiency in the Betting Market ... 17

3 Previous Literature and Research Gaps ... 19

3.1 Efficiency in the Home/Draw/Away Markets ... 19

3.2 Research Gaps in the Home/Draw/Away Markets ... 23

3.3 Efficiency in the Asian Handicap Markets ... 24

3.4 Research Gaps in the Asian Handicap Markets ... 25

4 Methodology ... 27

4.1 Finding the Biases ... 28
4.1.1 Logistic Regression ... 28
4.1.2 Home/Draw/Away ... 29
4.1.3 Asian Handicap ... 31

4.2 Return on Investment ... 32

4.3 Bootstrapping ... 34

4.4 Ignorance Score ... 35

5 Data **36**

5.1 Data Sources 36

5.1.1 Data on the Home/Draw/Away Markets 37

5.1.2 Data on the Asian Handicap Markets 38

5.2 Descriptive Statistics 39

6 Analysis **44**

6.1 The Home/Draw/Away Markets: Top Five Leagues vs. Nordic Leagues 44

6.1.1 Biases 44

6.1.2 Return on Investment 50

6.1.3 Bootstrapping 52

6.2 The Asian Handicap Markets: the Top Five Leagues 54

6.2.1 Biases 54

6.2.2 Return on Investment 57

6.2.3 Bootstrapping 58

7 Discussion **60**

7.1 Comparing the Home/Draw/Away Market Efficiency 60

7.1.1 Bias 60

7.1.2 Return of Investments 62

7.1.3 Comparison with Existing Literature 64

7.2 Comparing Efficiency in the Home/Draw/Away Markets with the Asian Handicap Markets 65

7.2.1 Biases and Return of Investments 65

7.2.2 Comparison with Existing Literature 67

7.3 Limitations 67

8 Conclusion **71**

8.1 Main Findings 71

8.2 Recommendation for Further Research 72

Declaration on the use of AI tools in the work on this master's thesis **74**

References **75**

Appendences **80**

List of Figures

Figure 1: Sports betting revenue by sport type in the United Kingdom 5

Figure 2: Average margin development in the top five leagues H/D/A and Asian Handicap markets over time. 8

Figure 3: Visualisation of Asian Handicap: +1 13

Figure 4: Visualisation of Asian Handicap: -1 13

Figure 5: Visualisation of Asian Handicap: +0.5 14

Figure 6: Visualisation of Asian Handicap: -1.5 14

Figure 7: Visualisation of Asian Handicap: -0.25 15

Figure 8: Relationship between implied probabilities and realized outcomes in the H/D/A market. 41

Figure 9: Relationship between implied probabilities and realized outcomes in the English H/D/A market across subsamples. 41

Figure 10: Relationship between implied probabilities and realized outcomes in the Asian Handicap market. 42

Figure 11: Relationship between implied probabilities and realized outcomes in the English A.H. market across subsamples. 43

Figure 12: Distribution of Handicap in the English Asian Handicap market 43

Figure 13: Bootstrap results of the top five league H/D/A markets. 53

Figure 14: Bootstrap results of the Nordic H/D/A markets. 53

Figure 15: Bootstrap results for the Asian Handicap markets. 59

List of Tables

Table 1: Overview of season allocation across subsamples. 37

Table 2: Home/Draw/Away outcome summary in the top five leagues. 39

Table 3: Home/Draw/Away outcome summary in the Nordic leagues. 39

Table 4: Asian Handicap match outcome distribution in the top five leagues. 40

Table 5: Asian Handicap betting outcome distribution in the top five leagues. 40

Table 6: Adjusted pseudo-R^2 values for the top five league H/D/A markets. 46

Table 7: Adjusted pseudo-R^2 values for the Nordic H/D/A markets. 46

Table 8: Logit regression results for the English and French H/D/A markets. 46

Table 9: Logit regression results for the German and Italian H/D/A markets. 47

Table 10: Logit regression results for the Spanish H/D/A market. 48

Table 11: Logit regression results for the Danish and Finnish H/D/A markets. 49

Table 12: Logit regression results for the Norwegian and Swedish H/D/A markets. 50

Table 13: ROI results for the top five league H/D/A markets. 51

Table 14: ROI results for the Nordic H/D/A markets. 52

Table 15: Adjusted pseudo-R^2 values for the top five league Asian Handicap markets. 54

Table 16: Logit regression results for the German Asian Handicap market. 55

Table 17: Logit regression results for the Spanish Asian Handicap market. 55

Table 18: Logit regression results for the English and the French Asian Handicap markets. 56

Table 19: Logit regression results for the Italian Asian Handicap market. 57

Table 20: ROI results for the top five league Asian Handicap markets. 57

1 Introduction

Sports betting refers to wagering on the outcomes of sporting events. Since the beginning of organized competitions, gambling and sports have been closely intertwined. The practice of sports betting has been traced back to Ancient Greece, Ancient China, and Ancient Egypt, where gambling was a regular part of athletic contests (Bulski 2020, Harris et al., 2024). Over centuries, these practices have gradually transformed from local traditions into institutionalized forms of betting.

Today, sports betting represents a central segment of the global gambling economy, generating more than €70 billion in 2025 (Statista Market Insights, 2025a), which is almost twice the size of Iceland's GDP (World Bank Group, 2025). Among all sports, football holds a dominant position in the global betting (Grand View Research, 2025).

Levitt (2004) argues that sports betting markets mirror financial markets in their underlying mechanisms of pricing and information aggregation. A central question concerns how efficiency should be understood and evaluated in sports betting markets, given the considerable amount of money involved. In financial markets, Fama's (1970) definition of market efficiency serves as a foundation and Thaler and Ziemba (1988) has adapted this framework to a betting context.

Thaler and Ziemba (1988) define weak efficiency in betting markets as the absence of betting strategies that yield positive expected returns. In practice, however, behavioural biases, informational frictions, and bookmaker pricing strategies may produce deviations from efficiency, potentially creating opportunities for systematic profit. Persistent inefficiencies would therefore suggest structural weaknesses in the market.

Research on football betting markets has traditionally concentrated on the major European leagues, such as the *Premier League*, *Bundesliga*, and *La Liga*. The market for these competitions operates under conditions of high liquidity, extensive information availability, and deep analytical coverage. The existing literature largely evaluates whether the betting odds in these markets accurately reflect the true probabilities of match outcomes, with particular focus on identifying biases such as the favourite-longshot bias.

Although these studies provide valuable insight, smaller markets such as the betting markets for the Nordic leagues have not received the same attention. Similarly, research on the impact of the Covid-19 pandemic on the betting market has been limited. The pandemic introduced disruptions to the football competitions, including matches played without spectators, congested schedules, unexpected player absences, and rapidly changing conditions. These factors may have altered both the information environment, and the behaviour of bookmakers and bettors.

This thesis aims to extend the existing literature by applying more recent data and examining how efficient the Home/Draw/Away betting markets were before, during and after the Covid-19 pandemic. We will also compare the efficiency of the top five European leagues with the efficiency of the Nordic leagues. The analysis investigates whether bookmaker odds, in each period, exhibited systematic biases and if they could be exploited to yield positive returns.

Accordingly, the research question guiding this thesis is:

How did weak efficiency in the Home/Draw/Away betting markets evolve before, during, and after Covid-19, and how did this differ between the top five European leagues and the Nordic leagues?

Within the betting market, several other betting formats remain relatively under-researched. One such format is the Asian Handicap market, which differs fundamentally from traditional Home/Draw/Away betting. While the Home/Draw/Away analysis includes both the top five European leagues and Nordic leagues, the Asian Handicap analysis is limited to the top five leagues due to data availability. The following research question is addressed:

Are Asian Handicap markets less efficient than Home/Draw/Away markets in the top five European leagues?

2 Theory

2.1 Background

A key turning point the modern betting industry came in 1960 with the legalizing of betting shops in the United Kingdom by the Betting and Gaming Act of 1960 (Betting and Gaming Act, 1960). This was the beginning of the regulated betting market as we know it today. Thousands of betting shops opened within the first years, bringing sports betting into the mainstream and laying the foundation for the professionalized industry that later spread globally (Delasport, 2025). When the internet became publicly accessible in 1991, the foundation of the online betting was laid and in 1996 the first online bet was placed in a FA cup tie between Tottenham and Hereford United (Delasport, 2025).

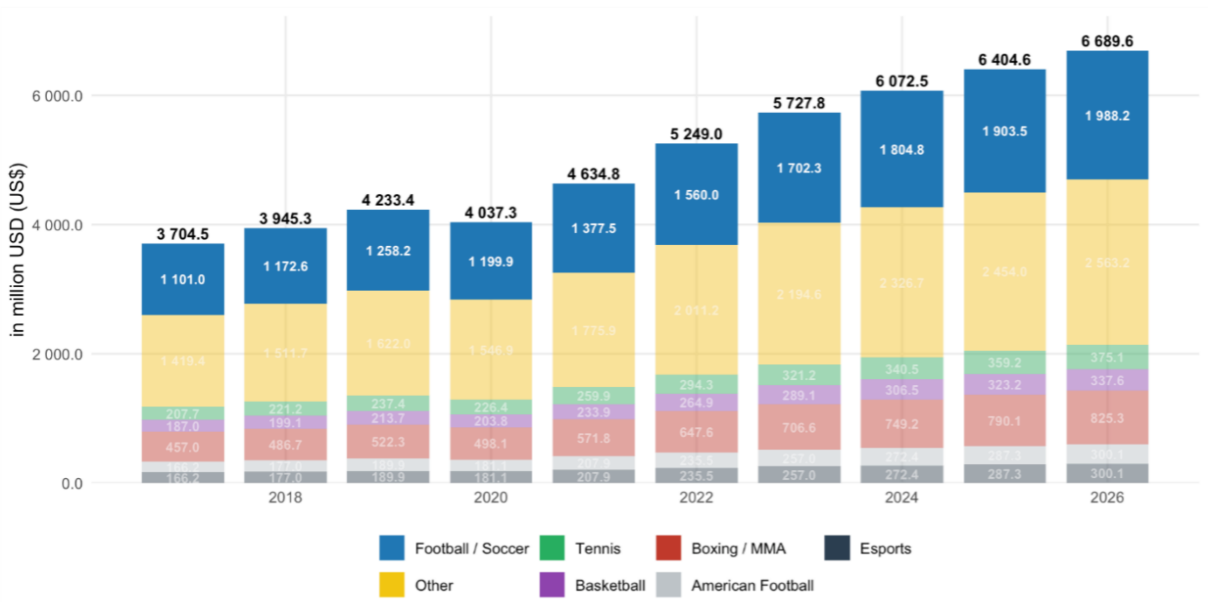
The gambling market has evolved both locally and globally, and global gambling revenue was projected to reach almost €410 billion in 2025 (Statista Market Insights, 2025a), including approximately €125 billion in the European Union (EGBA, 2025). Within this market, sports betting accounted for around 17% of total global gambling revenue in 2024, corresponding to an estimated €70 billion globally and roughly €28 billion in Europe (Statista Market Insights, 2025b).

The most established form of sports betting is bookmaker betting (Franck et al., 2010). In this form, bookmakers act as market makers by setting odds. To ensure profit, the bookmakers incorporate a margin in the odds rather than reflecting true probabilities (Levitt, 2004). Traditionally, bookmakers operated through physical shops with high control and informational asymmetry. With technological progress, bookmaking has shifted to online platforms that use data-driven models, algorithms, and real-time information to update odds dynamically. Online bookmakers offer a vast range of markets, enabling global participation, and increasing price transparency.

2.2 Betting in Football

Within the industry of sports betting, football betting occupies a central position. Football is one of the most popular sports in the world, with a huge global fan base and with matches played all year around providing continuous opportunities for wagering. Its popularity has made football one of the largest segments of the sports betting market, attracting both casual fans and professional bettors. According to the European Gambling and Betting Association (EGBA, 2025), football accounted for almost half of all online sports betting revenue in Europe, far surpassing other sports such as tennis and basketball.

Figure 1: Sports betting revenue by sport type in the United Kingdom (Statista Market Insights, 2025b).



In the United Kingdome, the football betting market has the biggest market share in the sports betting market, with about 30% of the market, as observed from Figure 1.

In the following section, we will explain key concepts such as odds, bookmaker margins, and implied probability.

2.2.1 Betting Odds

Odds are a way of representing the probability of an event occurring, for instance the probability of Liverpool beating Crystal Palace.

There are several different ways of odds notation (Cortis, 2015). This thesis will use the European style when referring to odds. The European style is also known as *decimal odds* and presented as the inverse of the probability. o_i are the odds of an outcome i, and p_i is the probability of the that outcome. The formula for the European style is:

$$o_i = \frac{1}{p_i} \tag{1}$$

(Cortis, 2015)

When betting on the decimal odds, the amount wagered multiplied by the odds is the total return one receives when winning. For example, if the probability of Manchester United beating Chelsea is 25%, then the European Style odds would be $\frac{1}{0{,}25} = 4.00$ without margin included. If €10 is wagered, and Manchester United beats Chelsea, then the payout would be €40.

2.2.2 Margin

Margins are included in the odds, which means that the odds do not reflect true probabilities but rather quasi-probabilities. For a set of values to constitute valid probabilities, certain conditions must be satisfied: each probability must lie between 0 and 1, and the sum of probabilities across a complete set of outcomes must equal 1 (i.e., 100%) (Ubøe, 2015, p. 45).

In order to make profit, the bookmakers set the odds so the sum of the quasi-probabilities, p_i^q, based on the odds, will be at 1 or more.

$$\sum_{i=0}^{n} p_i^q \geq 1 \tag{2}$$

The excess of over 1, is what's called the bookmaker's margin.

$$Margin = \sum_{i=0}^{n} p_i^q - 1 \tag{3}$$

This margin is what represent the profit for bookmakers. Conversely, it can also be expressed as the bookmaker's payout rate, which shows how much of the bets are returned to bettors on average.

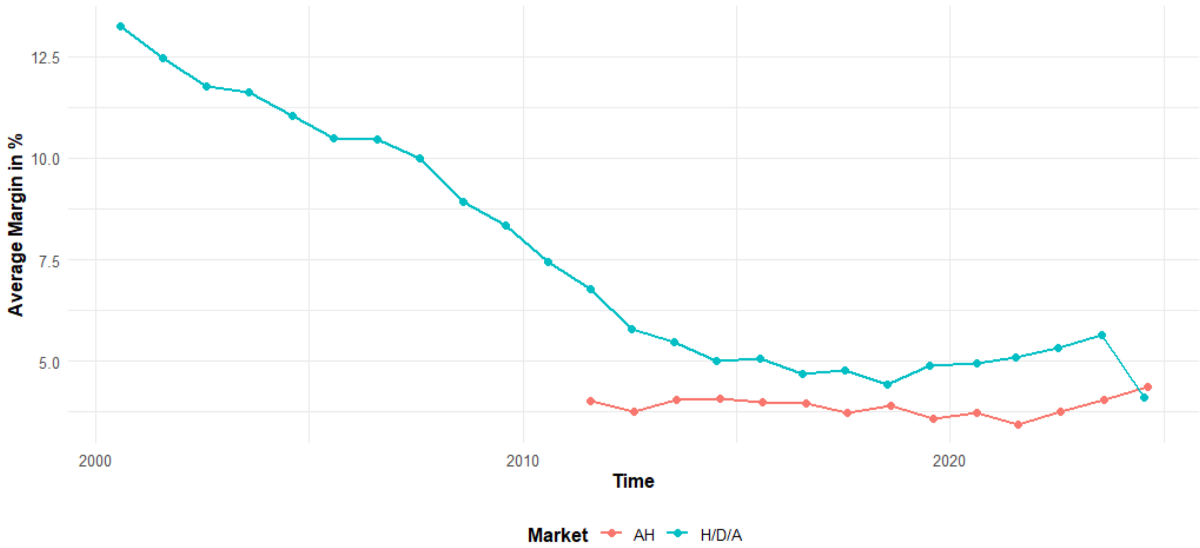
$$Payout = 1 - Margin \tag{4}$$

If the sum of the quasi-probabilities were to be less than 1, then the bookmaker would risk being arbitraged (Cortis, 2015).

The data from Football-Data[1] shows that the bookmaker's margin has evolved over time. As observed from Figure 2, the average margin in the total Home/Draw/Away market has declined from around 12,5% to below 5%. From standard microeconomic theory, an increase in the number of competitors leads to lower prices (Goolsbee et al., 2020, p. 13) Accordingly, one explanation for the decline in margins is the intensification of competition, both among traditional bookmakers and with the emergence of online betting.

[1] www.football-data.co.uk

Figure 2: Average margin development in the top five leagues H/D/A and Asian Handicap markets over time.



The margins also tend to differ substantially between top-tier leagues and lower divisions. Myrsten and Hwang (2021) argues that when bookmakers face stronger competition in the most popular leagues, where betting interest is highest, may drive margins down. They also mention another possible reason, is that top leagues benefit from more abundant and reliable information, making outcome predictions more accurate (Myrsten & Hwang, 2021). In such cases, bookmakers may not need to set margins as high to hedge against informational asymmetry.

2.2.3 Biases in Betting

In betting markets, bookmakers may incorporate systematic biases into their odds in order to increase profitability. Levitt (2004) shows that bookmakers do not simply set prices to balance their books; instead, they exploit predictable bettor biases by posting odds that deviate from the market-clearing price. Because bettors tend to favour certain outcomes, such as favourites or home teams, bookmakers can strategically shade odds away from true underlying probabilities and thereby earn higher expected profits than would be obtained from the margin alone. As a result, the odds observed by bettors often embed structural distortions, favouring some outcomes while undervaluing others.

One of the most well documented biases in betting is the favourite-longshot bias. The favourite-longshot bias describes the systematic tendency for bets on favourites to offer higher expected returns than bets on the disfavoured teams, also known as the longshots (Cain et al., 2003). Classical explanations attribute this bias either to behavioural explanations such as bettor preferences or to cognitive misperceptions (Snowberg & Wolfers, 2010). Bettor preferences indicate that some gamblers are locally risk-loving and therefore willingly accept lower expected returns in exchange for the excitement of backing longshots, while the misperceptions explanation points to where bettors overestimate small winning probabilities and therefore, bet more often on longshots (Snowberg & Wolfers, 2010).

A study of the favourite-longshot bias in tennis matches showed that the magnitude of the bias varies significantly with match characteristics, suggesting that behavioural preferences alone cannot account for the observed patterns (Lahvicka, 2013). This research supported an informational explanation as well, which implies that bookmakers strategically reduce the odds for the disfavoured outcome to protect themselves against better informed insiders, especially in matches with less information.

There is a possibility of a reversed favourite-longshot bias, where it is the longshots offering higher expected value. This thesis will reference the two possibilities as favourite bias and longshot bias.

Another commonly discussed bias is the home bias. In football betting, this bias refers to the tendency of bookmakers undervalue the probability of a home victory. Forrest and Simmons (2008) found in their research that the home teams won more often than the implied probability of the bookmakers would suggest.

A third bias known in football is the sentiment bias. The sentiment bias refers to the tendency for betting odds and wagering patters to be influenced by emotional preferences and team popularity rather than purely objective assessment of winning probabilities (Forrest & Simmons, 2008).

To illustrate the mechanisms of biases, consider a football match with two possible outcomes: a home win and an away win. Suppose the bookmaker's assessed probabilities are 60% and 40%, corresponding to odds of 1.67 and 2.50 without margin. If bettors strongly favour the home team, maintaining fair odds would expose the bookmaker to large losses if the home team wins. To manage this imbalance, the bookmaker lowers the odds on the home win and raises the odds on the away win. This discourages further betting on the popular outcome and attracts bets on the less popular one, thereby reducing risk and stabilising the bookmaker's expected margin across outcomes.

2.2.4 Implied Probability

As discussed in Section 2.2.1, the quoted odds embed a bookmaker's margin, resulting in quasi-probabilities that sum to more than 100%. Consequently, the quasi-probabilities inferred from the odds must therefore be adjusted to obtain the bookmaker's underlying probability assessments. This will be referred to as implied probabilities.

The easiest and most straightforward method is to assume that the overround of the quasi-probabilities is distributed equally among all the outcomes, as Franck et al. (2010) assumes. The implied probabilities are obtained by taking the inverse odds

$$p_i^q = \frac{1}{o_i} \tag{5}$$

and linearly normalizing them so that each outcome shares the bookmaker's margin evenly. This yields the implied probability:

$$p_i^{imp} = \frac{p_i^q}{\sum_{i=0}^{n} p_i^q} \tag{6}$$

(Franck et al., 2010)

While the equal distribution approach provides a straightforward way to obtain the implied probability, it is important to acknowledge that it may not fully capture such structural pricing patterns in betting markets.

A study by Myrsten and Hwang (2021), points out that bookmakers do not necessarily distribute their margin equally across outcomes. Instead, they may impose higher margins on longshots and lower margins on favourites, meaning that the total overround varies with the odds level. This asymmetric margin structure can itself generate a pattern resembling the favourite bias.

However, Myrsten and Hwang (2021) underline that such a pattern is challenging to establish empirically, given that the bookmaker's true margin allocation cannot be inferred from publicly available data. This thesis will therefore move forward with the assumption that the margins are equally distributed between the outcomes.

2.2.5 Formats of Betting

In football betting there exist several formats of betting. This thesis focuses on the Home/Draw/Away betting and Asian Handicap betting.

2.2.5.1 Home/Draw/Away Betting

The Home/Draw/Away betting, also referred to as H/D/A, is the most common way of betting on a football match. The term "Home/Draw/Away" refers to the three possible outcomes:

Home: *The home team wins*

Draw: *The match ends in a draw*

Away: *The away team wins*

In the Home/Draw/Away betting market, the bettor predicts the result of a match after 90 minutes including added injury time[2]. Outcomes from extra time and penalty shoot outs are normally excluded from this market, if not explicitly stated by the bookmaker.

2.2.5.2 Asian Handicap Betting

Asian handicap is a form of betting designed to eliminate the outcome of a draw in a match. The bookmaker assigns a handicap to one of the teams, which is added to the final score before determining the result for betting purposes. The handicap reflects the perceived difference in team strength and is intended to make both sides of the market equally attractive to the bettors. Asian handicap lines can take form of:

- whole goals, such as -1 or +1
- half goals, such as -0.5 or +0.5
- quarter goals, such as -0.75 or +0.25

(Smarkets, 2025)

In addition to these, some handicap lines can be 0. In these cases, the format acts as a Home/Draw/Away format, besides draws being refunded.

