



Munich Personal RePEc Archive

Forecasting Soccer Matches With Betting Odds: A Tale of Two Markets

Whelan, Karl and Hegarty, Tadgh

University College Dublin, University College Dublin

23 February 2023

Online at <https://mpra.ub.uni-muenchen.de/116925/>
MPRA Paper No. 116925, posted 06 Apr 2023 07:47 UTC

Forecasting Soccer Matches With Betting Odds: A Tale of Two Markets

Tadgh Hegarty* Karl Whelan[†]

University College Dublin

February 23, 2023

Abstract

We compare the properties of betting market odds set in two distinct markets for a large sample of European soccer matches. We confirm inefficiencies in the traditional market for bets on a home win, an away win or a draw as found in previous studies such as Angelini and De Angelis (2019), in particular that there is a strong pattern of favourite-longshot bias. Conversely, we document how a betting market that has emerged in recent years, the Asian handicap market, can generate efficient forecasts for the same set of matches using a new methodology for mapping its odds into probabilities.

Keywords: Sports Forecasting, Betting Markets, Market Efficiency, Asian Handicap

JEL Classification: G14, L83, Z20, Z21

*tadgh.hegarty@ucdconnect.ie. Thanks to Chris Jepsen for comments and Joseph Buchdahl for compiling the main data set and for assistance with information about the data.

[†]karl.whelan@ucd.ie. Corresponding author.

1. Introduction

Sports betting markets offer a good environment for testing the efficiency of markets in processing information and there is a large literature assessing the efficiency of various types of betting markets.¹ In this paper, we examine a large dataset of European soccer matches and compare the properties of odds set in the traditional market in which you can bet on either a home win, an away win or a draw with odds set in a large online betting market with a number of interesting features that has emerged in recent years—the Asian Handicap market.

Payouts on Asian Handicap bets depend on an adjustment of the match result that applies a deduction (known as a handicap) to the goals total of the team considered more likely to win. For example, if Manchester City play Everton at home and the Asian Handicap is quoted at -2 (meaning a two goal deduction is applied to City's total) then a bet on Everton would pay out even if they lost the game by one goal. If the result precisely matches the handicap (in the above example, City beat Everton by two goals) then all bets are refunded. Unlike in US spread betting, the handicaps do not usually equalise the chances of the two sides of bet winning, so differing payout odds are generally offered for the bets on the two teams.

From being almost unheard of outside Asia in the 1990s, the Asian Handicap betting market for soccer has become increasingly important around the world over the past 20 years.² While information on its betting volumes are not publicly available, it seems likely the Asian Handicap market now accounts for a high share of betting on European soccer matches. Much of the volume is placed with specialist online bookmakers such as Pinnacle who have low profit margins per bet and seek to offset this by taking high betting volumes. This market's low margins make it particularly attractive for well informed bettors and professional betting syndicates. This raises the question of whether the odds in this market have different properties to more traditional markets. However, despite its increased prominence in betting on European soccer, there has been almost no previous research on whether the Asian Handicap market for soccer operates in an efficient manner.³

We first illustrate some inefficiencies in the forecast probabilities that are generated by odds in the traditional home/away/draw market using a dataset with over 80,000 European professional soccer matches. This is a larger and updated version of the type of dataset previously used by Angelini and De Angelis (2019) and, like them, we find the odds show a strong favourite-longshot bias such that bets on longshots lose significantly more money than bets on favourites. We show that this pattern

¹In addition to the many studies on racetrack betting, as surveyed by Snowberg and Wolfers (2008), other examples include studies of betting on US professional football (Gray and Gray, 1997), US college football and basketball (Berkowitz et al. (2017), Moscovitz and Vasudevan (2022)) and UK soccer (Cain et al., 2000).

²The term Asian handicap was coined by journalist Joseph Saumarez Smith in 1998 when he was asked to give an English translation to describe a new type of betting he had encountered while visiting Indonesia.

³We note that Hegarty (2021) and Hegarty and Whelan (2022) study the effect the absence of spectators in stadiums had on the Asian Handicap market during the COVID lockdowns. Also, Constantinou (2021) presents a complex Bayesian econometric model using Asian Handicap odds to predict outcomes in a sample of English Premier League matches.

should directly imply other inefficiencies. Specifically, it should lead to the probabilities implied by market efficiency being too high for longshots and too low for favourites and it implies that the expected losses on a broad portfolio of bets in the home/away/draw market are larger than would be predicted based on the assumption that the market is efficient. We verify that these results hold for the home/away/draw market.

In contrast, we document a number of ways in which the Asian Handicap market's odds for the same sample of European soccer matches can be characterised as efficient. To derive probability estimates from these odds, we need to confront a technical problem. Traditionally, where there are N possible outcomes for sports events, the N odds plus the condition that probabilities sum to one allow you to solve for a unique set of N probabilities consistent with the market being efficient as well as a figure for the bookmaker's gross profit margin. However, for most Asian handicap bets, there are three possible outcomes (the bet on the stronger team wins, the bet on weaker team wins or the bet involves a refund) but only two odds are provided. To address this issue, we develop a new methodology using assumptions about the probability of a refund to estimate the market's probabilities of the bets on the stronger or weaker team winning and also the bookmaker's gross profit margin. We present evidence in favour of our approach to modelling refunds.

We show that the implied probabilities for match outcomes from Asian Handicap odds derived from our procedure do not exhibit favourite-longshot bias and that they are unbiased estimates of the win rate for predicted outcomes. We also show that average loss rates for bettors in this market are lower than in the home/away/draw market and can be predicted accurately from betting odds under the assumption of market efficiency.

Our paper is structured as follows. Section 2 provides a brief description of the structure of betting markets for European soccer, explains how Asian Handicap betting works and introduces our dataset. Section 3 presents results on the efficiency of forecasts derived from home/away/draw betting odds. Section 4 describes our methodology for calculating probabilities from Asian Handicap betting odds and presents our empirical analysis of the properties of these probability estimates. Section 5 offers some conclusions and suggestions for future research.

2. Betting Markets for European Soccer

Here we briefly describe the development of European soccer betting markets, explain how Asian Handicap betting works and describe the dataset that we use.

2.1. Background on European Soccer Betting

The market for betting on European soccer emerged from the legalisation of betting in the UK in 1961, which led to the emergence of retail betting shops. Other European countries followed suit in the following years. The odds generally had a high margin in favour of the bookmakers and bettors needed to physically attend the betting office to place a bet. The rise of the internet meant that many retail bookmakers turned their attentions to online betting. Online betting greatly increased the turnover of bookmakers but it also made it easier for well-informed bettors to find market inefficiencies and arbitrage opportunities across providers. How bookmakers have dealt with informed bettors has led to the emergence of two different business models for online bookmaking, namely the so-called “soft” and “sharp” models.⁴

The traditional retail bookmakers in Europe have adopted the “soft” bookmaker model. This model focuses on maintaining high gross profit margins and spending on marketing to attract and retain bettors who will take on high-margin bets. Well-informed bettors that consistently make profits are generally restricted in how much they can bet and can ultimately be cut off from placing bets.⁵ These bookmakers have moved away from investing money in odds compilation research in favour of spending on advertising, web site development, and customer profiling algorithms.

In contrast, “sharp” bookmakers such as Pinnacle have essentially the opposite model. These bookmakers are online only with no retail presence and have business headquarters usually located outside Europe. They focus on offering low margins with profits driven by attracting high betting volumes. They do not spend much money on advertising and accept bets from informed bettors and professional betting syndicates, using the information from these bets to shape their betting odds. Indeed, one of the business lines of sharp bookmakers is providing forecast probabilities for a fee to soft bookmakers. Sharp bookmakers do not have licenses to operate in some European markets but bettors can usually access them via brokers that act as intermediaries.

