Disagreement and Market Structure in Betting Markets: Theory and Evidence from European Soccer

Hegarty, Tadgh and Whelan, Karl

University College Dublin, University College Dublin

8 May 2023

Disagreement and Market Structure in Betting Markets: Theory and Evidence from European Soccer

Tadgh Hegarty[*] Karl Whelan[†]

University College Dublin

May 8, 2023

Abstract

Online sports betting is growing rapidly around the world. We describe how the competitive structure of the bookmaking market affects odds when bettors disagree about the probabilities of the outcomes of sporting events but are on average correct. We show that the demand for bets on longshots is less sensitive to the odds than bets on favorites. This means monopolistic bookmakers will set odds exhibiting favorite-longshot bias while competitive bookmaking markets will not have this feature. We develop a version of the model for soccer matches and use these results to explain empirical findings on odds for over 80,000 European soccer games from two different bookmaking markets.

Keywords: Favorite-Longshot Bias, Monopoly, Sports Betting

JEL Classification: D42, G14, L83

[*]tadgh.hegarty@ucdconnect.ie

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

1. Introduction

Sports betting markets have long been studied by economists. Their appeal has largely been because they provide a convenient setting for assessing how people take decisions involving risk rather than because these markets were substantively important.[1] However, mobile internet technology has turned online sports betting into a huge global business in recent years. It is also growing rapidly in the US after a 2018 Supreme Court decision voiding previous federal restrictions on sports betting. The increasing prominence of online sports betting increases the importance of understanding how prices are set in these markets.

Theoretical research on this topic has largely concentrated on pari-mutuel betting, which pools the funds placed by gamblers and pays them out (minus a fraction to cover costs and profits) to those who picked the winner in proportion to the size of their bet. This focus reflected the fact that pari-mutuel betting at horse racing tracks was exempted from US prohibitions on sports betting and was thus the best-known way to bet in the US. However, modern online sports betting uses fixed odds set by bookmakers: They make offers like "You take back $3 if your pick wins and lose your $1 bet otherwise" and if a bettor accepts this offer, their odds are not changed by subsequent actions of other bettors.

The literature on sports betting has often focused on the empirical pattern that bets on longshots in pari-mutuel markets tend to lose more than bet on favorites, meaning market odds do not imply unbiased measures of the underlying probabilities. Because pari-mutuel payouts are determined by betting volumes, theoretical explanations of this bias have focused on the behavior of bettors. For example, explanations have included that gamblers may have risk-loving preferences and enjoy the high-variance returns from betting on low-probability outcomes (Quandt, 1986), that they systematically overestimate probabilities of unlikely events (Thaler and Ziemba, 1988) or that there may be "noise gamblers" who pick their bets randomly (Hurley and McDonough, 1995).[2]

These explanations do not necessarily apply to modern fixed-odds betting markets. Odds in these markets are set by bookmakers and they do not need to be inversely-related to volumes. This means the behavioral biases of bettors are less likely to explain favorite-longshot bias in fixed-odds markets. And yet, favorite-longshot bias has also been reported for some fixed-odds betting markets.[3]

In this paper, we present a simple approach that can explain the pricing patterns observed in different fixed-odds betting markets. We assume that potential bettors disagree about the probabilities of the outcomes of sporting events and, as in Ali (1977), their distributions of beliefs centers around the true probabilities, so while essentially no individual bettors have the correct beliefs, the average

[1]See, for example, Sauer (1998) for a review.

[2]Ottaviani and Sørensen (2008) is a comprehensive review of the various theories on favorite-longshot bias.

[3]Studies finding evidence of favorite-longshot bias in fixed-odds betting markets, include Gabriel and Mardsen (1990) and Snowberg and Wolfers (2010) for UK horse races, Berkowitz, Depken and Gandarc (2017) for US college football and basketball and Forrest and McHale (2007) for tennis.

belief has a "wisdom of crowds" property. Ali showed this assumption implied favorite-longshot bias in pari-mutuel betting markets.[4] We show how disagreement among bettors generates favorite-longshot bias in fixed-odds markets when betting markets have a monopolistic structure but will not generate this bias when markets are competitive. We provide empirical evidence from betting on European soccer to support these results.

The paper is structured as follows. First, we consider a perfectly competitive fixed-odds betting market with a normal profit requirement. Odds in this market do not exhibit favorite-longshot bias and factors such as disagreement among bettors or other mechanisms highlighted in discussions of pari-mutuel markets do not influence the odds.

Second, we show that monopolistic bookmakers will set odds that exhibit favorite-longshot bias. Our model of a monopolist bookmaker shares some similarities with Shin (1991) but also has key differences. In Shin's model, some of the bettors are insiders who know the outcome of the game in advance while the rest have no useful information. In our model, there are no insiders who know the outcome in advance (a possibility we believe is generally unlikely to apply to the high-profile sporting events that dominate volumes in modern online betting) and the public are, on average, correct in their assessment of probabilities of various outcomes.

Our result stems from standard microeconomic foundations. As shown by Montone (2021), monopolist bookmakers set odds as a "markdown" on their perfectly competitive levels with the size of the markdown depending on the elasticity of demand of bettors: Bettors that are more price-sensitive get quoted better odds. We show that disagreement among bettors implies the demand for bets on favorites is more sensitive to odds than demand for bets on longshots and our model predicts that expected loss rates on bets rise nonlinearly as the probability of winning declines.

The result about price sensitivity of demand for bets can be explained as follows. Consider an event with two competitors where the probability of the favorite winning is $p = 0.7$ and bettors have beliefs $\tilde{p}$ about p. Let O_F and O_L be the so-called decimal odds for the favorite and longshot (the payout on a successful \$1 bet inclusive of the original bet). The "fair odds" at which bets break even on average are $O_F = \frac{1}{0.7} = 1.43$ and $O_L = \frac{1}{0.3} = 3.33$. Assume bettors make bets if they believe the expected payout is at least one, i.e. $\tilde{p}O_i \geq 1$, and suppose the bettors that are most optimistic about the favorite have $\tilde{p} = 0.8$ while those who are most optimistic about the longshot have $\tilde{p} = 0.6$. Those who are most optimistic about the favorite's prospects would accept odds of $O_F = \frac{1}{0.8} = 1.25$, which is one-eighth lower than competitive odds. Conversely, those who are most optimistic about the longshot would accept odds of $O_L = \frac{1}{0.4} = 2.5$ which is one-quarter lower than competitive odds.

Third, we present a version of our model designed to match betting odds on soccer games where there are three possible outcomes: Home win, away win or draw. The model predicts there should

[4]Gandhi and Serrano-Padial (2015) generalized these results to illustrate how disagreement generates favorite-longshot bias in pari-mutuel betting as long as the true probability is part of set of beliefs.

usually be less disagreement about the probability of a draw than the probability of a longshot win. This means a monopolistic bookmaker will set odds so that loss rates on draws will often be lower than for longshots, with losses similar to those for strong favorites if the probability of a draw is close to one third.

Fourth, we use these results to explain betting odds for over 80,000 European soccer games. We show that the market for betting on home/away/draw outcomes exhibits substantial favorite-longshot bias, as found in previous studies with smaller samples such as Buhagiar, Cortis and Newall (2018) and Angelini and de Angelis (2019). These losses closely match the nonlinear pattern predicted by our model. The data also confirm our monopolistic model's specific predictions about loss rates for bets on draws.

We also compare our results for home/away/draw bets with results from bets placed on the same set of games in the so-called Asian Handicap market, where payouts on bets depend on an adjustment of the match result that applies a deduction to the goals total of the team considered more likely to win. The Asian handicap market is known to be popular with professional gamblers who place large-volume bets with "sharp" bookmakers who compete against each other to offer the most attractive odds. These bookmakers make low profit margins per bet but offset this by generating high volume. In contrast, the market for home/away/draw bets is dominated by retail European bookmakers who have high margins and engage in various anti-competitive practices such as limiting or banning customers who appear to be actively seeking out the best odds. These firms rarely compete on odds offered, focusing instead on marketing and promotions to attract and retain high-revenue customers.

Our models suggest the market dominated by the bookmakers who operate in an essentially competitive manner should not exhibit favorite-longshot bias while the market dominated by firms with anti-competitive practices will feature this bias. The evidence confirms what the models suggest. The Asian handicap market does not exhibit favorite-longshot bias for the same set of games for which a significant bias is reported for home/away/draw betting. By its nature, the handicap market features fewer extreme longshots or favorites but within the probability ranges where you can compare both types of bets, the pattern for loss rates is very different.

Finally, we conclude with some discussion of our results including a comparison with the work of Moskowitz and Vasudevan (2022), which like our empirical evidence, compares the pattern of returns for "moneyline" bets on which team wins versus returns on spread-type bets across the same set of matches.

2. Competitive Fixed Odds Bookmakers

We first consider a competitive fixed odds bookmaking market in which there is an event with N possible outcomes, each with a probability p_i of happening. Bookmakers know the values of p_i and offer odds O_i so that a \$1 bet returns O_i back to the gambler if event i happens. There are a number of bookmakers who compete against each other for business and the competitive equilibrium requires each to make an expected profit rate of θ on each bet, reflecting the need to earn a "normal" rate of profit. If expected profit rates are higher than this, new bookmakers enter the market and offer lower odds to bring the profit rate back to θ. Costs for bookmakers are a fraction μ of the total dollar amount of bets placed.

