WHY ARE GAMBLING MARKETS ORGANISED SO DIFFERENTLY FROM FINANCIAL MARKETS?*

Steven D. Levitt

The market for sports gambling is structured very differently from the typical financial market. In sports betting, bookmakers announce a price, after which adjustments are small and infrequent. Bookmakers do not play the traditional role of market makers matching buyers and sellers but, rather, take large positions with respect to the outcome of game. Using a unique data set, I demonstrate that this peculiar price-setting mechanism allows bookmakers to achieve substantially higher profits. Bookmakers are more skilled at predicting the outcomes of games than bettors and systematically exploit bettor biases by choosing prices that deviate from the market clearing price.

There are many parallels between trading in financial markets and sports wagering. First, in both settings, investors with heterogeneous beliefs and information seek to profit through trading as uncertainty is resolved over time. Second, sports betting, like trading in financial derivatives, is a zero-sum game with one trader on each side of the transaction. Finally, large amounts of money are potentially at stake. The four major British bookmaking firms report turnover of almost £10 billion in 2002; estimates of wagering on sporting events in the US go as high as $380 billion annually (National Gambling Impact Study Commission, 1999).

In light of these similarities, it is surprising that these two types of markets are organised so differently. In most financial markets, prices change frequently. The prevailing price is that which equilibrates supply and demand. The primary role of market makers is to match buyers with sellers. With sports wagering and also horse racing in the UK, market makers (i.e. bookmakers) simply announce a 'price' (which can be odds to win a horse race, or for sporting events can be odds to win a game or a point spread, e.g., the home team to win an American football game by at least 3.5 points), after which adjustments are typically small and relatively infrequent.[1] If that price is not the market clearing price, then the

[1] In my data on American football, in the five days preceding a game, the posted price changes an average of 1.4 times per game. When the price does change, in 85% of the cases the line moves by the minimum increment of one-half of a point. Thus, the posted spread on Tuesday is within one point of the posted spread at kickoff on Sunday in 90% of all games. These calculations are based on information on changes in the casino lines reported at http://www.wagerline.com. In horse racing, the odds set by bookmakers change more frequently.

bookmakers may be exposed to substantial risk.[2] If bettors are able to recognise and exploit mispricing on the part of the bookmaker, the bookmaker can sustain large losses. The risk borne by bookmakers on sports betting is categorically different from the casino's risk on other games of chance such as roulette, keno, or slot machines. In those games of chance, the odds are stacked in favour of the casino and the law of large numbers dictates profits for the house. In contrast, however, if the bookmaker sets the wrong line on sporting events, it can lose money, even in the long run. The presence of a small number of bettors whose skills allow them to achieve positive expected profits could prove financially disastrous to the bookmaker. Such bettors could either amass large bankrolls or, in the presence of credit constraints, sell their information to others.

Although the mechanism used for price-setting in sports betting seems peculiar, there are at least three scenarios in which the bookmaker can sustain profits implementing it. In the first scenario, bookmakers are extremely good at determining in advance the price which equalises the quantity of money wagered on each side of the bet. If this occurs, the bookmaker makes money regardless of who wins the game since the bookmaker charges a commission (known as the 'vig') on bets.[3] Following this strategy, bookmakers do not have to have any particular skill in picking the actual outcome of sporting events, they simply need to be good at forecasting how bettors behave. Popular depictions of bookmaker behaviour have stressed this explanation.[4]

An alternative scenario under which this price-setting mechanism could persist is if bookmakers are systematically better than gamblers in predicting the outcomes of games. If that were the case, the bookmaker could set the 'correct' price (i.e. the one which equalises the probability that a bet placed on either side of a wager is a winner). Although the money bet on any individual game would not be equalised, on average the bookmaker will earn the amount of the commission

---

[2] There are many notorious examples of bookmakers suffering large losses. It is reported that over half of all British bookmakers were bankrupted when Airborne won the Epsom Derby in 1946 (the first running after World War II ended) at odds of 50 to 1 (Smith, 2002). In 1996, popular jockey Frankie Dettori won seven straight races, costing bookmakers an estimated £30 million (*Gambling Magazine*, 1999). Coral Eurobet reported losses of £12 million on internet betting in quarter-final round of Euro 2000 soccer championship (Smith, 2002).

[3] Typically, bettors must pay the casino 110 units if a bet loses, but are paid only 100 units if the bet wins.

[4] For instance, a website devoted to educating novice gamblers (http://www.nfl-betting.org) writes, 'A sports bettor needs to realize that the point spread on a game is NOT a prediction by an odds maker on the outcome of a game. Rather, the odds are designed so that equal money is bet on both sides of the game. If more money is bet on one of the teams, the sports book runs the risk of losing money if that team were to win. Bookmakers are not gamblers – they want to make money on every bet regardless of the outcome of the game.' Similarly, Lee and Smith (forthcoming) write, 'Bookies do not want their profits to depend on the outcome of the game. Their objective is to set the point spread to equalize the number of dollars wagered on each team and to set the total line to equalize the number of dollars wagered over and under. If they achieve this objective, then the losers pay the winners $10 and pay the bookmaker $1, no matter how the game turns out. This $1 profit (the "vigorish") presumably compensates bookmakers for making a market and for the risk they bear that the point spread or total line may be set incorrectly.'

charged to the bettors. Unlike the first scenario above, however, if prices are set in this manner, the bookmaker will lose if gamblers are actually more skilled in determining the outcome of games than is the bookmaker. The third possible scenario combines elements of the two situations described in the preceding paragraphs. If bookmakers are not only better at predicting game outcomes but also proficient at predicting bettors preferences, they can do even better in expectation than to simply collect the commission. By systematically setting the 'wrong' prices in a manner that takes advantage of bettor preferences, bookmakers can increase profits. For instance, if bookmakers know that local bettors prefer local teams, they can skew the odds against the local team. There are constraints on the magnitude of this distortion, however, since bettors who know the 'correct' price can generate positive returns if the posted price deviates too much from the true odds.[5]

In this paper, I attempt to understand the structure of the market for sports gambling better by exploiting a data set of approximately 20,000 wagers on the National Football League, the premier league of American football placed by 285 bettors at an online sports book as part of a high-stakes handicapping contest. Two aspects of this data set are unique. First, in contrast to previous studies of betting that only had information on prices, I observe both prices *and* quantities of bets placed. That information allows me to determine whether the bookmaker appears to be equalising the amount of bets on each side of a wager. Second, I am able to track the behaviour of individual bettors over time, which provides a means of determining whether some bettors are more skillful than others. Although my data are from one bookmaker, the patterns observed here are likely to generalise since all bookmakers offer nearly identical spreads on a given game.