In Asian Handicap notation, the minus sign denotes the favoured team, while the plus sign designates disfavoured team (Play the Percentage, 2025). The favourite team is effectively penalised by the handicap, meaning they must overcome a virtual goal deficit for the bet to succeed. Conversely, the disfavoured team receives a virtual advantage that compensates for their lower expected performance.

A whole-goal handicap assigns an integer goal adjustment to one of the team's final scores. For instance, if one team wins 1-0 and has as +1 handicap, then the betting result

[2] A football match consists of two halves of 45 minutes each. At the end of the halves, the referee has the authority to add time as a compensation for stoppage of play, also known as injury time or added time.

would be 2-0. If the team has a -1 handicap, then the betting result would be 0-0. In the last example, the betting result would be a draw, in which case the bet is refunded, as illustrated in Figure 3 and Figure 4.

Figure 3: Visualisation of Asian Handicap: +1

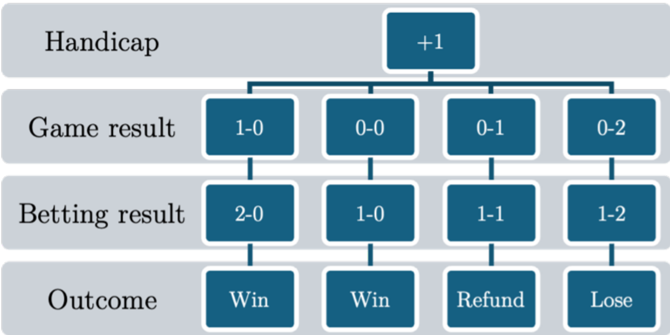
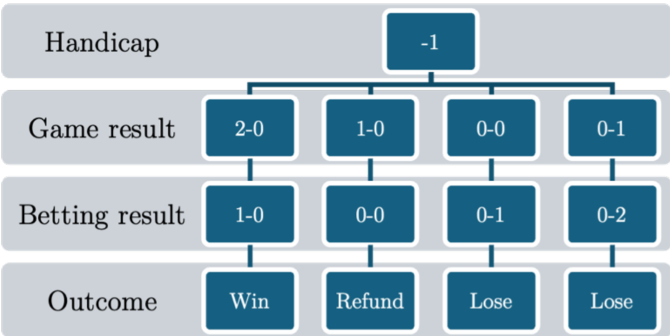


Figure 4: Visualisation of Asian Handicap: -1



A half-goal handicap applies a fractional adjustment of 0.5 goal, for instance -0.5 or +1.5. This effectively eliminates the possibility of a draw outcome, since it is not possible to score half a goal. For example, a team has a betting line at +0.5 and loses 0-1, the betting result would be 0.5-1 which would not affect the result as a loss. If the team has a -0.5 line and wins 1-0, their betting result would be 0.5-0, which is still a win.

If the line is -1.5, then in order to win the team must win by at least two goals, as illustrated in Figure 6.

Figure 5: Visualisation of Asian Handicap: +0.5

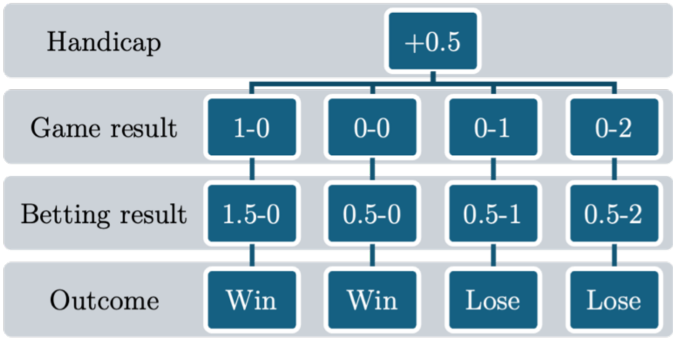
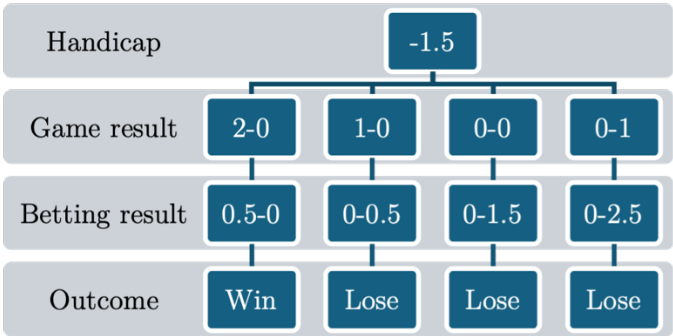


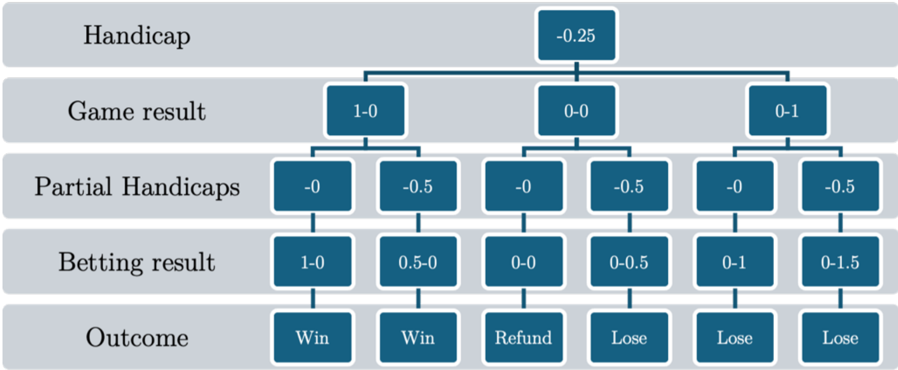
Figure 6: Visualisation of Asian Handicap: -1.5



A quarter-goal handicap represents a more nuanced way of betting. This type of handicap is also referred as a split handicap, due to the fact that it introduces the possibility of partial wins and losses (Play the Percentage, 2025). The bettor's bet is divided equally between two adjacent lines, one at the nearest whole or half a goal above, and on below the quoted handicap (Hegarty & Whelan, 2025).

For example, a -0.25 handicap divides the wager into two parts: half a wager on 0-handicap and the other half on -0.5. If the match ends with the favoured team winning, for instance 1-0, then both components win. If the match ends in a draw, 0-0, then the 0-handicap portion of the bet is refunded while the -0.5 portion loses. This results in a half loss. If the match ends 0-1, both bets lose.

Figure 7: Visualisation of Asian Handicap: -0.25



Similarly, a -0.75 handicap splits between -0.5 and -1.0. A one-goal victory yields half a win, since the -0.5 portion wins and the -1.0 is refunded. A two-goal victory would produce a full win.

On the other hand, if the line is +0.25, the wager is divided with half of the bet on 0-handicap and the other half on +0.5. In this case, both bets wins if the ends in a draw or the disfavoured team wins. If the game ends in a draw, then the 0-handicap portion would be refunded, but the +0.5 portion would win. For instance, if one were to wager £100 on Chelsea +0.25 at odds of 1.90 and the match ended 0–0, half of the wager (£50) would be refunded, while the remaining £50 would be settled at odds of 1.90, resulting in a total return of £95.

2.3 Market Efficiency

This section will introduce the theoretical framework of the thesis. First we will introduce the Efficient Market Hypothesis as formulated by Fama (1970), highlighting its key implications for price formation and returns in financial markets. Then we will explain the framework proposed by Thaler and Ziemba (1988) which provide the basis for the empirical analysis conducted in this thesis.

2.3.1 Efficient Market Hypothesis

In the paper "Efficient Capital Markets: A review of Theory and Empirical Work", Eugene Fama provided a formulation of the efficient market. He defined an it as a market in which prices "fully reflect available information" (Fama, 1970). Meaning that any new information is immediately incorporated into prices once it becomes known. Because new information arrives randomly and unpredictably, price changes should also follow a random pattern.

Fama illustrated this relationship through what is called the fair-game model. In its simplest form, the expected future price of an asset P_{t+1} equals its current price P_t adjusted for the normal, risk-based expected return given all information, ϕ_t at time t:

$$E(P_{t+1}|\ \phi_t) = \Big(1 + E(r_{t+1}|\phi_t)\Big)P_t \tag{7}$$

Here, r_{t+1} is the return between periods t and $t+1$. The key implication is that, after accounting for risk, the expected change in price based on known information is zero. Any actual deviation from this expectation is a random shock of new information. In other words, markets form rational expectations about value, and subsequent price movements simply reflect surprises. Intuitively, this can be expressed using returns:

$$r_{t+1} = E(r_{t+1}|\phi_t) + \epsilon_{t+1} \tag{8}$$

where ϵ_{t+1} is a random, unpredictable component with an expected value of zero. The Efficient Market Hypothesis therefore implies that tomorrow's prices equal today's prices plus an unpredictable innovation, known as a martingale process.

Fama also identified three forms of efficiency, depending on how much information is assumed to be reflected in prices: weak form, semi-strong form and strong form efficiency (Fama, 1970). In weak form efficiency the information set contains all past prices:

$$\phi^w = \{Past\ prices\} \tag{9}$$

Semi-strong form efficiency includes all public information

$$\phi^{ss} = \{All\ public\ information\} \tag{10}$$

, and strong form all information, both public and private:

$$\phi^s = \{All\ information\} \tag{11}$$

Despite the differences in information assumptions, all three forms imply that the expected return conditional on the information set is equal to the unconditional expected return:

$$E(r_{t+1}|\phi_t) = E(r_{t+1}) \tag{12}$$

2.3.2 Efficiency in the Betting Market

Building on the broader concept of market efficiency introduced by Fama (1970), Thaler and Ziemba (1988) and Kuypers (2000) propose two different frameworks tailored to betting markets. Kuypers (2000) adapts Fama's classification of weak form, semi-strong form and strong form efficiency to betting markets by interpreting odds as asset prices.

Thaler and Ziemba's (1988) divides efficiency into two classifications: weak- and strong efficiency. Weak efficiency requires that no strategy based solely on publicly observable information, such as historical results, posted odds, or simple performance indicators, yields a positive expected return on investment (ROI).

$$E(ROI) \leq 0 \tag{13}$$

Systematic biases may exist, but they do not constitute inefficiency unless they can be exploited to earn profitable returns.

Strong efficiency imposes a stricter condition, stating that every wager has an expected value equal to $(1-t)$ times the bet, where t represents the bookmaker's margin or transaction cost. Under this assumption, no strategy can earn higher returns relative to random betting.

While Thaler and Ziemba (1988) use Fama as a baseline for their framework, they adopt a different terminology. What they categorize as weak efficiency is closer to Fama's semi-strong form efficiency.

This thesis follows Thaler and Ziemba's (1988) weak efficiency return-based framework, as it aligns directly with the empirical analysis of betting strategies, return on investment, conducted in this study.

3 Previous Literature and Research Gaps

3.1 Efficiency in the Home/Draw/Away Markets

Research on the Home/Draw/Away betting market presents a complex and conflicting picture of the market efficiency. Previous studies have examined a range of topics, including biases, bookmaker strategies, and market structures, employing a variety of methodological approaches. This thesis wants to add to the existing literature by including updated and new data, while looking at other countries.

Pope and Peel (1989) laid the foundation for the research field by investigating whether the fixed odds football betting market in the United Kingdom operated efficiently, given that bookmakers set divergent odds several days before matches. They found that the posted odds for draw outcomes, appeared to not fully reflect all available information, while outcomes for home and away were unbiased.

However, Kuypers (2000) found no systematic difference between implied probabilities and actual probabilities, contradicting what Pope and Peel (1989) found. He also tested for semi-strong form efficiency, by testing if a model combining publicly available performance data with odds could consistently generate positive abnormal returns, even out of sample. This demonstrated that the fixed-odds market failed Fama's semi-strong form efficiency criterion (Kuypers, 2000).

While the studies discussed above have only focused on the betting market in the United Kingdom, Vlastakis et al. (2009) broaden this perspective by examining fixed-odds football betting across 26 European countries. Using a Poisson regression, they documented that betting on the away favourite has a significantly higher average return, and in some cases positive returns (Vlastakis et al., 2009). These findings are consistent with violations of weak efficiency as defined Thaler and Ziemba's (1988).

In contrast to Vlastakis et al.'s (2009) findings of an away favourite bias, Albertsen (2017) reported largely contrasting patterns across European leagues. He found a home bias in Spain, Germany, and France, with home bets outperforming draw and away bets. This was most pronounced in Spain. Italy deviates from this pattern, as draw bets perform best. England was the exception, in where Albertsen identified an away bias consistent with Vlastakis et al (Albertsen, 2017; Vlastakis et al., 2009).

Albertsen (2017) also documented a favourite- and a longshot bias, with bets on very low and very high implied probabilities outperforming those in the mid-probability range. This pattern was stable across seasons and was especially pronounced in Spain. Finally, Albertsen (2017) found that statistical models can generate profits, violating weak efficiency.

Angelini and De Angelis (2019) researched the efficiency of online football betting markets, and their results showed clear evidence of the favourite bias across most leagues. Using the average odds offered by all bookmakers, they found that forecast errors tend to increase with the implied probability, indicating that favourites were systematically undervalued and longshots overvalued. Although the bias was generally too small to generate profitable opportunities for bettors, Angelini and De Angelis (2019) identified certain leagues in which betting on favourites yielded positive expected returns.

Based on the biases, the authors constructed data-driven betting strategies that selected wagers only within the probability ranges where the favourite bias created positive expected value. Tested both in-sample and out-of-sample, these strategies generated consistent profits, around 2% in-sample for all three leagues, and positive out-of-sample returns for Greece and Spain (Angelini & De Angelis, 2019).

Hegarty and Whelan (2024a) also found evidence of favourite bias only appearing in some of the European leagues. They showed that the leagues with lower profile and lower betting volumes exhibit a statistically significant favourite bias, while the more prominent leagues, such as the English Premier League, the German Bundesliga and Spain's La Liga, did not show strong evidence of this pattern.

Favourite bias is not the only bias documented in previous literature. Forrest and Simmons (2008) found that markets may also be influenced by sentiment bias, meaning that odds reflect the relative popularity of the teams rather than purely objective win probabilities. In their study of the Spanish football betting market, they found that matches with large differences in supporter base, systematically yield more favourable odds for the more popular team (Forrest & Simmons, 2008). They found that bettors could reduce losses by exploiting the observed biases, but they did not achieve sustained positive returns. Therefore, the authors concluded that the market was not strong efficiency, after Thaler and Ziemba's (1988) definition.

Another bias examined by Forrest and Simmons (2008) is the home bias, where their results indicated that bookmakers tend to undervalue the probability of a home victory. Their descriptive statistics showed that home teams in La Liga won 48% of the matches in their sample, while bookmakers priced home wins at an implied probability of only 45.6%. Their model confirmed the *Home*-variable as positive and statistically significant, controlling for other factors. This implied that bookmakers still undervalued home teams. In contrast, when the authors replicated the analysis for the Scottish Premier League, the home coefficient remained positive but was no longer statistically significant, suggesting that home bias could be present in some markets but not universal.

Myrsten and Hwang (2021) analysed almost 200 000 matches from 41 countries to evaluate whether the Home/Draw/Away market exhibits systematic mispricing. By using OLS, LOWESS and quantile regression, they found no evidence of a longshot bias and only a very small favourite bias, indicating that implied probabilities closely align with realised frequencies across most of the probability distribution (Myrsten & Hwang, 2021). Importantly, they showed that the degree of efficiency varies substantially across leagues. The top European leagues, appear highly efficient, whereas lower-tier leagues display larger deviations, particularly for draws. These deviations, however, remain modest and do not translate into profitable betting opportunities, and does not violate Thaler and Ziemba's (1988) weak efficiency.

Winkelmann et al. (2024) examined odds from the top five leagues and employ Monte Carlo simulations to assess whether the frequently reported favourite bias and other biases truly reflect market inefficiency. Their results revealed that many patterns interpreted as evidence of bias in earlier studies can emerge purely by chance when analyses rely on short time periods or small samples. When evaluating the full dataset, they found only limited and inconsistent deviations from efficiency, with apparent biases occurring sporadically rather than persistently across seasons (Winkelmann et al., 2024). Crucially, their simulation framework demonstrated that an efficient market could generate return patterns resembling favourite bias in isolated years, implying that earlier literature may have overinterpreted random variation as systematic mispricing. Overall, their findings supported the view that the European Home/Draw/Away market violates weak efficiency. However, these inefficiencies are short-lived and occur unsystematically (Winkelmann et al., 2024). For the profitable opportunities they found, the suggested to use a bootstrap procedure to examine the robustness of the returns.

Deutscher et al. (2018) adopted a within-season perspective that examined how temporary deviations from efficiency arise and subsequently disappear as information accumulates.

They examined 14 seasons of Bundesliga odds and showed that inefficiencies can emerge temporarily within a season, especially in the early rounds when uncertainty about team strength is highest (Deutscher et al., 2018). They found that bookmakers systematically underestimate the strength of newly promoted teams at the beginning of the season, producing a window in which bettors who back these sides can achieve positive returns (Deutscher et al., 2018). However, as the season progresses and more information becomes available, this advantage disappears, and pricing converges toward efficiency. In line with the conclusions of Winkelmann et al. (2024), their results suggested that apparent biases in the Home/Draw/Away market are not stable structural features but instead reflect short-term informational frictions that dissipate once uncertainty declines.

3.2 Research Gaps in the Home/Draw/Away Markets

Despite the extensive research on efficiency in the Home/Draw/Away football betting market, there are gaps that justify further investigation.

First, the geographical scope of the existing literature is uneven. While a number of studies focus on the United Kingdom (Kuypers, 2000; Pope & Peel, 1989) and the major European leagues (Albertsen, 2017; Winkelmann et al., 2024), the Nordic countries have not been investigated[3]. These leagues differ structurally from the major European competitions in terms of season timing, competitive balance, betting volume and market liquidity, all of which may influence the efficiency. To date, no study provides a systematic cross-country comparison including the Nordic leagues, leaving an important research gap in the European betting market literature.

Second, the literature is fragmented with respect to time periods. Many of the classic studies rely on data from the 1990s and early 2000s, when bookmaker margins, market structure and information environments were substantially different from today. More recent contributions, such as Hegarty and Whelan (2024a), provide updated evidence, but only for selected leagues and relatively short periods. As a result, there is limited understanding of how efficiency has evolved in more recent years.

Third, existing studies have not systematically examined how the pandemic period affected betting market efficiency. The pandemic fundamentally altered match conditions scheduling density, team performance, and potentially bettor behaviour. No study to date

[3] Vlastakis et al. (2009) do not specify which European leagues are investigated.

provides a comprehensive comparison of market efficiency before, during and after Covid-19 across countries. This represents a research gap that may reveal whether structural inefficiencies are amplified under conditions of heightened uncertainty in the betting markets.

Lastly, although recent studies such as Winkelmann et al. (2024) emphasise the importance of accounting for sampling variation and recommend simulation-based approaches, the majority of the literature relied on classical statistical inference. Very few studies employ bootstrap procedures to formally assess the robustness of estimated returns. Bootstrap techniques provide a powerful tool for distinguishing genuine inefficiencies from patterns that could plausibly arise by chance, addressing exactly the methodological concern raised by Winkelmann et al. (2024)

3.3 Efficiency in the Asian Handicap Markets

The Asian Handicap market is newer than the Home/Draw/Away market, and therefore the research is more limited.

Constantinou (2022) presented a model specifically developed for prediction of the Asian Handicap betting market. He found that the Asian Handicap market contained inefficiencies that could be exploited. The model consistently identified betting opportunities on favourites, and he also found that Asian Handicap returns were less volatile than the Home/Draw/Away market. However, this lower volatility disappears once wagers are scaled to equalize expected profits across markets.

In contrast to Constantinou (2022) research, Hegarty and Whelan (2025) approached the question of market efficiency from a different angle by analysing whether the odds themselves provided unbiased forecasts of match outcomes. Instead of testing a model against the market, they assessed the internal efficiency of the betting lines, examining whether implied probabilities accurately reflected realised results.

Their findings showed that the Asian Handicap market behaves largely as an efficient market, with implied probabilities closely matching realised outcomes and no evidence of favourite- or longshot bias (Hegarty & Whelan, 2025).

They also conducted another study, analysing the market efficiency from a different analytical perspective by focusing on the return structure of the Asian Handicap market (Hegarty & Whelan, 2024b)[4]. Instead of evaluating predictive accuracy, they investigated whether bettors faced equal expected losses across the different handicap lines offered on the same match. Effectively they tested whether the Asian Handicap markets were in violation of Thaler and Ziemba's (1988) strong efficiency. Their results showed clear and systematic deviations from this benchmark. Bets without the possibility of refunds, such as half-goal handicaps, consistently generated the highest realised losses. Whereas bets with full refund potential, such as integer handicaps, yield significantly lower losses, with quarter-goal handicaps falling in between (Hegarty & Whelan, 2024b).

3.4 Research Gaps in the Asian Handicap Markets

Although recent studies by Constantinou (2022) and Hegarty and Whelan (2024b; 2025) have begun to address this gap, there are some limitations that persist.

Firstly, as already mentioned, the overall volume of research in the Asian Handicap market is limited. While the Home/Draw/Away market had been analysed for more than three decades, the Asian Handicap market is comparatively new and less documented. Even though, Hegarty and Whelan (2025) has compared the two markets, there are

[4] While this study was published earlier than the other two article referred in this thesis, the two articles are referred to as working papers in *Returns on complex bets: evidence from Asian Handicap betting on soccer* (2024)

opportunities to test other methods and apply them to test the Asian Handicap betting form.

Similar to the Home/Draw/Away market, the impact of the Covid-19 on the Asian Handicap market have not been extensively examined. This lack of empirical evidence presents an interesting avenue for further research, particularly in light of potential differences in how the pandemic affected the Asian Handicap markets compared the Home/Draw/Away markets.