The competitive equilibrium profit rate of θ implies the following equality for revenue minus costs on a \$1 bet

$$1 - \mu - p_i O_i = \theta \tag{1}$$

so the equilibrium odds are

$$O_i = \frac{1 - \mu - \theta}{p_i} \tag{2}$$

This means that ratio of the odds for two bets reflects the ratio of the underlying probabilities. Bettors can also figure out the value of $\mu + \theta$ from these odds because

$$\sum_{i=1}^{N} \frac{1}{O_i} = \sum_{i=1}^{N} \frac{p_i}{1 - \mu - \theta} = \frac{1}{1 - \mu - \theta} \tag{3}$$

This means

$$\mu + \theta = 1 - \frac{1}{\sum_{i=1}^{N} \frac{1}{O_i}} \tag{4}$$

In other words, gamblers can use the "overround" from the quoted odds—the sum of the inverse of the odds—to calculate the bookmaker's gross profit margin and with this value calculated, bettors can also figure out the underlying probabilities from the odds using equation 1. As discussed in Hegarty and Whelan (2023), most practical advice on sports betting recommends bettors do this calculation to assess the gross profit rate of bookmakers on bets on an event. Also, once they have calculated $\mu + \theta$, bettors can back out the values of the underlying probabilities from the odds.

Without describing how potential gamblers behave, we can see that the theoretical devices that have been used to explain favorite-longshot bias in pari-mutuel markets will not influence odds in this market. Gamblers may have locally risk-loving preferences and enjoy making high-variance low-probability bets, as suggested by Quandt (1986). Alternatively, gamblers may systematically overestimate the probabilities of unlikely events, as suggested by Kahenman and Tversky's (1979) prospect theory. There could be "noise bettors" who pick their bets randomly, as modeled by Hurley and McDonough (1995) or perhaps gamblers who choose to back a team because they are a fan.

Each of these features drives betting volume towards longshot bets in a way that is unrelated to the fundamental chance the bet has of winning. Because the pari-mutuel structure requires winning odds to be inversely related to betting volumes, this reduces the winning pari-mutuel odds on the longshot team and increases them on the favorite. However, these factors would have no influence on odds in a competitive market because these odds are not determined by betting volumes. In each case, these mechanisms imply the existence of gamblers who would back the longshot at lower odds than the competitive prices. But getting those gamblers to accept these poorer odds would generate rates of return for bookmakers above θ and incentivize other bookmakers to undercut those making these excess profits.

The form of disagreement among bettors that we discuss below—which follows Ali (1977) in assuming that bettors have symmetric beliefs about probabilities that center on their true values—will also not induce favorite-longshot bias in competitive markets. Ali explained how pari-mutuel betting odds for favorites needed to be more attractive in terms of expected returns for the average bettor than for longshots to attract the additional volumes required for odds on favorites to be lower than for longshots. In contrast, in competitive markets, odds do not depend on betting volumes so this kind of disagreement would have no impact.

3. Disagreement with Monopolistic Bookmakers

We now consider the case where the bookmaker is a monopolist. Our model shares similarities with Shin (1991) in that it has a monopolistic bookmaker and disagreement among bettors generates demand functions for bets on each outcome. There are, however, two important differences. First, we do not incorporate "insiders" who know the outcome of the sporting event before it has happened. This motivation for this feature stemmed from the early literature on sports betting which was focused largely on horse racing, where sometimes those who are involved in training a horse will know it has a much higher chance of winning than the general public. In the high-profile sports that dominate modern online betting, this level of inside information is generally unlikely to apply.

Second, Shin's model assumes that non-insider bettors have no useful knowledge about the sporting event. In two-team event, he assumes these bettors have beliefs about the probability that one of the teams will win that are uniform on $[0, 1]$ no matter what the true value of this probability is. In contrast, we will assume the betting public is, on average, correct in its assessment of the underlying probabilities.

3.1. How the monopolist bookmaker sets odds

Consider again a contest with N possible outcomes, each with a probability p_i of happening. A monopolistic bookmaker, who knows the probabilities, has a processing cost for each bet equal to a fraction μ of the amount placed and faces a demand function for bets $D(O_i)$ such that $D'(O_i) > 0$.

At first sight, bookmaking seems a very different business to the typical monopolist in economics. The prices set (the betting odds) have only an indirect influence on its revenues but have a direct influence on its costs. However, Montone (2021) has shown that profit-maximizing odds for monopoly bookmakers follow the same logic as the traditional economic theory of monopoly pricing. This result is derived as follows. The bookmaker's expected profit on bets placed on outcome i is

$$E(\Pi_i) = (1 - \mu - p_i O_i) D(O_i) \tag{5}$$

The condition for maximizing expect profit is

$$\frac{dE(\Pi)}{dO_i} = (1 - \mu - p_i O_i) D'(O_i) - p_i D(O_i) = 0 \tag{6}$$

This solves to give

$$O_i = \frac{1-\mu}{p_i} - \frac{D(O_i)}{D'(O_i)} \tag{7}$$

We can see immediately that the odds offered by a monopolist will be lower than in a zero-profit competitive market. Defining the elasticity of demand for the bet as

$$\epsilon_i = \frac{O_i D'(O_i)}{D(O_i)} \tag{8}$$

optimal odds become

$$O_i = \frac{\epsilon_i}{\epsilon_i + 1} \frac{1-\mu}{p_i} \tag{9}$$

The bookmaker chooses to set odds as a "markdown" on the zero-profit competitive odds and bets with higher elasticities of demand will have lower markdowns, with odds tending towards their zero-profit-margin competitive level as the elasticity of demand becomes infinite. If we interpret the price set by the monopolist as the inverse of the odds, we recover the standard result that the monopoly price is a markup over the competitive price with the usual markup formula.

3.2. Disagreement and elasticity of demand

So what determines elasticity of demand and thus betting odds? We use a simple model of bettors who disagree about the probabilities of outcomes and show that the demand for longshot bets is less sensitive to prices than demand for bets on favorites.

We assume there is a continuum of risk-neutral potential gamblers of size 1 who can chose to bet or not. When they do bet, they place equal sized bets, normalized to equal one. Potential gamblers have a subjective belief $\tilde{p_i}$ about the true probability p_i that a sporting event ends with the i-th specific outcome. As in Ali (1977), beliefs are characterized by a cumulative distribution function $F(\tilde{p_i})$ with mean p_i, meaning the gamblers are on average correct about the probability of the event occurring, and beliefs are symmetrically distributed around p. Potential bettors are risk-neutral and only bet if the odds imply an expected profit greater than or equal to zero, so they take the bet on outcome i if

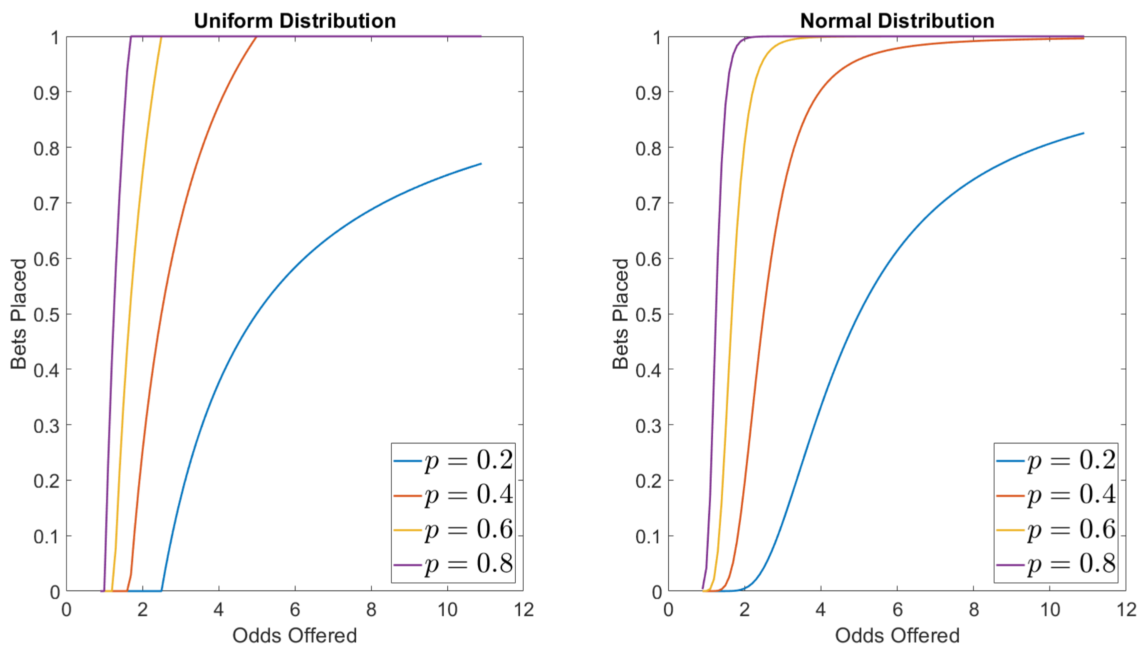
$$\tilde{p_i} O_i \geq 1 \tag{10}$$

In the introduction, we described the intuition for why demand for longshot bets is less sensitive to changes in odds than demand for favorite bets. A concrete example of a distribution of beliefs can illustrate this result better. Consider again an event with two competitors where the probability of the favorite winning is $p = 0.7$. Assume now that potential bettors have beliefs $\tilde{p}$ about p that are uniformly distributed on $[0.6, 0.8]$ so that, on average, the public's belief about this probability is correct. As noted above, the "fair odds" in a market with no gross profits for bookmakers would be $O_F = \frac{1}{0.7} = 1.43$ and $O_L = \frac{1}{0.3} = 3.33$. At these odds, half the potential bettors (those with $\tilde{p} \geq 0.7$) will take the bet on the favorite and the other half will take the bet on the longshot.