A number of results emerge from the analysis. First, I demonstrate that the bookmaker does not appear to be trying to set prices to equalise the amount of money bet on either side of a wager. In almost one-half of all games, at least two-thirds of the bets fall on one side of the gamble. Moreover, the spread chosen systematically fails to incorporate readily available information (e.g. which team is the home team) that would help in equalising the money bet on either side of a wager.[6] For instance, in games where the home team is an underdog, *on average* two-thirds of the wagers are on the visiting team. These findings argue strongly against the first scenario presented above and the popular depictions of book-maker behaviour. A rationale for this failure to equalise the money emerges in the

---

[5] This assumes that bookmakers are unable to offer different prices to different bettors. Indeed, there is evidence that local bookmakers who deal repeatedly with the same clients are able to exercise some degree of price discrimination. See Strumpf (2002) for empirical evidence that bookmakers both shade the odds against the home team and offer different odds to bettors with different past betting histories.

[6] Of course, it is not the number of bets on either side of a wager that the bookmaker wants to equalise, but rather, the total dollars bet on either side. In my data, however, all wagers are constrained to have the same dollar value, so the two are equivalent. If there are large bankroll bettors outside my sample who systematically bet against the prevailing sentiment of other bettors, conclusions based on my sub-sample may be erroneous.

© Royal Economic Society 2004

paper's second finding: bookmakers appear to be strategically setting prices in order to exploit bettors' biases, just as DellaVigna and Malmendier (2003) demonstrate health clubs do with their clients. Bettors exhibit a systematic bias toward favourites and, to a lesser extent, towards visiting teams.[7] Consequently, the bookmakers are able to set odds such that favourites and home teams win less than 50% of the time, yet attract more than half of the betting action. By choosing these prices, it appears bookmakers increase their gross profit margins by 20–30% over a price-setting policy that attempts to balance the amount of money on either side of the wager. On the dimension of favourites versus underdogs, bookmakers appear to have distorted prices as much as possible without allowing a simple strategy of always betting on underdogs to become profitable. The fact that home teams and underdogs cover the spread a disproportionate share of the time has been well established in the literature (Golec and Tamarkin, 1991; Gray and Gray, 1997). My findings provide an explanation for that empirical regularity: it is profit maximising for the bookmaker who sets the spread. Third, there is little evidence that there are individual bettors who are able to beat the bookmaker systematically.[8] The distribution of outcomes across bettors is consistent with data randomly generated from independent tosses of a 50-50 coin. Moreover, how well a bettor has done up to a certain point in time has no predictive value for future performance. Finally, the evidence is mixed as to whether aggregating bettor preferences has any predictive value in helping to beat the spread. In my sample, there is some weak and ultimately statistically insignificant, evidence that bettors are more likely to predict correctly when there is agreement among them as to which team looks attractive. Altogether, the results are consistent with the conclusion that the bookmakers are at least as good at predicting the outcomes of games than are even the most skilled gamblers in the sample and the bookmakers exploit their advantage by strategically setting prices to achieve profits that are likely higher than would be possible if they simply acted as market makers letting supply and demand equilibrate prices.

The remainder of the paper is organised as follows. Section 1 provides some background on wagering on professional football in the US. Section 2 describes the data set used in the paper. Section 3 presents the empirical findings. Section 4 concludes.

1. Background on American Football Wagering

American football is the most popular sport for wagering in the US, generating 40% of sports-betting revenue for legal bookmakers in Nevada (Nevada Gaming Control Board, 2002). *USA Today* reports that half of all Americans have a wager on

[7] This finding is not to be confused with the bias towards longshots that has been observed in parimutuel horse-race betting (Ali, 1977; Golec and Tamarkin, 1998; Jullien and Salanie, 2000; Shin, 1991, 1992, 1993; Thaler and Ziemba, 1988; Vaughan Williams and Paton, 1997). In football, the odds are set to make the chance of each team covering the spread about 50%, making this consideration irrelevant.

[8] While only tangentially related to the issues addressed in this paper, it is worth noting that there is an extensive academic literature devoted to the question of testing for market efficiency in wagering markets (Asch et al., 1984; Sauer et al., 1988; Woodland and Woodland, 1994; Zuber et al., 1985).

© Royal Economic Society 2004

the outcome of the Super Bowl. The most common type of bet in pro football involves picking the winner of a game against a point spread (a so-called 'straight bet').[9] For instance, if the casino posts a betting line with the home team favoured by 3 points, a bettor can choose either (1) the home team to win by more than that amount, or (2) the visiting team to either lose by less than three points or to win outright. In the event the game ends exactly on the point spread, all bets are refunded. Regardless of which team is chosen, the bettor typically pays the casino 110 units if they lose the bet and collects 100 units when victorious. The difference in the amount paid on a loss versus the amount won for a victory is the casino's commission, known as the 'vig'. Because the bettor can take either side of the wager at the same payout rate, the casino needs to pick a betting line that roughly equalises the probability of the two events occurring (in this case home team winning by more than three or more points, or failing to do so). The bettor receives the spread in force at the time a wager is placed, regardless of later adjustments made by the bookmaker.

Define terms as follows: p is the probability that the favourite wins a particular game, f is the fraction of the total dollars bet on the game that go to the favourite, and v is the vig or commission charged by the bookmaker, which is paid only on losing bets.[10] The bookmaker's expected gross profit per unit bet[11] is given by

$$E(\text{Bookmaker profit}) = [(1 - p)f + p(1 - f)](1 + v) - [(1 - p)(1 - f) + pf]. \quad (1)$$

The terms inside the left set of brackets is the fraction of dollars bet in which the bookmaker wins. That amount is multiplied by $1 + v$ to reflect the bookmakers vig. The terms in the right set of brackets are the cases in which the bookmaker loses and has to payout to the bettor. Rearranging terms, (1) simplifies to

$$E(\text{Bookmaker profit}) = (2 + v)(f + p - 2pf) - 1. \quad (2)$$

If either the probability the two teams win is equal ($p = 0.5$) or the money bet on both teams is equal ($f = 0.5$), the bookmakers gross profit simplifies to $v/2$. In either of these instances, the bookmaker is indifferent about the outcome of the game and earns a profit proportional to the size of the commission charged. As noted in the introduction, therefore, the bookmaker does not need to be able to predict the outcome of the games more accurately than the bettors to ensure a profit. The bookmaker just needs to be able to predict bettor preferences so as to balance out the money on each side of the wager.

---

[9] There are many other types of bets available. For instance, one can bet on whether the total number of points scored in a game is above or below a certain level. One can also bet on which team will win the game (not against the spread), with the payouts appropriately adjusted to reflect the probability of these outcomes. It is also possible to parlay bets on a series of games such that the bettor receives a large payout if correctly picking all the games and receives zero otherwise. Because my data only cover straight bets, I do not focus on these other bet types. Woodland and Woodland (1991) argue that the use of point spreads as opposed to odds that depend only on which teams wins or loses is a bookmaker profit-maximising response to risk aversion on the part of bettors.

[10] I define the model in terms of favourites and underdogs simply because this is the most salient dimension empirically. Game outcomes could be characterised along any relevant set of dimensions.

[11] For simplicity, I treat the number of wagers placed as fixed in the analysis. To the extent that changing the odds affects the total volume of bets, the bookmaker's overall gross profit would be a function of both the number of bets and the gross profit per unit bet.