4 Methodology

In this section, we outline the empirical framework used in this thesis. This thesis explores a range of biases documented in the betting market literature by applying a logistic regression framework introduced by Winkelmann et al. (2024) extended with additional explanatory variables.

We further analyse the return on investment to see if any economic gains can be achieved by exploiting potential biases and violate weak efficiency. To estimate empirical confidence intervals and assess the stability of the ROI results, we also apply a bootstrap procedure. Taken together, ROI and bootstrap provide a formal test of market efficiency in the sense of Thaler and Ziemba (1988).

The data is divided into three time periods: before Covid-19, during Covid-19 and after Covid-19. The decision to split the data specifically into periods based on Covid-19 is motivated by the substantial change caused by the pandemic. Covid-19 led to major disruptions in football, most notably restrictions or bans on spectators, which research has shown to affect the home performance (Fischer & Haucap, 2022). Consistent with this, Lyhagen (2025) found that the absence of spectators significantly affected home performance in the English Championship, while no corresponding effect is observed in the Premier League.

The data is grouped into larger sample sizes, which helps reduce the likelihood of false positives, by reducing the number of regression. Winkelmann et al. (2024) found that when conducting season-by-season analyses, there was a high probability of observing at least one significant effect due solely to random variation.

4.1 Finding the Biases

This thesis uses logistic regression to identify systematic biases in the bookmaker odds. Under the framework of Thaler and Ziemba (1988), such biases do not imply weak market inefficiency unless they can be exploited to generate positive expected returns.

4.1.1 Logistic Regression

Logistic regression is a modelling technique used to analyse the relationship between one or more explanatory variables and a binary dependent variable (Wooldridge, 2020). This makes the model particularly suitable when the dependent variable captures the occurrence of a discrete event, such as if a bet is won or lost.

Let y denote the binary outcome variable,

$$y \in \{0,1\} \tag{14}$$

Logistic regression models the conditional probability that $y = 1$ given a vector of explanatory variables $\boldsymbol{x}$ and a vector of coefficients $\boldsymbol{\beta}$, as

$$P(y = 1|\boldsymbol{x}) = G(\beta_0 + \beta_1 x_1 + \beta_2 x_2 \dots + \beta_i x_i) = G(\beta_0 + \boldsymbol{x^T \beta}) \tag{15}$$

, where $G(\cdot)$ denotes the cumulative distribution function of the logistic distribution:

$$G(z) = \frac{e^z}{1 + e^z} \tag{16}$$

This form ensures that predicted values are strictly bounded between zero and one for all real values of z,

$$0 < G(z) < 1 \tag{17}$$

(Wooldridge, 2020).

In this thesis, the estimated coefficients indicate whether systematic biases are present in bookmaker odds, holding all other variables constant. Statistically significant coefficients provide evidence of biases, but not violation of weak efficiency. In this logistic regression model the significance levels are denoted as *** for 1%, ** for 5%, and * for 10%.

The sign of the coefficient determines the direction of the bias, indicating whether outcomes are for instance associated with a favourite or a longshot bias. A larger coefficient magnitude reflects a stronger effect of the bias on the probability of the dependent variable through the logistic distribution function, $G(\cdot)$.

A measure of goodness of fit for logistic regression models is Adjusted McFadden Pseudo-R^2. Following Long (1997, p. 104), the Adjusted McFadden Pseudo-R^2 penalizes the log-likelihood of the fitted model for the number of estimated parameters and is defined as:

$$\overline{R}^2 = 1 - \frac{\ln \hat{L}(M_a) - K}{\ln \hat{L}(M_{null})} \tag{18}$$

, where $\ln \hat{L}(M_a)$ is the log-likelihood of the fitted model. $\ln \hat{L}(M_{null})$ is the log-likelihood of the null model, and K denotes the number of estimated parameters. This method is used to measure the model fit in our analysis.

The next two sections will discuss how the logistic regression is implemented in this research and how the analysis of the two markets differs.

4.1.2 Home/Draw/Away

To operationalise this, the logistic regression is performed at team level, where each observation represents one of the two clubs participating in a match. The dependent variable Won_i takes the value 1 if team i wins the match, implying that the bet yields an economic gain, and 0 otherwise. In context of the Home/Draw/Away market, this implies that both losses and draws are coded as non-wins. This allows match outcomes to be modelled using binary logistic regression, making the results directly comparable to the Asian Handicap market, which offers only two betting options.

Since both teams from the same match enter the dataset separately, we cluster standard errors at match level to account for within-match correlation, ensuring consistent inference.

Our baseline specification is given by:

$$P(Won_i = 1|X_{\boldsymbol{i}}) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 Home_i + \beta_3 Favourite_i + \beta_4 Promoted_i + \beta_5 Success + \beta_6 DiffAttend_i) \tag{19}$$

Consistent with the betting market literature, implied probabilities and the included control variables are assumed to summarise the information reflected in bookmaker odds.

The variable $ImProb_i$ represents the bookmakers implied probability of a win for team, as inferred from the odds. In line with standard betting market research, implied probabilities serve as a measure of the market's expectation of match outcomes. A higher implied probability should therefore be associated with higher likelihood of winning.

$Home_i$ is a dummy variable, which takes the value 1 if the team $\boldsymbol{i}$ plays at home, and 0 if team $\boldsymbol{i}$ plays away. This variable is used to document potential home bias.

Another dummy variable included is $Favourite_i$, which has the value 1 if the team $\boldsymbol{i}$ has the lowest odds, in other words the highest probability of winning, and 0 otherwise. This variable accounts for the favourite bias, and if the coefficient is negative, a longshot bias.

Winkelmann et al. (2024) also included the variable $DiffAttend_i$. This captures the difference in home attendance between the two competing clubs based on last year's average. Due to it being a large number, the output is divided by 1,000. For instance, if Barcelona had an average attendance on their home games of 70,000 last year and Real Madrid had an average of 50,000, then $DiffAttend_{Barcelona}$ would be 20, and for Real Madrid it would be -20. This variable aims to proxy club sentiment and popularity by comparing the fan bases of the clubs involved in the match.

A key challenge for this variable arises for the seasons affected by the Covid-19, where the attendance was not allowed and only partially allowed. In this study, the attendance during the Covid-years is substituted with the attendance from the most recent season with normal conditions. The implications of this will be further discussed in Section 7.3.

In addition to the variables introduced by Winkelmann et al. (2024), this study includes the dummy variable $Promoted_i$ which indicate if $team_i$ was promoted from a league below in the previous season. This variable explores if there is a potential bias against newly promoted. This study also includes a variable $Success_i$, which is used to capture team momentum. The variable is constructed following the same approach as in Franke (2020). The value of this variable is:

$$Number\ of\ wins\ for\ team_i\ in\ the\ last\ 4\ league\ match -$$
$$Number\ of\ wins\ for\ the\ opposing\ team\ over\ the\ same\ period$$

For teams that competed in the top division in the previous season, match outcomes from that season remain relevant for $Success_i$. On the other hand, promoted teams, enter the new season without recent top-league match history and therefore begin with zero recorded wins, which is the amount of wins they have in the top division. As for the implication with $DiffAttend_i$, the implication with $Success_i$ will be discussed in Section 7.3.

4.1.3 Asian Handicap

The Asian Handicap specification adapts the empirical framework to reflect the payoff structure of handicap betting rather than the binary win-loss outcome used for the Home/Draw/Away model. In this setting, the dependent variable $AHGain_i$ captures whether the bet yields a positive economic return. The variable is coded as 1 when the wager results in an economic gain, and 0 otherwise. An economic gain is defined as either a full win or a half-win under the Asian Handicap rules, meaning that any outcome yielding a positive monetary return is coded as success in the dependent variable.

To incorporate the structural asymmetry in handicap betting, the model includes the dummy PH_i, which identifies whether $team_i$ receives a favourable handicap line. The variable takes value of 1 if $team_i$ gets the handicap advantage, or 0 if not. Apart from this, the logistic regression is the same as for the Home/Draw/Away model.

The logistic regression is specified as:

$$\begin{gathered} P(AHGain_i = 1|X_i) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 PH_i + \beta_3 Home_i + \\ \beta_4 Favourite_i + \beta_5 Promoted_i + \beta_6 Success + \beta_7 DiffAttend_i) \end{gathered} \tag{20}$$

4.2 Return on Investment

To evaluate whether any betting strategies generate positive economic returns over a time period the analysis employs return on investment as the primary performance measure. Consistent with this framework, market efficiency is assessed by testing the null hypothesis:

$$H_0\colon E(ROI) \leq 0 \tag{21}$$

If the null hypothesis is rejected, this suggests a violation of weak efficiency in line with Thaler and Ziemba's (1988) framework.

ROI is a widely used metric in the betting and finance literature, as it standardises profitability relative to the original wager and allows for direct comparison across different strategies and time periods (Botchkarev & Andru, 2011). The measure is defined as:

$$ROI = \frac{Payout - Wager}{Wager} \tag{22}$$

For this analysis, all bets are placed using a fixed wager of 1, meaning that the ROI for each strategy reflects the net profit or loss per unit wagered.

A positive ROI indicates that the strategy yields an economic gain on average, while a negative ROI implies that the strategy leads to loss of money. This approach aligns with the methodology applied by Winkelmann et al. (2024) who similarly used ROI to assess the financial performance of betting strategy in their study.

The ROI analysis in this thesis is based on betting strategies derived from statistically identified biases. A bias is deemed relevant if it is statistically significant in at least one period within a given market. When multiple biases are significant within the same period, additional strategies combining these biases are considered. However, to avoid excessively small samples, combinations are limited to at most two biases. All strategies identified in this manner are subsequently evaluated across all time periods within the same market, regardless of whether the underlying bias is significant in each individual period.

4.3 Bootstrapping

This thesis will also include an analysis of the statistical uncertainty surrounding the estimated returns on investment by applying bootstrapping techniques. Winkelmann et al. (2024) explicitly recommended this approach, emphasising that point estimates of betting returns should be supplemented with confidence intervals to provide a more comprehensive assessment.

> For further analyses regarding the returns of the different betting strategies, one might consider confidence intervals, as the returns reported throughout this article are only point estimates. In particular, one could consider resampling methods, such as bootstrapping, to obtain confidence intervals for the returns.
>
> (Winkelmann et al., 2024, p. 79)

Bootstrapping is a technique used to measure the uncertainty of an estimator by repeatedly sampling the data with replacement and recalculating the estimator each time (James et al., 2022). In this thesis, the estimator is the ROI, and the bootstrapping produces a distribution reflecting the variability of the ROI.

To maintain the integrity of the match structure, resampling is carried out at the match level, ensuring that both observations from a single match are included or excluded together. The procedure is performed for 1,000 iterations, generating 1,000 ROI estimates. A 95% confidence interval is then constructed by taking the 2.5th and 97.5th percentiles of the bootstrap distribution.

The bootstrap analysis is conducted only for strategies that yield positive ROIs in at least one period, to determine whether such positive returns provide statistically robust evidence against market efficiency.

4.4 Ignorance Score

For descriptive purposes we will provide a way to evaluate the accuracy of the bookmakers implied probabilities. The Ignorance Score is a measure of how accurate a prediction is. The score is calculated as

$$IG = -\log_2(p(x)) \tag{23}$$

$p(x)$ represents the probability of the of the observed outcome (Wheatcroft, 2021).

Wheatcroft argues that the Ignorance Score is a more informative measure than the Brier Score. This stands in contrast to Winkelmann et al. (2024), who rely on the Brier Score in their descriptive analysis. In our context, however, the Ignorance Score is preferable because it can be normalized, allowing us to compare predictive accuracy across different formats, with different number of betting options.

The Normalized Ignorance Score is given by

$$NIG = -\frac{\log_2 p(x)}{\log_2 K} \tag{24}$$

where K is the number of possible betting options. For the Home/Draw/Away market, $K = 3$, while for the Asian Handicap market, $K = 2$. Uninformed betting, in which equal probabilities are assigned to all betting options, serves as the benchmark case and yields a Normalized Ignorance Score of 1. The corresponding Ignorance Score benchmarks are 1.58 for the Home/Draw/Away market and 1 for the Asian Handicap market. In the latter case, the Normalized Ignorance Score and the Ignorance Score coincide because $log_2(2) = 1$.

5 Data

5.1 Data Sources

The data gathered for this thesis is mainly sourced from Football-Data[5], which is a website containing historical football betting data for multiple football leagues worldwide (Football-Data, 2025d). This data source is commonly used in the football betting market research, for instance by Winkelmann et al. (2024). Among the data collected are games from the English *Premier League*, the French *Ligue 1*, the German *Bundesliga*, the Italian *Serie A* and the Spanish *La Liga*, referred to as the top five leagues in football in this thesis. Data from the Danish *Superligaen*, the Finnish league *Veikkausliigaen*, the Norwegian *Eliteserien*, and the Swedish *Allsvenskan*, referred to as the Nordic leagues, have also been included. Throughout the thesis, we refer to each league by the nation in which it is based. The data spans from the season 2013/2014 through the 2024/2025 season, covering a total of 32,123[6] matches.

The odds data differs in structure and detail across leagues. The top five leagues contain data for both the Home/Draw/Away- and the Asian Handicap markets. While the Nordic leagues only include the Home/Draw/Away market. The data coverage is also more limited in the Nordic leagues, with fewer bookmakers and no match statistics are available other than the result.

Note that for the odds, this thesis will examine the closing odds. The closing odds are the last odds quoted before the game starts, consequently meaning that these are the odds that contain the most information.

[5] https://www.football-data.co.uk/

[6] The uneven number of matches is due to Ligue 1 canceling its 2019/2020 season (Freeborn, 2020).

The data of the number of attendances at each game is collected from European Football Statistics (2025). The reason for using an external source stem from a lack of consistency in the dataset provided from Football-Data.

5.1.1 Data on the Home/Draw/Away Markets

Even within the top five league markets, differences in how odds are reported exist. In particular, the number of bookmakers offering closing Home/Draw/Away odds varies between leagues and seasons. This study calculates its own average closing Home/Draw/Away odds by taking the average of all bookmaker-specific odds available for each match. The reason for this choice is that the variable containing average odds incorporates bookmakers that are not explicitly listed in the dataset.

For the Nordic league markets, where bookmaker coverage is not as diverse as in the top five league markets, the quoted average closing odds supplied in the dataset are used instead.

Another difference is that only Denmark among the Nordic leagues have the same season structure as the top five leagues. Due to the winter climate, the Norwegian, Swedish and Finnish leagues, start their season in the spring part and ends before Christmas. As a consequence, these leagues have one fewer season in the *post-Covid* period. Generally, the Nordic leagues also have less games in a season due to less teams in the division. How the structure of the season affects the allocation across the subsamples *Pre-Covid*, *Covid* and *Post-Covid* is displayed in Table 1.

Table 1: Overview of season allocation across subsamples.

	PRE-COVID	COVID	POST-COVID
STANDARD SEASON FORMAT	13/14 - 18/19	19/20 - 21/22	22/23 - 24/25
FINLAND, NORWAY, SWEDEN	13 - 19	20 - 22	23 - 24

5.1.2 Data on the Asian Handicap Markets

The dataset includes information on Asian Handicap betting odds and is also sourced from Football-Data. For the seasons 2013/2014 to 2018/2019, the odds used in this study are the BetBrain average Asian Handicap odds. Football-Data specifies that the averages are calculated across a wide range of bookmakers aggregated within the BetBrain platform (Football-Data, 2025i). It should be noted that the first match week of the 2013/2014 season lacks Asian Handicap observations across all leagues, creating a small gap in the dataset.

From the 2019/2020 season onwards, the BetBrain-based variables are no longer included in Football-Data's files. Instead, Football-Data provides its own market average Asian Handicap odds, which are used in this study as a direct replacement to ensure continuity in the dataset. These are reported in the variables $AvgAHH$, average home-team Asian Handicap odds, and $AvgAHA$, average away-team Asian Handicap odds. Although these averages are not methodologically identical to the BetBrain averages, these market averages represent a consolidated estimate across available bookmakers. Therefore, they are the closest functional equivalent.

These cases where the handicap line is 0 are removed, because this thesis aims to examine biases in cases with handicap. As explained in Section 2.2.5.2, a 0-handicap bet is the same as a Home/Draw/Away bet, except the outcome of draw would provide a refund.

5.2 Descriptive Statistics

In Table 2 and Table 3, the summary of outcomes in the Home/Draw/Away market are presented. The tables also report the average margin, the Ignorance Score, and the Normalized Ignorance Score. They indicate that the Nordic leagues have higher average margins, and their Ignorance Scores are generally higher as well. All Ignorance Scores are below 1.58, indicating that the implied probabilities provide better predictive performance than uninformed betting.

Table 2: Home/Draw/Away outcome summary in the top five leagues.

Top 5 H/D/A Summary

	MATCHES	HOME WINS	DRAWS	AWAY WINS	AVG MARGIN	IG	NIG
ENG	4,560	44.8%	23.2%	32.0%	3.50	1.37	0.87
FRA	4,311	44.0%	25.7%	30.4%	3.90	1.42	0.90
GER	3,672	44.8%	24.6%	30.7%	3.85	1.41	0.89
ITA	4,560	42.9%	25.8%	31.4%	3.95	1.37	0.86
SPA	4,560	45.5%	25.7%	28.8%	3.81	1.38	0.87

Table 3: Home/Draw/Away outcome summary in the Nordic leagues.

Nordic H/D/A Summary

	MATCHES	HOME WINS	DRAWS	AWAY WINS	AVG MARGIN	IG	NIG
DNK	2,550	43.2%	26.2%	30.6%	6.25	1.45	0.92
FIN	2,150	43.3%	25.9%	30.8%	6.96	1.45	0.91
NOR	2,880	46.7%	24.6%	28.7%	6.36	1.43	0.90
SWE	2,880	44.2%	24.8%	31.0%	6.22	1.41	0.89

Table 4 exhibits the summary of outcomes in the Asian Handicap markets. A home win here is defined as:

$$Adjusted\ goals\ for\ the\ home\ team\ >\ Adjusted\ goals\ for\ the\ away\ team$$

As observed, the share of draws is lower, as many handicaps use decimal values and therefore cannot result in a draw. The average margin in the Asian Handicap markets is close to what we observe in the top five leagues in the Home/Draw/Away market.

Table 4: Asian Handicap match outcome distribution in the top five leagues.

Top 5 AH Summary							
	MATCHES	HOME WINS	DRAWS	AWAY WINS	AVG MARGIN	IG	NIG
ENG	4,102	46.6%	5.2%	48.2%	3.69	0.94	0.94
FRA	3,684	46.5%	4.3%	49.2%	4.09	0.95	0.95
GER	3,189	47.8%	3.9%	48.4%	3.85	0.96	0.96
ITA	4,053	44.8%	5.8%	49.4%	3.92	0.94	0.94
SPA	4,054	46.4%	5.3%	48.2%	3.87	0.94	0.94

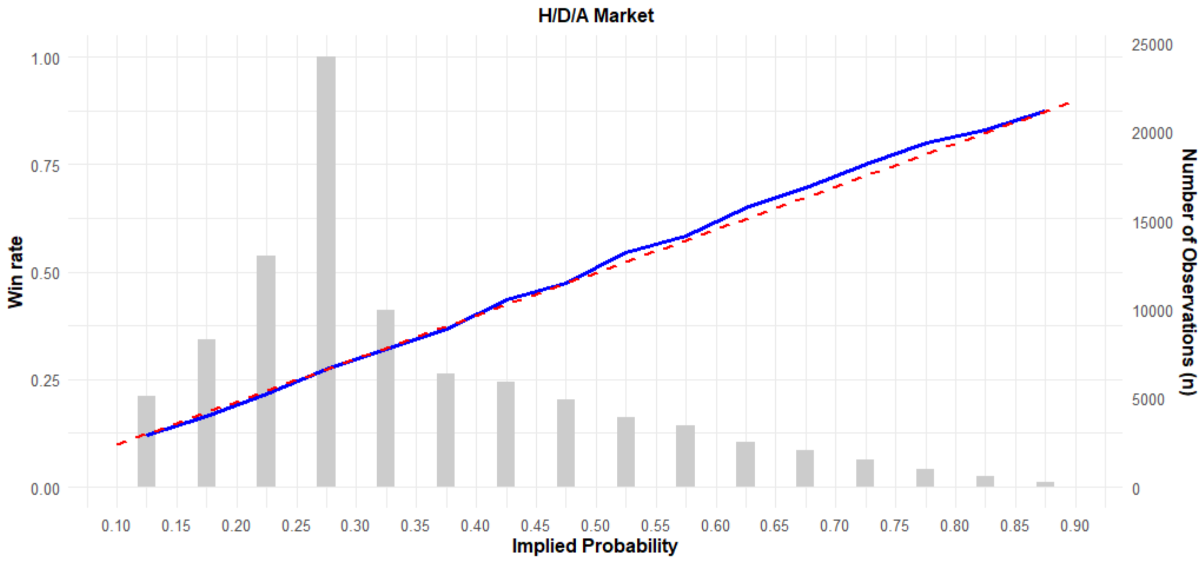
It should be noted that the Normalized Ignorance Score is higher than in the Home/Draw/Away market. This suggests that the implied probability performs worse in predicting whether a bet will yield a positive economic gain (i.e. $AHGain = 1$, as defined in Section 4.1.3). As noted in Section 4.4, a score higher than Normalized Ignorance Score of 1, would imply that the implied probability is worse than uninformed betting.

Table 5: Asian Handicap betting outcome distribution in the top five leagues.