The distinction between soft and sharp bookmakers matters for our analysis because soft bookmakers dominate the traditional market for betting on home/away/draw outcomes while the sharp bookmakers take most of their bets on soccer in the form of Asian Handicaps. In recent years, soft bookmakers have also begun offering Asian Handicap bets but their odds largely follow those set by the “sharp” bookmakers. Asian Handicap odds movements tend not to differ much across the two types of providers, with the soft bookmakers adding a higher margin and not putting any effort

⁴Buchdahl (2016) provides a more detailed discussion of how the various business models for bookmaking operate.

⁵The soft bookmaker practices of customer profiling and stake restrictions are discussed by Davies (2022).

into promoting this business. Given the difference in how odds are set in these two markets as well as the different profiles of their participants, it is interesting to investigate whether their odds have different properties.

2.2. How Asian Handicap Betting Works

The Asian Handicap features four types of bets with handicaps that change in increments of 0.25 goals. Obviously, teams can't score a quarter of a goal, so bets at quarter-goal handicaps are actually "hybrids" in which money is split between bets at other handicaps. We will explain how this type of betting works by illustrating four cases in which a stronger team has different handicaps applied to it—0.75, 1, 1.25 and 1.5. Asian Handicap bookmakers use the decimal odds convention. This means an odds quote of O_S on the strong team means $\$O_S$ is the payout on a \$1 bet (inclusive of the original \$1 stake) when the team beats the handicap. We assume decimal odds on the weak team of O_W .

To explain how Asian Handicap betting works, we will start with the simpler bets and then explain the more complex hybrid bets. Consider first the case in which the Asian handicap is 1.5. There are only two possible outcomes:

- The stronger team wins by 2 or more. In this case, the bet on the stronger team pays out O_S and the bet on weaker team loses in full.
- The stronger team fails to win by 2 or more. In this case, the bet on the weaker team pays out O_W and the bet on stronger team loses in full.

For the case in which the Asian handicap is 1, there are three possible outcomes:

- The stronger team wins by 2 or more. In this case, the bet on the stronger team pays out O_S and the bet on weaker team loses in full.
- The stronger team wins by 1. In this case, bets on both teams are refunded.
- The stronger team fails to win. In this case, the bet on the weaker team pays out O_W and the bet on stronger team loses in full.

Bets with an Asian handicap of 1.25 place half the money on a bet with a handicap of 1 and the other half on a bet with a handicap of 1.5. Again, there are three possible outcomes:

- The stronger team wins by 2 or more. In this case, both halves of the bet on the stronger team are successful and there is a pay out O_S while the bet on the weaker team loses in full.
- The stronger team wins by 1. In this case, the half-bet on the stronger team with the handicap of 1.5 loses and the half-bet on the weaker team wins $\frac{O_W}{2}$. The half bets on both teams with the handicap of one are refunded.

- The stronger team fails to win. In this case the bet on the stronger team is lost and the bet on the weaker team pays out O_W .

The final example is an Asian handicap is 0.75. This puts half the money on a bet with a handicap of 1 and the other half on a bet with a handicap of 0.5. There are again three possible outcomes:

- The stronger team wins by 2 or more. In this case, both halves of the bet on the stronger team are successful and there is a full pay out O_S while the bet on the weaker team loses in full.
- The stronger team wins by 1. In this case, the half-bet on the stronger team with the handicap of 1 gives a refund and the half-bet on the stronger team at 0.5 pays out $\frac{O_S}{2}$. The half bets on weaker team at 1 gives a refund and the half bet on the stronger team at 0.5 loses.
- The stronger fails to win. In this case, the bet on the stronger team is lost and the bet on the weaker team pays out O_W .

All bets in the Asian Handicap market work in a similar fashion to these four cases, with handicaps that are either integers or else numbers ending in .25, .5 or .75.

2.3. Data Description

Our data comes from www.football-data.co.uk, a website maintained by gambling expert and author, Joseph Buchdahl. The dataset has information on outcomes and odds for both home/away/draw and Asian Handicap betting markets for 84,230 matches spanning the 2011/12 to 2021/22 seasons for 22 prominent European soccer leagues across 11 different nations as described in Table 1. Our measure of betting odds is the average closing odds (posted just before kickoff) across the various online bookmakers surveyed by www.football-data.co.uk.⁶ In an efficient market, the closing odds should incorporate all relevant information. For Asian Handicap betting, it is possible to find different handicaps quoted for the same match but our sample lists only one handicap per match, generally the one that is offered by the most bookmakers, and it reports the average odds associated with that handicap.⁷

Our data source also lists the the maximum odds quoted across providers for each match but we do not use these data. There are a number of reasons why we use average odds rather than maximum odds. First, bookmakers will occasionally run “loss leaders” by posting generous odds on specific matches with the intention of attracting new customers, usually with restrictions on how

⁶From the 2019/2020 season onwards, the odds data come from the sample of providers available at www.oddsportal.com. For previous seasons, the sample was made up of those providers listed on www.betbrain.com.

⁷In personal communication, Joseph Buchdahl informed us “The one I select is a combination of two methods ... closest to 50-50 and with the most contributing bookmakers. Usually both criteria apply together, but sometimes if the line with the most bookmakers is far from 50-50, I will choose the one closest to 50-50.”

much money can be placed. These odds are not based on the bookmaker's assessment of the probabilities of the relevant outcomes. Since we are attempting to check the market's ability to assess the underlying probabilities correctly, these odds would not be appropriate. Second, even if one was focusing only on whether it was possible to make profits due to bookmakers posting inefficient odds, those bettors who choose to only place bets at the best available odds will generally find themselves cut off by soft bookmakers, so this is more a theoretical strategy than a practical possibility.

It is worth emphasising that, despite some obvious similarities, the Asian Handicap market differs from spread betting markets on US sports along a couple of dimensions. Spread bets offered on high scoring sports such as basketball and American football are generally set to equate the odds of each side of the bet winning. This means the odds offered on each bet are typically the same so the implied probabilities of success of each bet are equal. This is not the case with the Asian Handicap market. Figure 1 shows a histogram of average decimal odds on Asian Handicap bets in our sample. The average decimal odds is 1.92 (meaning a \$1 bet pays out \$1.92 if fully successful) but there is a wide variation in odds offered: The 10th percentile of odds offered is 1.77 while the 90th percentile is 2.08. Our calculations below suggest that bets in the bottom decile for probabilities of a full payout have an average probability of such a payout of 0.26 while the corresponding average probability for bets in the top decile is 0.53.