Now suppose the odds on both teams were cut by 10%. At the new favorite odds of $O_F = 1.29$ only those with $\tilde{p} > 0.777 = \frac{1}{1.29}$ will take the bet, meaning only 11% of potential bettors now bet on the favorite. In contrast, at the new longshot odds of $O_L = 3$, those with $1 - \tilde{p} > \frac{1}{3}$, meaning $\tilde{p} < 0.667$ will take the bet, meaning 33% of potential bettors will still take the longshot bet. This illustrates the weaker price sensitivity of demand for longshot bets.

This intuition about demand for low probability bets being less elastic does not stem from assuming a uniform distribution for beliefs. Figure 1 shows the demand for bets as a function of odds on an outcome with a probability of p_i using different values of p_i and using both uniform distributions for beliefs on $[p_i - 0.2, p_i + 0.2]$ and normal distributions with mean p_i and standard deviation 0.115 (the same as the standard deviation for the uniform distribution). We see similar patterns for the demand functions for both normal and uniform distributions. Demand functions for bets on high probability events are much steeper than for bets on low probability events.

Figure 1: Total bets placed as a function of odds and underlying probability of bet win p



Notes: The uniform belief distributions are of the form $U\,[p - 0.2, p + 0.2]$. The normally distributed beliefs are of the form $N\,(p, 0.115)$.

3.3. Odds with a uniform distribution

To provide a concrete illustration of how the monopolist sets odds, we will stick with beliefs about the probability p_i being uniformly distributed over $[p_i - \sigma_i, p_i + \sigma_i]$.[5] This implies a demand function for bets on outcome i of the form

$$D(O_i) = 1 - \frac{1}{2\sigma_i}\left[\frac{1}{O_i} - p_i + \sigma_i\right] \tag{11}$$

The elasticity of demand with respect to odds is given by

$$\epsilon_i = \frac{1}{\sqrt{1-\mu}\sqrt{1+\frac{\sigma_i}{p_i}} - 1} \tag{12}$$

As σ_i rises, the elasticity of demand falls while as p_i rises, the elasticity of demand increases. Inserting this expression for the elasticity into the pricing formula, equation 9, the profit-maximizing odds are

$$O_i = \sqrt{\frac{1-\mu}{p_i(p_i+\sigma_i)}} \tag{13}$$

This formula generates a favorite-longshot bias pattern of returns on longshot bets being worse than for favorite bets. To illustrate the magnitudes, suppose $p_i = 0.8$, $\mu = 0$ and $\sigma_i = 0.1$. A competitive zero-profit bookmaker will set odds of $O_i = \frac{1}{0.8} = 1.25$ while the monopolist will set odds of $O_i = 1.18$, which are 6 percent worse. In contrast, if $p_i = 0.2$, the zero-profit competitive odds are $O_i = \frac{1}{0.2} = 5$ while the monopolist odds are 4.08, which are 18 percent worse.

The expected profit rate per $1 bet for the bookmaker is

$$1 - \mu - p_i O_i = (1-\mu)\left(1 - \sqrt{\frac{1}{1-\mu}\frac{p_i}{p_i+\sigma_i}}\right) \tag{14}$$

The profit rate rises as the probability of the team winning declines and also depends positively on σ. Inserting the odds into the demand function, the total amount of bets is

$$D(O_i) = 1 - \frac{1}{2\sigma}\left[\sqrt{\frac{p_i(p_i+\sigma_i)}{1-\mu}} - p_i + \sigma_i\right] \tag{15}$$

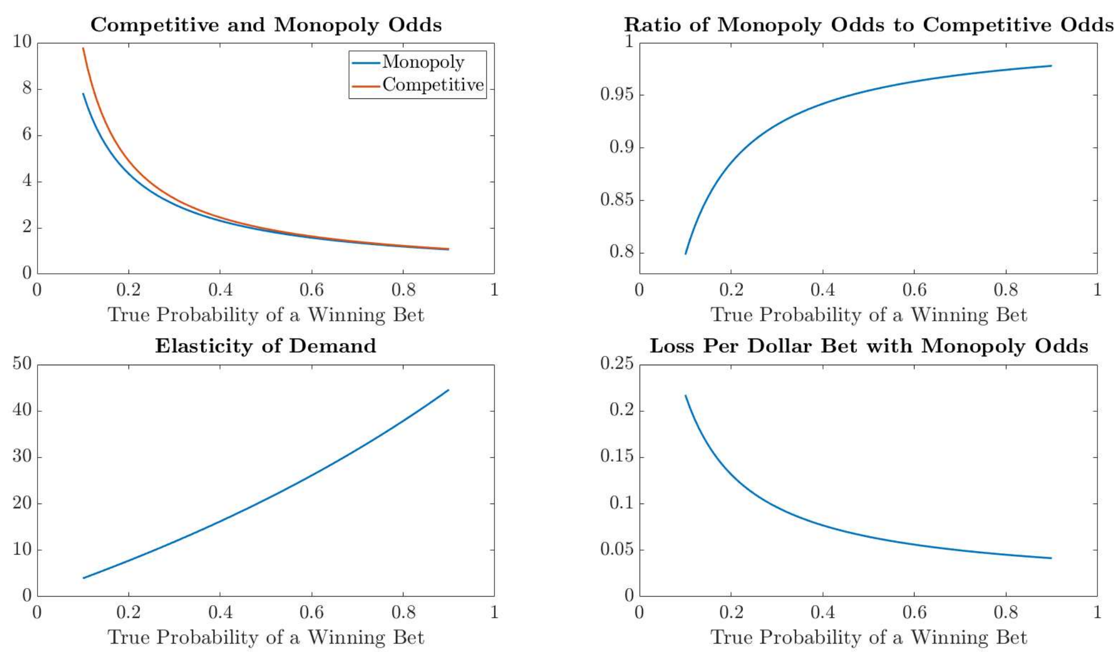
This declines as p_i rises because the effect on the square root term is smaller than the effect on the linear term being subtracted. So, despite being offered comparatively worse odds with greater losses per bet, the quantity of bets on longshots is greater than on favorites. Betting volumes also rise with

[5] Because the probabilities of all possible outcomes must sum to one, we cannot specify separate distributions of beliefs for all N probabilities. Once $N-1$ belief distributions have been specified, the beliefs for the other probability are then a function of them. We provide an example of how to deal with this below in the context of soccer matches that can end in a home win, away win or draw.

σ_i because the more disagreement there is, the more people there are who believe they are going to beat the bookmaker's margin and make a profit.

The upper panels of Figure 2 illustrate how the monopolist's odds differ from competitive odds for various values of p_i ranging from 0.1 to 0.9 and for uniform beliefs on $[p_i - \sigma_i, p_i + \sigma_i]$ with $\sigma_i = 0.06$, $\mu = 0.02$ and $\theta = 0$ (we will explain our choices of $\sigma_i = 0.06$ and $\mu = 0.02$ later). The monopolist's odds decline substantially relative to competitive odds as the true probability of the bet winning falls. The bottom left panel shows how the equilibrium elasticity of demand rises sharply as p_i goes up, meaning smaller markdowns on competitive odds for bets that are more likely to win. The bottom right panel shows the loss rates for bets as a function of the bet's probability of winning. There is a highly nonlinear pattern of rising loss rates as the probability of winning falls. For this calibration, longshot bets with win probabilities as low as ten percent have expected loss rates of over 20 percent.

Figure 2: Comparing competitive and monopoly odds ($\sigma = 0.06, \mu = 0.02, \theta = 0$)



3.4. Two outcome events

The derivation just presented assumed the extent of disagreement about probabilities differed for each possible outcome. However, with two possible outcomes, if σ represents the amount of disagreement about the probability of the favorite winning, it also represents the amount of disagreement about the probability of the longshot winning. Thus, for a game in which one team is the favorite with a probability of winning $p > 0.5$ and the public's beliefs are uniform on $[p - \sigma, p + \sigma]$, the ratio of favorite to longshot odds will be

$$\frac{O_F}{O_L} = \sqrt{\frac{(1-p)(1-p+\sigma)}{p(p+\sigma)}} \tag{16}$$

Note that if there is no disagreement in beliefs, so $\sigma = 0$, the formula above gives the same odds as the competitive market. However, in this model, the betting market would collapse in this case because all bettors would have a correct assessment of the value of p and none of them would believe they could make a profit given the bookmaker's need to cover costs.[6]

The square-root formula in equation (16) looks similar to a formula in Shin's (1991) paper for the ratio of odds on the favorite to odds on the longshot

$$\frac{O_F}{O_L} = \sqrt{\frac{1-p}{p}} \tag{17}$$

but the underlying assumptions generating the formulas are different. The bettors in our model are correct, on average, about the probability of each team winning whereas the "outsiders" in Shin's model have uniform beliefs over $[0, 1]$. Only when $p = 0.5$ and $\sigma = 0.5$ do the two formulas imply the same odds.