AH Home Bet Results						
	MATCHES	WINS	HALF WINS	REFUNDS	HALF LOSSES	LOSSES
ENG	4,102	40.3%	6.3%	5.2%	7.9%	40.3%
FRA	3,684	40.6%	5.9%	4.3%	8.6%	40.6%
GER	3,189	42.0%	5.8%	3.9%	7.0%	41.4%
ITA	4,053	38.2%	6.6%	5.8%	8.4%	41.0%
SPA	4,054	39.5%	6.9%	5.3%	8.4%	39.9%

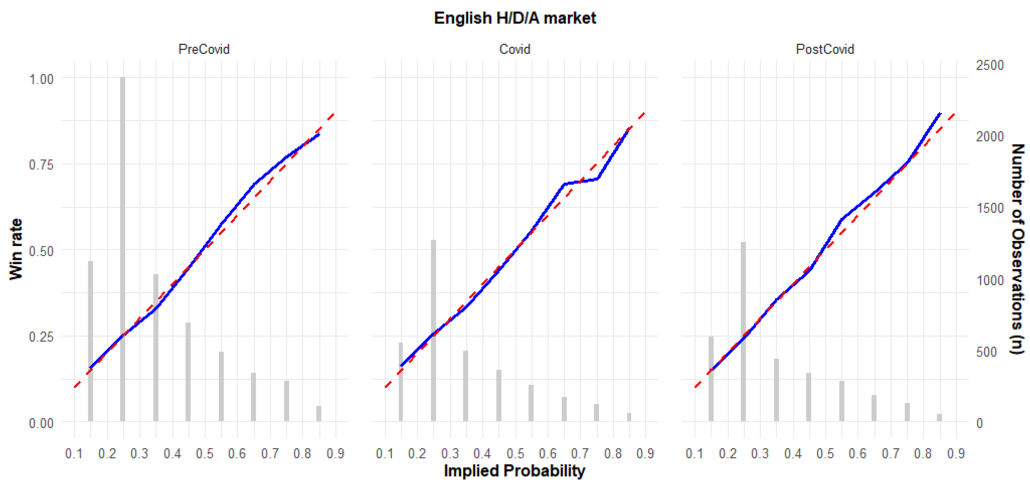
Table 5 show the distribution of outcomes if one where to bet on the home team in the Asian Handicap market.

Figure 8: Relationship between implied probabilities and realized outcomes in the H/D/A market.



In Figure 8, the grey bars indicate the distribution of implied probability. The red line in figure represents the theoretical relationship between implied probability and the rate of economic wins (i.e. $Won_i = 1$ and $AHGain_i = 1$), which is a 45° line. The blue line shows the observed relationship. Based on this, we observe that in the Home/Draw/Away market, the relationship between the win rate and the implied probability aligns closely with the theoretical expectation. This holds for Figure 9 as well.

Figure 9: Relationship between implied probabilities and realized outcomes in the English H/D/A market across subsamples.



The Asian Handicap market behaves differently. Figure 10 shows that the relationship between *AHGain* and implied probability deviates somewhat from the theoretical line. This deviation becomes even more pronounced when the data is split into time periods and restricted to the English league, as illustrated in Figure 11. A reason for this could be that the implied probabilities are heavily compressed within the 44% to 56% interval, with very few observations outside this range. Since nearly all probabilities are close to 50%, more observations are needed to identify the true underlying distribution. This could also explain why the Normalized Ignorance Score is worse in the Asian Handicap markets compared to the Home/Draw/Away market, the variance in implied probabilities is very low.

Figure 10: Relationship between implied probabilities and realized outcomes in the Asian Handicap market.

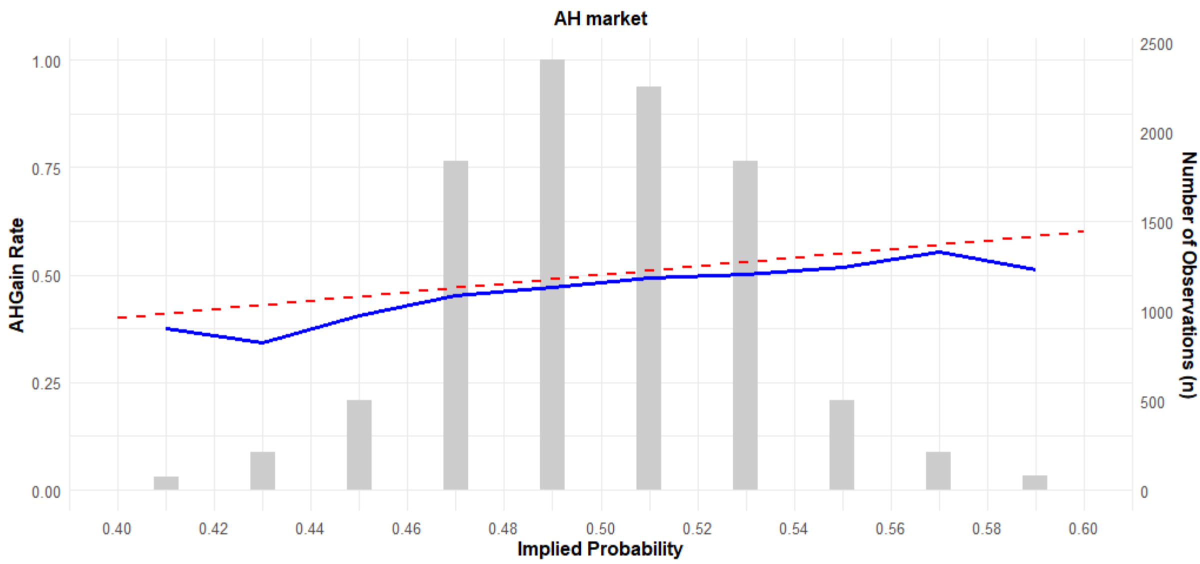


Figure 11: Relationship between implied probabilities and realized outcomes in the English A.H. market across subsamples.

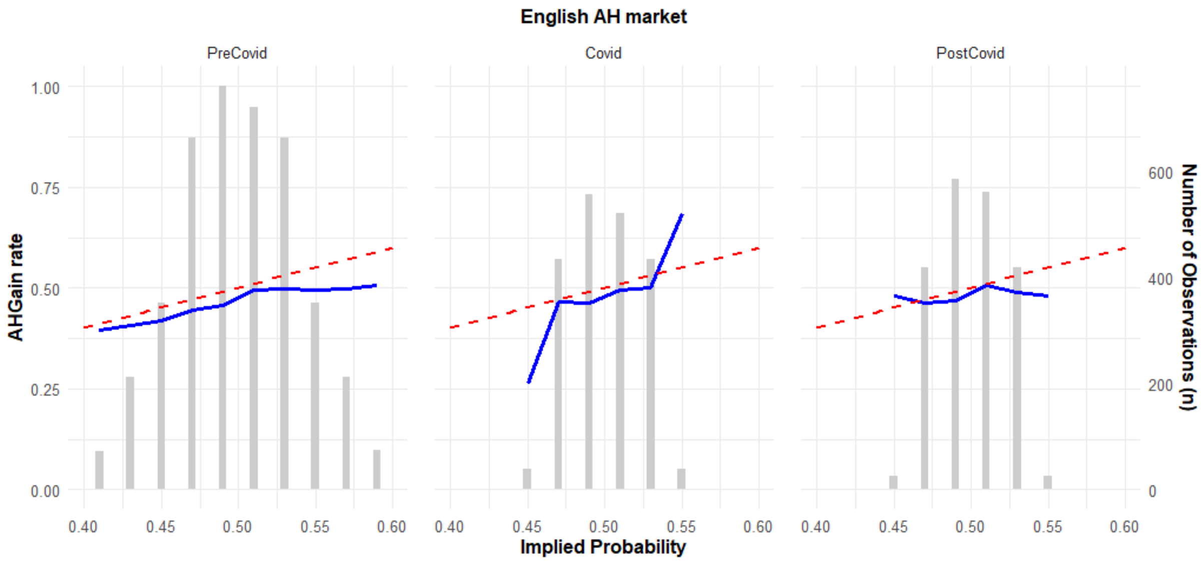
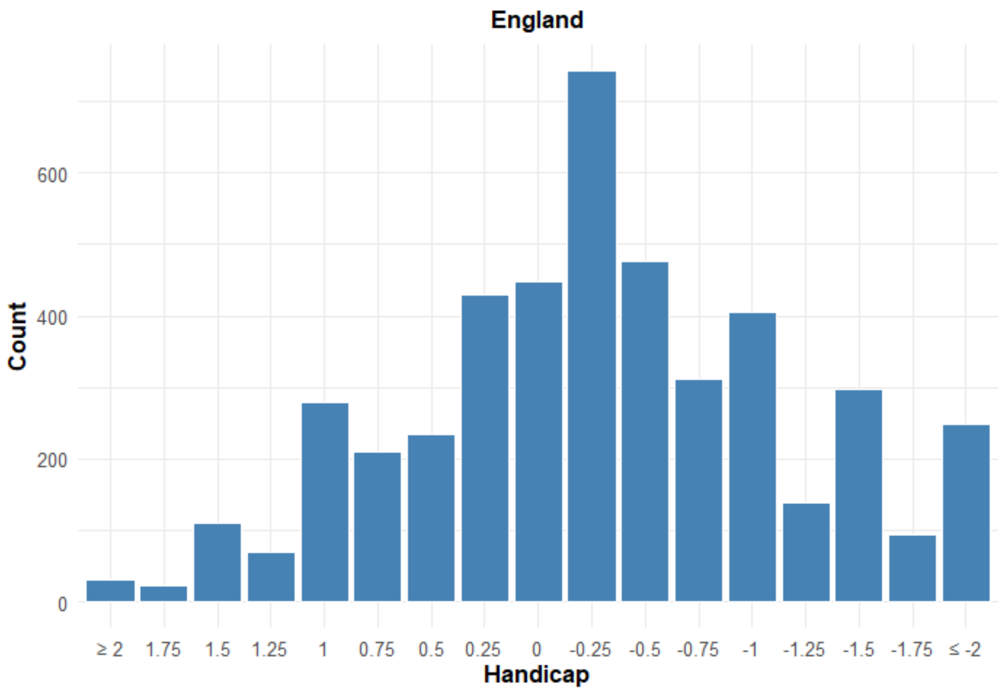


Figure 12 presents the distribution of handicaps applied to the home team in the English league. The distribution appears approximately normal and is centred around -0.25. This centre reflects the home advantage, with the handicap mechanism designed to offset this advantage. Distributions for each league are provided in Appendix C.

Figure 12: Distribution of Handicap in the English Asian Handicap market.



6 Analysis

Our analysis employs a logistic regression model to identify systematic biases in both the Home/Draw/Away betting market and the Asian Handicap market. First, we examine each league to find biases and then compare the findings with each other.

6.1 The Home/Draw/Away Markets: Top Five Leagues vs. Nordic Leagues

This section is divided into three parts; first we will go through the results from the logistic regression, looking at what biases we find and in which leagues. From there we will look at the return on investments, based on strategies exploiting the biases found in the league. Finally, a bootstrap procedure is used to assess whether a ROI strategy yields statistically significant returns and therefore constitute evidence against weak efficiency.

6.1.1 Biases

The general result of the logistic regression is that biases vary considerably across leagues. There is no single bias that persists consistently across leagues or time periods. However, the implied probability seems to be a good predictor of the match outcome. The coefficient for $ImProb_i$ is consistently positive and significant at 1% level for each of the top 5 leagues before, during and after the pandemic. The same applies to the Nordic leagues, with the only exception being the Finnish league in the period after Covid-19, where the implied probability-coefficient is significant at 10% level. As multiple coefficients are tested within a logistic regression framework, some of the observed significant biases may represent false positives. This is discussed further in Section 7.1.1.

Based on the logistic regression results, the pandemic appears to be associated with a temporary increase in the prevalence of biases in several leagues. In particular, the top five leagues appear to have been more affected during the pandemic than the Nordic leagues, as statistically significant biases are observed in all of the top five leagues. In contrast, biases in the Nordic leagues are more diverse across the subsamples.

Before analysing the logistic regression in more detail, it is important to note that the magnitude of the coefficients is not central for determining whether biases exist. Rather, the presence of bias is identified by whether variables other than the implied probability are statistically significant and whether their coefficients are positive or negative. The magnitude of the coefficients instead reflects how strongly changes in the explanatory variables shift the predicted probabilities through the logistic distribution function.

To illustrate effect of the magnitudes, suppose the home team has an $ImProb_i$ of 50% and the away team has 30%. Assume a constant term of -2.5 and a coefficient of 5 on $ImProb_i$.

The resulting predicted probabilities are 50% for the home team and 26.9% for the second team.

$$p_{home}(x) = \frac{e^{-2.5+0.5\cdot 5}}{e^{-2.5+0.5\cdot 5}+1} = 0.5 \tag{25}$$

$$p_{away}(x) = \frac{e^{-2.5+0.3\cdot 5}}{e^{-2.5+0.3\cdot 5}+1} = 0.269 \tag{26}$$

As for the goodness of fit, the adjusted pseudo-R^2 values for the logistic regressions models are reported in Table 6 and Table 7. They indicate that, in general, the top five leagues exhibit a higher degree of model fit than the Nordic markets.

Table 6: Adjusted pseudo-R^2 values for the top five league H/D/A markets.

Top 5 H/D/A Adj. Pseudo-R^2				
	ALL	PRE-COVID	COVID	POST-COVID
ENG	14.37%	14.97%	12.38%	14.95%
FRA	11.52%	12.02%	10.25%	11.18%
GER	11.79%	11.67%	11.43%	12.58%
ITA	15.74%	16.36%	15.02%	14.78%
SPA	14.87%	15.79%	13.16%	14.76%

Table 7: Adjusted pseudo-R^2 values for the Nordic H/D/A markets.

Nordic H/D/A Adj. Pseudo-R^2				
	ALL	PRE-COVID	COVID	POST-COVID
DNK	9.28%	9.31%	9.63%	8.12%
FIN	9.82%	8.11%	14.22%	9.54%
NOR	10.23%	9.55%	12.79%	8.74%
SWE	11.86%	13.96%	9.42%	8.08%

6.1.1.1 *The Top Five Leagues*

As observed from the logistic regression tables, there is limited evidence of systematic biases when considering the full sample for each league. The only league showing signs of a bias for the overall sample is the Italian market.

Table 8: Logit regression results for the English and French H/D/A markets.

England

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.419***	-2.501***	-2.182***	-2.495***
	(0.0884)	(0.1255)	(0.1733)	(0.1826)
ImProb	4.984***	5.252***	4.508***	4.992***
	(0.3077)	(0.4401)	(0.5919)	(0.6389)
Home	-0.0243	0.0406	-0.1494	0.0006
	(0.0671)	(0.0991)	(0.1269)	(0.1332)
Favourite	0.0146	-0.1036	0.0643	0.1205
	(0.0991)	(0.1413)	(0.1964)	(0.2028)
Promoted	-0.0942	-0.0044	-0.1400	-0.2341
	(0.0727)	(0.1021)	(0.1402)	(0.1552)
Success	-0.0353	-0.0208	-0.0965**	-0.0077
	(0.0234)	(0.0336)	(0.0467)	(0.0462)
DiffAttend	0.0006	0.0005	0.0054	-0.0029
	(0.0016)	(0.0024)	(0.0033)	(0.0030)
Observations	9,120	4,560	2,280	2,280

France

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.452***	-2.522***	-2.449***	-2.281***
	(0.0959)	(0.1292)	(0.2037)	(0.2080)
ImProb	5.133***	5.247***	5.399***	4.534***
	(0.3438)	(0.4714)	(0.7155)	(0.7313)
Home	-0.1149	-0.0672	-0.1516	-0.1637
	(0.0727)	(0.1077)	(0.1399)	(0.1413)
Favourite	0.0422	0.0419	-0.1197	0.2060
	(0.0983)	(0.1349)	(0.2001)	(0.2139)
Promoted	0.0057	-0.0118	0.0387	0.0337
	(0.0751)	(0.1003)	(0.1662)	(0.1577)
Success	-0.0233	-0.0086	-0.0887*	0.0230
	(0.0236)	(0.0326)	(0.0476)	(0.0494)
DiffAttend	-0.0005	-0.0014	0.0031	-0.0015
	(0.0022)	(0.0034)	(0.0042)	(0.0038)
Observations	8,622	4,560	2,078	1,984

As observed from Table 8, the English market exhibits a generally a bias-less market, with mainly the $ImProb_i$-coefficient being statistically significant.

The only bias exhibited in the results from the English market negative momentum bias. The $Success_i$-variable becomes significant at 5% significance level during the *Covid*-period and taking a negative sign. This illustrates a bias where bookmakers tend to overestimate the effect of having better recent form than the opposition. This variable is only significant during the pandemic, indicating that this is just a temporary bias. The absence of a statistically significant home bias is also noted.

The result for the French market is similar to the English market for the overall sample. The only statistically significant variable is the $ImProb_i$. During the pandemic, $Success_i$ becomes marginally significant at 10% level and negative.

Table 9: Logit regression results for the German and Italian H/D/A markets.

Germany

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.472***	-2.517***	-2.593***	-2.295***
	(0.0998)	(0.1418)	(0.2051)	(0.1983)
ImProb	4.976***	5.039***	6.010***	3.970***
	(0.3437)	(0.4893)	(0.7149)	(0.6726)
Home	0.0425	0.1649	-0.0106	-0.1710
	(0.0741)	(0.1067)	(0.1432)	(0.1510)
Favourite	-0.0384	-0.1044	-0.5157**	0.5210**
	(0.1054)	(0.1492)	(0.2210)	(0.2073)
Promoted	0.0322	0.1321	-0.0259	-0.0924
	(0.0864)	(0.1229)	(0.1618)	(0.1911)
Success	-0.0105	-0.0234	-0.0049	0.0061
	(0.0262)	(0.0376)	(0.0516)	(0.0529)
DiffAttend	-0.0022	-0.0015	-0.0044	-0.0011
	(0.0015)	(0.0023)	(0.0030)	(0.0029)
Observations	7,344	3,672	1,836	1,836

Italy

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.544***	-2.643***	-2.263***	-2.681***
	(0.0966)	(0.1373)	(0.1992)	(0.2058)
ImProb	5.211***	5.550***	4.503***	5.470***
	(0.3412)	(0.4990)	(0.6678)	(0.7327)
Home	-0.1307*	-0.1048	-0.2373*	-0.1127
	(0.0686)	(0.1067)	(0.1279)	(0.1332)
Favourite	0.0653	0.0069	0.1357	0.0900
	(0.1013)	(0.1440)	(0.2055)	(0.2035)
Promoted	0.0051	0.0289	0.0416	-0.0752
	(0.0748)	(0.1061)	(0.1509)	(0.1486)
Success	0.0013	0.0118	-0.0310	0.0107
	(0.0240)	(0.0347)	(0.0477)	(0.0473)
DiffAttend	0.0018	-0.0017	0.0084**	0.0004
	(0.0021)	(0.0040)	(0.0040)	(0.0034)
Observations	9,120	4,560	2,280	2,280

Table 9 displays the logistic regression results for Germany and Italy. In Germany, no biases are found in the overall sample, nor in the period before the pandemic. However, in the *Covid*-sample, the coefficient of $Favourite_i$ is negative and statistically significant at 5% significance level. This indicates the presence of a longshot bias. Interestingly, the coefficient changes from negative to positive in the period after the pandemic, suggesting that the direction of the bias changes across periods rather than representing a persistent structural feature of the market.

The Italian betting market is the only top five league where a statistically significant bias appears in the full sample. The coefficient on $Home_i$ is negative and statistically significant at the 10% level, indicating that bookmaker odds systematically undervalue away teams. When splitting the data into subperiods, this away bias is only present during the *Covid*-period and significant at the 10% level. During the pandemic the $DiffAttend_i$-coefficient is positive and becomes significant, although the magnitude is small. This indicates that bookmakers tend to undervalue the more supported teams in Italy.

Table 10: Logit regression results for the Spanish H/D/A market.

Spain

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.504***	-2.639***	-2.302***	-2.592***
	(0.0998)	(0.1461)	(0.2100)	(0.1986)
ImProb	4.933***	5.845***	3.570***	5.010***
	(0.3593)	(0.5282)	(0.7536)	(0.7176)
Home	0.0115	-0.0654	0.1719	-0.0617
	(0.0767)	(0.1141)	(0.1447)	(0.1518)
Favourite	0.0522	-0.2353*	0.3337*	0.2706
	(0.0973)	(0.1406)	(0.1942)	(0.1961)
Promoted	0.0813	-0.0105	0.2728*	0.1093
	(0.0717)	(0.1029)	(0.1394)	(0.1420)
Success	-0.0327	-0.0445	-0.0582	-0.0106
	(0.0243)	(0.0351)	(0.0458)	(0.0511)
DiffAttend	0.0029	-0.0018	0.0096***	0.0021
	(0.0018)	(0.0029)	(0.0034)	(0.0033)
Observations	9,120	4,560	2,280	2,280

As for the Spanish market, none of the variables except $ImProb_i$ are statistically significant in the full sample. The coefficient for $Favourite_i$ becomes marginally statistically significant at 10% and negative before the pandemic, leading to an observed longshot bias. This coefficient turns positive at the same significance level during the pandemic, indicating a shift in the bias from longshot to favourite.

Two other biases occurring during covid is a promoted bias and a sentiment bias. The coefficients on $Promoted_i$ and $DiffAttend_i$ becomes statistically significant and positive at the 10% level and 1%, respectively. This suggests that promoted teams and more supported teams were priced more favourably during the pandemic period, potentially creating exploitable betting opportunities.

6.1.1.2 The Nordic Leagues

Table 11: Logit regression results for the Danish and Finnish H/D/A markets.