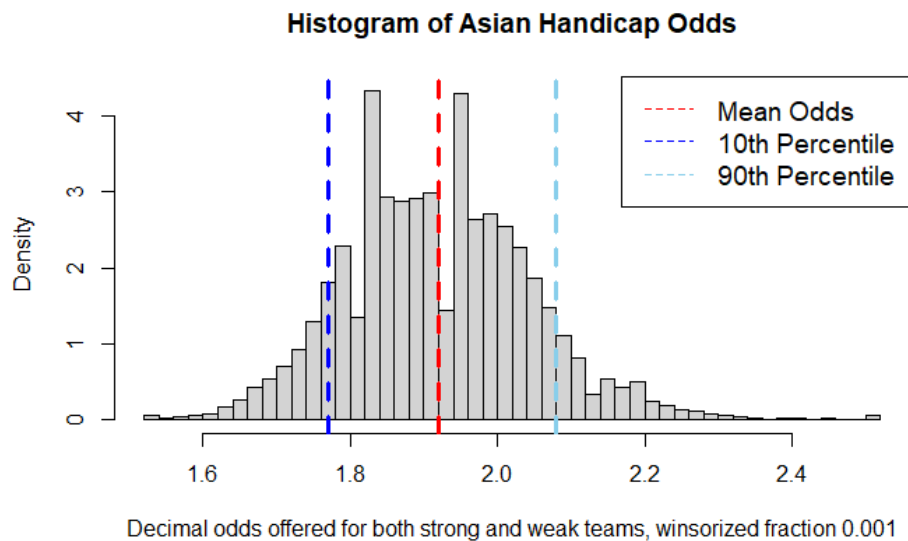
There are several reasons for the wide variation in probabilities of full payouts implied by Asian Handicap odds. Handicaps are only set in quarter-goal increments and these will rarely correspond precisely to the market's expected goal difference. This means bettors will generally think a bet on one of the teams in a match is more likely to win than the other, which will be reflected in differing odds. Also, bookmakers that offer spread bets in US sports respond to incoming betting volumes by adjusting the spread while maintaining equal odds on both sides of the bet. In contrast, Asian Handicap bookmakers keep the handicap the same and alter the betting odds. So a handicap that may be associated with equal odds when first offered can end up with differing closing odds if betting volumes favour one of the bets more than the other.

Finally, the hybrid quarter-point handicap bets have the feature that one of the bets earns a profit in two of the three possible outcomes while the other only makes a profit in one of the three outcomes. For both sides of such bets to be equally attractive, the expected payouts must be the same. We show below that this compensation occurs via bets that only make a profit in one outcome tending to have a higher probability of a full payout.

Table 1: Description of the 22 football leagues included in the dataset

| Nation | Number of Divisions | Division(s) |
|-------------|---------------------|--|
| England | 5 | Premier League, Championship, League 1 & 2, Conference |
| Scotland | 4 | Premier League, Championship, League 1 & 2 |
| Germany | 2 | Bundesliga 1 & 2 |
| Spain | 2 | La Liga 1 & 2 |
| Italy | 2 | Serie A & B |
| France | 2 | Ligue 1 & 2 |
| Belgium | 1 | First Division A |
| Greece | 1 | Super League Greece 1 |
| Netherlands | 1 | Eredivisie |
| Portugal | 1 | Primeira Liga |
| Turkey | 1 | Super Lig |

Figure 1: Distribution of Odds Offered for Asian Handicap Bets



3. The Home, Away & Draw Betting Market

Here we describe how to calculate probabilities from the home/away/draw betting markets under the assumption of market efficiency and describe the forecasting properties of these probabilities.

3.1. Calculating Efficient Market Probabilities

Consider a sporting event with N possible outcomes, each with probability P_i . An efficient betting market will have the property that the expected return to betting on each outcome will be the same. Bookmakers make profits on average and have to cover costs, so the expected payout on a \$1 bet must be some value $\mu < 1$. Characterising the odds O_i as the total payout from betting \$1 on outcome i when this outcome occurs, the hypothesis of a common expected payout across all bets implies

$$P_i O_i = \mu \quad i = 1, \dots, N \quad (1)$$

Combined with the condition that the probabilities sum to one, this provides $N + 1$ linear equations for each sporting event that can be solved to obtain a unique set of $N + 1$ unknown values, namely the N probabilities and the expected return μ . Specifically, μ is given by

$$\mu = \frac{1}{\sum_{i=1}^N \frac{1}{O_i}} \quad (2)$$

The expected payout is determined by the sum of the inverses of the odds. This sum, known in bookmaking as the “overround”, is commonly used by gamblers to estimate the gross profit margin being taken by bookmakers. Once μ has been calculated, the so-called “normalised” probabilities can then be derived directly from equation 1.

3.2. Favourite-Longshot Bias

The simplest way to illustrate the favourite-longshot bias pattern in the home/away/draw odds is to look at average returns on bets sorted by their estimated probability of success under the assumption of market efficiency. The chart in Figure 2 shows the results from dividing all 252,690 bets in our sample into deciles of probability estimate and calculating the average payout on these bets.

A clear pattern of favourite-longshot bias is evident. For bets in the lowest decile, the average estimated probability of success is 14% and the average payout on a \$1 bet is only \$0.83 (meaning an average loss of 17%). In contrast, for bets in highest decile, the average estimated probability of success is 63% and the average payout on a \$1 bet is \$0.98 (a 2% average loss rate). The pattern of the bias is strongly nonlinear, with average payouts dropping sharply for the lowest deciles. Similar findings of higher payout rates by estimated probability of bet success are obtained when we look at payout rates focusing only on bets on a single outcome, such as bets on favourites only, bets on

longshots only or bets only on a home win, a loss or a draw.

It is possible, of course, that this pattern could perhaps be driven by some kind of composition effect. For example, if longshot bets had tended to underperform during seasons where the book-making market was less competitive and margins were higher, then there could be a correlation between average payouts and ex ante probabilities that was not due to favourite-longshot bias. Table 2 addresses this issue by reporting results for the following regression

$$\Pi_{ijk} = \sum_{j=1}^{22} \alpha_j L_j + \sum_{k=1}^{11} \beta_k S_k + \sum_{n=1}^{10} \gamma_n D_n + v_{ijk} \quad (3)$$

where Π_{ijk} is the payout from bet i in league j in season k , L_j are dummy variables for the 22 leagues, S_k are dummy variables for each season and the D_n are dummies for which decile of estimated probability values the bet is in.

We estimate the regression using Weighted Least Squares (WLS). This is because, as has been recognised in the literature on forecasting soccer games since Pope and Peel (1989), regressions explaining the outcomes of sporting contests feature heteroskedasticity. In this case, under market efficiency, the payout on a bet with odds of O that has a probability p of winning has a variance of $O^2 p(1-p)$. To account for this issue, we follow Pope and Peel in using WLS with the variances approximated by $O_{ijk}^2 P_{ijk}(1-P_{ijk})$ where O_{ijk}^2 and P_{ijk} are the odds and estimated probabilities of bet success under the assumption of market efficiency. Because each match shows up three times in the full-sample regression (as a bet on home win, a bet on away win and a bet on draw) there are correlations between the errors for each individual match so standard errors were clustered at the match level.

The results show the coefficients on the decile dummies steadily increasing with the estimated probability of success of the bets, consistent with the pattern for the raw averages in Figure 2. One question is whether this pattern is driven by bookmakers mis-pricing the home advantage effect. Home teams are more likely to be favourites, so under-estimating their advantage could drive a pattern of payouts on longshots being over-estimated. The second column shows, however, that with bets on the home team as a baseline, a dummy variable for bets on the away team is not significant, so the estimated pattern is not related to mis-estimating the extent of home advantage. Interestingly, however, there is some evidence that bets on draws actually have a slightly higher payout than expected once one controls for probability deciles. The expected payout (the estimated value of μ) also shows up as significant if added to this regression but it does not change the estimated pattern of favorite-longshot bias.