We believe our assumptions are a more reasonable representation of bettors in modern online sports markets. When considering Manchester City versus Everton, there may be many people who do not have an accurate opinion on the likely outcome. However, the population of people who are willing to consider making a bet on this game will surely have opinions that are somewhat informed by reality (which, at the time of writing, is that City are much better than Everton). Given this, we view the assumption that the set of potential bettors are on average correct in their assessment as an appropriate modeling approach. One can also show that Shin's formula implies a much larger favorite-longshot bias than is observed in datasets such as the one we use for European soccer later in this paper.

In the two outcome case, bookmakers make higher profits on games that are mis-matches than on games that are closer to toss-ups. The nonlinear pattern of increased loss rates for longshots (and thus increased profits for bookmakers) means the additional gains from bets on more extreme longshots

[6]The use of a fixed value of σ in this example does not imply that σ has to be fixed across different values of p.

offsets the slightly lower profits obtained from bets on a favorite with a high win probability. One question about this result is whether gamblers can tell from the odds that the bookmakers are making larger profits on bets on games with a strong favorite and thus that at least one of the bets may offer poor value. As discussed above, gamblers tend to use the overround implied by the odds to measure a bookmaker's gross profit margin. In a competitive market, the overround is

$$R = \frac{\mu + \theta}{1 - \mu - \theta} \tag{18}$$

which depends positively on bookmakers' costs and the profit margin requirement and would be zero if these didn't exist.

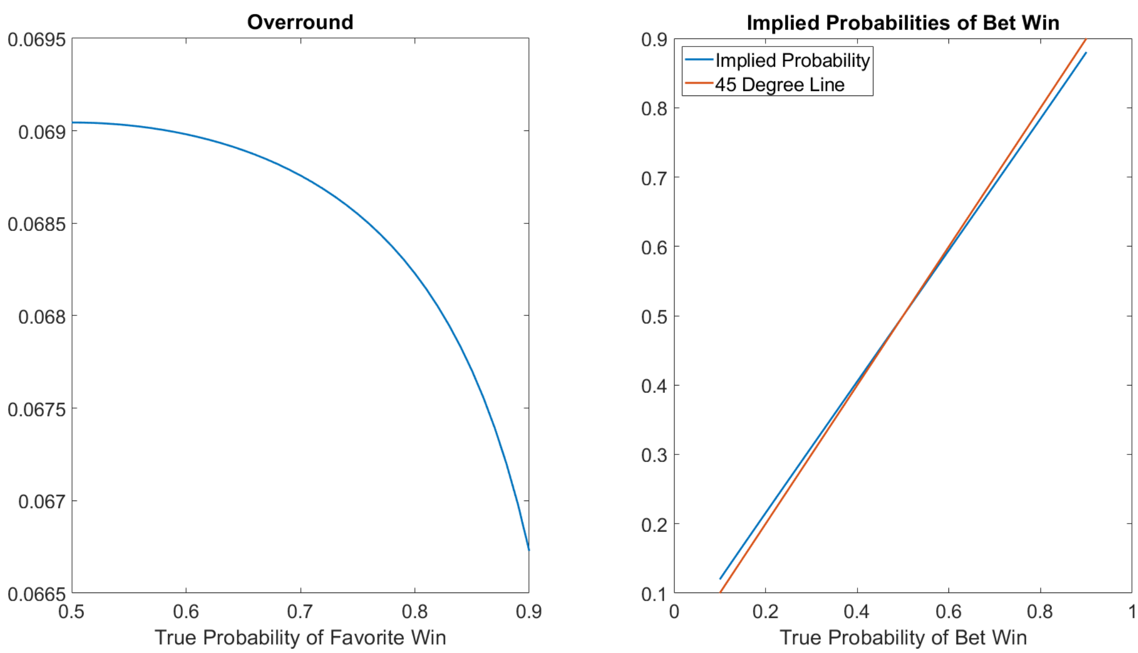
In this model, the overround is

$$\begin{aligned} R &= \frac{1}{O_F} + \frac{1}{O_L} - 1 && (19) \\ &= \sqrt{\frac{p\,(p+\sigma)}{1-\mu}} + \sqrt{\frac{(1-p)\,(1-p+\sigma)}{1-\mu}} - 1 && (20) \end{aligned}$$

This depends positively on μ and σ. Calculations show that the overround increases almost one-for-one with the value of σ and for each value of σ there is very little variation in the overround with the value of p. Indeed, perhaps surprisingly, the model predicts that overrounds for games where there is a strong favorite will be slightly lower than for games that are effectively toss ups: See the left panel of Figure 3. This means the overround does not send the correct signal to gamblers about the profit rates being earned by the bookmaker.

The implicit model behind using the overround to calculate the bookmaker's gross profit margin—that odds start as the inverse of the probabilities and are then marked down to cover a margin—does not describe how the monopolist bookmaker sets odds. The right panel of Figure 3 shows the problem with using the overround to estimate gross profit margins and probabilities. The probabilities implied by the standard way of extracting the bookmaker's margin from odds (described above in Section 2) do not match with the actual probabilities. Longshots win less often than this method suggests and favorites win more often.

Figure 3: Overround with $\sigma = 0.06$ and comparison of true and implied probabilities ($\mu = 0.02, \sigma = 0.06$)



4. A Model of Betting on Soccer

The monopolistic model with two outcomes can be applied to many sporting events, such as tennis matches or "moneyline" bets on Major League Baseball or NBA basketball games. However, we will use the model to explain the odds on European soccer matches where people can bet on three possible outcomes: A home win, an away win and a draw. As well providing an empirical example of a huge betting market, this larger set of outcomes helps to generate some unique predictions of our model that can then be assessed against the data.

Calibrating a model with three outcomes requires specifying beliefs about more than one probability. One way to do this would be to assume a fixed number for the probability of a draw, so the disagreement is only about the probabilities of the home and away wins. Pope and Peel (1989), for example, used data from British soccer matches to argue that while the overall percentage of matches ending in draws was predictable, there was no reliable way to forecast which kinds of matches might end in a draw. That result, however, does not hold in our comprehensive dataset of European soccer matches. In fact, contingent on the the probability of one team winning, there is a clear pattern for the two other probabilities.

Figure 4 shows the evidence on this from our sample of over 80,000 European soccer games (more details on the dataset are below). The left panel shows the percentage of draws and longshot wins charted against the percentage of favorite wins, with the data having been sorted by deciles of the ex ante probability of the favorite winning, calculated using the traditional "overround" method. It shows that as the probability of the favorite winning rises, the probability of both draws and longshot wins decline but the probability of the longshot win falls more than the probability of a draw. The right panel shows that as the probability of the favorite winning increases, the share of the other possible results accounted for by draws rises according to an essentially linear pattern. Intuitively, the more likely the favorite is to win, the more likely it is that the best result the longshot can get is a draw.

In light of this evidence, we calibrate beliefs about the probabilities of draws ($\tilde{p}_D$) and longshot wins ($\tilde{p}_L$) as functions of beliefs about the probability of the favorite winning ($\tilde{p}_F$). Those with subjective belief $\tilde{p}_F$ about the probability that the favorite will win are assumed to have subjective beliefs about the probability of draws and longshot wins given by

$$\tilde{p}_D = (\delta_0 + \delta_1 \tilde{p}_F)(1 - \tilde{p}_F) = \phi_D(\tilde{p}_F) \tag{21}$$

$$\tilde{p}_L = (1 - \delta_0 - \delta_1 \tilde{p}_F)(1 - \tilde{p}_F) = \phi_L(\tilde{p}_F) \tag{22}$$

We use $\delta_0 = 0.338$ and $\delta_1 = 0.405$ to calibrate our model—these are the values from a regression using the data in the right panel of Figure 4.

Technically, the functions $\phi_D(\tilde{p}_F)$ and $\phi_L(\tilde{p}_F)$ are quadratics but, given our calibration, both are decreasing functions over the range of favorite probabilities observed in our model. This means the decision rules for making bets on draws is to make the bet if

$$\tilde{p}_D \geq \frac{1}{O_D} \Longrightarrow \phi_D\left(\tilde{p}_F\right) \geq \frac{1}{O_D} \Longrightarrow \tilde{p}_F < \phi_D^{-1}\left(\frac{1}{O_D}\right) \tag{23}$$

with the corresponding rule for longshots is to make the bet if

$$\tilde{p}_F < \phi_L^{-1}\left(\frac{1}{O_L}\right) \tag{24}$$

These inverse functions, and thus the threshold levels of $\tilde{p}_F$ are uniquely defined over the probability range of $\tilde{p}_F$ being considered. Further details on deriving these decision rules are contained in an appendix.

Because we want to match this theoretical model with the evidence from home/away/draw betting, we need a realistic calibration of the parameters μ and σ. For μ, which measures non-payout costs per bet for the bookmakers, we can point to some bookmaking markets having margins as low as 3%. As such, we think $\mu = 0.02$ represents a realistic cost rate. For σ, we can use observed values of the overround to calibrate its value. As with the two-outcome version of the model, the average overround in this version moves almost one for one with the value of σ. The average overround for home/away/draw bets in our dataset is 6.9% and this matches well with the model if we assume the value of $\sigma = 0.06$ that we have already used in previous illustrative charts.