	Denmark					**Finland**			
	Won					**Won**			
	All	PreCovid	Covid	PostCovid		All	PreCovid	Covid	PostCovid
Constant	-2.616***	-2.855***	-2.401***	-2.432***	Constant	-2.425***	-2.374***	-2.705***	-2.049***
	(0.1570)	(0.2298)	(0.3231)	(0.3276)		(0.1531)	(0.1904)	(0.3251)	(0.4564)
ImProb	5.438***	6.374***	4.212***	5.069***	ImProb	4.907***	4.746***	6.270***	2.813*
	(0.5397)	(0.7821)	(1.125)	(1.132)		(0.5432)	(0.6785)	(1.144)	(1.636)
Home	0.0247	-0.0176	0.1017	0.0066	Home	-0.0464	0.0593	-0.3510*	0.1418
	(0.0873)	(0.1261)	(0.1739)	(0.1814)		(0.0984)	(0.1275)	(0.2116)	(0.2458)
Favourite	-0.0267	-0.1975	0.3856	-0.0892	Favourite	0.1109	-0.0187	0.0628	0.7201*
	(0.1280)	(0.1765)	(0.2612)	(0.2745)		(0.1378)	(0.1713)	(0.3119)	(0.3824)
Promoted	0.0145	0.1251	-0.1021	-0.2550	Promoted	-0.0380	0.0079	-0.1276	-0.1130
	(0.0911)	(0.1216)	(0.2025)	(0.1986)		(0.1056)	(0.1289)	(0.2427)	(0.2806)
Success	-0.0266	-0.0410	0.0119	-0.0431	Success	0.0337	0.0255	-0.0027	0.1104
	(0.0305)	(0.0428)	(0.0608)	(0.0636)		(0.0335)	(0.0416)	(0.0783)	(0.0846)
DiffAttend	0.0015	-0.0089	0.0065	0.0056	DiffAttend	-0.0355	-0.0249	-0.0763	-0.0209
	(0.0074)	(0.0145)	(0.0160)	(0.0104)		(0.0371)	(0.0474)	(0.0754)	(0.1025)
Observations	5,100	2,686	1,258	1,156	Observations	4,300	2,710	922	668

In one of the Nordic leagues, there is limited evidence of persistent biases. The logistic regression results for Denmark do not indicate the presence of any statistically significant biases. However, in Finland, two notable differences arise. First, there is marginal evidence of an away bias during the pandemic period, as well as marginal evidence of a favourite bias in the post-pandemic period. Second, the implied probability in the post-pandemic period is statistically significant at the 10% level, which stands in contrast to the results obtained for all other logistic regressions within the Home/Draw/Away market.

Table 12: Logit regression results for the Norwegian and Swedish H/D/A markets.

Norway

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.433***	-2.340***	-2.814***	-2.209***
	(0.1168)	(0.1599)	(0.2476)	(0.2608)
ImProb	5.163***	4.510***	6.969***	4.867***
	(0.4375)	(0.6027)	(0.9144)	(0.9701)
Home	-0.0071	0.0489	-0.0854	-0.1105
	(0.0912)	(0.1232)	(0.1848)	(0.2184)
Favourite	-0.1323	0.1016	-0.6827***	-0.1746
	(0.1235)	(0.1624)	(0.2491)	(0.3070)
Promoted	-0.0312	-0.0318	-0.1527	0.1031
	(0.0866)	(0.1185)	(0.1722)	(0.1978)
Success	0.0260	-0.0049	0.1100**	-0.0079
	(0.0278)	(0.0376)	(0.0527)	(0.0697)
DiffAttend	0.0065*	0.0083**	-0.0108	0.0111
	(0.0039)	(0.0042)	(0.0189)	(0.0203)
Observations	5,760	3,360	1,440	960

Sweden

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.623***	-2.644***	-2.637***	-2.471***
	(0.1224)	(0.1644)	(0.2424)	(0.2963)
ImProb	5.992***	5.907***	5.678***	6.355***
	(0.4484)	(0.6122)	(0.8496)	(1.087)
Home	-0.1269	-0.1497	-0.0844	-0.0655
	(0.0878)	(0.1206)	(0.1626)	(0.2160)
Favourite	-0.2563**	-0.1763	-0.1244	-0.6523**
	(0.1241)	(0.1639)	(0.2451)	(0.3001)
Promoted	-0.0238	0.0084	-0.0493	-0.0493
	(0.0900)	(0.1191)	(0.1786)	(0.2227)
Success	-0.0150	-0.0022	-0.0590	-0.0051
	(0.0284)	(0.0391)	(0.0552)	(0.0661)
DiffAttend	0.0019	0.0133*	-0.0128	0.0012
	(0.0048)	(0.0074)	(0.0091)	(0.0091)
Observations	5,760	3,360	1,440	960

As illustrated in Table 12, Norway and Sweden display broadly similar patterns of bias in their betting markets. In Norway, we identify three forms of bias: a sentiment bias, a longshot bias, and a momentum bias. The longshot bias and the momentum bias appear only during the pandemic, indicating that bookmakers overestimated the favourite while underestimating the team in better form during this period. The sentiment bias, however, is present both before Covid-19 and in the overall sample. This sentiment bias implies that teams with larger supporter bases tend to be priced more favourably in the odds. Sweden exhibits the same types of biases as Norway with respect to the longshot bias and the sentiment bias. The longshot bias is statistically significant in both the full sample and the *post-Covid* period. As in Norway, a sentiment bias is also observed in the pre-pandemic period.

6.1.2 Return on Investment

While statistically significant biases are observed, there is no guarantee that they will translate into positive strategy returns. However, our analysis suggests there are cases in which the strategies yield a positive ROI, which could indicate a potential violation of weak efficiency. Table 13 and Table 14 display the returns of strategies that had at least one period with a positive ROI. Notably, the markets in the top five leagues had the most opportunities for profit during the pandemic period, while during the seasons after the pandemic only one strategy yielded a positive ROI. In contrast, four of the six profitable opportunities in the Nordic markets occurred in the *post-Covid* period.

Across the full sample, three strategies exhibit positive average ROI estimates, in England, Italy and Spain. The strategy in the Spanish market is particularly notable because it shows a positive ROI in all periods, including 22.43% during the pandemic. However, this result is driven by a very limited number of betting opportunities, with only 68 bets in total.

It is noteworthy that some strategies yield profits even when the bias is not statistically significant. For example, betting on the worse-form team in England yields positive returns in both the overall period and the *pre-Covid* period, despite the bias not being statistically significant in those periods. This could be because of high standard errors, and low sample size to show the true bias, which will be discussed in Section 7.1.

Table 13: ROI results for the top five league H/D/A markets.

Top 5 H/D/A ROI				
	OVERALL	PRE-COVID	COVID	POST-COVID
ENG: WORSE FORM	0.13	0.49	11.48	−11.68
FRA: WORSE FORM	−4.70	−5.48	1.63	−9.54
GER: LONGSHOT	−5.69	−4.41	2.77	−16.69
ITA: AWAY	−7.20	−8.28	3.03	−15.26
ITA: AWAY & SENTIMENT	0.36	−1.29	4.62	−0.59
SPA: PROMOTED	−4.97	−4.45	1.15	−12.15
SPA: PROMOTED & SENTIMENT	6.05	1.07	22.43	1.52

Table 14: ROI results for the Nordic H/D/A markets.

Nordic H/D/A ROI	OVERALL	PRE-COVID	COVID	POST-COVID
FIN: FAVOURITE	−2.75	−6.49	2.50	5.13
NOR: LONGSHOT	−10.31	−12.58	−12.89	1.51
SWE: LONGSHOT	−11.73	−17.85	−9.16	5.83
SWE: SENTIMENT	−2.10	2.05	−15.17	3.00

6.1.3 Bootstrapping

To assess whether the strategies yield returns significantly greater than zero and thereby constitute evidence against weak efficiency, a bootstrap procedure is applied to quantify uncertainty in the estimated return on investment.

Across all leagues and the subsamples, the results demonstrate that none of the estimated 95% confidence intervals are entirely above zero. Consequently, there is no statistical evidence to suggest that these strategies yield systematically positive returns, as can be seen in Figure 13 and Figure 14. In some cases, certain strategies display confidence intervals fully below zero, suggesting statistically significant negative expected returns under repeated sampling.

During the *Covid*-period, some strategies within the top five league market display point estimates marginally exceeding zero. However, the evidence remains insufficient to assert the presence of a systematic exploitable inefficiency, since the bootstrap intervals spans both positive and negative values.

In Figure 13 the bootstrap clearly shows that the Spanish strategy of betting on newly promoted teams with a larger following than their opponents have very high variance, which can likely be explained by the limited number of observations.

While the bootstrap results provide no evidence of systematically positive returns, confidence intervals overlapping zero do not rule out economically relevant differences across strategies. Some strategies appear consistently loss-making, while others yield estimates closer to breakeven, providing a baseline for further research. This also applies for the bootstrap analysis for the Asian Handicap markets in Section 6.2.3.

Figure 13: Bootstrap results of the top five league H/D/A markets.

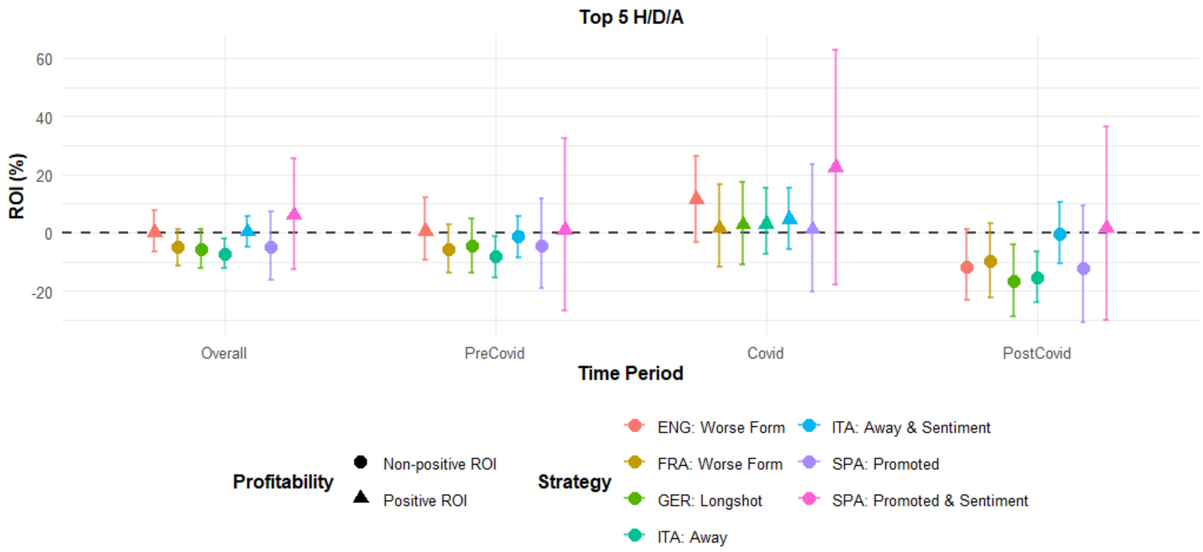
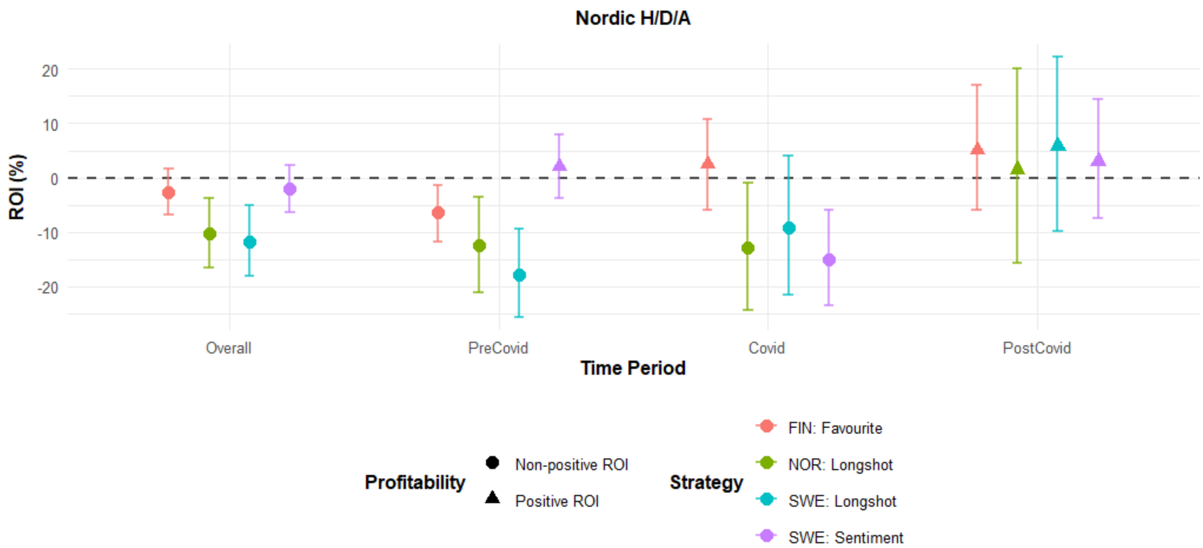


Figure 14: Bootstrap results of the Nordic H/D/A markets.



6.2 The Asian Handicap Markets: the Top Five Leagues

As the last section, the following is divided into three parts. First part explains the logistic regression results, while the second and third addresses the ROI results and the bootstrap, respectively.

6.2.1 Biases

An interesting finding from the logistic regression analysis of the Asian Handicap market is that both the constant term and the implied probability are often statistically insignificant in the majority of leagues. One possible explanation for this is illustrated in Section 5.2 where deviations of implied odds from their theoretical assumptions are observed. This pattern is also reflected in Table 15, which reports the adjusted pseudo-R^2 values for the logistic regressions in the Asian Handicap market.

Table 15: Adjusted pseudo-R^2 values for the top five league Asian Handicap markets.

Top 5 AH Adj. Pseudo-R^2				
	ALL	PRE-COVID	COVID	POST-COVID
ENG	0.44%	0.29%	0.57%	0.18%
FRA	0.52%	0.52%	0.21%	0.09%
GER	0.48%	0.56%	0.47%	0.27%
ITA	0.37%	0.31%	0.32%	0.41%
SPA	0.54%	0.81%	0.37%	−0.09%

A clear contrast emerges when comparing these results to those from the Home/Draw/Away market; while the latter exhibits adjusted pseudo-R^2 values of approximately 10%, the Asian Handicap models achieve values of only about 0.4%. Moreover, in one case the adjusted pseudo-R^2 is negative, indicating that the model performs worse than the null model. The only market in which the constant term and coefficient are consistently significant is in the German market.

Table 16: Logit regression results for the German Asian Handicap market.

Germany

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.845***	-2.913***	-5.785*	-6.678**
	(0.7723)	(0.9034)	(2.987)	(2.820)
ImProb	5.319***	5.434***	11.32*	13.40**
	(1.630)	(1.944)	(6.200)	(5.852)
PH	0.2925***	0.2929**	0.4287**	0.1915
	(0.0971)	(0.1422)	(0.1894)	(0.1943)
Home	0.1076	0.1999*	0.0172	0.0094
	(0.0783)	(0.1152)	(0.1495)	(0.1567)
Favourite	-0.1848*	-0.3399**	-0.2360	-0.3426
	(0.1038)	(0.1544)	(0.2574)	(0.2367)
Promoted	0.0210	0.0613	-0.1776	0.1321
	(0.0813)	(0.1182)	(0.1563)	(0.1632)
Success	0.0101	-0.0058	0.0786	-0.0228
	(0.0260)	(0.0374)	(0.0524)	(0.0514)
DiffAttend	4.56×10^{-5}	-0.0012	0.0001	0.0015
	(0.0015)	(0.0022)	(0.0029)	(0.0029)
Observations	6,378	3,122	1,612	1,644

Germany exhibits a positive handicap bias that appears in the full sample, as well as before and during the Covid-19 period. This bias reflects bookmakers' undervaluation of teams receiving a positive handicap. In addition to this bias, there is also statistical significance of a longshot bias in the period before the pandemic. This bias is not significant in the other periods; however, it is slightly present in the full sample size.

Table 17: Logit regression results for the Spanish Asian Handicap market.

Spain

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.553***	-2.687***	-2.794	-2.667
	(0.6750)	(0.7728)	(2.392)	(2.452)
ImProb	4.632***	4.652***	5.020	5.468
	(1.436)	(1.664)	(5.019)	(5.135)
PH	0.1903**	0.3598***	0.1526	-0.0825
	(0.0961)	(0.1378)	(0.1938)	(0.1916)
Home	0.0342	0.0997	0.1121	-0.1572
	(0.0726)	(0.1042)	(0.1442)	(0.1451)
Favourite	0.0088	-0.0605	0.0751	0.0126
	(0.0957)	(0.1353)	(0.2372)	(0.2291)
Promoted	0.0823	0.1139	0.1897	-0.0511
	(0.0640)	(0.0895)	(0.1296)	(0.1296)
Success	0.0034	0.0322	-0.0758	0.0139
	(0.0236)	(0.0323)	(0.0494)	(0.0493)
DiffAttend	0.0018	0.0025	0.0028	-0.0002
	(0.0015)	(0.0021)	(0.0030)	(0.0029)
Observations	8,108	4,142	1,964	2,002

In Spain, as in Germany, we find clear evidence of a positive handicap bias. However, in Spain it is present in the full sample and the period before the pandemic. Interestingly, both during the pandemic and after, none of the coefficients are statistically significant, suggesting a poor model fit. This is observed in Table 15 as well. The lack of significance of the implied probability will be discussed in Section 7.2.1.

Table 18: Logit regression results for the English and the French Asian Handicap markets.

England

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-1.626**	-1.527*	-4.460**	-0.3709
	(0.6777)	(0.7863)	(2.240)	(2.685)
ImProb	3.008**	2.639	9.085*	0.5142
	(1.437)	(1.692)	(4.676)	(5.590)
PH	0.0704	0.1030	-0.0235	0.0997
	(0.0876)	(0.1262)	(0.1749)	(0.1735)
Home	-0.0438	0.0020	-0.1165	-0.0601
	(0.0671)	(0.0968)	(0.1327)	(0.1331)
Favourite	0.0366	0.0668	-0.1604	0.1010
	(0.0956)	(0.1396)	(0.2153)	(0.2357)
Promoted	-0.1007	-0.0798	-0.0628	-0.2020
	(0.0647)	(0.0916)	(0.1298)	(0.1296)
Success	-0.0738***	-0.0376	-0.1349***	-0.0958**
	(0.0230)	(0.0320)	(0.0491)	(0.0456)
DiffAttend	0.0019	0.0035	0.0015	-0.0005
	(0.0016)	(0.0022)	(0.0033)	(0.0030)
Observations	8,204	4,152	2,022	2,030

France

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.199***	-1.737*	-5.112*	-2.788
	(0.8075)	(0.9225)	(2.708)	(2.790)
ImProb	4.116**	2.979	10.28*	5.514
	(1.706)	(1.973)	(5.638)	(5.817)
PH	0.0592	0.1042	-0.0012	0.0480
	(0.0972)	(0.1411)	(0.1974)	(0.1907)
Home	-0.0642	-0.0150	-0.0487	-0.1657
	(0.0774)	(0.1149)	(0.1527)	(0.1474)
Favourite	0.1116	0.1843	-0.1282	0.0424
	(0.1037)	(0.1448)	(0.2494)	(0.2470)
Promoted	-0.0006	-0.0057	0.0920	-0.0372
	(0.0714)	(0.0962)	(0.1650)	(0.1435)
Success	-0.0295	-0.0004	-0.0674	-0.0582
	(0.0247)	(0.0340)	(0.0502)	(0.0529)
DiffAttend	0.0026	0.0045	0.0021	0.0010
	(0.0021)	(0.0032)	(0.0040)	(0.0038)
Observations	7,368	3,824	1,796	1,748

In England there is a clear presence of negative momentum bias, as it is significant in all periods, except before the pandemic. In France, the implied probability is only significant during the pandemic and in the overall sample, suggesting a similarly poor goodness of fit as for Spain.

The result from the Italian market exhibits an away bias in the full sample as well as in the during and after the pandemic. There is also a sentiment bias present, appearing briefly in the seasons during the pandemic.

Table 19: Logit regression results for the Italian Asian Handicap market.

Italy

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.444***	-2.631***	0.4611	-6.291**
	(0.6781)	(0.7560)	(2.180)	(2.558)
ImProb	4.822***	5.032***	-1.053	12.99**
	(1.444)	(1.635)	(4.549)	(5.354)
PH	0.0587	0.0727	0.0536	0.0610
	(0.0969)	(0.1446)	(0.1896)	(0.1905)
Home	-0.1533**	-0.0254	-0.2902**	-0.2660**
	(0.0677)	(0.0995)	(0.1301)	(0.1354)
Favourite	-0.0950	-0.1232	0.1409	-0.3859*
	(0.0945)	(0.1347)	(0.2105)	(0.2329)
Promoted	0.0674	-0.0108	0.1023	0.1947
	(0.0650)	(0.0922)	(0.1299)	(0.1316)
Success	-0.0080	0.0194	-0.0749	0.0029
	(0.0237)	(0.0332)	(0.0504)	(0.0468)
DiffAttend	0.0028	0.0004	0.0079**	0.0019
	(0.0019)	(0.0033)	(0.0038)	(0.0030)
Observations	8,106	4,048	2,018	2,040

6.2.2 Return on Investment

Table 20: ROI results for the top five league Asian Handicap markets.

Top 5 AH ROI

	OVERALL	PRE-COVID	COVID	POST-COVID
ENG: HANDICAP ADVANTAGE & WORSE FORM	1.13	−2.14	6.32	2.60
ENG: WORSE FORM	1.92	−0.65	7.03	2.13
GER: HANDICAP ADVANTAGE	−0.68	−0.92	1.04	−1.89
GER: HANDICAP ADVANTAGE & LONGSHOT	2.60	5.49	−0.10	0.51
GER: LONGSHOT	−0.87	1.98	−4.47	−2.78
ITA: AWAY	−0.12	−2.91	3.96	1.38
ITA: AWAY & LONGSHOT	0.66	−3.09	4.69	3.89
ITA: AWAY & SENTIMENT	1.39	−2.20	7.54	2.24
ITA: SENTIMENT	−2.08	−2.97	0.37	−2.73

There are multiple strategies in the Asian Handicap market that leads to positive ROIs and therefore indicates potential violations of weak efficiency. Table 20 contains the strategies that have at least one period of positive ROI. The table also shows that returns in the post-pandemic period are generally lower than those observed during the Covid-19 period. Moreover, no strategy exhibits consistently positive returns across all sub-samples, although several strategies are profitable when considering the full sample.

Interestingly, in Italy, three different strategies using the away bias exhibits positive ROIs both during and after the pandemic. Two of these strategies also yields positive ROI in the full sample.