© Royal Economic Society 2004

Equations (1) and (2) take f and p as given. Of course, the fraction of money bet on the favourite will be a function of the probability the favourite actually wins, i.e. $f = f(p)$, with $\partial f / \partial p > 0$.[12] Taking the derivative of (2) with respect to p, an optimising bookmaker will set p such that

$$[1 - 2f(p)] + (1 - 2p)\partial f / \partial p = 0. \tag{3}$$

The term in brackets is the benefit the bookmaker would achieve from distorting the odds if gamblers did not respond to changes in prices. The remaining term captures the impact on profits of the behavioural response of bettors who switch towards the team with better odds. Note that if bettors' preferences are unbiased in the sense that $f(p = 0.5) = 0.5$, then the bookmaker's optimum is to choose $p = 0.5$, which implies $f = 0.5$ as well.

If, on the other hand, bettors' preferences are biased so that $f(p = 0.5) > 0.5$, as is true empirically in the data set, then the bookmaker can increase profits by reducing p below 0.5 (Kuypers (2000) makes this same point).[13] Intuitively, if bettors prefer favourites at fair odds, the bookmaker can offer odds slightly worse than fair on favourites and still attract more than half of the wagers on the favourite, yielding profits that are strictly higher than is the case at $p = 0.5$. Mathematically, at $p = 0.5$, the term in brackets in (3) is negative but the other term on the left-hand-side is equal to zero, demonstrating that the bookmaker is not at an optimum. The bookmaker will not want to push p too far away from 0.5 for two reasons, however. First, as p diverges from 0.5, it becomes increasingly costly to the bookmaker when a bettor switches from the favourite to the underdog. This is because the bet on the favourite is at (increasingly) unfair odds, whereas the bet on the underdog is at (increasingly) better than fair odds. Note that when p is lowered to the point where $f(p) = 0.5$, gross profits are back to the level attained when $p = 0.5$. Thus, the bookmaker would never want to distort prices to that point. The second reason that the bookmaker cannot distort prices too much is that if some subset of bettors do not have biased preferences, those bettors can exploit the distorted prices. With the standard vig, a bettor must win 52.4% of bets to make profit.[14] One could imagine that the volume of capital available to bettors with positive expected profits could be enormous, both because their bankrolls would grow over time and because the availability of such profits would attract new investors. Thus, it would be surprising to observe price distortions so large that simple strategies (e.g. always bet the underdog) could yield a positive profit.

The discussion above assumes that bookmakers have some market power. In a perfectly competitive market, $\partial f / \partial p$ will be near infinity and competition will drive

---

[12] Throughout this analysis, I treat v (the commission charged by the bookmaker) as a parameter rather than a decision variable for the bookmaker. Commissions are virtually always 10%. Gaining a better understanding of the reasons for the uniformity of commissions across bookmakers and over time presents an interesting puzzle for future research.

[13] Although I use the term 'bias' to describe bettor preferences, I do not necessarily imply irrationality on their part. If there is more consumption value associated with betting on favourites, then bettor preferences for favourites could be completely rational.

[14] A bettor breaks even when $p - (1 + v)(1 - p) = 0$. The solution to that expression is $p = 0.5238$.

© Royal Economic Society 2004

the spread back to the point where $f(p) = 0.5$ and no excess profits are obtained by bookmakers. Empirically, competitive pressure does not appear to be strong enough to eliminate excess profits. Understanding why this is the case is an important unanswered question of this research.

2. The Data Set

The data used in this paper are wagers placed by bettors as part of a handicapping contest offered at an online sports book during the 2001–2 NFL season. In the contest, bettors were required to pick five games per week against the spread for each of the 17 weeks of the NFL regular season (a total of 85 games). Bettors could choose those five games from any of the 13–15 games being played in a given week. One point was given for each correct pick, and one-half point if a game ended exactly on the spread. The entry fee was $250 per person, and there were 285 entrants. All of the entry fees were returned as prize money, so participants were competing for a total pool of $71,250.[15] 60% (or $42,750) went to the bettor with the most correct picks. Second to fifth place finishers received declining shares of the pool. The last-place finisher (conditional on making 85 picks) received 5% of the pool. In the data, I observe the ID number of each bettor, all wagers placed as part of this contest, the spread at which the bet was placed and the outcome of the game.

There are a number of potential shortcomings with these data. First, these are not wagers in the traditional sense. The bettor does not receive a direct payoff from winning any particular game; the payoff is only based on the cumulative number of wins. Nonetheless, the presence of large monetary rewards to the winners provides strong incentives to the participants. Presumably, the picks made by bettors in the contest closely parallel the actual betting wagers they were making; anecdotally that is true among the contest participants known to the author.[16] Second, the nature of the data make it impossible to ascertain the intensity of preferences across games since all selections receive equal weighting. Bettors may have much stronger preferences for their most favoured pick of the week than would be the case for the fifth-favourite pick. Third, there is substantial attrition in the sample. Of the 285 bettors who entered the contest, 100 (a little more than one-third) made their entries all 17 weeks of the season. More than 60% participated in at least 15 of the 17 weeks of the season. Less than 10% of the contestants recorded data for fewer than eight weeks. For bettors in the middle of the pack as the season progresses, the incentive to continue participating decreases substantially. Bettors who miss a week receive zero points, greatly reducing their likelihood of winning the contest, and

[15] The apparent purpose of the contest was to ensure that the bettors had a reason to visit the website each week. The fact that the worst-place finisher (conditional on having bet each week) received a payoff, reinforces this point.

[16] One important way in which the contest wagers might be expected to differ from actual wagers placed is that there is an incentive in the contest to pick outcomes that the bettor believes will be unpopular with other gamblers. That is because of the tournament structure of the contest in which rewards are great in the extreme right-hand tail but no differentiation is made elsewhere in the distribution.

disqualifying themselves from eligibility for the last-place prize. It should be noted, however, that attrition in this context adversely affects my ability to test only one of the hypotheses: what the overall distribution of bettor success rates looks like. Because attrition is non-random, the set of bettors who continue to the end will be skewed. Fourth, the spreads used in the handicapping contest are fixed on the Tuesday preceding the game and do not fluctuate with the actual spread, even though a bettor's contest picks are not due until the Friday before the game. As a consequence, in some games, the actual spread and the contest spread differ at the time an entry is made. All of the results of the paper, however, are robust to dropping games in which there are substantial fluctuations in the spread between Tuesday and Friday. Finally, these data are not the universe of bets placed at the sports book (although they are the universe of bets in this contest), much less at sports books in general. Nor, as discussed below, are the bettors who participated in the contest likely to be a random subset of all bettors.