The top left panel of Figure 5 shows the odds generated by this model for favorites, longshots and draws as a function of the probability of each bet's success. Over the full range of probabilities, the most apparent prediction is the obvious one that odds on bets fall as their probability of success rises. However, the apparently tiny differences between the yellow line (odds on longshots) and the orange line (odds on draws) actually generate significant differences in loss rates on these two kinds of bets. The top right panel also shows that, for most of the range of relevant probabilities, loss rates on draws are lower than loss rates on longshots. Indeed, when draws are most likely—when the match is effectively considered a toss-up between the three outcomes—the loss rates on draws are as low as on extreme favorites.

This result occurs because loss rates depend on the level of disagreement as well as the true probability. The bottom left panel shows the level of disagreement about each type of bet: For each value of p_F it reports the gap between the highest and lowest beliefs about the probabilities of the three outcomes. Only when the probability of the favorite winning is very high does disagreement about the probability of a draw become larger than disagreement about the probability of a longshot win and, at this point, the model predicts that the loss rate on a draw becomes larger than the loss rate on longshots.

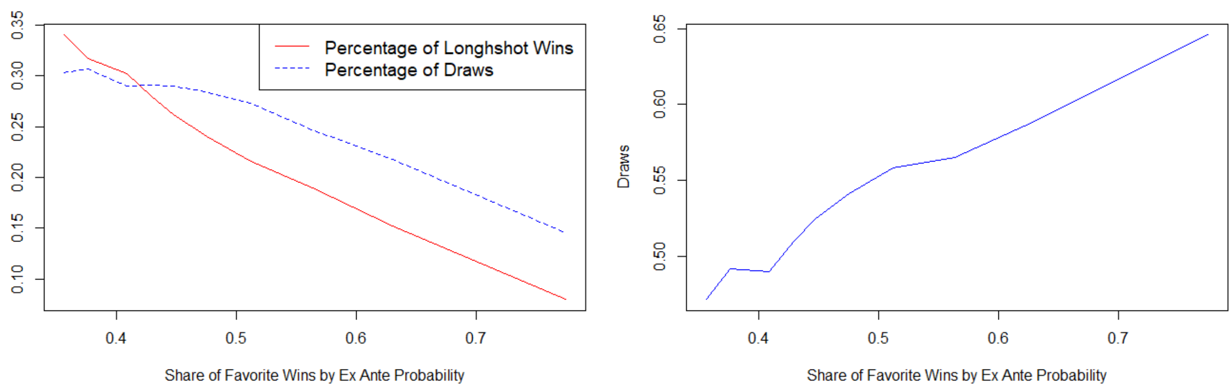
The different levels of disagreement for draws and longshot bets can be explained as follows. Suppose a game has equal probabilities for the three possible outcomes. There will be some people who think the home team has a higher chance of winning and some who think the away team has a higher chance of winning. These bettors will disagree about the probabilities of the two teams winning but they will agree on how likely the draw is: Both will think that because their preferred team is a favorite, the draw is somewhat less likely than its true one-third probability. As the probability of the favorite winning increases to a high level, however, there is not much disagreement about whether the longshot can win (almost nobody thinks they can) so disagreement focuses on the probability of a draw and loss rates for those bets rise.

This prediction about loss rates from betting on draws is an interesting one because there are many models that predict favorite-longshot bias. For example, there could be differences in elasticity of demand between favorites and longshots stemming from a preference for risk that makes longshot bettors less sensitive to odds. However, our model's prediction about loss rates on draws is not related to their riskiness and arises specifically from disagreement among bettors.

Another interesting prediction here is that while loss rates for bets on longshots should generally be greater than loss rates for favorites, loss rates on teams that are only very marginal longshots are lower than for some of the favorite bets with higher probabilities of winning. Again, this is because the amount of disagreement about the probability of these marginally less fancied teams is lower than for favorites.

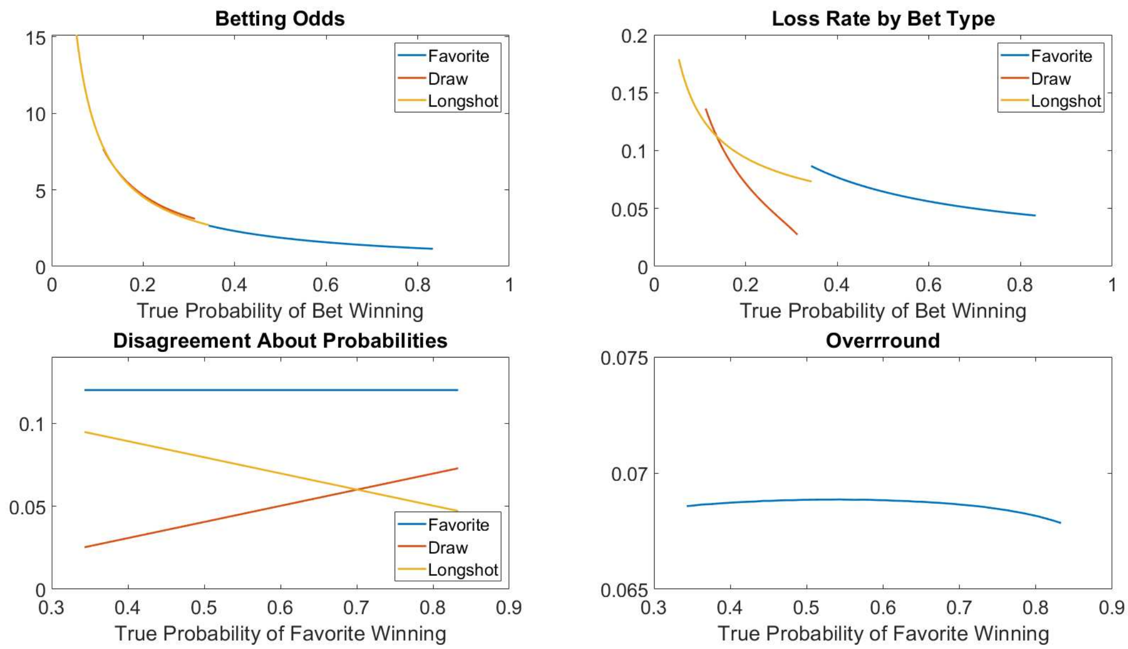
The bottom right panel shows the overrounds implied by the odds generated by the model. These are relatively flat across all probabilities of a favorite win and, as noted, σ has been set to match the average overround seen in our data. Finally, Figure 5 does not show the betting volumes predicted by the model but it is worth noting that the model predicts bets on draws will be the least popular option across almost the full range of favorite-win probabilities seen in our data. This appears consistent with anecdotal evidence on sports betting patterns and from volumes of bets placed on draws reported on betting exchanges such as BetFair.

Figure 4: Draws rise relative to longshot wins as the probability of the favorite winning rises



Notes: Estimates based on the deciles of our dataset of European soccer games, organised by the ex ante probability of the favorite winning, as measured by the overround method.

Figure 5: Overround and ratio of monopoly odds to competitive odds in a home/away/draw monopolistic market ($\sigma = 0.06$, $\mu = 0.02$, $\theta = 0$)



5. Application: European Soccer Betting

We will apply our model to understanding betting odds on European professional soccer matches. We start with a discussion of the structure of the market for betting on soccer and then introduce our data.

5.1. Market Structure in Online Sports Betting

The internet has transformed sport betting and turned it into a huge global business. As betting moved to become predominately online, two different business models emerged.[7] The traditional retail bookmakers in Europe have adopted what is known as the "soft" bookmaker model. This model focuses on maintaining high gross profit margins and spending on marketing to attract bettors to 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.[8] Importantly for our purposes, these bookmakers dominate the long-established market for betting on home wins, away wins or draws.

In contrast, "sharp" bookmakers have essentially the opposite model. They have low profit margins per bet with a focus on making larger profits 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. For our purposes, it is useful to note that most of the betting on European soccer taken by these bookmakers in the form of so-called Asian Handicap bets, which we will describe below.

Given the low margins of sharp bookmakers, why do people still place bets with the soft bookmakers? The explanation is that the sharp bookmakers do not have a significant retail presence in Europe. They do not engage in traditional promotional advertising and are not licensed in most European countries so that people can only place bets with them indirectly via brokers. In addition, sharp bookmakers such as Pinnacle also earn fees as "market makers", selling their estimated probabilities for events to retail European bookmakers, allowing them to profit from retail gambling without having to participate directly in this market. These frictions mean that most regular gamblers have either not heard of the sharp bookmakers or consider the hoops to be jumped through to bet with them to be not worth the bother.

Our theoretical sections presented the two extreme cases of perfect competition and monopoly. We believe the low-margin sharp bookmakers that dominate the Asian Handicap betting market come close to the competitive ideal. In contrast, the market for home/away/draw bets that is dominated by the soft bookmakers comes closer to having a monopolistic structure. This claim may seem

[7] An excellent introduction to these two models is David Hill's 2019 article "Requiem for a Sports Bettor" in The Ringer. He provides an informative and entertaining account of how the arrival of European bookmakers was changing the face of US sports betting only one year after the repeal of its federal prohibition.