Figure 2: Average Payouts for the Probability Deciles of Home/Away/Draw Bets

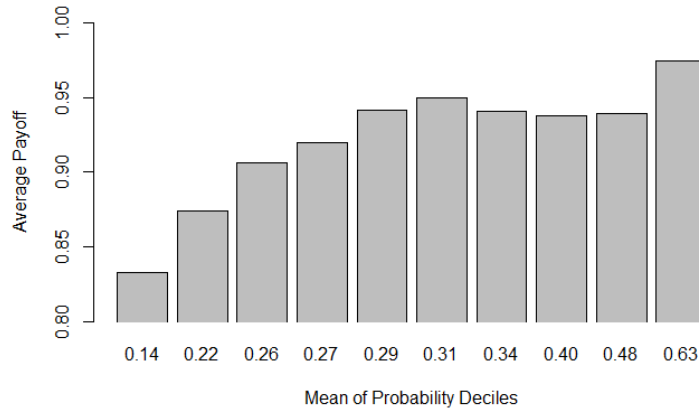


Table 2: WLS Regression of Payouts on Home/Away/Draw Bets on Probability Decile Dummies

| | | | | |
|---------------|-----------|----------|-----------|----------|
| Constant | 0.8374*** | (0.0148) | 0.8287*** | (0.0162) |
| Decile 2 | 0.0255 | (0.0182) | 0.0222 | (0.0182) |
| Decile 3 | 0.0575** | (0.0174) | 0.0500** | (0.0176) |
| Decile 4 | 0.0727*** | (0.0168) | 0.0620*** | (0.0171) |
| Decile 5 | 0.0938*** | (0.0166) | 0.0826*** | (0.0170) |
| Decile 6 | 0.1013*** | (0.0163) | 0.0951*** | (0.0164) |
| Decile 7 | 0.0938*** | (0.0153) | 0.0983*** | (0.0154) |
| Decile 8 | 0.0916*** | (0.0156) | 0.0985*** | (0.0163) |
| Decile 9 | 0.0920*** | (0.0152) | 0.0995*** | (0.0160) |
| Decile 10 | 0.1295*** | (0.0167) | 0.1373*** | (0.0174) |
| Bet Type Away | | | 0.0053 | (0.0073) |
| Bet Type Draw | | | 0.0246* | (0.0110) |
| <i>N</i> | 252,690 | | 252,690 | |

Standard errors in parentheses are clustered at the match level.

Specification also includes dummy variables for each league and season.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

3.3. Implications of Favourite-Longshot Bias

The pattern of favourite-longshot bias shown here implies that the home/away/draw market is inefficient. It is also worth noting that the presence of this bias invalidates the standard calculations of both probabilities and the expected return based on the assumption of market efficiency. To illustrate this, assume that odds are determined by bookmakers according to

$$O_i = \frac{\mu_i}{P_i} \quad i = 1, \dots, N \quad (4)$$

where the average payout rates μ_i depend positively on the P_i .

Consider first the estimated probabilities based on the assumption of market efficiency. These are calculated by using the overround to estimate the expected payout rate, which we will now denote as $\hat{\mu}$. With varying payout rates across bets, the calculation for the expected payout rate under the assumption of market efficiency becomes

$$\hat{\mu} = \frac{1}{\sum_{i=1}^N \frac{P_i}{\mu_i}} \quad (5)$$

The estimated probabilities can be re-expressed as follows:

$$\hat{P}_i = \frac{\hat{\mu}}{O_i} = \frac{\hat{\mu}}{\frac{\mu_i}{P_i}} = \frac{\hat{\mu}}{\mu_i} P_i = \left(\frac{1}{\mu_i \sum_{j=1}^N \frac{P_j}{\mu_j}} \right) P_i \quad (6)$$

The term in the denominator of the fraction multiplying P_i can be written as

$$\mu_i \sum_{j=1}^N \frac{P_j}{\mu_j} = P_i + \sum_{\substack{j=1 \\ j \neq i}}^N P_j \frac{\mu_i}{\mu_j} \quad (7)$$

Now consider the implications of favorite-longshot bias for this calculation. It calculates a probability weighted average of 1 and a set of terms of the form $\frac{\mu_i}{\mu_j}$. Suppose outcome i has the lowest probability and thus the lowest value of μ_i . Then the terms in the $\frac{\mu_i}{\mu_j}$ will all be less than one and the overall sum in equation 7 will be less than one. This will imply $\hat{P}_i > P_i$. The same logic says that $\hat{P}_i < P_i$ for the outcome with the highest probability and that the size and sign of the bias in probability estimates will depend monotonically on the size of the underlying probability.

Second, consider the accuracy of $\hat{\mu}$ as a measure of the expected loss rate. Hegarty and Whelan (2023) show that the following approximation works well for samples of betting odds such as the one studied in this paper

$$\hat{\mu} \approx \sum_{i=1}^N P_i \mu_i \quad (8)$$

When there is favorite-longshot bias, the probability weighted sum of payouts on the right hand side places more weight on high payouts than the equally weighted average payout across all bets. This means the standard overround-based calculations of the expected payout will be higher than the average payout across all bets.

3.4. Ex Ante versus Ex Post Outcomes

We now show that the predictions just described hold in the home/away/draw dataset. Consider first the accuracy of the probabilities implied by market efficiency. Figure 3 divides all 252,690 bets in our sample into 20 probability estimate quantiles and calculates the fraction of winning bets for each quantile. There is a systematic pattern in which the estimated probabilities of bet success implied by market efficiency are too high for low estimated values and too low for high estimated values. A WLS regression of match outcomes on the estimated probabilities strongly rejects the hypotheses that the slope of the blue line is one, a result also reported by Angelini and De Angelis (2019) for their earlier dataset. The deviations of these probability estimates from the 45 degree line may seem small but, for low values, these deviations are a big percentage of the estimated probabilities, consistent with the large average loss estimates above. Ultimately, the favourite-longshot bias in payouts occurs because longshot bets don't win as often as the odds suggest they should.

Second, consider the average payouts across all bets in the sample. Table 3 reports that the average expected loss rate across all matches under the assumption of market efficiency (i.e. the average value of $1 - \hat{\mu}$) is 6.5%. However, the actual average loss from placing an equal-sized bet on all possible outcomes for all matches (i.e. betting on the home win, the away win and the draw) is 7.8%. The average loss rate if draws are excluded (as is done in many studies) is 7.7%. *t*-tests strongly reject the hypotheses that the means of the two payout distributions are equal to the average expected loss rate.

Figure 4 further illustrates this finding by sorting the data into 20 quantiles by ex ante expected loss rate and calculating the actual ex post average loss rates from the strategy of betting an equal amount on all matches in each quantile. It shows that across the full range of ex ante expected loss rates implied by market efficiency (with the exception of the bottom quantile) the actual loss rates are larger than the expected loss rates.

Figure 3: Actual Fraction of Wins on Home/Away/Draw Bets Sorted by Estimated Probability of a Win