   On the other hand, these data do offer enormous advantages over that which is typically available. Virtually all previous analyses of sports wagering have focused exclusively on prices but have not had access to any information about the quantity of bets on each side of the wager (Avery and Chevalier, 1999; Golec and Tamarkin, 1991; Gray and Gray, 1997; Kuypers, 2000).[17] In my sample, there are a total of 19,770 bets in the data set covering 242 different games. An average of 80.5 different bettors make a selection on a game, with the minimum and maximum number of bets on a game ranging from 28 to 146. Although the bettors included in my sample are not a random selection of all gamblers, they represent a particularly interesting subset. These bettors are likely to be relatively sophisticated, serious bettors. Because they are betting at an online sports book, they are likely to be geographically quite diverse. In signing up for the contest, they are indicating an expectation that they plan to visit (and presumably bet at) the internet sports book every week of the season. In addition, to the extent there are differences in skill across bettors, this contest should attract the most skilful players because it rewards exceptional long-run performance.

3. Empirical Results

I begin the analysis by addressing the issue of whether the spread is set so as to equalise the wagers on either side, as well as testing the predictions of a model of profit-maximising bookmakers who exploit biases on the part of bettors. The analysis then turns to the question of whether bettors differ in their skill at picking winners. Finally, I examine whether aggregating information across bettors provides any valuable information.

---

[17] In his innovative work, Strumpf (2002) does have actual bets based on seized bookmaker records. Although his data are extremely informative on a number of questions, they are less than ideal for the questions posed in this paper both because they cover a short time period and are geographically localised.

© Royal Economic Society 2004

3.1. *How Are Prices Set?*

The first question addressed is whether bookmakers set prices so as to equalise the amount of money on either side of the wager. Although I have data for only one bookmaker, it is important to note that the prices (i.e. spreads) offered by this sports book are virtually identical to those at any bookmaker online or at Las Vegas casinos. Thus, in practice individual bookmakers are not actively setting prices but, rather, following the lead of a handful of influential odds makers who are paid by large Las Vegas casinos for their services.

Figure 1 presents a histogram of the fraction of the total wagers placed on the team that bettors most prefer. By definition, this fraction must lie between 0.5 and 1. If the bookmaker balances bets, these values will be concentrated near 0.50. In the data, however, this is clearly not the case. In only 20% of the games are 50–55% of the wagers placed on the preferred team. In the median game, almost two-thirds of the bets fall on one side. In almost 10% of the games, more than 80% of the bets go one direction.

The dispersion in Figure 1 is not simply the result of sampling error due to the fact that I observe only 80 bets per game on average. If bettor choices were independent and each bettor had a 50% chance of picking either team, then one would expect the preferred team to garner between 50% and 55% of the wagers in nearly two-thirds of the games, compared to only 20% in the data. Furthermore, if the sample is divided in half based on the number of bets placed, the dispersion of the fraction of wagers placed on the preferred team is actually greater in the games with more total bets.[18]

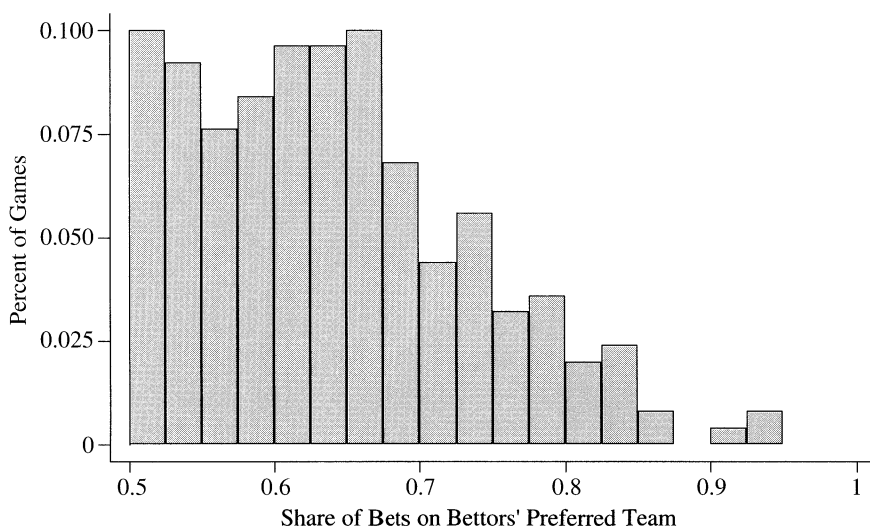
Fig. 1. *Share of Bets on the Team that Bettors Prefer*

[18] Further confirmation of the patterns in Figure 1 come from data available at http://www.wagerline.com. At that website, visitors make hypothetical wagers on game outcomes as part of small-payoff contests. The breakdown of bets on each team is available. The patterns in that data are strikingly similar.

© Royal Economic Society 2004

Two alternative hypotheses could explain the failure of the bookmaker to equalise wagers on the two sides of the spread. The first possibility is that the bookmaker would like to balance the bets but is unable to do so because it is difficult to predict accurately what team bettors will prefer. The second hypothesis is that balancing the wagers is not the objective of the bookmaker. Indeed, as demonstrated earlier, if bettors exhibit systematic biases, a profit maximising bookmaker does not want to equalise the money bet on both sides. Rather, the bookmaker intentionally skews the odds such that the preferred team attracts more wagers but wins less than half of the time.

If bookmakers are attempting to balance the money bet on each side of the wager, one would expect that observable characteristics of a team or game would have no power in predicting the fraction of bettors preferring that team. Otherwise, the bookmaker could have used that information to set a spread that would have better equalised the distribution of bets. If the bookmaker is attempting to exploit bettor biases by setting skewed odds, however, the opposite is true. Dimensions along which bettors exhibit bias should be systematically positively related to bet shares (and as demonstrated below, will also be systematically negatively related to win percentages).

Figures 2 and 3 provide some initial evidence on the question of whether observable characteristics are correlated with the distribution of wagers on a game. Figure 2 presents a histogram of the fraction of bets placed on the home team for the subset of games in which the home team is the favourite.[19] Since the identity of
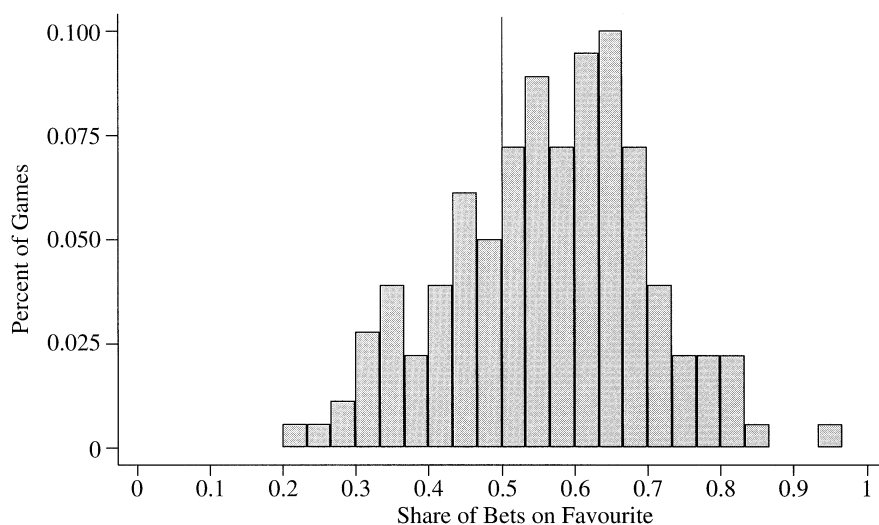
Fig. 2. *Share of Bets on the Favourite when the Home Team is the Favourite*

[19] Note that despite the similarity in the word 'favourite' and the phrase 'bettors' preferred team', these are two completely different concepts. 'Favourite' refers to the team judged most likely to win the game by the bookmaker. The 'bettors' preferred team', on the other hand, is the team that the bettors think is most likely to cover the spread. In other words, the bettors are making their choices conditional on the bookmaker already setting a spread that ostensibly equalises the chance of the favourite and the underdog covering.