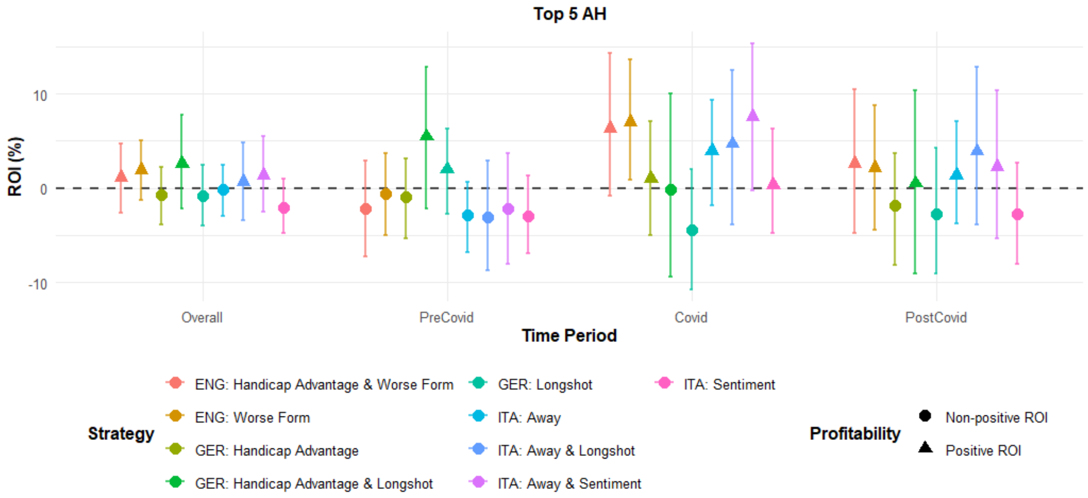
6.2.3 Bootstrapping

To quantify the uncertainty around the estimated returns and to assess the robustness of the Asian Handicap betting strategies across leagues and time periods, we apply the same bootstrapping method as in the Home/Draw/Away analysis.

Across the Asian Handicap markets, the bootstrapped results show that only one strategy produces a 95% confidence interval entirely above zero. The strategy of betting on teams with worse recent form in the English market under the pandemic. Since the confidence interval lies entirely above zero, the strategy yields a statistically significant positive return on investment, which constitutes a statistically significant violation of weak efficiency.

For all other Asian Handicap based betting strategies, the confidence intervals include zero. Hence, while several biases are identified in the Asian Handicap markets, only one strategy yields robust evidence of positive expected returns. This is illustrated in Figure 15.

Figure 15: Bootstrap results for the Asian Handicap markets.



7 Discussion

In this thesis, we have used a logistic regression to identify biases in two types of football betting markets and to assess whether such biases translate into economically exploitable inefficiencies. This discussion evaluates the findings within the framework of Thaler and Ziemba (1988), and relates it with existing literature.

7.1 Comparing the Home/Draw/Away Market Efficiency

7.1.1 Bias

Based on the logistic regression results for the Home/Draw/Away markets, the top five leagues, besides Italy, exhibit no evidence of persistent biases when considering the full sample period. The logistic regression results indicate that the implied probability is the only variable that remains consistently statistically significant, suggesting that the implied probability is a strong predictor of match outcomes.

In the Italian market, a negative $Home_i$-coefficient appears marginally significant in the full sample. However, this effect might be driven by the Covid-19 subsample, because it only exists in this subperiod. This suggests that the observed full-sample effect does not represent a persistent bias, but rather a temporary deviation during an extraordinary period. As argued by Winkelmann et al. (2024), the presence of biases in isolated subperiods does not imply that the biases exist in the full sample, even if the full-sample logistic regression displays statistically significant coefficients.

Based on this, the results indicate that none of the markets exhibit persistent biases across the entire sample period. However, several markets exhibit temporary home/away-biases, as well favourite bias, longshot bias, sentiment bias, and momentum bias. Our results also suggest that promoted teams are, in certain periods, priced more favourably than implied by realised outcomes. This bias has been discussed earlier by Deutscher et al. (2018). While their analysis is limited to the German market, we observe this effect only in the Spanish market during the pandemic.

In Germany and Spain, we find evidence of a reversal of the longshot bias. In both markets, a longshot bias is observed in one period, which subsequently shifts into a favourite bias in the following period. Notably, the timing of this reversal coincides with a possible Covid-19 disruption. In Spain, the shift occurs from the pre-pandemic to the pandemic period, while in Germany it occurs from the pandemic to the post-pandemic period. Why the reversal of bias occurs remains unclear and would be interesting to study further. However, one possibility could be that the bookmakers became observant of the longshot bias. Therefore, they might have overcorrected the odds, leading to the reversal.

The Danish is the only country in this study where no biases appear to exist. In the other Nordic markets, multiple biases are observed. Both a favourite bias and an away bias is present in Finland while the Norwegian and Swedish league exhibits biases, such as longshot-, momentum- and sentiment bias.

In Norway, there is strong evidence of biases occurring before and during Covid-19, indicating that the Norwegian market exhibited pronounced pricing distortions during these periods. However, all these biases seem to disappear *post-Covid*, suggesting that the market has been corrected in. In the full sample, a small sentiment bias is observed at 10% significance level.

Although, biases are statistically significant based on the logistic regression results, the significance level could be of importance. Winkelmann et al. (2024) discusses that higher significance level increases the likelihood of this being random noise. Considering this, there are some of the biases we have found that could just be random noise. In particular, biases identified at the 10% significance level should be interpreted with caution. Consequently, some biases observed in markets such as Italy, France, and Spain may not be robust. Similarly, in the Nordic leagues, the sentiment bias identified in Norway may be questioned. In the Finnish market, both observed biases may be subject to the same concern. Moreover, given that we estimate a large number of logistic regression models, it is expected that some statistically significant results arise by chance, increasing the likelihood of false positives.

The Danish market appears to be the most plausible candidate for efficiency, as no biases are detected. However, this should not be interpreted as proof of the absence of biases, since not all possible sources of inefficiency are examined.

7.1.2 Return of Investments

While several markets exhibit temporary biases, economically exploitable inefficiencies seem considerably rarer. The results show eleven strategies yielding a positive ROI, four of them being in the Nordic markets.

A notable observation is that positive returns sometimes arise in periods where the logistic regression does not identify statistically significant biases. This is in line with what Winkelmann et al. (2024) found in their paper. For example, in both England and Norway, strategies based on biases identified in one period yield positive returns in another period, even when the corresponding biases are not statistically significant. This may indicate that the logistic regression model lacks sufficient statistical power to detect all biases. Large standard errors may obscure small biases, resulting in false negatives. These false negatives may correspond with a strategy yielding positive return on investment.

Winkelmann et al. (2024) noted that betting markets are known to exhibit small but persistent pricing inefficiencies that are difficult to detect econometrically without very large samples. Consequently, some biases may be present but too subtle to reach statistical significance within individual estimation windows, while still generating positive ROI estimates in the return analysis.

In England, we also observe a positive average ROI for the full sample period. However, a positive point estimate alone does not imply violation of weak efficiency unless the return is statistically robust. In this case, the positive full-sample ROI appears to be largely driven by returns during the Covid-19 period, suggesting that extraordinary market conditions rather than persistent mispricing may explain the result.

Although some strategies yield positive ROI estimates, the bootstrap analysis shows that none of these returns are statistically robust, as all confidence intervals include zero or negative values. Consequently, there is no clear evidence that any of the markets exhibit violations of weak efficiency.

Overall, while both Nordic- and top five league markets exhibit indications of biases, none of the strategies based on these are statistically robust a 95% level. Consequently, we cannot determine whether Nordic markets are systematically more or less weak efficient than the top five league markets. However, examining confidence intervals at a lower significance level, such as 80%, may provide additional insights and should be explored in future research.

Examining how markets reacted to the Covid-19 pandemic, a higher prevalence of biases is observed during and after this period. This is also observed analysing the ROIs, a large share of strategies in the top five leagues generate returns exceeding zero. Although these confidence intervals overlap zero, the pattern is suggestive of a temporary weakening in market efficiency. This is not surprising, as the pandemic posed a significant challenge for bookmakers. During Covid-19, bookmakers had to assess how the absence of spectators affected home advantage, how sudden quarantines influenced team strength, and how broader economic disruptions impacted clubs' performance. Following the pandemic, bookmakers may have need to readjust to normal conditions, potentially leading to temporary biases in certain markets, particularly in the Nordic leagues.

7.1.3 Comparison with Existing Literature

Relating to the existing literature, the most comparable study to this thesis is Winkelmann et al. (2024). Our findings are broadly in line with their results, particularly with respect to the occurrence of positive ROI estimates in periods where the logistic regression analysis does not identify statistically significant biases.

As for the other research discussed in Section 3.1, even though they use other methods than Winkelmann et al. (2024) and us, we also find evidence of biases. Some previous studies, such as Vlastakis et al. (2009), employ forecasting models that generate positive returns, whereas our thesis do not identify similar opportunities. While much of the earlier literature documents evidence of bias, it is worth noting that these studies rely on older data. As discussed in the Section 2.1, the betting market has evolved quite a lot since the early 2000s, with the increasing access to data and information, this could be a reason for them finding consistent biases and our study does not.

Hegarty and Whelan (2024a) do not find strong evidence of a favourite bias in England, Germany, or Spain, which contrasts with some of our findings. However, their analysis is based on the full sample period, which may mask short-lived biases. In contrast, our results indicate that such biases primarily emerge in specific sub-periods, most notably during the Covid-19 period.

Angelini and De Angelis (2019) report results that largely align with ours, identifying bias across most of the leagues they examine, similarly to us. However, they document a pronounced favourite bias, this effect appears less dominant in our results. One possible explanation is that their study does not explicitly account for other types of biases, which may instead be subsumed within the favourite effect. In contrast, our model incorporates a broader set of explanatory variables, potentially diluting the estimated impact of any single bias. This could be further investigated using subset selection methods to identify the combination of explanatory variables that best captures the underlying biases in betting markets.

As noted earlier, methodological approaches and theoretical frameworks used to identify betting market biases vary across studies. These differences impose clear limitations on the extent to which results can be directly compared. Nevertheless, examining whether biases emerge across different settings can still provide valuable insights, even if the magnitude or statistical significance of the effects is not perfectly comparable.

7.2 Comparing Efficiency in the Home/Draw/Away Markets with the Asian Handicap Markets

7.2.1 Biases and Return of Investments

A clear difference between the Home/Draw/Away market and the Asian Handicap market is the lack of consistent statistical significance in both the constant and the $ImProb_i$ coefficient in the Asian Handicap market. A possible reason for this is that the implied probability calculated from the odds in Asian Handicap only varies between about 40% and 60%, whereas in the Home/Draw/Away market it varies from about 10% to 90%. Since the range of implied probabilities in the Asian Handicap market is substantially narrower, this variation may be insufficient variation in the explanatory variable to identify its effect in the logistic regression. As a consequence, standard errors may be larger, causing the variable to be statistically non-significant, even if implied probabilities are in fact informative. This interpretation is consistent with the low adjusted pseudo-R^2 values observed for the Asian Handicap models, indicating limited overall explanatory power. In this setting, a larger number of observations may be necessary to reliably identify the underlying relationships. Figure 11 in Section 5.2 shows that the relationship between implied probabilities and economic gains does not behave as expected.

Differences in sample composition may contribute to variation in the presence of biases across the two markets. Matches excluded from the Asian Handicap dataset due to a zero handicap may be systematically related, for example by involving teams of relatively equal strength. Consequently, direct comparisons between the Asian Handicap and Home/Draw/Away markets should be interpreted with caution.

As discussed in the previous section, we find no evidence of persistent biases in either the Home/Draw/Away or the Asian Handicap market when considering the full sample period. Nevertheless, several temporary biases are identified. In particular, the positive handicap bias arises exclusively in the Asian Handicap market and is observed in both Spain and Germany. This bias cannot occur in the Home/Draw/Away market due to the absence of a handicap structure. By contrast, a promoted bias is present in the Home/Draw/Away market but does not appear in the Asian Handicap market.

When evaluating economic outcomes using return on investment, we find that several strategies in the Asian Handicap market yield positive returns, particularly during the pandemic period. However, all strategies, except one, lack sufficient robustness to conclude that expected returns are statistically greater than zero. The sole exception is the strategy of betting on teams with worse form than their opponents in England during the pandemic. In this case, we identify one short-lived bias in the Asian Handicap market that yields positive returns in a specific market and period. This strategy also yields a positive ROI in the Home/Draw/Away market. However, in that case the bootstrap confidence interval includes zero. We also observe that the variance in the bootstrap results is substantially lower for the Asian Handicap market. This is expected given the greater range of payoff outcomes, which include refunds, in the Asian Handicap market relative to the simple win–loss structure of the Home/Draw/Away market.

7.2.2 Comparison with Existing Literature

Relating to existing literature, Hegarty and Whelan (2025) find that the Asian Handicap market exhibits limited evidence of systematic biases, including no evidence of a favourite- or a longshot bias. In contrast, our results reveal the presence of a longshot bias in Germany within the Asian Handicap market. Moreover, we find that over short time periods, the Asian Handicap markets exhibit the presence of temporary biases and, in one case, economically exploitable returns. One possible explanation for this discrepancy is that Hegarty and Whelan (2025) analyse the full sample as a single aggregate and only investigates favourite bias.

Constantinou (2022) finds inefficiencies in the Asian Handicap market, where betting strategies generate profits that vary over time. He also finds that the Asian Handicap market exhibits a lower variance in return on investment and argues that this is due to the different possible outcomes in this market, which is in line with our findings. However, our analysis only finds one strategy yielding statistically positive returns. A possible reason for the differences is the methods used in the analyses

7.3 Limitations

There are certain weaknesses in our analysis. One limitation is how we handled the attendance data during the *Covid*-period. We assumed that attendance figures from the season before the lockdown could be used as a proxy for attendance during the pandemic years. This approach may introduce measurement errors, as attendances could potentially change substantially across the season. In particular, team performance, promotion, or player transfers may affect the attendance and the supporter interest. In some cases, a team could upgrade their stadium, leading to a spike in attendance, or even a decrease if the stadium is under construction.

Despite this limitation, we have included $DiffAttend_i$ as a measure of sentiment. We assume that relative differences in fan support across clubs are sufficiently persistent over time to serve as a reasonable proxy. For instance, Real Madrid would still have a large fan base even though fans were not allowed to attend matches.

We have also tested the implications of including the variable, by conducting an additional logistic regression where the variable $DiffAttend_i$ was excluded. In the Home/Draw/Away markets, changes in the significance levels are observed for the $Home_i$-variable in Finland and Italy. In addition, both the favourite- and the longshot bias is no longer present in Spain. This additional logistic regression can be found in Appendix H.

In the Asian Handicap markets, the changes affect the leagues in England, Spain, and Italy. In England, $Promoted_i$ gains a small level of significance; in Italy, the level of significance increases from 5% to 1% for Home and in Spain, PH_i loses its significance for the overall period. These are found in Appendix J.

A similar limitation concerns the construction of the variable $Success_i$, which requires two assumptions. First, it assumes that end-of-season form provides a reasonable proxy for form at the beginning of the subsequent season. This assumption may be violated due to changes in squad composition, coaching staff, or injury status between seasons. For example, if Manchester United ends a season with four straight wins, it may be unrealistic to assume they will start the next season with similarly strong form.

The second assumption is that promoted teams have been assigned zero wins at the start of the season. This could understate their recent performance, given that promotion typically reflects strong prior form. However, this issue affects only the first four matchdays and therefore diminishes quickly. Moreover, the inclusion of a promoted-team dummy is expected to capture much of the remaining variation associated with promotion status. Despite these limitations, $Success_i$ is included as a coarse measure of momentum, and its role is evaluated through robustness checks excluding the variable.

The robustness analysis of $Success_i$ is presented in Appendix G and shows no major differences for the Home/Draw/Away markets. The only changes are that the coefficient for the $Favourite_i$ variable in Finland shifts from being significant at the 10% level to the 5% level, while in Spain the $Favourite_i$variable loses its 10% significance.

In the Asian Handicap markets, the coefficient most affected by the removal of the $Success_i$ explanatory variable is PH_i. In England, where $Success_i$ was previously significant, PH_i becomes significant. In Germany, the significance of PH_i also increases. This suggests that $Success_i$ correlate with $Favourite_i$ in Home/Draw/Away markets, and PH_i in Asian Handicap. The result of this can be found in Appendix I.

Another potential weakness concerns the assumptions made when calculating the implied probabilities. Specifically, we assume that the bookmaker's margin is evenly distributed across all odds. However, this assumption may not hold in practice, as discussed by Myrsten and Hwang (2021). Bookmaker maybe be incentivized to adjust the margin on the longshot, since a change in implied probability from 5% to 7% corresponds to odds shifting from 20.0 to 14.3. This reduction greatly lowers the potential payout, where the longshot to win. Without access to bookmakers' proprietary pricing strategies or margin allocation mechanisms, it is not possible to adjust for this potential distortion.

Consequently, by maintaining the equal distribution assumption, the analysis may understate the true probabilities of longshots and thereby partially amplify patterns consistent with favourite- or longshot bias. As a result, part of the observed favourite- and longshot biases in the data may reflect bookmakers' pricing strategies rather than genuine mispricing or systematic errors in the underlying probability assessments.

A limitation specific to the Asian Handicap analysis is the variation in handicap lines across matches. In our dataset, only a single handicap is observed for each match, which restricts the ability to examine how outcomes would differ under alternative handicap lines. This limitation may affect the biases we identify, particularly for the positive handicap variable, as it is constructed as a dummy based on the observed handicap.

8 Conclusion

8.1 Main Findings

This thesis has explored the European football betting market, with a particular focus on biases and the extent to which such biases translate into economically exploitable inefficiencies. This has been done by examining bookmaker odds for several types of biases and assessing whether these biases can be exploited to generate positive economic returns. Our research also examined whether the degree of exploitability changed before, during, and after the Covid-19 pandemic.

Our results show that many of the findings reported by Winkelmann et al. (2024) continue to hold, with biases fluctuating across periods, and appearing not only in the top five leagues but also in the Nordic markets. This thesis also builds on their analysis by extending the methodology to the Asian Handicap markets and by incorporating a bootstrap procedure to assess whether positive ROI estimates are statistically robust.

Using Thaler and Ziemba's (1988) weak efficiency criterion, we find that the Home/Draw/Away markets exhibit weak efficiency, as no systematically profitable betting strategies are observed before, during, or after Covid-19. The logistic regression analysis suggests that biases were more prevalent during the pandemic in the top five leagues, consistent with a temporary disruption in market conditions. Nevertheless, once uncertainty is accounted for through bootstrapping, none of the strategies deliver confidence intervals entirely above zero. Therefore, while the Covid-19 period is associated with more frequent temporary biases, these patterns do not translate into statistically robust positive expected returns. Likewise, we do not find conclusive evidence that the Nordic league markets violate weak efficiency and therefore cannot conclude that they are either more or less efficient than the top five leagues.

Compared with the Home/Draw/Away markets, the Asian Handicap markets exhibit similar instances of positive ROI point estimates across leagues, particularly around the *Covid*-period. In the *post-Covid* period, the two markets exhibit notable differences, with the Asian Handicap markets contain several profitable strategies. However, the bootstrap analysis shows that only one strategy in the Asian Handicap market yields a 95% confidence interval entirely above zero: betting on teams with worse recent form than their opponents in England during the pandemic. Under Thaler and Ziemba's (1988) framework, this constitutes evidence of a short-lived violation of weak efficiency in that specific market and period. Even though we identify a single temporary instance of a violation of weak efficiency, we cannot conclude that the Asian Handicap markets are less efficient than the Home/Draw/Away markets.

Taken together, this study finds that although biases in European football betting markets occur across leagues and time periods, they generally do not provide a basis for systematically profitable betting strategies. Overall, the findings support the conclusion that European football betting markets are largely weak efficient, with only limited and temporary deviations.

8.2 Recommendation for Further Research

Further research should build on addressing some of the assumptions and limitations of this thesis. One assumption is that the odds used are averages of all bookmaker odds provided in the dataset. Analysing odds from a single bookmaker, or comparing pricing behaviour across bookmakers, could provide deeper insights into market efficiency and bookmaker-specific strategies.

In addition, this thesis did not perform a season-by-season analysis during the pandemic. Winkelmann et al. did put forward the argument that of short periods affecting a larger period. Therefore, conducting an in-season analysis during the *Covid*-period is a natural step to take assessing the effects of the pandemic in even more detail. Some biases also appear to reverse between periods. Investigating the underlying reasons for this would be an interesting avenue for future research.

Our bootstrap analysis discovers that the confidence intervals for some of the positive and negative point estimates are both above and below zero. Future research could build on this by moving beyond a purely significance-based interpretation and instead provide a qualitative classification of betting strategies based on the sign, magnitude, and stability of their estimated returns. This would help identify which strategies appear consistently loss-making and which may warrant further investigation, as well as provide a clearer profile of their associated risks.

Beyond this, focusing on investigating semi-professional league betting markets could provide valuable insights into how market quality varies across competition levels. These competitions often have much lower liquidity, limited media coverage, and fewer professional participants, which may increase their vulnerability to mispricing and strategic behaviour.

Furthermore, the methodology used in this thesis could be applied to sports beyond football. Comparative studies across different sports may help determine whether the findings are sport-specific or driven by broader market mechanisms.

Finally, this thesis restricts the analysis of the Asian Handicap market to the handicaps available in the dataset. However, bookmakers usually provide multiple handicap alternatives with corresponding odds adjustments. Examining market responses across these alternative handicaps represents a promising step for further research.

Declaration on the use of AI tools in the work on this master's thesis

Declaration on the use of AI tools in the work on this master's thesis

Name (and version) of the AI tool: ChatGPT, 5.1 and 5.2

Purpose of using the tool: The AI tool was used in improvement of grammar and sentences, proposing relevant articles and websites, and improvement of code used to generate logistic regressions, plots and tables.

We are aware that we are responsible for all content of this master's thesis, including the parts where AI tools are used. We are responsible for ensuring that the thesis complies with ethical rules for privacy and publication.