[8] The soft bookmaker practices of customer profiling and stake restrictions are documented by Davies (2022a, 2022b).

surprising because there is an appearance of competition for business among these firms. For example, you can find odds comparison sites that will tell you the best odds on a game and there are often small differences in the odds offered. However, those who take up the best odds offers generally find the amount they can bet at those odds is small. Indeed, if these bookmakers profile you as being someone who only places bets at the best available odds, you may be banned altogether. This means the reality of this market may be closer to the monopoly case modeled here but rather than one firm, there are a number of firms setting similar prices with their market share determined by non-odds-related competition in areas such as marketing or the provision of other services like streaming sports events on betting websites.

5.2. 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 game outcomes and odds for bets on a home win, an away win and a draw for 84,230 matches spanning the 2011/12 to 2021/22 seasons for 22 European soccer leagues across 11 different nations as described in Table 1. The betting odds are the average closing odds across the various online bookmakers surveyed by Buchdahl.

Table 1: The 22 soccer leagues 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

6. Empirical Results

Here, we present our empirical findings for loss rates from bets on home wins, away wins and draws and then describe our results for Asian Handicap bets on the same sample of matches.

6.1. Loss rates for home/away/draw bets

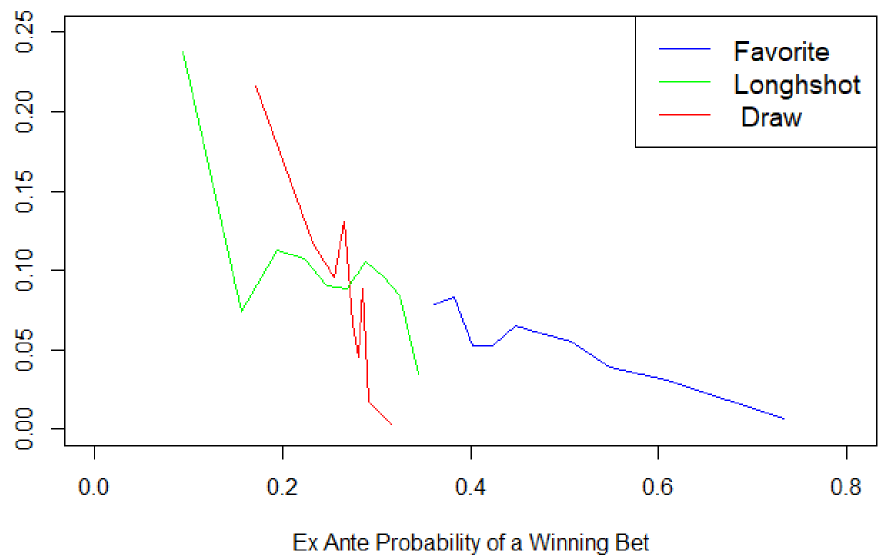
Figure 6 shows the results for loss rates for bets on favorites, on longshots and on draws in our datasets, with each type of bet sorted into ten deciles by their implied probability of success, as calculated using the usual "overround" method. The most obvious pattern to note is that this betting market exhibits a significant favourite-longshot bias pattern, with bets on favourites losing less than bets on longshots. This confirms the existing findings of Buhagiar, Cortis and Newall (2018) and Angelini and de Angelis (2019) using similar but smaller datasets.

The specific pattern of the favourite-longshot bias also fits well with the general pattern predicted by the model as illustrated in Figure 5. Losses rise in a sharp nonlinear pattern as the estimated probability of bet success falls. The magnitudes do not fully match our calibrated version of the model—actual losses are a little bit higher than predicted when the probability of success is relatively low and a little bit lower when the probability of success is relatively high—but the size of loss rates generally fit well with the model's predictions.

More importantly, however, the results also conform well with the model's specific predictions about loss rates from betting on draws and longshots. The data conform to the model's prediction that bets on draws with a comparatively high probability of occurring will have some of the lowest loss rates but that bets on draws will have higher loss rates than bets on longshots when draws are relatively unlikely to occur. The data also match the model's predictions that longshot bets with the highest chance of winning have slightly lower loss rates than bets on the weakest favorites.

We should note here that we have set the parameters σ and μ to match the average overround rather than to match any specific feature of the odds on favorites, longshots or draws—the close relationship between the loss rates in the theoretical Figure 5 and the empirical Figure 6 is not the result of any attempt to match how loss rates on different types of bets vary by ex ante probability of success. However, a simulated method of moments estimator applied to estimating σ and μ based on having the model match the odds for each bet in the sample as best as possible, given the estimated value of the probability of the favorite winning, produced similar estimates of our calibrated values and predicted essentially the same pattern for loss rates.

Figure 6: Average loss rates for bets in the home/away/draw market organized by type of bet and ex ante probability of bet win



6.2. Asian Handicap bets

Our dataset also contains Asian Handicap odds for the same set of 84,230 matches. As noted above, payouts on these bets depend on an adjustment of the match result that applies a deduction to the goals total of the team considered more likely to win.. The Asian Handicap market features 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 "hybrids" in which money is split between bets at other handicaps. To see how this kind of betting works, consider first the case in which the Asian handicap is 1.5 and O_S is the decimal odds for the strong team and O_W is the decimal odds on the weak team. In this case, 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 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.

Alternatively, 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 a 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. while bets with a handicap is 0.75 put half the money on a bet with a handicap of 1 and the other half on a bet with a handicap of 0.5. Both of these hybrid bets feature the possibility of half the money being refunded, so the only bets that do not feature the possibility of at least a partial refund are those with half-goal handicaps

To compare the behavior of Asian Handicap odds with those from the home/away/draw market, we need to calculate probabilities of the different outcomes. For handicaps ending in .5, where there are only two possible outcomes, one can use the standard "overround" method described in Section 2. However, when refunds are possible, calculating probabilities is more complex. Consider the case where the Asian handicap is 1. In this case, we want to calculate the probabilities for the three different outcomes

$$P_{S2} = \text{Probability the stronger team wins by 2 or more} \tag{25}$$

$$P_{S1} = \text{Probability the stronger team wins by 1} \tag{26}$$

$$P_W = \text{Probability of a draw or the weaker team winning} \tag{27}$$

Again assuming the expected payoff for all \$1 bets is μ, then we have the following conditions hold.

$$P_{S2}O_S + P_{S1} = \mu \tag{28}$$

$$P_W O_W + P_{S1} = \mu \tag{29}$$

$$P_W + P_{S1} + P_{S2} = 1 \tag{30}$$

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

Hegarty and Whelan (2023) approach this problem by setting P_{S1} equal to the sample average of the fraction of matches than end in a refund for each kind of handicap. They show that for each of the three types of Asian handicap bets with a refund element, the fraction of bets that end in refunds is stable over time and does not depend on observable match-specific factors such as the betting odds quoted.[9] Using this approach, for example for the case with handicap of 1, conditional on a specific

[9] See an appendix for details.

value of the probability of a refund, P_{S1}, the other unknowns can be solved to give

$$P_{S2} = \frac{(1 - P_{S1})\, O_W}{O_S + O_W} \tag{31}$$

$$P_W = \frac{(1 - P_{S1})\, O_S}{O_S + O_W} \tag{32}$$

$$\mu = P_{S_1} + \frac{(1 - P_{S1})\, O_S O_W}{O_S + O_W} \tag{33}$$

The same method can also be used to calculate probabilities of full bet wins for the "hybrid" bets that split odds between integer and half-goal handicaps.

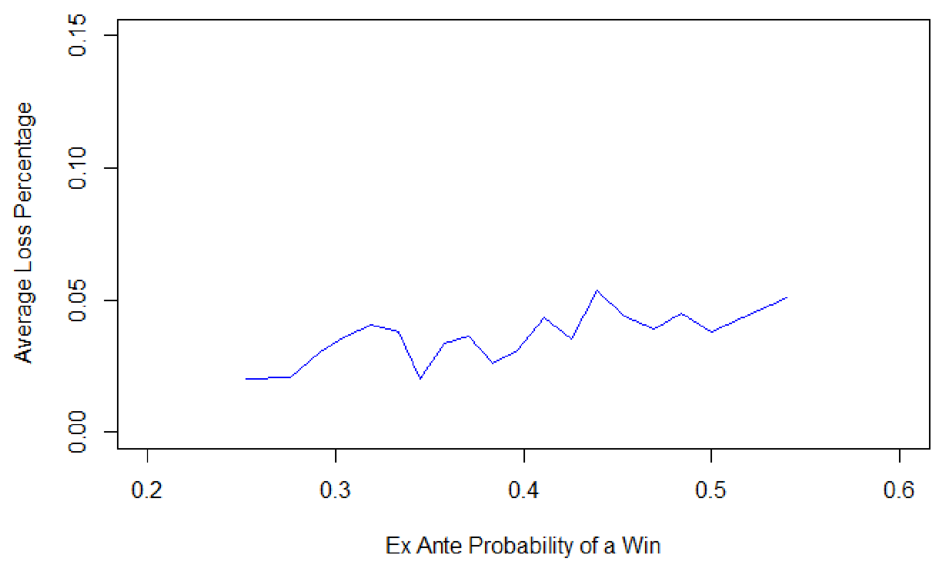
It is worth emphasizing that, despite some obvious similarities, the Asian Handicap market differs from spread betting markets on US sports along a couple of dimensions that are important for the question we are examining. Spread bets offered on high scoring sports are generally set to equate the odds of each side of the bet winning. This means the implied probabilities of success of each bet are equal and the odds offered on each bet are typically the same. This is not the case with the Asian Handicap market and the implied probabilities of full bet success that we calculate range from about 0.25 to about 0.55.