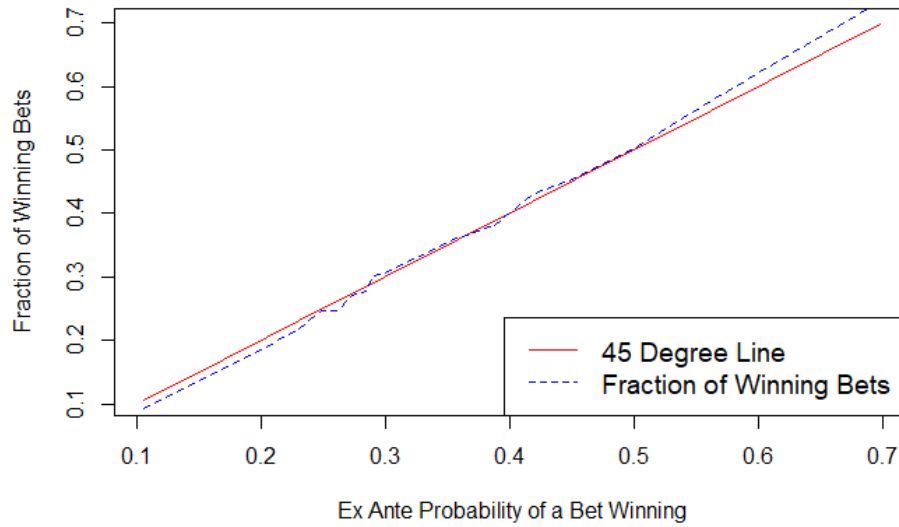
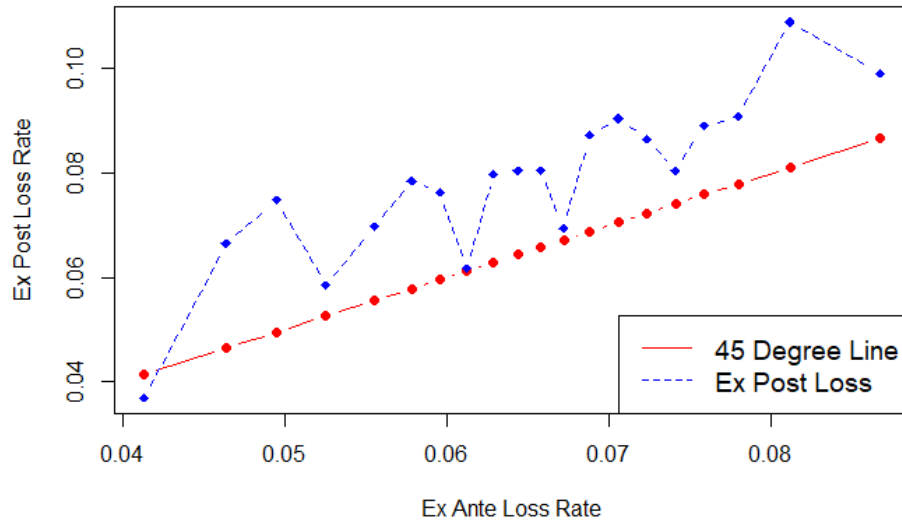


Table 3: Average Expected Ex Ante Loss Rates Compared with Actual Average Loss Rates for the Home/Away/Draw Market

| | Mean | <i>N</i> |
|---|--------|----------|
| Expected Ex Ante Average Loss on All Bets | 0.0646 | 84,230 |
| <i>Equally Weighted Portfolios</i> | | |
| Loss Rate from Betting on All Home, Away & Draw | 0.0783 | 252,690 |
| Loss Rate from Betting on All Home & Away | 0.0777 | 168,460 |

Figure 4: Actual Ex Post Loss Rates on an Equally-Weighted Portfolio of Home, Away & Draw Bets Sorted By Ex Ante Expected Loss Rate



4. The Asian Handicap Market

We will now describe our methodology for translating Asian Handicap odds into probability estimates and then present evidence on the properties of these estimates.

4.1. Calculating Probabilities from Asian Handicap Odds

There are four different types of Asian Handicap bets, depending on whether the handicap is an integer or ends in 0.5 or ends in either 0.25 or 0.75. We will take each type of handicap in turn.

Asian Handicap Ends in .5

Consider the case in which the Asian handicap ends in .5. In this case, either the bet on the stronger team wins or the bet on the weaker team wins. Refunds do not occur. Because there are two possible outcomes and two betting odds and a condition that the probabilities sum to one, this means we have a system of 3 linear equations in 3 unknowns (the two probabilities and the expected return) so the method described in Section 3.1. can be used to calculate both the probabilities of each bet winning and the expected payout rate. Recall that O_S and O_W are the odds for the strong and weak teams winning the handicap-adjusted match, this method gives the following values for the probability of each bet winning and the expected payout consistent with market efficiency:

$$P_S = \frac{1}{O_S} \frac{1}{\frac{1}{O_S} + \frac{1}{O_W}} = \frac{O_W}{O_S + O_W} \quad (9)$$

$$P_W = \frac{1}{O_W} \frac{1}{\frac{1}{O_S} + \frac{1}{O_W}} = \frac{O_S}{O_S + O_W} \quad (10)$$

$$\mu = \frac{1}{\frac{1}{O_S} + \frac{1}{O_W}} = \frac{O_S O_W}{O_S + O_W} \quad (11)$$

Asian Handicap Is An Integer

Now suppose the Asian handicap is 1 so we want to calculate the probabilities for three different outcomes

$$P_{S2} = \text{Probability the stronger team wins by 2 or more} \quad (12)$$

$$P_{S1} = \text{Probability the stronger team wins by 1} \quad (13)$$

$$P_W = \text{Probability of a draw or the weaker team winning} \quad (14)$$

Again assuming the expected payout for all \$1 bets is μ , then market efficiency implies

$$P_{S2}O_S + P_{S1} = \mu \quad (15)$$

$$P_W O_W + P_{S1} = \mu \quad (16)$$

$$P_W + P_{S1} + P_{S2} = 1 \quad (17)$$

This is a system of three linear equations in four unknowns (the three probabilities and the expected return) so there is no unique solution.

One possible approach to calculating the probabilities would be to assume that the expected payout μ equalled some fixed number across all bets. However, the evidence from the home/away/draw market suggests that expected payouts vary considerably from match to match. When we implemented this approach, we also found that the probabilities it implied were often not sensible, with values sometimes below zero or greater than one. Instead, the approach we took was to specify a value of the probability of the refund outcome (in this case P_{S1}) for each type of handicap based on the historical average frequencies of refunds for that type of handicap. We will provide empirical justification for this approach below.

With this assumption made, we can then calculate the other two probabilities and the expected payout for each match. Conditional on a specific value of the probability of a refund, P_{S1} , we can solve for the other unknowns as

$$P_{S2} = \frac{(1 - P_{S1}) O_W}{O_S + O_W} \quad (18)$$

$$P_W = \frac{(1 - P_{S1}) O_S}{O_S + O_W} \quad (19)$$

$$\mu = P_{S1} + \frac{(1 - P_{S1}) O_S O_W}{O_S + O_W} \quad (20)$$

Asian Handicap Ends in .25

Now suppose the Asian handicap was 1.25. Market efficiency implies the probabilities satisfy

$$P_{S2}O_S + \frac{P_{S1}}{2} = \mu \quad (21)$$

$$P_W O_W + P_{S1} \left(\frac{1 + O_W}{2} \right) = \mu \quad (22)$$

$$P_W + P_{S1} + P_{S2} = 1 \quad (23)$$

Again taking P_{S1} as given, we can solve these equations to give

$$P_{S2} = \frac{(1 - P_{S1}) O_W}{O_S + O_W} + \frac{P_{S1}}{2} \left(\frac{O_W}{O_S + O_W} \right) \quad (24)$$

$$P_W = \frac{(1 - P_{S1}) O_S}{O_S + O_W} - \frac{P_{S1}}{2} \left(\frac{O_W}{O_S + O_W} \right) \quad (25)$$

$$\mu = \frac{P_{S1}}{2} + \frac{(1 - P_{S1}) O_S O_W}{O_S + O_W} \quad (26)$$

These probabilities are adjusted relative to the integer handicap to reflect the asymmetric outcome when the stronger team wins by 1. In that case, the bet on the weak team gets a “half win” with the other half refunded while the bet on the strong team gets half the bet refunded while the rest is lost. If the probability formulas were not adjusted from the integer case, then the return on the bet on the weak team would be higher than the return from betting on the strong team. The adjustment raises the probability of the strong team winning by two or more and lowers the probability of them failing to win.