References

Albertsen, K. (2017). *Efficiency of the Football Betting Market*. Copenhagen Business School. https://research.cbs.dk/en/studentProjects/b52ed495-87b0-4044-b5e9-cb201d49603f/

Angelini, G., & De Angelis, L. (2019). Efficiency of online football betting markets. *International Journal of Forecasting*, *35*(2), 712–721. https://doi.org/10.1016/j.ijforecast.2018.07.008

Betting and Gaming Act 1960. Retrieved 16 December 2025, from https://www.legislation.gov.uk/ukpga/Eliz2/8-9/60/contents/enacted

Botchkarev, A., & Andru, P. (2011). A Return on Investment as a Metric for Evaluating Information Systems: Taxonomy and Application. *Interdisciplinary Journal of Information, Knowledge, and Management*, *6*, 245–269. https://doi.org/10.28945/1535

Bulski, I. (2020). *The Earliest Forms of Gambling in Ancient Egypt*.

Cain, M., Law, D., & Peel, D. (2003). The Favourite-Longshot Bias, Bookmaker Margins and Insider Trading in a Variety of Betting Markets. *Bulletin of Economic Research*, *55*(3), 263–273. https://doi.org/10.1111/1467-8586.00174

Constantinou, A. C. (2022). Investigating the efficiency of the Asian handicap football betting market with ratings and Bayesian networks. *Journal of Sports Analytics*, *8*(3), 171–193. https://doi.org/10.3233/JSA-200588

Cortis, D. (2015). EXPECTED VALUES AND VARIANCES IN BOOKMAKER PAYOUTS: A THEORETICAL APPROACH TOWARDS SETTING LIMITS ON ODDS. *The Journal of Prediction Markets*, *9*(1), 1–14. https://doi.org/10.5750/jpm.v9i1.987

Delasport. (2025). *The History of Sports Betting: From Antiquity to 2025*. https://www.delasport.com/history-of-sports-betting/

Deutscher, C., Frick, B., & Ötting, M. (2018). Betting market inefficiencies are short-lived in German professional football. *Applied Economics*, *50*(30), 3240–3246. https://doi.org/10.1080/00036846.2017.1418082

EGBA. (2025). EGBA - EU Market. *EGBA*. https://www.egba.eu/eu-market/

European Football Statistics. (2025). *EFS Attendances*. https://www.european-football-statistics.co.uk/attn.htm

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

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

Football-Data. (2025a). *Denmark Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/denmark.php

Football-Data. (2025b). *England Football Results Betting Odds | Premiership Results & Betting Odds* [Data set]. https://www.football-data.co.uk/englandm.php

Football-Data. (2025c). *Finland Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/finland.php

Football-Data. (2025d). *Football Results, Statistics & Soccer Betting Odds Data*. https://www.football-data.co.uk/data.php

Football-Data. (2025e). *France Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/francem.php

Football-Data. (2025f). *Germany Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/germanym.php

Football-Data. (2025g). *Italy Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/italym.php

Football-Data. (2025h). *Norway Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/norway.php

Football-Data. (2025i). *Notes for Football Data*. https://www.football-data.co.uk/notes.txt

Football-Data. (2025j). *Spain Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/spainm.php

Football-Data. (2025k). *Sweden Football Results and Betting Odds* [Data set]. https://www.football-data.co.uk/sweden.php

Forrest, D., & Simmons, R. (2008). Sentiment in the betting market on Spanish football. *Applied Economics*, *40*(1), 119–126. https://doi.org/10.1080/00036840701522895

Franck, E., Verbeek, E., & Nüesch, S. (2010). Prediction accuracy of different market structures—Bookmakers versus a betting exchange. *International Journal of Forecasting*, *26*(3), 448–459. https://doi.org/10.1016/j.ijforecast.2010.01.004

Franke, M. (2020). Do market participants misprice lottery-type assets? Evidence from the European soccer betting market. *The Quarterly Review of Economics and Finance*, *75*, 1–18. https://doi.org/10.1016/j.qref.2019.05.016

Freeborn, J. (2020, April 29). France's Ligue 1 cancels remainder of 2019-20 season. *The Sports Daily*. https://thesportsdaily.org/news/france-cancels-the-remainder-of-their-professional-soccer-season/

Goolsbee, A., Levitt, S. D., & Syverson, C. (2020). *Microeconomics* (Third edition). Macmillan International Higher Education.

Grand View Research. (2025). *Sports Betting Market Size & Share | Industry Report, 2030*. https://www.grandviewresearch.com/industry-analysis/sports-betting-market-report

Harris, B., Holden, J. T., & Fried, G. (2024). *The Business of Sports Betting*. Human Kinetics.

Hegarty, T., & Whelan, K. (2024a). Comparing two methods for testing the efficiency of sports betting markets. *Sports Economics Review*, *8*, 100042. https://doi.org/10.1016/j.serev.2024.100042

Hegarty, T., & Whelan, K. (2024b). Returns on complex bets: Evidence from Asian Handicap betting on soccer. *Review of Behavioral Finance*, *16*(5), 904–924. https://doi.org/10.1108/RBF-11-2023-0314

Hegarty, T., & Whelan, K. (2025). Forecasting soccer matches with betting odds: A tale of two markets. *International Journal of Forecasting*, *41*(2), 803–820. https://doi.org/10.1016/j.ijforecast.2024.06.013

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2022). *An introduction to statistical learning: With applications in R* (Second edition). Springer.

Kuypers, T. (2000). Information and efficiency: An empirical study of a fixed odds betting market. *Applied Economics*, *32*(11), 1353–1363. https://doi.org/10.1080/00036840050151449

Lahvicka, J. (2013). What Causes the Favorite-Longshot Bias? Further Evidence from Tennis. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.2287335

Levitt, S. D. (2004). Why are Gambling Markets Organised so Differently from Financial Markets? *The Economic Journal*, *114*(495), 223–246. https://doi.org/10.1111/j.1468-0297.2004.00207.x

Long, J. S. (1997). *Regression models for categorical and limited dependent variables*. Sage Publications.

Lyhagen, J. (2025). The home advantage and COVID-19: The crowd support effect on the english football premier league and the championship. *Computational Statistics*, *40*(4), 1919–1932. https://doi.org/10.1007/s00180-025-01600-x

Myrsten, M., & Hwang, I. (2021). *Favorite-Longshot Bias and Market Efficiency in the Soccer Betting Market*. https://www.researchgate.net/publication/351985837_Favorite-Longshot_Bias_and_Market_Efficiency_in_the_Soccer_Betting_Market

Play the Percentage. (2025). *What Is Asian Handicap, Corners & Goal Lines Betting?* Inplay Football Stats: Get Better Live Inplay Predictions. https://playthepercentage.com/blog/asian-betting-explained?utm_source=chatgpt.com

Pope, P. F., & Peel, D. A. (1989). Information, Prices and Efficiency in a Fixed-Odds Betting Market. *Economica*, *56*(223), 323–341. https://doi.org/10.2307/2554281

Smarkets. (2025). *What is Asian Handicap betting?* Smarkets Help Centre. https://help.smarkets.com/hc/en-gb/articles/115001324169-What-is-Asian-Handicap-betting

Snowberg, E., & Wolfers, J. (2010). Explaining the Favorite–Long Shot Bias: Is it Risk-Love or Misperceptions? *Journal of Political Economy*, *118*(4), 723–746. https://doi.org/10.1086/655844

Statista Market Insights. (2025a). *Gambling—Worldwide | Statista Market Forecast*. Statista. http://frontend.xmo.prod.aws.statista.com/outlook/amo/gambling/worldwide?currency=EUR

Statista Market Insights. (2025b). *Sports Betting—United Kingdom | Statista Market Forecast*.Statista. http://frontend.xmo.prod.aws.statista.com/outlook/amo/gambling/sports-betting/united-kingdom

Thaler, R. H., & Ziemba, W. T. (1988). Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries. *Journal of Economic Perspectives*, *2*(2), 161–174. https://doi.org/10.1257/jep.2.2.161

Ubøe, J. (2015). *Statistikk for økonomifag* (5. utg). Gyldendal akademisk.

Vlastakis, N., Dotsis, G., & Markellos, R. N. (2009). How efficient is the European football betting market? Evidence from arbitrage and trading strategies. *Journal of Forecasting*, *28*(5), 426–444. https://doi.org/10.1002/for.1085

Wheatcroft, E. (2021). Evaluating probabilistic forecasts of football matches: The case against the ranked probability score. *Journal of Quantitative Analysis in Sports*, *17*(4), 273–287. https://doi.org/10.1515/jqas-2019-0089

Winkelmann, D., Ötting, M., Deutscher, C., & Makarewicz, T. (2024). Are Betting Markets Inefficient? Evidence From Simulations and Real Data. *Journal of Sports Economics*, *25*(1), 54–97. https://doi.org/10.1177/15270025231204997

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (Seventh edition). Cengage Learning.

World Bank Group. (2025). *World Bank Open Data*. World Bank Open Data. https://data.worldbank.org

Appendences

Appendix A: Data

The data used in this master's thesis are obtained from the website https://www.football-data.co.uk/data.php.

To access data for a specific league, the user must first select the corresponding nation under

To access data for a specific league, the user must first select the corresponding nation under which the league is categorized. Each league page contains downloadable CSV files for every available season.

In addition, the file Notes.txt, which is available within each nation's section. This file contains detailed explanations of the variables in data.

Attendance data are obtained from https://www.european-football-statistics.co.uk/attn.htm. To access data for a specific league, the corresponding nation must first be selected. Within each nation's page, attendance figures for a given season are available by selecting the relevant season.

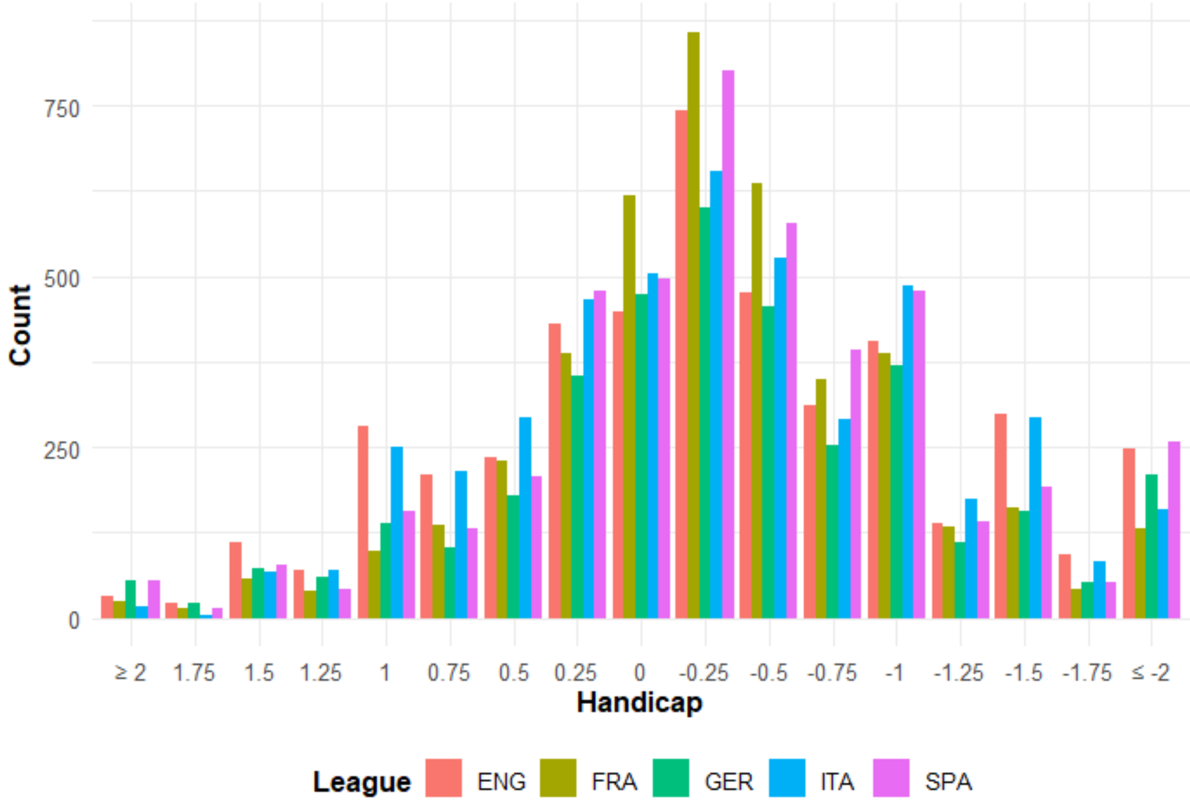
One point of note is that, although attendance data are available for Finland for the 2020 season, the reported average attendance figures are substantially lower than the season before and is therefore assumed to be missing data.

Appendix B: Asian Handicap Outcome Overview

Handicap	Team result	Bet result	Handicap	Team result	Bet result
	Win by 3+	Win		Win	Win
	Win by 2	Stake refund		Draw	Win
-2.00	Win by 1	Lose	**+2.00**	Lose by 1	Win
	Draw	Lose		Lose by 2	Stake refund
	Lose	Lose		Lose by 3+	Lose
	Win by 3+	Win		Win	Win
	Win by 2	50% Refunded		Draw	Win
-1.75		50% Win	**+1.75**		
	Win by 1	Lose		Lose by 1	Win
	Draw	Lose		Lose by 2	50% Refunded
					50% Lose
	Lose	Lose		Lose by 3+	Lose
	Win by 2+	Win		Win	Win
	Win by 1	Lose		Draw	Win
-1.5	Draw	Lose	**+1.5**	Lose by 1	Win
	Lose	Lose		Lose by 2+	Lose
	Win by 2+	Win		Win	Win
	Win by 1	50% Refunded		Draw	Win
-1.25		50% Lose	**+1.25**		
	Draw	Lose		Lose by 1	50% Refunded
					50% Win
	Lose	Lose		Lose by 2+	Lose
	Win by 2+	Win		Win	Win
	Win by 1	Stake refund		Draw	Win
-1.0	Draw	Lose	**+1.0**	Lose by 1	Stake refund
	Lose	Lose		Lose by 2+	Lose
	Win by 2+	Win		Win	Win

-0.75	Win by 1	50% Refunded 50% Win	**+0.75**	Draw	Win
	Draw	Lose		Lose by 1	50% Refunded 50% Lose
	Lose	Lose		Lose by 2+	Lose
	Win	Win		Win	Win
-0.5	Draw	Lose	**+0.5**	Draw	Win
	Lose	Lose		Lose	Lose
	Win	Win		Win	Win
-0.25	Draw	50% Refunded 50% Lose	**+0.25**	Draw	50% Refunded 50% Win
	Lose	Lose		Lose	Lose
	Win	Win		Win	Win
0	Draw	Stake refund	**0**	Draw	Stake refund
	Lose	Lose		Lose	Lose

Appendix C: Asian Handicap Distribution by Leagues



Appendix D: ROI for Home/Draw/Away Strategies in the Top Five Leagues

Top 5 H/D/A ROI	OVERALL	PRE-COVID	COVID	POST-COVID
ENG: WORSE FORM	0.13	0.49	11.48	−11.68
FRA: WORSE FORM	−4.70	−5.48	1.63	−9.54
GER: FAVOURITE	−4.26	−2.83	−10.71	−0.69
GER: LONGSHOT	−5.69	−4.41	2.77	−16.69
ITA: AWAY	−7.20	−8.28	3.03	−15.26
ITA: AWAY & SENTIMENT	0.36	−1.29	4.62	−0.59
ITA: SENTIMENT	−2.15	−1.14	−2.79	−3.55
SPA: FAVOURITE & PROMOTED	−8.17	−3.28	−2.54	−22.41
SPA: LONGSHOT	−9.62	−1.63	−15.47	−19.76
SPA: LONGSHOT & SENTIMENT	−5.73	−0.76	−7.46	−13.66
SPA: PROMOTED	−4.97	−4.45	1.15	−12.15
SPA: PROMOTED & SENTIMENT	6.05	1.07	22.43	1.52
SPA: SENTIMENT	−3.09	−4.31	−1.78	−1.95

Appendix E: ROI for Home/Draw/Away Strategies in the Nordic Leagues

Nordic H/D/A ROI				
	OVERALL	PRE-COVID	COVID	POST-COVID
FIN: AWAY	−7.04	−7.81	−3.27	−9.15
FIN: FAVOURITE	−2.75	−6.49	2.50	5.13
NOR: BETTER FORM	−6.30	−8.16	−1.24	−7.78
NOR: BETTER FORM & LONGSHOT	−12.38	−17.82	−2.38	−5.42
NOR: LONGSHOT	−10.31	−12.58	−12.89	1.51
NOR: SENTIMENT	−5.70	−3.60	−11.69	−4.05
SWE: LONGSHOT	−11.73	−17.85	−9.16	5.83
SWE: SENTIMENT	−2.10	2.05	−15.17	3.00

Appendix F: ROI for Asian Handicap Strategies in the Top Five Leagues

Top 5 AH ROI				
	OVERALL	PRE-COVID	COVID	POST-COVID
ENG: HANDICAP ADVANTAGE & WORSE FORM	1.13	−2.14	6.32	2.60
ENG: HANDICAP ADVANTGE	−2.85	−5.15	−0.84	−0.15
ENG: WORSE FORM	1.92	−0.65	7.03	2.13
GER: HANDICAP ADVANTAGE	−0.68	−0.92	1.04	−1.89
GER: HANDICAP ADVANTAGE & LONGSHOT	2.60	5.49	−0.10	0.51
GER: LONGSHOT	−0.87	1.98	−4.47	−2.78
ITA: AWAY	−0.12	−2.91	3.96	1.38
ITA: AWAY & LONGSHOT	0.66	−3.09	4.69	3.89
ITA: AWAY & SENTIMENT	1.39	−2.20	7.54	2.24
ITA: LONGSHOT	−2.61	−3.12	−2.13	−2.06
ITA: SENTIMENT	−2.08	−2.97	0.37	−2.73
SPA: HANDICAP ADVANTGE	−3.21	−1.08	−4.13	−6.72

Appendix G: Logistic Regressions Home/Draw/Away Betting without $Success_i$

Formula:

$$P(Won_i = 1|X_i) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 Home_i + \beta_3 Favourite_i + \beta_4 Promoted_i + \beta_6 DiffAttend_i)$$

Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01 for all regressions.

Denmark

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.573***	-2.789***	-2.420***	-2.355***
	(0.1492)	(0.2216)	(0.3084)	(0.3015)
ImProb	5.305***	6.170***	4.272***	4.824***
	(0.5186)	(0.7591)	(1.085)	(1.058)
Home	0.0432	0.0123	0.0936	0.0358
	(0.0847)	(0.1216)	(0.1690)	(0.1774)
Favourite	-0.0316	-0.2092	0.3872	-0.0876
	(0.1278)	(0.1757)	(0.2609)	(0.2744)
Promoted	0.0131	0.1223	-0.1017	-0.2544
	(0.0910)	(0.1213)	(0.2021)	(0.1984)
DiffAttend	0.0016	-0.0082	0.0066	0.0057
	(0.0075)	(0.0145)	(0.0160)	(0.0105)
Observations	5,100	2,686	1,258	1,156

Finland

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.477***	-2.412***	-2.701***	-2.248***
	(0.1458)	(0.1840)	(0.2906)	(0.4348)
ImProb	5.084***	4.878***	6.256***	3.445**
	(0.5196)	(0.6584)	(1.036)	(1.573)
Home	-0.0733	0.0373	-0.3494*	0.0518
	(0.0945)	(0.1217)	(0.2059)	(0.2353)
Favourite	0.1135	-0.0171	0.0628	0.7486**
	(0.1377)	(0.1712)	(0.3118)	(0.3808)
Promoted	-0.0400	0.0063	-0.1273	-0.1126
	(0.1056)	(0.1289)	(0.2421)	(0.2795)
DiffAttend	-0.0375	-0.0266	-0.0762	-0.0328
	(0.0370)	(0.0473)	(0.0753)	(0.1015)
Observations	4,300	2,710	922	668

Norway

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.466***	-2.334***	-2.978***	-2.200***
	(0.1111)	(0.1506)	(0.2364)	(0.2522)
ImProb	5.292***	4.485***	7.587***	4.834***
	(0.4156)	(0.5667)	(0.8708)	(0.9460)
Home	-0.0346	0.0543	-0.2141	-0.1038
	(0.0870)	(0.1170)	(0.1739)	(0.2131)
Favourite	-0.1336	0.1020	-0.6780***	-0.1746
	(0.1235)	(0.1623)	(0.2503)	(0.3068)
Promoted	-0.0341	-0.0313	-0.1693	0.1040
	(0.0866)	(0.1185)	(0.1715)	(0.1975)
DiffAttend	0.0066*	0.0082**	-0.0136	0.0111
	(0.0039)	(0.0042)	(0.0187)	(0.0202)
Observations	5,760	3,360	1,440	960

Sweden

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.603***	-2.641***	-2.565***	-2.464***
	(0.1164)	(0.1565)	(0.2333)	(0.2780)
ImProb	5.923***	5.897***	5.428***	6.331***
	(0.4284)	(0.5832)	(0.8225)	(1.033)
Home	-0.1145	-0.1478	-0.0452	-0.0610
	(0.0850)	(0.1163)	(0.1593)	(0.2072)
Favourite	-0.2571**	-0.1764	-0.1245	-0.6534**
	(0.1240)	(0.1639)	(0.2445)	(0.2997)
Promoted	-0.0220	0.0086	-0.0365	-0.0484
	(0.0900)	(0.1191)	(0.1778)	(0.2225)
DiffAttend	0.0019	0.0133*	-0.0127	0.0013
	(0.0048)	(0.0074)	(0.0090)	(0.0089)
Observations	5,760	3,360	1,440	960

England

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.371***	-2.473***	-2.051***	-2.484***
	(0.0824)	(0.1174)	(0.1592)	(0.1708)
ImProb	4.818***	5.153***	4.068***	4.956***
	(0.2880)	(0.4123)	(0.5448)	(0.6042)
Home	0.0015	0.0575	-0.0939	0.0063
	(0.0646)	(0.0949)	(0.1241)	(0.1278)
Favourite	0.0159	-0.1024	0.0762	0.1199
	(0.0990)	(0.1412)	(0.1954)	(0.2026)
Promoted	-0.0931	-0.0036	-0.1345	-0.2346
	(0.0728)	(0.1021)	(0.1408)	(0.1552)
DiffAttend	0.0006	0.0005	0.0051	-0.0029
	(0.0016)	(0.0024)	(0.0033)	(0.0030)
Observations	9,120	4,560	2,280	2,280