There are for a few reasons for this wide spread in implied probabilities. First, 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. Second, for US spread bets, bookmakers normally react to new information by adjusting the handicap while leaving the odds fixed while Asian Handicap bookmakers normally adjust the odds and leave the handicap fixed. Third, the hybrid quarter-point handicap bets have the feature that one of side the bets earns a profit in two of the three possible outcomes while the other side 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. This compensation occurs via bets that only make a profit in one outcome tending to have a higher probability of a full payout and this contributes to the opposing sides of these bets having different chances of success.

Figure 7 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 home/away/draw bets. 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 different between the pattern for loss rates in the two markets. For home/away/draw bets, those 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%.

These results thus confirm our prediction that a competitive market, in which bookmakers compete with each other on odds and do not engage in restrictive practices, will not feature favorite-longshot bias.

Figure 7: Average loss rates for bets in the Asian handicap market organized by ex ante probability of bet win



7. Discussion

Here we provide a further discussion of a few aspects of our model and findings.

7.1. Risk aversion?

We have described a model with risk-neutral bookmakers and bettors. There are several reasons why an assumption of risk-neutrality for bettors makes sense.

First, there is the scale of risk involved. Most gamblers place small bets that do not have life-altering consequences. Rabin (2000) pointed out that people should be approximately risk-neutral when the stakes are small. If utility functions are already concave in the region of the changes in wealth generated by winning or losing small bets, then they should be even more risk-averse when it comes to making decisions with big financial consequences. Rabin shows that people turning down positive expected value risky small bets would imply counter-factual predictions for the pricing of insurance of various types and for decisions such as whether to invest in stocks.

Second, the empirical studies used to estimate risk aversion consider wider variations in income or assets than would be generated by the betting decisions made by most gamblers. For example, Layard, Mayraz and Nickell (2008) and Gandelman and Hernandez-Murillo (2013) use self-reported happiness to assess the curvature of utility functions across a wide range of incomes. It is questionable whether these estimates are relevant for assessing the risk of a gambler placing a $20 bet.

In contrast, bookmakers are taking a large volume of bets and can be at risk of losing a lot of money, so Rabin's arguments suggest they could be more plausibly considered risk averse. In reality, it appears that the soft and sharp bookmakers take different attitudes to risk, with soft bookmakers appearing to be risk averse, adopting strategies to limit various types of exposures, while sharp bookmakers appear to act in an approximately risk-neutral manner, setting accurate probabilities and not worrying about the variance of profits on individual games.

These considerations re-enforce our results. In an appendix, we describe the home/away/draw model when the monopolist is risk averse. Profits earned from longshot bets have a higher variance than profits earned on favorite bets, so risk aversion provides another reason for monopolists to set odds with a favorite-longshot bias. For this reason, this version of the model predicts somewhat larger loss rates, particularly for longshot bets. However, once the disagreement parameter, σ, has been set to again match the average value of the overround in the data, the predictions of version of the model with risk aversion are almost identical to the risk-neutral version, though the predicted loss rates for longshot bets and draws are a bit higher and thus match the data somewhat better.

7.2. Comparison with Moscowitz and Vasudevan (2022)

Our study shares some similarities with the work of Moscowitz and Vasudevan (2022). Like us, they compare the odds for the same set of games for two different types of bets, one based on an adjusted outcome and the other based on the actual outcome. Specifically, they compare spread betting odds with odds for money line bets (bets on whether a team will win or lose) for a sample of US basketball and American football games. They compare the odds for these two types of bets cross the same set of games and, like us, they find a pattern of favorite-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 higher risk money line bets on longshots and still find takers. However, our disagreement-based model can also explain their findings. The spread bets in their sample offer equal odds on both teams, so in relation the bet itself, neither team is a favorite. In these circumstances, our model predicts the average return for bets on each team will be the same. One team being the favorite to win the actual game doesn't have any influence on the odds of the handicap-adjusted bet. Also, market structure still needs to be considered when evaluating Moscowitz and Vasudevan's evidence. They show that US bookmakers 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 for these bets.

There is also an important contrast between our results and those of Moscowitz and Vasudevan because in the score-adjusted market that we examine, there is still a wide range of estimated probabilities of bet success and thus a wide range of risk but we don't find evidence of lower returns for higher risk bets in this market.

7.3. On beliefs and betting

An obvious question to ask about the gamblers in this model is why they don't use the market price to infer the correct odds. While the monopolistic market differs from the competitive case where you can calculate the probabilities from the odds, there is still a monotonic relationships between the probability of a bet winning and the odds quoted by the bookmaker. Shouldn't those who have a strong belief that one team will win change their mind when the bookmaker's odds show the team is unlikely to?

This question is perhaps best answered by the very existence of betting markets. Bookmakers make profits. Using the market odds as a guide to estimate the true probabilities will generally mean accepting that all bets have a negative expected return. And most people do not bet, most likely for that reason. But some people do bet. This may partly be fueled by the utility they get from placing bets, perhaps in a way that increases their enjoyment of watching sporting events. But even then, this still leaves the question of why they choose to bet on one team rather than another. When one person bets on the home team and another person bets on the away team, the explanation for their differing decisions is likely that they disagree with each other about the probability of their picks winning.

8. Conclusions

Given the increasing size and importance of the online sports betting industry, it is important to understand how it sets prices. We have provided theoretical and empirical evidence that the formation of fixed betting odds varies depending on the structure of the betting market. We have shown that monopolistic betting markets will set odds featuring favorite-longshot bias as long as bettors disagree about the probabilities of various outcomes. Our model does not rely on many of the features previously invoked to explain this bias such as the presence of insiders who know the outcome of the event or bettors who have a systematic inability to estimate small probabilities. Instead, it relies on basic microeconomic reasoning: If people disagree about probabilities but the public is on average correct, the elasticity of demand will be greater for bets on favorites than for bets on longshots. We applied our model to soccer and described some specific predictions that emerge from our disagreement-based models but not from other models of the favorite-longshot bias, most notably that odds on draws would behave differently from other odds. A large dataset of odds and outcomes from European soccer leagues supports the model's predictions.

Our findings have a number of policy implications. A 2021 review of the evidence by Public Health England estimated that one in two hundred of the British adult population had a problem with gambling and about one in twenty five were gambling at levels placing them "at risk". Academic research such as Muggleton et al (2021) has found that gambling raises the risk of financial distress and lowers health and well-being outcomes. The risk of financial distress could be reduced if steps were taken to ensure betting markets were more competitive and to minimize the profiling of customers and restrictive practices that currently characterize European retail betting.

These results are also relevant for the emerging US sports betting market. While many states have welcomed the opportunity that legal sports betting gives them for earning additional tax revenues, the legalization of sports betting and its easy availability via cell phones is also likely to result in a significant increase in problem gambling. The US market has attracted heavy involvement European bookmakers, either via direct entry such as Paddy Power owner Flutter's acquisition of FanDuel or via selling services to newly-licensed US bookmakers to allow them to copy the European business model. Bloomberg have reported that the two leading firms (FanDuel and DraftKings) have 65% of the current market, so price setting seems unlikely to meet the ideal of a perfectly competitive market.[10] Our model suggests this market is likely to become characterized by high profit rates and large loss rates for bettors, particularly those taking on long-odds bets.

[10] https://www.bloomberg.com/news/articles/2021-09-21/how-fanduel-gained-more-fantasy-sports-gamblers-than-draftkings-dkng

References

[1] Ali, Mukhtar (1977). "Probability and Utility Estimates for Racetrack Bettors," *Journal of Political Economy*, Volume 85, pages 803-815.

[2] Angelini, Giovanni and Luca De Angelis (2019). "Efficiency of Online Football Betting Markets," *International Journal of Forecasting*, Volume 35, pages 712–721.

[3] Berkowitz, Jason, Craig Depken and John Gandar (2017). "A Favorite-Longshot Bias in Fixed-Odds Betting Markets: Evidence from College Basketball and College Football," *Quarterly Review of Economics and Finance*. Volume 63, pages 233-239

[4] Buhagiar, Ranier, Dominic Cortis, and Philip Newall (2015). "Why Do Some Soccer Bettors Lose More Money Than Others?" *Journal of Behavioural and Experimental Finance*, Volume 18, pages 85-93.

[5] Davies, Rob (2022a). *Jackpot*, Faber.

[6] Davies, Rob (2022b). Revealed: how bookies clamp down on successful gamblers and exploit the rest, The Guardian. www.theguardian.com/society/2022/feb/19/stake-factoring-how-bookies-clamp-down-on-successful-gamblers

[7] Gabriel, Paul and James Marsden (1990). "An Examination of Market Efficiency in British Racetrack Betting," *Journal of Political Economy*, Volume 98, pages 874–885.