Asian Handicap Ends in 0.75

If the Asian handicap is 0.75, then market efficiency implies the probabilities satisfy

$$P_{S2} O_S + P_{S1} \left(\frac{1 + O_S}{2} \right) = \mu \quad (27)$$

$$P_W O_W + \frac{P_{S1}}{2} = \mu \quad (28)$$

$$P_W + P_{S1} + P_{S2} = 1 \quad (29)$$

Again taking P_{S1} as given, we can solve these equations to give

$$P_{S2} = \frac{(1 - P_{S1}) O_W}{O_S + O_W} - \frac{P_{S1}}{2} \left(\frac{O_S}{O_S + O_W} \right) \quad (30)$$

$$P_W = \frac{(1 - P_{S1}) O_S}{O_S + O_W} + \frac{P_{S1}}{2} \left(\frac{O_S}{O_S + O_W} \right) \quad (31)$$

$$(32)$$

while the formula for μ is identical to the 1.25 handicap. These probability formulas are symmetric with the 1.25 case, with a higher probability of the weaker team getting a draw or win and a lower probability of the stronger team winning by more than 2.

4.2. Evidence on Predictability of Refunds

Before documenting the properties of our calculated probabilities and expected payouts, we first provide evidence to explain our approach of setting the probability of refunds equal to a fixed number for each type of handicap. If the probability of a refund varied systematically across matches, then our approach could be flawed and a correct calculation of the probabilities would require a match-by-match adjustment for the refund probability.

To test whether refunds were predictable, we estimated the following specification for all three types of bets where refunds are possible

$$R_{ijkq} = \sum_{j=1}^{22} \alpha_j L_j + \sum_{k=1}^{11} \beta_k S_k + \sum_{n=1}^{11} \beta_n S_n + \sum_{q=1}^3 \delta_q H_q + \eta_1 O_{iH} + \eta_2 O_{iA} + u_{ijkq} \quad (33)$$

where R_{ijkq} equals 1 if a refund was issued for match i in league j and season n with handicap type q and equals zero otherwise and O_{iH} and O_{iA} are the Asian Handicap odds for the bets on the home and away teams. The H_q are dummies for the three handicap types featuring refunds.

Table 4 reports the results from estimation of this regression via WLS for the 63,468 matches that had the possibility of a refund occurring, where the estimated handicap-specific average rate of refund is used to construct match-specific variances for weighting purposes.⁸ None of the year dummies are significant, implying the probability of refunds occurring has been stable across seasons. We also do not find any significant effect of either the home or away odds. We do find evidence that refunds are most likely for bets with handicaps ending in .25 and least likely for bets with handicaps ending in .75. For this reason, to generate our probability estimates, we estimate the probabilities of a refund separately for each of the three relevant handicap types as the sample average fractions of bets that end in refunds for each type.⁹

We can summarise the evidence on refunds as follows: The fraction of refunds that occur for each type of handicap is stable and predictable over time but there is no information available in the betting odds that help predict which specific matches will generate refunds.

⁸Similar results are obtained from Probit estimation.

⁹One concern with this procedure is that it uses data from the full sample, so information about future matches is being used to “forecast” matches occurring at a time when this information is not available. However, we obtain the same results if we only use estimates of the probability of a refund from seasons prior to when matches occurred.

Table 4: WLS Regression Predicting Refunds

| | Coefficients | Standard Errors |
|--------------------------|--------------|-----------------|
| Home Odds | 0.0316 | (0.0539) |
| Away Odds | 0.0235 | (0.0546) |
| 2012 Season | -0.00235 | (0.00831) |
| 2013 Season | -0.00360 | (0.00866) |
| 2014 Season | 0.00493 | (0.00855) |
| 2015 Season | -0.00215 | (0.00851) |
| 2016 Season | 0.00493 | (0.00855) |
| 2017 Season | -0.00449 | (0.00836) |
| 2018 Season | -0.00210 | (0.00826) |
| 2019 Season | 0.00353 | (0.00856) |
| 2020 Season | 0.000989 | (0.00836) |
| 2021 Season | 0.00952 | (0.00838) |
| Handicap Type ending .25 | 0.00884* | (0.00400) |
| Handicap Type ending .75 | -0.0260*** | (0.00521) |
| <i>N</i> | 63,468 | |
| <i>R</i> ² | 0.003 | |

The baseline bet here relates to a match in 2011 with an integer handicap.

Specification also includes dummy variables for each league.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.3. Favourite-Longshot Bias?

As we did above for home/away/draw odds, Figure 5 shows the average payouts for \$1 bets sorted by the estimated probability of the bet winning a full payout. There is no clear pattern of bias across estimated probability ranges and the average returns vary much less than for the home/away/draw market.

Because of the handicap adjustment, this market features fewer extreme longshot or extreme favorite bets. This can be seen in the narrower range of estimated probabilities of bet success. The average probabilities of full bet success range from 0.26 in the bottom decile to 0.53 in the top decline, compared with a range of 0.14 to 0.63 for the home/away/draw market. This narrower range of probabilities, however, is not the explanation for the difference in payout rates between the handicap market and the traditional market. The handicap market still features a fairly wide range of ex ante probabilities of bet success and when comparisons are made over the same probability range, there is a notable difference between the pattern for loss rates in the two markets.

For the home/away/draw market, bets in the decile with an estimated average probability of success of 0.26 have a loss rate of 9.4% while bets in the decile with an estimated average probability of success of 0.48 have a loss rate of 6.1%, so across this range of probabilities, loss rates are over 50% higher for the longshot bets than for the favourite bets. Across the same range of probabilities for the Asian Handicap, the longshot bets have an average loss rate of 3.5% and the favourite bets have a loss rate of 4.1%.

The absence of a favourite-longshot bias is further confirmed by Table 5 which reports results for the following regression

$$\Pi_{ijkq} = \sum_{k=1}^{22} \alpha_j L_j + \sum_{k=1}^{11} \beta_k S_k + \sum_{q=1}^4 \delta_q H_q + \sum_{n=1}^{10} \gamma_n D_n + v_{ijkq} \quad (34)$$

where Π_{ijkq} is the payout from bet i in league j in season k of handicap type q and the various dummy variables are as defined before. With the bottom decile as the baseline, positive coefficients would be evidence of favourite-longshot bias. In fact, the reported coefficients are all negative, albeit with small values, and only the coefficients for the top two deciles are statistically significant. This suggests some weak evidence for a reverse of the favourite-longshot bias operating for Asian Handicap bets. Again, the expected payout (the estimated value of μ) also shows up as significant if added to this regression but it does not change the estimated pattern of favorite-longshot bias.

Our study shares some similarities with the work of Moscowitz and Vasudevan (2022) who analyse the different properties of spread betting odds and odds for money line bets (bets on whether a team wins or loses a match) for a sample of US basketball and American football games. Like us, they compare the odds for two different markets across the same set of games and, like us, they find a pattern of favourite-longshot bias in bets on outright outcomes but not in the spread betting market

that pays out based on an adjusted scoreline.

Moscowitz and Vasudevan explain their results as being due to bettors having a preference for risk, so bookmakers can offer inferior odds on high risk money line bets and still find takers. In contrast, with spread betting on US sports, each side of the bet is priced the same and considered equally risky, so those with a preference for risk treat each side of the bet equivalently. However, there is an important contrast between our results and those of Moscovitz and Vasudevan because in our score-adjusted market, there is still a wide range of estimated probabilities of success and thus a wide range of risk but we don't find evidence of lower returns for higher risk bets in this market.