© Royal Economic Society 2004

the home team is readily observable, the histogram in Figure 2 should be centred on 50% if the bookmaker is attempting to equalise bets on either side of the wager. The distribution is clearly skewed to the right, implying that home teams systematically attract more than half of the bets in games in which they are favoured. In almost three-quarters of the games, more bets are placed on home favourites than on their opponents. In the median game, roughly 58% of the bets go to the home favourite. Figure 3 is identical to Figure 2, except that the sample is games in which the visiting team is favoured, and the values in the Figure are the fraction of bets placed on the visiting team. In these games, the distribution of bets is even more skewed. In more than 90% of games with a visiting favourite, more bets are placed on the visitor than the home team. In the median game in Figure 3, two-thirds of the money is wagered on the visitor. Thus, Figures 2 and 3 demonstrate quite definitively that the spreads are set such that substantially more than half of the bets are placed on favourites.

Table 1 further explores the issue of whether observable characteristics are correlated with betting patterns. The dependent variable in Table 1 is the percentage of bettors who choose the favourite. The unit of observation is a game. The method of estimation is weighted least squares, with the weights proportional to the total number of bets placed on the game. The first column of the Table demonstrates that, not controlling for anything else, 60.6% of the bets accrue to the favourites. The standard error on that point estimate is 0.009, so the null hypothesis that 50% of the bets are placed on favourites is strongly rejected. Column (2) adds five indicator variables corresponding to which team is favoured and by how much (the omitted category is games in which the visitor is favoured by more than 6 points). Consistent with the figures presented earlier, home favourites do not attract as high a fraction of the bets as do visiting favourites. These five variables capturing the spread are jointly highly statistically significant, as reported
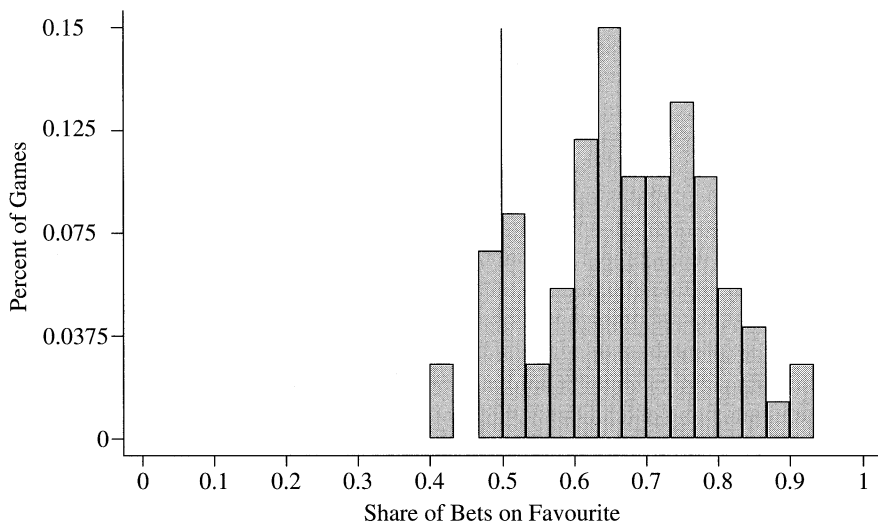

Fig. 3. *Share of Bets on the Favourite when the Visiting Team is the Favourite*

© Royal Economic Society 2004

Table 1

Predicting the Fraction of Bets Placed on the Favourite

Variable	Dependent variable: % of bettors placing bets on the team that is favoured			
	(1)	(2)	(3)	(4)
Constant	0.606 (0.009)	0.689 (0.025)	–	–
Home team favoured by more than 6 points	–	−0.129 (0.031)	−0.131 (0.031)	−0.144 (0.031)
Home team favoured by 3.5 to 6 points	–	−0.127 (0.033)	−0.123 (0.032)	−0.136 (0.037)
Home team favoured by 3 or fewer points	–	−0.126 (0.031)	−0.126 (0.031)	−0.123 (0.043)
Visiting team favoured by 3 or fewer points		−0.005 (0.030)	−0.026 (0.030)	−0.057 (0.033)
Visiting team favoured by 3.5 to 6 points		−0.016 (0.035)	−0.002 (0.034)	−0.002 (0.034)
Week of season dummies included?	No	No	Yes	Yes
Team dummies included?	No	No	No	Yes
R-squared	–	0.165	0.299	0.484
p-value of test of joint significance of:				
Spread variables	–	< 0.01	< 0.01	< 0.01
Week dummies	–	–	< 0.01	< 0.01
Team dummies	–	–	–	< 0.01

Notes: Omitted category for the spread variables are games in which the visiting team is favoured by ten or more points. The unit of observation is a game. The number of observations is equal to 242 in all columns. Standard errors are in parentheses. The method of estimation is weighted least squares, with the weights proportional to the total number of bets placed on the game.

in the bottom of the Table. The third column adds dummy variables for each week of the season. These week dummies are also jointly statistically significant at the 0.01 level. Although not shown individually, the dummies suggest that a greater fraction of the bets are placed on favourites early in the season. Finally, the last column adds 31 indicator variables corresponding to each team in the league. These variables take the value of one if a team is favoured and −1 if the team is an underdog. Once again, the team variables are highly statistically significant.[20] The R^2 in the final column of the Table is 0.484, implying that these observable characteristics explain a great deal of the variation in the fraction of money bet on the favourite.

The results in Table 1 uniformly argue against the hypothesis that bookmakers are doing the best they can to even out the bets on each game, suggesting instead that the imbalance is intentional. If that is true, then the model presented earlier makes strong predictions about the expected pattern of winning percentages for bets of different kinds. In particular, teams with attributes that attract a disproportionate share of the money from bettors (e.g., being the favourite) should cover the spread less than 50% of the time. But, the deviation from 50-50 cannot

---

[20] Based on these estimates, the five teams that attracted the most bettors, controlling for other factors, were Green Bay, Kansas City, New Orleans, Oakland, and San Francisco. The least popular teams were Baltimore, Chicago, Jacksonville, Pittsburgh, and Saint Louis.