[8] Gandhi, Amit and Ricardo Serrano-Padial (2015). "Does Belief Heterogeneity Explain Asset Prices: The Case of the Longshot Bias," *Review of Economic Studies*, Volume 82, pages 156-186.

[9] Forrest, David and Ian McHale (2007). "Anyone For Tennis (Betting)?" *European Journal of Finance*, Volume 13, pages 751-768.

[10] Hegarty, Tadgh and Karl Whelan (2023). Forecasting Soccer Matches With Betting Odds: A Tale of Two Markets. CEPR discussion paper No. 17949.

[11] Hill, David (2019). "Requiem for a Sports Bettor," *The Ringer*. www.theringer.com/2019/6/5/18644504/sports-betting-bettors-sharps-kicked-out-spanky-william-hill-new-jersey

[12] Hurley, William and Lawrence McDonough (1995). "A Note on the Hayek Hypothesis and the Favorite-Longshot Bias in Parimutuel Betting," *American Economic Review*, Volume 85, pages 949-955.

[13] Kahneman, Daniel and Amos Tversky (1979). "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, Volume 47, pages 263-291.

[14] Layard, Richard, Guy Mayraz and Steve Nickell (2008). "The Marginal Utility of Income," *Journal of Public Economics*, Volume 92, pages 1846-1857.

[15] Montone, Maurizio (2021). "Optimal Pricing in the Online Betting Market," *Journal of Economic Behavior and Organization*, Volume 186, pages 344-363.

[16] Moskowitz, Tobias and Kaushik Vasudevan (2022). Betting Without Beta, working paper, Yale University.

[17] Muggleton, Naomi, Paula Parpart, Philip Newall, David Leake, John Gathergood and Neil Stewart (2021). "The Association between Gambling and Financial, Social, and Health outcomes in Big Financial Data," *Nature Human Behaviour*, Volume 5, pages 319–326

[18] Ottaviani, Marco and Peter Norman Sørensen (2008). "The Favorite-Longshot Bias: An Overview of the Main Explanations," in *Handbook of Sports and Lottery Markets*, Elsevier.

[19] Pope, Peter and David Peel (1989). "Information, Prices and Efficiency in a Fixed-Odds Betting Market," *Economica*, Volume 56, pages 223-241.

[20] Quandt, Richard (1986). "Betting and Equilibrium," *Quarterly Journal of Economics*, Volume 101, pages 201-208.

[21] Rabin, Matthew (2000). "Risk-Aversion and Expected Utility Theory: A Calibration Theorem," *Econometrica*, Volume 68, pages 1281-1292.

[22] Sauer, Raymond (1998). "The Economics of Wagering Markets," *Journal of Economics Literature*, Volume 36, pages 2021-2064.

[23] Shin, Hyun Song (1991). "Optimal Betting Odds Against Insider Traders," *Economic Journal*, Volume 101, pages 1179-1185.

[24] Thaler, Richard and William Ziemba (1988). "Anomalies: Parimutuel Betting Markets: Racetracks and Lotteries," *Journal of Economic Perspectives*, Volume 2, pages 161-174.

A Belief thresholds for betting on soccer

Here we derive the decisions rules for betting on draws and longshots in the home/away/draw market. Gamblers will bet on the draw if

$$\tilde{p}_D O_D \geq 1 \Longrightarrow (\delta_0 + \delta_1 \tilde{p}_F)(1 - \tilde{p}_F) \geq \frac{1}{O_D} \tag{34}$$

Multiplying out and dividing by δ_1, we get

$$p_F^2 - \frac{\delta_1 - \delta_0}{\delta_1} p_F - \frac{\delta_0}{\delta_1} + \frac{1}{\delta_1 O_D} \leq 0 \tag{35}$$

The roots of the quadratic function are

$$[\tilde{p}_{F1}^*, \tilde{p}_{F2}^*] = 0.5 \left[\frac{\delta_1 - \delta_0}{\delta_1} \pm \sqrt{\left(\frac{\delta_1 - \delta_0}{\delta_1} \right)^2 + 4 \frac{\delta_0}{\delta_1} - \frac{4}{\delta_1 O_D}} \right] \tag{36}$$

Assuming, there are two real roots, the condition becomes

$$(\tilde{p}_F - \tilde{p}_{F1}^*)(\tilde{p}_F - \tilde{p}_{F2}^*) \leq 0 \tag{37}$$

One of the terms on the left must be positive and one must be negative.

Using our calibrated values of $\delta_0 = 0.33$ and $\delta_1 = 0.40$, the roots are

$$[\tilde{p}_{F1}^*, \tilde{p}_{F2}^*] = 0.5 \left[0.175 \pm \sqrt{3.33 - \frac{10}{O_D}} \right] \tag{38}$$

Real solutions only exist if $O_D > 3$, so nobody will bet on a draw if the odds are lower then 3. For a concrete example, consider $O_D = 5$. In this case, the roots of the quadratic are $[-0.489, 0.664]$ so the decision rule is to bet on the draw if $\tilde{p}_F < 0.664$. To check that this is correct, note that at this probability, the chance of a draw is

$$(0.33 + 0.4\,(0.664))\,(1 - 0.664) = (0.5956)\,(0.336) = 0.2 \tag{39}$$

So the $O_D = 5$ bet is just worth taking.

Similarly, gamblers will bet on the longshot winning if

$$p_L O_L \geq 1 \Longrightarrow (1 - \tilde{p}_F)(1 - \delta_0 - \delta_1 \tilde{p}_F) \geq \frac{1}{O_L} \tag{40}$$

This becomes

$$\tilde{p}_F^2 + \frac{(\delta_0 - \delta_1 - 1)}{\delta_1} \tilde{p}_F + \frac{1}{\delta_1} - \frac{\delta_0}{\delta_1} - \frac{1}{\delta_1 O_L} \geq 0 \tag{41}$$

The roots of the quadratic function are

$$[\tilde{p}^*_{F1}, \tilde{p}^*_{F2}] = 0.5 \left[\frac{1-\delta_0+\delta_1}{\delta_1} \pm \sqrt{\left(\frac{1-\delta_0+\delta_1}{\delta_1}\right)^2 - \frac{4}{\delta_1} + \frac{4\delta_0}{\delta_1} + \frac{4}{\delta_1 O_L}} \right] \tag{42}$$

With $\delta_0 = 0.33$ and $\delta_1 = 0.40$, this becomes

$$[\tilde{p}^*_{F1}, \tilde{p}^*_{F2}] = 0.5 \left[2.675 \pm \sqrt{0.455625 + \frac{10}{O_L}} \right] \tag{43}$$

The term in the square root is positive so there will always be two real roots. The condition for taking the bet is

$$(\tilde{p}_F - \tilde{p}^*_{F1})\,(\tilde{p}_F - \tilde{p}^*_{F2}) > 0 \tag{44}$$

So the two terms must have the same sign. To give a concrete example, consider $O_L = 4$. This implies roots of $[0.4779, 2.201]$. The term $(\tilde{p}_F - 2.21)$ will be negative so the other term needs to be negative also, meaning the condition for taking the bet reduces to $\tilde{p}_F < 0.4779$. To illustrate that this is the correct figure, at this probability we have

$$\begin{aligned} p_L &= (1-\tilde{p}_F)\,(1-\delta_0-\delta_1\tilde{p}_F) \\ &= (1-.4779)\,(1-0.33-0.4\,(0.4779)) \\ &= 0.25 \end{aligned} \tag{45}$$

so again the bet on the longshot winning is just worth taking.

B Evidence on predictability of refunds

Here we provide evidence to support 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 expected losses 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

$$R_{ijkq} = \alpha_1 + \sum_{j=2}^{22} \alpha_j L_j + \sum_{n=2}^{11} \beta_n S_n + \sum_{q=2}^{3} \delta_q H_q + \eta_1 O_{iH} + \eta_2 O_{iA} + u_{ijkq} \tag{46}$$

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 2 reports the results from estimation of this regression via Weighted Least Squares 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.[11] 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. One 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.

To summarize, 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.

[11] Similar results are obtained from Probit estimation.

Table 2: Weighted least squares regression for 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)
N	63,468	
R^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.
The weights used are $\frac{1}{(PS_1)(1-PS1)}$ where PS_1 is the sample average fraction of refunds for each type of handicap.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

C Incorporating risk aversion for home/away/draw bookmakers

We also considered optimal pricing for bookmakers who instead of maximizing expected profit, instead maximize the mean-variance utility function

$$U(\pi) = E(\pi) - \frac{\gamma}{2} Var(\pi) \tag{47}$$

This version of the model does not generate analytically neat formulas but can be solved numerically. As noted above, profits earned from longshot bets have a higher variance than profits earned on favorite bets, so discouraging these bets provides another reason for risk-averse monopolists to set odds with a favourite-longshot bias. For the same set of parameters used in our baseline calibration, the model with a risk-averse monopolist generates higher average odds. The overround is about 9% rather than 7% because some bets that a risk-neutral bookmaker takes are less attractive to a risk-averse bookmakers. However, re-setting the disagreement parameter to $\sigma = 0.04$, the risk-averse model again matches the over-round and the overall pattern of loss rates roughly matches what we saw in Figure 5.

Figure 8: Loss rates for home/away/draw bets with a risk-averse monopolists ($\sigma = 0.04$, $\mu = 0.02$, $\gamma = 1$)