France

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.420***	-2.511***	-2.317***	-2.314***
	(0.0902)	(0.1229)	(0.1880)	(0.1925)
ImProb	5.024***	5.206***	4.980***	4.642***
	(0.3253)	(0.4475)	(0.6708)	(0.6916)
Home	-0.0960	-0.0591	-0.0892	-0.1791
	(0.0698)	(0.1029)	(0.1356)	(0.1371)
Favourite	0.0401	0.0415	-0.1378	0.2078
	(0.0982)	(0.1349)	(0.1987)	(0.2138)
Promoted	0.0055	-0.0118	0.0342	0.0336
	(0.0750)	(0.1003)	(0.1665)	(0.1576)
DiffAttend	-0.0004	-0.0013	0.0032	-0.0016
	(0.0022)	(0.0034)	(0.0042)	(0.0038)
Observations	8,622	4,560	2,078	1,984

Germany

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.456***	-2.483***	-2.585***	-2.304***
	(0.0919)	(0.1294)	(0.1923)	(0.1812)
ImProb	4.925***	4.924***	5.988***	4.000***
	(0.3199)	(0.4501)	(0.6833)	(0.6214)
Home	0.0502	0.1827*	-0.0078	-0.1760
	(0.0717)	(0.1032)	(0.1393)	(0.1459)
Favourite	-0.0387	-0.1035	-0.5164**	0.5213**
	(0.1053)	(0.1490)	(0.2205)	(0.2072)
Promoted	0.0321	0.1293	-0.0256	-0.0931
	(0.0865)	(0.1230)	(0.1617)	(0.1911)
DiffAttend	-0.0021	-0.0014	-0.0044	-0.0012
	(0.0015)	(0.0023)	(0.0030)	(0.0029)
Observations	7,344	3,672	1,836	1,836

Italy

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.546***	-2.659***	-2.219***	-2.696***
	(0.0905)	(0.1295)	(0.1834)	(0.1928)
ImProb	5.217***	5.609***	4.373***	5.522***
	(0.3218)	(0.4702)	(0.6277)	(0.6920)
Home	-0.1317**	-0.1156	-0.2210*	-0.1205
	(0.0665)	(0.1031)	(0.1251)	(0.1287)
Favourite	0.0654	0.0061	0.1296	0.0924
	(0.1013)	(0.1440)	(0.2051)	(0.2035)
Promoted	0.0051	0.0288	0.0414	-0.0761
	(0.0748)	(0.1061)	(0.1507)	(0.1486)
DiffAttend	0.0018	-0.0018	0.0083**	0.0004
	(0.0021)	(0.0040)	(0.0040)	(0.0034)
Observations	9,120	4,560	2,280	2,280

Spain

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.463***	-2.577***	-2.245***	-2.578***
	(0.0952)	(0.1378)	(0.2051)	(0.1875)
ImProb	4.786***	5.613***	3.373***	4.960***
	(0.3429)	(0.4971)	(0.7360)	(0.6830)
Home	0.0399	-0.0207	0.2076	-0.0518
	(0.0742)	(0.1103)	(0.1421)	(0.1440)
Favourite	0.0514	-0.2299	0.3306*	0.2683
	(0.0972)	(0.1405)	(0.1937)	(0.1955)
Promoted	0.0818	-0.0071	0.2683*	0.1089
	(0.0716)	(0.1028)	(0.1392)	(0.1419)
DiffAttend	0.0029	-0.0016	0.0092***	0.0021
	(0.0018)	(0.0029)	(0.0034)	(0.0033)
Observations	9,120	4,560	2,280	2,280

Appendix H: Logistic Regressions for Home/Draw/Away Betting without $DiffAttend_i$

Formula:

$$P(Won_i = 1|X_i) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 Home_i + \beta_3 Favourite_i + \beta_4 Promoted_i + \beta_5 Success_i)$$

Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01 for all regressions.

Denmark

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.630***	-2.772***	-2.448***	-2.497***
	(0.1402)	(0.1887)	(0.2983)	(0.3009)
ImProb	5.485***	6.100***	4.362***	5.273***
	(0.4909)	(0.6588)	(1.054)	(1.057)
Home	0.0197	0.0119	0.0853	-0.0157
	(0.0842)	(0.1169)	(0.1706)	(0.1757)
Favourite	-0.0262	-0.1978	0.3911	-0.0862
	(0.1280)	(0.1762)	(0.2613)	(0.2743)
Promoted	0.0117	0.1403	-0.1240	-0.2552
	(0.0907)	(0.1213)	(0.1967)	(0.1985)
Success	-0.0267	-0.0396	0.0127	-0.0433
	(0.0305)	(0.0428)	(0.0609)	(0.0637)
Observations	5,100	2,686	1,258	1,156

Finland

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.361***	-2.330***	-2.564***	-2.005***
	(0.1372)	(0.1699)	(0.2978)	(0.3953)
ImProb	4.687***	4.586***	5.853***	2.677*
	(0.4924)	(0.6061)	(1.082)	(1.472)
Home	-0.0215	0.0779	-0.3087	0.1559
	(0.0944)	(0.1215)	(0.2061)	(0.2357)
Favourite	0.1134	-0.0131	0.0432	0.7159*
	(0.1377)	(0.1712)	(0.3114)	(0.3819)
Promoted	-0.0208	0.0204	-0.1082	-0.0988
	(0.1042)	(0.1274)	(0.2407)	(0.2735)
Success	0.0353	0.0267	-0.0006	0.1119
	(0.0334)	(0.0415)	(0.0784)	(0.0842)
Observations	4,300	2,710	922	668

Norway

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.467***	-2.399***	-2.778***	-2.249***
	(0.1151)	(0.1570)	(0.2369)	(0.2497)
ImProb	5.286***	4.724***	6.844***	5.006***
	(0.4318)	(0.5927)	(0.8817)	(0.9388)
Home	-0.0278	0.0154	-0.0603	-0.1406
	(0.0904)	(0.1223)	(0.1794)	(0.2097)
Favourite	-0.1361	0.0873	-0.6907***	-0.1616
	(0.1235)	(0.1621)	(0.2484)	(0.3057)
Promoted	-0.0234	-0.0027	-0.1382	0.0786
	(0.0864)	(0.1176)	(0.1704)	(0.1953)
Success	0.0277	-0.0006	0.1120**	-0.0072
	(0.0278)	(0.0375)	(0.0525)	(0.0698)
Observations	5,760	3,360	1,440	960

Sweden

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.642***	-2.767***	-2.497***	-2.486***
	(0.1109)	(0.1468)	(0.2180)	(0.2740)
ImProb	6.057***	6.344***	5.226***	6.404***
	(0.4128)	(0.5527)	(0.7859)	(1.019)
Home	-0.1366	-0.2200*	-0.0299	-0.0731
	(0.0837)	(0.1130)	(0.1577)	(0.2056)
Favourite	-0.2554**	-0.1753	-0.1328	-0.6504**
	(0.1241)	(0.1644)	(0.2453)	(0.2995)
Promoted	-0.0282	-0.0079	-0.0111	-0.0561
	(0.0900)	(0.1187)	(0.1787)	(0.2190)
Success	-0.0152	-0.0007	-0.0588	-0.0063
	(0.0284)	(0.0390)	(0.0551)	(0.0649)
Observations	5,760	3,360	1,440	960

England

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.431***	-2.512***	-2.273***	-2.427***
	(0.0811)	(0.1136)	(0.1640)	(0.1658)
ImProb	5.023***	5.289***	4.776***	4.760***
	(0.2849)	(0.3998)	(0.5743)	(0.5800)
Home	-0.0298	0.0350	-0.1858	0.0288
	(0.0651)	(0.0950)	(0.1249)	(0.1295)
Favourite	0.0150	-0.1036	0.0964	0.1267
	(0.0991)	(0.1413)	(0.1948)	(0.2024)
Promoted	-0.0979	-0.0076	-0.1826	-0.2173
	(0.0726)	(0.1021)	(0.1395)	(0.1550)
Success	-0.0355	-0.0211	-0.0932**	-0.0054
	(0.0234)	(0.0335)	(0.0465)	(0.0462)
Observations	9,120	4,560	2,280	2,280

France

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.443***	-2.498***	-2.513***	-2.242***
	(0.0853)	(0.1144)	(0.1849)	(0.1811)
ImProb	5.100***	5.155***	5.614***	4.406***
	(0.3097)	(0.4199)	(0.6605)	(0.6515)
Home	-0.1099	-0.0513	-0.1811	-0.1480
	(0.0689)	(0.1007)	(0.1347)	(0.1357)
Favourite	0.0421	0.0415	-0.1154	0.2074
	(0.0983)	(0.1348)	(0.2001)	(0.2138)
Promoted	0.0081	-0.0041	0.0172	0.0379
	(0.0748)	(0.1000)	(0.1648)	(0.1577)
Success	-0.0231	-0.0080	-0.0889*	0.0241
	(0.0236)	(0.0326)	(0.0475)	(0.0492)
Observations	8,622	4,560	2,078	1,984

Germany

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.414***	-2.480***	-2.467***	-2.262***
	(0.0910)	(0.1299)	(0.1844)	(0.1818)
ImProb	4.782***	4.907***	5.605***	3.863***
	(0.3149)	(0.4475)	(0.6561)	(0.6195)
Home	0.0661	0.1810*	0.0302	-0.1559
	(0.0723)	(0.1037)	(0.1401)	(0.1485)
Favourite	-0.0326	-0.0988	-0.5038**	0.5213**
	(0.1052)	(0.1489)	(0.2205)	(0.2072)
Promoted	0.0442	0.1453	-0.0047	-0.0930
	(0.0864)	(0.1224)	(0.1616)	(0.1910)
Success	-0.0078	-0.0222	0.0033	0.0077
	(0.0261)	(0.0376)	(0.0510)	(0.0526)
Observations	7,344	3,672	1,836	1,836

Italy

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.588***	-2.608***	-2.461***	-2.694***
	(0.0822)	(0.1119)	(0.1709)	(0.1729)
ImProb	5.361***	5.426***	5.149***	5.515***
	(0.2987)	(0.4160)	(0.5966)	(0.6336)
Home	-0.1522**	-0.0836	-0.3098**	-0.1179
	(0.0650)	(0.0954)	(0.1251)	(0.1281)
Favourite	0.0678	0.0037	0.1436	0.0895
	(0.1013)	(0.1436)	(0.2053)	(0.2035)
Promoted	-0.0056	0.0364	-0.0201	-0.0780
	(0.0740)	(0.1047)	(0.1497)	(0.1469)
Success	0.0011	0.0125	-0.0293	0.0106
	(0.0240)	(0.0347)	(0.0478)	(0.0473)
Observations	9,120	4,560	2,280	2,280

Spain

	Won			
	All	PreCovid	Covid	PostCovid
Constant	-2.599***	-2.579***	-2.639***	-2.663***
	(0.0802)	(0.1083)	(0.1748)	(0.1681)
ImProb	5.280***	5.623***	4.793***	5.261***
	(0.2918)	(0.3963)	(0.6302)	(0.6167)
Home	-0.0431	-0.0293	0.0089	-0.1021
	(0.0688)	(0.0997)	(0.1307)	(0.1400)
Favourite	0.0450	-0.2269	0.3020	0.2712
	(0.0974)	(0.1403)	(0.1944)	(0.1963)
Promoted	0.0714	-0.0038	0.2316*	0.1064
	(0.0717)	(0.1030)	(0.1400)	(0.1419)
Success	-0.0330	-0.0432	-0.0476	-0.0114
	(0.0243)	(0.0350)	(0.0454)	(0.0511)
Observations	9,120	4,560	2,280	2,280

Appendix I: Logistic Regressions for Asian Handicap Betting without $Success_i$

Formula:

$$P(AHGain_i = 1|X_i) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 PH_i + \beta_3 Home_i + \beta_4 Favourite_i + \beta_5 Promoted_i + \beta_7 DiffAttend_i)$$

Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01 for all regressions.

England

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-1.616**	-1.532*	-4.120*	-0.3074
	(0.6759)	(0.7859)	(2.230)	(2.679)
ImProb	2.841**	2.569	8.137*	0.1899
	(1.432)	(1.690)	(4.650)	(5.575)
PH	0.1795**	0.1592	0.1709	0.2441
	(0.0804)	(0.1163)	(0.1599)	(0.1579)
Home	-0.0035	0.0253	-0.0560	-0.0117
	(0.0658)	(0.0947)	(0.1303)	(0.1306)
Favourite	0.0337	0.0654	-0.1512	0.1016
	(0.0955)	(0.1396)	(0.2144)	(0.2355)
Promoted	-0.0901	-0.0742	-0.0462	-0.1871
	(0.0647)	(0.0916)	(0.1294)	(0.1294)
DiffAttend	0.0012	0.0031	0.0002	-0.0012
	(0.0015)	(0.0022)	(0.0032)	(0.0030)
Observations	8,204	4,152	2,022	2,030

France

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.156***	-1.736*	-4.952*	-2.664
	(0.8059)	(0.9222)	(2.707)	(2.787)
ImProb	3.967**	2.978	9.796*	5.140
	(1.699)	(1.970)	(5.628)	(5.808)
PH	0.1001	0.1047	0.1130	0.1247
	(0.0908)	(0.1340)	(0.1782)	(0.1777)
Home	-0.0447	-0.0148	-0.0014	-0.1391
	(0.0757)	(0.1123)	(0.1488)	(0.1457)
Favourite	0.1142	0.1843	-0.1215	0.0614
	(0.1036)	(0.1448)	(0.2492)	(0.2470)
Promoted	0.0004	-0.0057	0.0861	-0.0295
	(0.0714)	(0.0962)	(0.1650)	(0.1432)
DiffAttend	0.0022	0.0045	0.0016	0.0004
	(0.0021)	(0.0032)	(0.0040)	(0.0038)
Observations	7,368	3,824	1,796	1,748

Germany

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.852***	-2.905***	-5.730*	-6.692**
	(0.7720)	(0.9030)	(2.977)	(2.814)
ImProb	5.353***	5.405***	11.35*	13.38**
	(1.629)	(1.939)	(6.182)	(5.840)
PH	0.2772***	0.3018**	0.3032*	0.2257
	(0.0881)	(0.1295)	(0.1686)	(0.1775)
Home	0.1022	0.2032*	-0.0108	0.0238
	(0.0770)	(0.1131)	(0.1482)	(0.1532)
Favourite	-0.1848*	-0.3389**	-0.2184	-0.3452
	(0.1038)	(0.1543)	(0.2565)	(0.2364)
Promoted	0.0202	0.0612	-0.1837	0.1368
	(0.0813)	(0.1182)	(0.1557)	(0.1631)
DiffAttend	9.46×10^{-5}	-0.0012	0.0004	0.0015
	(0.0015)	(0.0022)	(0.0029)	(0.0029)
Observations	6,378	3,122	1,612	1,644

Italy

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.440***	-2.642***	0.4921	-6.289**
	(0.6772)	(0.7563)	(2.153)	(2.558)
ImProb	4.800***	5.090***	-1.221	12.99**
	(1.439)	(1.631)	(4.490)	(5.353)
PH	0.0693	0.0484	0.1501	0.0567
	(0.0909)	(0.1376)	(0.1765)	(0.1769)
Home	-0.1496**	-0.0359	-0.2704**	-0.2674**
	(0.0667)	(0.0977)	(0.1292)	(0.1336)
Favourite	-0.0949	-0.1242	0.1326	-0.3859*
	(0.0945)	(0.1346)	(0.2088)	(0.2329)
Promoted	0.0680	-0.0126	0.0989	0.1944
	(0.0650)	(0.0922)	(0.1294)	(0.1315)
DiffAttend	0.0026	0.0008	0.0066*	0.0019
	(0.0019)	(0.0033)	(0.0037)	(0.0030)
Observations	8,106	4,048	2,018	2,040

Spain

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.556***	-2.703***	-2.553	-2.701
	(0.6747)	(0.7728)	(2.381)	(2.448)
ImProb	4.643***	4.734***	4.413	5.562
	(1.434)	(1.664)	(4.988)	(5.122)
PH	0.1868**	0.3283**	0.2231	-0.0988
	(0.0931)	(0.1341)	(0.1877)	(0.1828)
Home	0.0325	0.0834	0.1491	-0.1637
	(0.0717)	(0.1028)	(0.1421)	(0.1434)
Favourite	0.0085	-0.0618	0.0940	0.0099
	(0.0957)	(0.1352)	(0.2364)	(0.2288)
Promoted	0.0819	0.1081	0.1840	-0.0528
	(0.0640)	(0.0896)	(0.1289)	(0.1297)
DiffAttend	0.0018	0.0030	0.0017	-1.6×10^{-5}
	(0.0014)	(0.0020)	(0.0029)	(0.0028)
Observations	8,108	4,142	1,964	2,002

Appendix J: Logistic Regressions for Asian Handicap Betting without $DiffAttend_i$

Formula:

$$P(AHGain_i = 1|X_i) = G(\beta_0 + \beta_1 ImProb_i + \beta_2 PH_i + \beta_3 Home_i + \beta_4 Favourite_i + \beta_5 Promoted_i + \beta_6 Success)$$

Significance levels: * p < 0.10, ** p < 0.05, *** p < 0.01 for all regressions.

England

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-1.619**	-1.507*	-4.436**	-0.3525
	(0.6785)	(0.7875)	(2.243)	(2.682)
ImProb	3.056**	2.717	9.083*	0.4612
	(1.438)	(1.694)	(4.685)	(5.579)
PH	0.0280	0.0227	-0.0561	0.1110
	(0.0791)	(0.1135)	(0.1581)	(0.1572)
Home	-0.0558	-0.0250	-0.1239	-0.0574
	(0.0663)	(0.0951)	(0.1317)	(0.1321)
Favourite	0.0357	0.0660	-0.1619	0.1019
	(0.0956)	(0.1397)	(0.2155)	(0.2355)
Promoted	-0.1176*	-0.1119	-0.0764	-0.1978
	(0.0647)	(0.0915)	(0.1293)	(0.1299)
Success	-0.0700***	-0.0301	-0.1320***	-0.0966**
	(0.0228)	(0.0316)	(0.0484)	(0.0453)
Observations	8,204	4,152	2,022	2,030

France

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.228***	-1.787*	-5.118*	-2.812
	(0.8075)	(0.9218)	(2.711)	(2.788)
ImProb	4.256**	3.229	10.36*	5.592
	(1.703)	(1.965)	(5.640)	(5.810)
PH	0.0055	0.0138	-0.0472	0.0258
	(0.0858)	(0.1231)	(0.1751)	(0.1698)
Home	-0.0838	-0.0539	-0.0641	-0.1716
	(0.0757)	(0.1109)	(0.1503)	(0.1456)
Favourite	0.1078	0.1746	-0.1307	0.0416
	(0.1036)	(0.1446)	(0.2494)	(0.2470)
Promoted	-0.0187	-0.0373	0.0731	-0.0426
	(0.0713)	(0.0959)	(0.1648)	(0.1435)
Success	-0.0256	0.0063	-0.0652	-0.0562
	(0.0245)	(0.0336)	(0.0499)	(0.0522)
Observations	7,368	3,824	1,796	1,748

Germany

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.846***	-2.906***	-5.789*	-6.778**
	(0.7723)	(0.9028)	(2.981)	(2.834)
ImProb	5.321***	5.381***	11.34*	13.65**
	(1.629)	(1.941)	(6.180)	(5.874)
PH	0.2915***	0.3190**	0.4261**	0.1490
	(0.0899)	(0.1329)	(0.1754)	(0.1758)
Home	0.1072	0.2090*	0.0166	-0.0042
	(0.0774)	(0.1139)	(0.1483)	(0.1543)
Favourite	-0.1848*	-0.3394**	-0.2360	-0.3459
	(0.1038)	(0.1543)	(0.2573)	(0.2377)
Promoted	0.0206	0.0743	-0.1787	0.1302
	(0.0812)	(0.1181)	(0.1562)	(0.1629)
Success	0.0101	-0.0083	0.0787	-0.0216
	(0.0259)	(0.0371)	(0.0523)	(0.0514)
Observations	6,378	3,122	1,612	1,644

Italy

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.454***	-2.630***	0.3622	-6.426**
	(0.6791)	(0.7559)	(2.259)	(2.545)
ImProb	4.941***	5.043***	-0.6018	13.35**
	(1.443)	(1.631)	(4.712)	(5.313)
PH	-0.0137	0.0635	-0.1368	0.0015
	(0.0808)	(0.1164)	(0.1621)	(0.1596)
Home	-0.1718***	-0.0280	-0.3273**	-0.2807**
	(0.0664)	(0.0964)	(0.1288)	(0.1332)
Favourite	-0.0952	-0.1236	0.1388	-0.3923*
	(0.0946)	(0.1347)	(0.2150)	(0.2324)
Promoted	0.0446	-0.0136	0.0270	0.1768
	(0.0649)	(0.0921)	(0.1298)	(0.1315)
Success	-0.0015	0.0201	-0.0503	0.0079
	(0.0234)	(0.0327)	(0.0493)	(0.0462)
Observations	8,106	4,048	2,018	2,040

Spain

	AHGain			
	All	PreCovid	Covid	PostCovid
Constant	-2.580***	-2.687***	-3.030	-2.650
	(0.6754)	(0.7733)	(2.377)	(2.440)
ImProb	4.771***	4.771***	5.643	5.423
	(1.432)	(1.661)	(4.971)	(5.097)
PH	0.1307	0.2747**	0.0525	-0.0759
	(0.0809)	(0.1151)	(0.1589)	(0.1657)
Home	0.0131	0.0660	0.0833	-0.1550
	(0.0703)	(0.0999)	(0.1407)	(0.1413)
Favourite	0.0089	-0.0555	0.0589	0.0137
	(0.0958)	(0.1352)	(0.2367)	(0.2286)
Promoted	0.0708	0.0960	0.1695	-0.0502
	(0.0639)	(0.0894)	(0.1296)	(0.1297)
Success	0.0106	0.0425	-0.0647	0.0132
	(0.0227)	(0.0310)	(0.0480)	(0.0478)
Observations	8,108	4,142	1,964	2,002