© Royal Economic Society 2004

be too large (more than a few percentage points), or bettors who do not suffer from biases can profitably exploit the price distortion.

Table 2 presents evidence consistent with those predictions. The unit of observation in the Table is an individual bet. The first three columns capture the fraction of the bets placed on a team; columns 4–6 are the percentage of those wagers in which the bettor correctly picks the winners. Bets are categorised according to whether they are placed on games in which the home team is favoured (top row) or the visiting team is favoured (second row), and whether the bet is for the favourite or the underdog (the columns in Table). Mirroring the results presented earlier, visiting favourites attract a disproportionate share of the bets placed in such games: 68.2% of the total.[21] To a lesser degree, home favourites also attract excess bets (56.1%). Shifting focus to columns 4 and 5, the model predicts that a high fraction of bets in columns 1 and 2 will be associated with low winning percentages in columns 4 and 5. Indeed, the results confirm the prediction. Across the four categories considered, the rank-order correlation between the percentage of bets placed and the win percentage is −1. Bets placed on visiting favourites win only 47.8% of the time. Bets on home favourites are successful in 49.1% of cases. Bets on underdogs, which are under-represented in the data, win more than half of the time. Notably, bets on home underdogs have win rates of 57.7%, well above the threshold required for a bettor to break even.[22]

One might be tempted to discount the observed relationship between a high fraction of bets made and low winning percentages since the results are based on

Table 2
Bets Placed and Won on Favourites and Underdogs

Which team is favoured in the game?	% of total bets on the game that are placed on:			% of bets placed that win (i.e. cover the spread) when a team is:		
	(1) Favourite	(2) Underdog	(3) Total, favourite and underdog	(4) Favourite	(5) Underdog	(6) Total, favourite and underdog
Home team	56.1 [12,011]	31.8 [7,190]	47.0 [19,201]	49.1 [6,741]	57.7 [2,286]	51.2 [9,027]
Visiting team	68.2 [7,190]	43.9 [12,011]	53.0 [19,201]	47.8 [4,904]	50.4 [5,270]	49.1 [10,174]
Total, home and visiting team	60.6 [19,201]	39.4 [19,201]	50.0 [19,201]	48.5 [11,645]	52.6 [7,556]	50.1 [19,201]

Notes: The values reported in the first three columns of the Table are the percentage of total bets placed on the named team (e.g. home favourite in row 1, column 1). The values reported in the last three columns of the Table are the fraction of bets placed that win. The unit of analysis is a bet. The number in brackets is the total number of bets placed in each cell. The results in this Table exclude the six games where the spread was equal to zero, i.e. neither team was favoured.

[21] By definition, the sum of the fraction of bets on visiting favourites (row 2, column 1) and home underdogs (row 1, column 2) must add up to 100%. The same holds for home favourites and visiting underdogs.

[22] The win percentages on visiting favourites and home underdogs need not sum to 100% because the fraction of the bets on the favourite and the underdog varies across games.

© Royal Economic Society 2004

Table 3

A Comparison of Betting Frequencies in my Sample and Game Outcomes against the Spread over the Last 21 Years

Type of game	% of bets in my sample that are placed on the favourite	% of games in which favourite wins over last 21 NFL seasons
All games	60.6	48.2 [4,793]
Games in which home team is favoured	56.1	48.8 [3,310]
Games in which visiting team is favoured	68.2	46.7 [1,483]
Games played in first half of season	63.4	47.7 [2,209]
Games played in second half of season	56.3	48.5 [2,584]
Games in which the point spread is greater than six points	60.4	48.5 [1,759]
Games in which the point spread is 3.5 to six points	59.5	48.1 [1,475]
Games in which the point spread is less than or equal to three points	61.5	47.8 [1,559]

Notes: Values in the first column are based on bets placed for the sample of 285 bettors for the 2001 season used throughout the paper. Values in the second column are game outcomes for the 21 NFL seasons covering the period 1980–2001. Total number of games included in the calculation in brackets.

the outcomes of only 236 games of a single season. Although it is impossible to obtain the data on quantities for earlier years, outcomes of games relative to the spread are readily available. Assuming that the strong tendencies towards betting on favourites are persistent across years, one would expect the patterns in winning percentages to also be persistent. Indeed, the winning percentage patterns have previously been documented by Golec and Tamarkin (1991) and Gray and Gray (1997). They are also present in Table 3, which includes results on game outcomes for 21 seasons of data (1980–2001), covering almost 5,000 games. The first column of the Table presents the percentage of bets made on teams of a particular type in the 2001–2 data; the second column is the win percentage over the past 21 years. The results are quite consistent with those of Table 2. Overall, favourites win less than half their games. The null hypothesis of a 50% win rate for favourites is rejected at roughly the 0.01 level. As predicted, visiting favourites (who attract a greater share of the bets) do especially poorly, winning only 46.7%, again rejecting the null of a 0.50 win percentage at approximately the 0.01 level. Given this win percentage, a naive strategy of always betting against visiting favourites would actually have yielded positive profits over these two decades (as was also true in 2001–2).[23] In light of the consistent lack of success of favourites, especially visiting favourites, it is remarkable that the strong bettor bias towards such bets persists. The bettor bias is not concentrated among a small fraction of bettors, either. Three-fourths of the contestants chose favourites more often than underdogs. Only 2% of bettors chose visiting underdogs for at least half of their picks.

[23] Although not shown in tabular form, the year-by-year patterns in the data confirm the overall findings. In only 4 of the last 21 years have favourites covered the spread in as many of 50% of the games. The likelihood of that occurring if the true win likelihood is 0.50 is less than 1 in 300. Only 3 times in 21 years have visiting favourites won against the spread in 50% of games.

© Royal Economic Society 2004

Similar results are also obtained from betting on other American sporting leagues. Home underdogs covered the spread in 53.2% of the National Collegiate Athletic Association (NCAA) college football games played in 2002, as well as 53.0% of professional basketball games in the 2002 season of the National Basketball Association (NBA).

Just how much do bookmakers increase their profits by exploiting bettor biases in professional football? Assuming that (1) the total distribution of bets in this sample is representative of overall betting and (2) there is no information in aggregate bettor preferences (the evidence presented below cannot reject this), it is straightforward using the values in Table 2 to calculate that the way spreads are currently set, bettors should win 49.45% of their bets.[24] Given the standard 'vig' of bettors risking 110 units to win 100, a bookmaker who wins half his bets has a gross profit rate of 5.0%. If bettors win only 49.45% of their bets, the expected gross profit rate jumps to 6.16% ($0.5055 \times 110 - 0.4945 \times 100$). Thus, in expectation, this seemingly minor distortion of the win rate increases gross profits by 23%. It is true, of course, that the bookmaker must bear some risk when the bets are not balanced on both sides of the wager. Because game outcomes are likely to be independent, however, the risk is minimised as the number of games played increases. For instance, in the case where 63% of the money is one side of each wager and that team wins 48% of the time, over the course of the NFL season (roughly 250 games) the bookmaker's expected gross profit rate is 6.1% with a standard deviation of 2.5%. Thus, the bookmaker would be expected to make negative gross profits less than once every one hundred seasons. If one looks over a five-year time frame, the standard deviation drops to 1.1%. So the probability of a bookmaker losing over any given five-year period in this scenario is less than one in 10,000. Relative to a 23% increase in gross profit, the costs associated with bearing this level of risk appear minimal.[25] One cost to bookmakers of bearing risk, however, is the need to have substantial liquid capital available to them in case of an adverse shock.