Figure 5: Average Payouts By Probability Deciles For Asian Handicap Bets

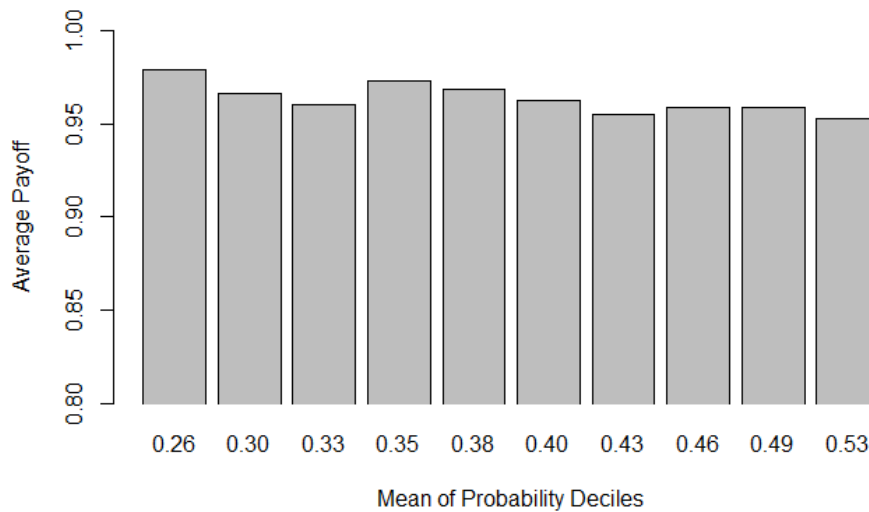


Table 5: WLS Regression for Payouts for Asian Handicap Bets on Probability Decile Dummies

| | | |
|-----------|-----------|-----------|
| Constant | 0.9763*** | (0.00860) |
| Decile 2 | -0.0114 | (0.00944) |
| Decile 3 | -0.0164 | (0.00945) |
| Decile 4 | -0.0035 | (0.00835) |
| Decile 5 | -0.0077 | (0.00927) |
| Decile 6 | -0.0124 | (0.00965) |
| Decile 7 | -0.0200 | (0.01203) |
| Decile 8 | -0.0177 | (0.01166) |
| Decile 9 | -0.0184* | (0.00915) |
| Decile 10 | -0.0248* | (0.00998) |
| <i>N</i> | 168,460 | |

Standard errors in parentheses.

Standard errors are clustered at the match level.

Specification also includes dummy variables for each league, season and handicap.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

4.4. Accuracy of Probability Estimates

As a visual check for bias in the calculated probabilities, Figure 6 divides all 168,460 Asian handicap bets in our sample into 20 quantiles and calculates the average probability that these bets result in a full payout. The chart shows that ex post full payout rates align well with the estimated probabilities with no evidence of a systematic deviations of actual success rates from the 45 degree line. A WLS regression of match outcomes on the estimated probabilities cannot reject the hypotheses that the slope of the blue line is one.

Given the difference in how the probability estimates were calculated for each handicap type, it is also interesting to examine separately for each type of handicap how well the averages of our calculated probabilities match with ex post average frequencies. Table 6 reports these results. In all cases, the average probabilities based on the Asian Handicap odds match closely with the ex post percentages of outcomes of each type, with t tests not rejecting the hypothesis that the samples were drawn from distributions with identical means.

In the case of integer and half-goal handicaps, where the process of determining payouts is symmetric for bets on the stronger and weaker teams, it is unsurprising to find that predicted probabilities and ex post outcomes show almost equal average probabilities of bet success as well as well as equally likely ex post successes.¹⁰ More interesting are the outcomes for the bets with handicaps ending in .25 or .75. Our calculated probabilities based on the Asian Handicap odds predict that for handicaps ending in .25, the bet on the strong team will earn a full payout 42% of the time and the bet on the weak team will earn a full payout 28% of the time, while for handicaps ending in .75, the bet on the strong team earns a full payout 30% of the time and the weak team earns a full payout 44% of the time. These highly asymmetric predictions match almost precisely with the average outcomes for these kinds of bets. In all cases, t -tests cannot reject the null hypothesis of equality of means of predicted and actual series.

¹⁰For the integer handicaps, for the calculations in Table 6 we adopt the convention that the home team is “the strong team” if the handicap is zero.

Figure 6: Actual Fraction of Full Wins on Asian Handicap Bets Sorted by Estimated Probability of a Full Win

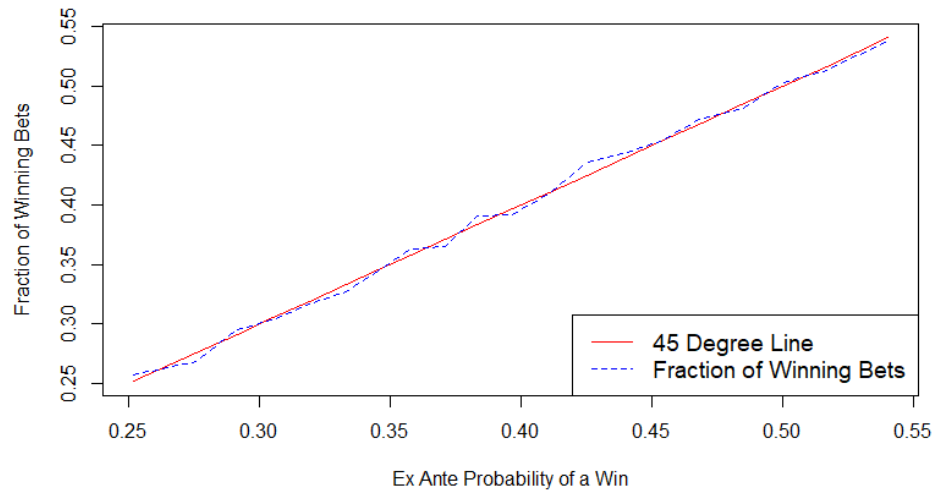


Table 6: Comparing Calculated Probabilities with Actual Outcomes for Each Handicap Type

| | Mean of Calculated Probabilities | Fraction of Actual Outcomes | <i>N</i> |
|---|-------------------------------------|--------------------------------|----------|
| <i>Handicap Type Integer</i> | | | 23,730 |
| Strong team win bet | 0.3585 | 0.3550 | |
| Weak team win bet | 0.3582 | 0.3617 | |
| Refund | | 0.2833 | |
| <i>Handicap Type .25</i> | | | 29,250 |
| Strong team bet wins | 0.4231 | 0.4224 | |
| Weak team bet wins | 0.2843 | 0.2850 | |
| Partial refund/Half-win for weak team bet | | 0.2926 | |
| <i>Handicap Type .5</i> | | | 20,762 |
| Strong team bet wins | 0.4946 | 0.4889 | |
| Weak team bet win | 0.5054 | 0.5111 | |
| <i>Handicap type .75</i> | | | 10,488 |
| Strong team bet wins | 0.3042 | 0.3031 | |
| Weak team bet wins | 0.4396 | 0.4407 | |
| Partial refund/Half-win for strong team bet | | 0.2562 | |

4.5. Ex Ante versus Ex Post Loss Rates

As we did for home/away/draw bets, we also calculate the average ex ante expected loss rates implied by market efficiency and compare them with the mean ex post loss rates from an equally weighted portfolio of all Asian handicap bets. The average loss rate for this portfolio is 3.6 percent, which is a lot lower than the 7.8 percent loss rate for the home/away/draw market. Another difference shown in Table 7 is that the realised average loss rate is essentially identical to the average expected loss rate from our calculations. *t*-tests cannot reject the null hypothesis of equality of means for the two series. This is true for both the full sample of bets and for various sub-samples, such as bets on home teams only, bets on away teams only and bets only on teams either favoured or not favoured by the handicap (strong or weak teams).