The last two panels of Table 3 explore the relationship between other factors and betting on favourites. In the 2001–2 data, a higher fraction of bets were placed on favourites in the first half of the season than in the second half (63.4% versus 56.3%), and the week of the season was highly statistically significant in predicting bet shares in Table 1. Whether this is simply an idiosyncracy of the 2001–2 season is uncertain. Consistent with the theory, the win percentages for favourites over 20 years are higher in the second half of the season (48.5 versus 47.7), although the differences is not statistically significant. Finally, the bottom panel of the Table demonstrates that the size of the spread has little impact on the distribution of bets on the favourite; correspondingly,

---

[24] The number 49.45 is obtained by multiplying the probability that the home favourite wins a game times the percentage of total bets on home favourites plus the probability that the visiting underdog wins times the percentage of overall bets on visiting underdogs etc.

[25] A major puzzle in this industry is the rarity of price competition, i.e. the vig is almost universally 10%. It is possible that the bearing of risk somehow supports this equilibrium. One website, http://www.tradesports.com, acts as a traditional financial market-maker, matching buyers and sellers but taking no positions on game outcomes. The commission charged for this match-making service is less than 1% of the bet – far smaller than the traditional vig.

© Royal Economic Society 2004

there is little apparent difference in win percentages across these games in column 2.

3.2. *Is There Evidence that Some Bettors Are Especially Skillful in Picking Winners?*

In order for the current system of price-setting (in which the bookmakers set a price and do not adjust that price to equilibrate supply and demand) to survive, there cannot exist a sufficient number of bettors with an ability to pick winners that exceeds that of the bookmaker. This is particularly true when the bookmaker distorts prices to exploit the subset of bettors with biases. In that case, a sophisticated bettor only needs to be slightly better than the bookmaker in determining the true odds to turn a profit.[26] Indeed, Strumpf (2002) argues that much of the internal structure of bookmaker organisations is designed to protect the bookmaker against adverse selection by these talented bettors.

   Testing for bettor skill is complicated in my data set by the fact that there is a great deal of attrition over the course of the sample, and the attrition is not random. Bettors who have performed poorly up to that point in time are much more likely to leave the sample since the chances of receiving a prize are very low for these contestants.

   I consider two possible approaches for testing for heterogeneity in skill across bettors. The first approach is to look at the overall distribution of games won over the course of the season and to test whether that distribution is consistent with that which would have been generated by homogeneous bettors.[27] Because of attrition, however, these data are incomplete. Approximately 82% of possible bets were actually placed. Under the assumption that the outcomes of bets on the missing games can be modelled as being generated by independent coin tosses with probability 0.5, it is possible to simulate what the distribution of wins would have been without attrition. This approach has the obvious drawback that the simulated portion of the data is generated by the process that I have defined as the null hypothesis against which to test. Thus, such a test is biased against rejecting the null of no heterogeneity.[28] Figure 4 presents a representative histogram of the distribution of these simulated final win totals. Superimposed on the histogram is the corresponding normal distribution which the data would be expected to approximate if generated by i.i.d. coin tosses with a win probability of 0.50. Visually, the observed distribution closely mirrors the normal distribution. p-values for the three generally applied tests of normality (skew test, Shapiro-Francis, and Shapiro-Wilk) are well within the acceptable range. Thus, with the caveat that the test is biased against rejection due to the simulated data, there is no evidence to reject the null hypothesis of no differences in skill across bettors in the sample.

---

[26] And, as demonstrated above, even a naive strategy of betting against all visiting favourites has been marginally profitable.

[27] Because the worst-place finisher gets a payoff, those near the bottom have an incentive to try to pick losers intentionally. If they have some ability to do this, that will exaggerate the bottom tail, exacerbating deviations from normality.

[28] In defence of the manner in which the missing data are generated, other results presented below suggest that there is no evidence of serial correlation across weeks in a given bettors ability to pick winners.
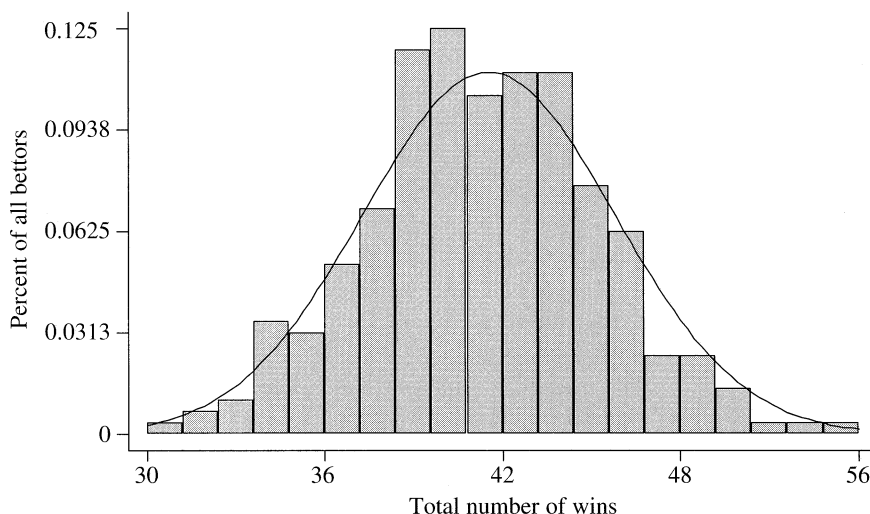
© Royal Economic Society 2004

Fig. 4. *Simulated Distribution of Total Wins Relative to Null Hypothesis of Random Coin Flips*

The second approach to testing whether there is heterogeneity across bettors in ability to pick winners is to look for persistence in win rates. *A priori*, it is not known who the skillful bettors are. Success early in the season, however, is likely to be a (possibly noisy) signal of talent. Thus, in the presence of heterogeneity in skill, one would expect those who do better at the beginning of the season also to outperform later in the season. Like the test of homogeneity of skill presented above, this approach is not robust to particular sources of attrition. If high-skilled bettors who have been unlucky in the early part of the season are less likely to quit than similarly placed low-skilled bettors, then this approach will be biased against finding heterogeneity in skill across bettors. The poor-performing bettors who persist will be disproportionately drawn from the high-skilled group and thus will be expected to perform well on average later in the season.