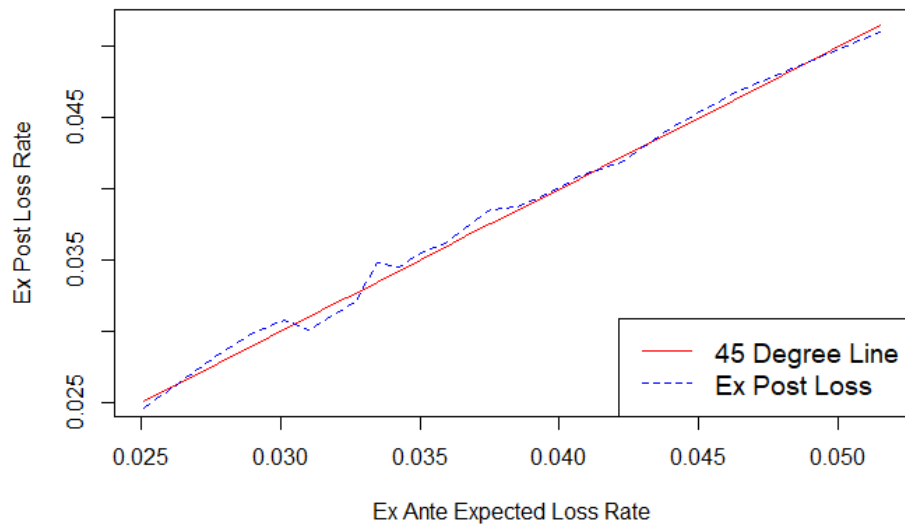
To visually illustrate the close matching between ex ante and ex post losses, we plot the ex post losses (blue line) for each ex ante loss quantile in Figure 7. The 45-degree line (red) indicates when ex ante and ex post losses are equal, and in the case of the Asian handicap, the two lines fit closely. This result is in stark contrast to the traditional market (Figure 2), where ex post losses are almost always higher than ex ante expected losses.

Table 7: Mean Expected Ex Ante Loss Rates Compared with Mean Actual Loss Rates for Different Betting Strategies in the Asian Handicap Market

| | Mean | <i>N</i> |
|--|--------|----------|
| Ex Ante Expected Loss Rate | 0.0361 | 84,230 |
| <i>Actual Ex Ante Loss Rates</i> | | |
| Loss Rate From All Bets on Home & Away | 0.0363 | 168,460 |
| Loss Rate From All Bets on Home | 0.0411 | 84,230 |
| Loss Rate From All Bets on Away | 0.0316 | 84,230 |
| Loss Rate From All Bets on Strong* | 0.0417 | 69,910 |
| Loss Rate From All Bets on Weak* | 0.0328 | 69,910 |

*Zero handicap is not considered for strong and weak bets

Figure 7: Losses Rates on an Equally-Weighted Portfolio of Home and Away Asian Handicap Bets
By Ex Ante Expected Loss Rate



5. Conclusions

The evolution of online betting on soccer has led to the emergence of two distinct betting markets. Using a large sample of European soccer matches, we show that odds in the traditional market for bets on a home win, away win or draw are systematically biased with bets on favourites likely to lose less than bets on longshots. We also provide evidence that in this market, bettors cannot easily establish their expected loss rate from calculations using the bookmaker's odds under the assumption of market efficiency. In contrast, we find that the Asian handicap betting market behaves in an efficient manner for the same set of matches. This market shows no pattern of favorite-longshot bias, its implied probabilities are unbiased and its implied ex ante expected loss rates accurately predict the actual ex post loss rates.

What explains these results? One explanation is that the population of bettors is different across the two markets we have examined. The low-margin "winners welcome" ethos of the bookmakers that dominate the Asian handicap market attracts professional syndicates and some of the sharpest minds in sports betting. These bookmakers, however, do not have a retail presence in Europe and opening accounts with them is tricky in many European countries. This leaves those who bet smaller stakes and are perhaps less informed to place their bets with the "traditional" bookmakers who do not promote Asian Handicap bets. There is also perhaps a difference in attitudes to risk across the customer base of the two markets, with bettors making smaller bets in the home/away/draw market perhaps having more preference for high risk longshot bets than those who have large amounts of money at stake in the Asian Handicap market.

Beyond the differences in their customer bases, another likely explanation of our results is the competitive structure of the two markets. The low margins offered in the Asian Handicap market are indicative of a high level of competition. This market also operates in a transparent manner with bookmakers willing to take very large bets from customers. In contrast, the traditional bookmakers have higher gross profit margins and place restrictions on who can bet and how much can be placed. While websites exist with odds comparisons that suggest these bookmakers compete to offer the best odds, those who selectively choose the best available odds usually find the amounts that can be placed at those odds are small and those with a record of making profits tend to be banned.

These restrictive practices suggest the traditional market is not a particularly competitive one. Indeed, the evidence we have presented shows that bookmakers in the home/away/draw market are making large average profits on bets on longshots which suggests a lack of competition because these high profits are not being competed away by some bookmakers choosing to offer more attractive odds on longshot bets. The roles played in generating these outcomes by the differences in customer base and the differences in competitive structures are likely to be useful areas for future research.

References

- Angelini, G. and De Angelis, L. (2019). Efficiency of online football betting markets. *International Journal of Forecasting*, 35(2):712–721.
- Berkowitz, J. P., Depken II, C. A., and Gandar, J. M. (2017). A favorite-longshot bias in fixed-odds betting markets: Evidence from college basketball and college football. *The Quarterly Review of Economics and Finance*, 63:233–239.
- Buchdahl, J. (2016). *Squares & sharps, suckers & sharks: The science, psychology & philosophy of gambling*. Oldcastle Books.
- Cain, M., Law, D., and Peel, D. (2000). The favourite-longshot bias and market efficiency in UK football betting. *Scottish Journal of Political Economy*, 47(1):25–36.
- Constantinou, A. C. (2021). Investigating the efficiency of the Asian handicap football betting market with ratings and bayesian networks. *Journal of Sports Analytics*, (Preprint):1–23.
- Davies, R. (2022). Revealed: how bookies clamp down on successful gamblers and exploit the rest. <https://www.theguardian.com/society/2022/feb/19/stake-factoring-how-bookies-clamp-down-on-successful-gamblers>.
- Gray, P. K. and Gray, S. F. (1997). Testing market efficiency: Evidence from the NFL sports betting market. *The Journal of Finance*, 52(4):1725–1737.
- Hegarty, T. (2021). Information and price efficiency in the absence of home crowd advantage. *Applied Economics Letters*, 28(21):1902–1907.
- Hegarty, T. and Whelan, K. (2022). The wisdom of no crowds: The reaction of betting markets to lockdown soccer games. University College Dublin Working Paper Series WP2022/13.
- Hegarty, T. and Whelan, K. (2023). Calculating the bookmaker’s margin: Why bets lose more on average than you are warned. University College Dublin Working Paper.
- Moscowitz, T. and Vasudevan, K. (2022). Betting without beta. Yale University Working paper.
- Pope, P. F. and Peel, D. A. (1989). Information, prices and efficiency in a fixed-odds betting market. *Economica*, pages 323–341.
- Snowberg, E. and Wolfers, J. (2008). Examining explanations of a market anomaly: Preferences or perceptions? In *Handbook of sports and lottery markets*, pages 103–136. Elsevier.