Bearing in mind this caveat, I estimate equations of the form

$$WIN_{bwg} = \alpha + \beta_1 H_{bw} + \beta_2 (H_{bw})^2 + \beta_3 (H_{bw})^3 + \beta_4 (H_{bw})^4 + \gamma \mathbf{X}_{bw} + \varepsilon_{bwg} \qquad (4)$$

where b, w, and g denote bettors, weeks of the season, and specific games respectively. *WIN* is a variable equal to one if the bettor picks the game and covers the spread, 0.5 if the game is a push, and zero otherwise. The variable H represents the bettors cumulative historical winning percentage across all games played thus far in the season. The vector $\mathbf{X}$ captures other predictors of whether or not the bet is won, for example if the team chosen is a visiting favourite. The quartic in the cumulative winning percentage is designed to non-parametrically capture the serial correlation across betting performances. The first week of the season is omitted from the regression because there is no bettor history. The equation is estimated using weighted least squares, with weights determined by the number of games making up the history. The reported standard errors have been corrected through clustering to account for the fact that H is correlated across games for a given contestant.

© Royal Economic Society 2004

Table 4 reports results of the estimation. In the first column, the cumulative betting success rate is constrained to enter linearly. Although not statistically significant, the point estimate implies that bettors who have been more successful up until that point in the season are predicted to do slightly *worse* in the current week. This argues against heterogeneity in skill across bettors, which would lead to a positive coefficient. The bottom panel of the Table reports the predicted success rate for bettors with varying win percentages up to this point in the season. The second equation adds the quartic in betting history. Although the history variables are jointly statistically significant, the R^2 is very low (0.0004). Most bettors averaging are predicted to perform right around 50%, bettors in the top quartile prior to this week are projected to win only 49.0% of games. Once again, these results argue against persistent differences in skill. Adding covariates in column 3 has little impact on the conclusions. Columns 4 and 5 restrict the sample to exclude the first five and ten weeks respectively, on the rationale that cumulative win percentages early in the season may not be very informative. The results provide no evidence that strong past performance predicts wins today.

In summary, there is little in the data to suggest that, at least in this particular sample, there is heterogeneity in skill across bettors. This result may be due partially to the relative sophistication of bettors in the sample – perhaps the most naive bettors are unlikely to frequent internet bookmakers.[29]

3.3. Does Pooling Information Across Bettor Preferences Help in Predicting the Outcome of Games?

In other contexts, it has been argued that aggregating information across agents provides valuable information in predicting future outcomes. For example, Clemen and Winkler (1986) and Fomby and Samant (1991) find that the consensus estimate of future GNP growth is a better predictor than any one individual's estimate. One might also expect such a pattern to be present in sports betting, especially because price is set unilaterally by the bookmaker. To the extent that the bookmaker sometimes makes mistakes, one would expect that many bettors will simultaneously recognise the presence of the mistake and disproportionately pick one team.

There is one simple result in my data which suggests that aggregating opinions across bettors may carry valuable information: despite the fact that more than half the money is bet on favourites and the bookmaker set the odds so that favourites win less than half the games, the overall winning percentage for bets placed is 50.1%. As noted earlier, based on the odds offered by the bookmaker and the distribution of money bet, one would expect 49.45% of all bets to win if there was no correlation between the percentage of bettors choosing a game and the game's outcome. The difference between 50.1% and 49.45% implies that games in which a greater frac-

---

[29] Strumpf (2002), for instance, reports the existence of a fraction of New York bettors who always bet on the Yankees, even though the bookmakers, knowing their preferences, systematically offer these bettors substantially worse odds than other clients.

© Royal Economic Society 2004

Table 4

Bettor Performance as a Function of Cumulative Win Percentage up to that Point in the Season

Variable	Dependent variable = bettor wins this bet				
	(1)	(2)	(3)	(4)	(5)
Bettor's cumulative ranking on win percentage (0 = last place, 1 = first place)	−0.012 (0.015)	−0.06 (0.32)	0.03 (0.34)	0.10 (0.36)	0.07 (0.47)
Cumulative rank 2	–	0.54 (1.31)	0.25 (1.36)	−0.07 (1.47)	0.21 (1.91)
Cumulative rank 3	–	−1.04 (1.97)	−0.69 (2.04)	−0.18 (2.20)	−0.68 (2.86)
Cumulative rank 4	–	0.56 (0.98)	0.41 (1.01)	0.15 (1.09)	0.41 (1.41)
Bettor bets on home favourite	–	–	−0.048 (0.010)	−0.064 (0.014)	−0.037 (0.016)
Bettor bets on home underdog	–	–	0.046 (0.012)	0.026 (0.014)	−0.010 (0.018)
Bettor bets on visiting favourite	–	–	−0.005 (0.010)	−0.002 (0.013)	0.121 (0.016)
Constant	0.506 (0.009)	0.499 (0.024)	0.497 (0.026)	0.534 (0.030)	0.598 (0.036)
R-squared	0.0001	0.0002	0.0041	0.0048	0.0131
Number of observations	18,345	18,345	18,345	12,985	6,785
Sample of games used in estimation	Weeks 2–17	Weeks 2–17	Weeks 2–17	Weeks 6–17	Weeks 11–17
p-value of test of joint significance of cumulative rank variables	–	0.64	0.51	0.73	0.72
Predicted win percentage if cumulative ranking on win percentage falls in:					
Top quartile	49.4	49.0	49.0	49.7	49.0
Second quartile	49.8	50.2	50.1	50.4	50.5
Third quartile	50.1	50.7	50.8	51.0	51.2
Bottom quartile	50.4	49.9	49.9	50.1	49.4

Notes: The dependent variable in all regressions is equal to one if the bettor wins the bet, zero otherwise. The unit of observation is a bet. A bettor's cumulative ranking on win percentage is calculated by taking an unweighted average of that bettor's win percentage on all bets up to that point in time, and then computing the bettor's rank among all bettors. The method of estimation is weighted least squares, with the weights proportional to the total number of games making up the year-to-date history. Standard errors (in parentheses) are corrected to take into account correlation in the cumulative rank variable across bets for a given bettor.

© Royal Economic Society 2004

tion of bettors choose the favourite (or alternatively the underdog) are more likely to be won by the favourite (underdog). Thus, in principle one might believe that knowledge of aggregate bettor preferences might be useful in prediction, making access to quantity data (which is in general very difficult to obtain, but is available prior to the start of the games through this contest) valuable.[30]

Table 5 provides additional suggestive evidence on this issue. Bets are divided into four categories: home favourites, visiting favourites, home underdogs, and visiting underdogs. Within each category, bets are divided into quartiles according to how popular that team was with bettors. For instance, the top quartile of bets on home favourites would capture those bets in which an especially high share of the people making selections in that game chose the favourite.

Results presented in the Table show the total for all games (the first column) and then broken down by quartile (columns 2–5). If aggregating bettor preferences is valuable in predicting performance against the spread, then one would expect to observe a monotonic, declining win percentage moving from the top quartile to the bottom quartile of bets by popularity. In practice, bets in the second quartile yield the highest winning percentage (52.9%). The top and third quartiles also win more than 50% of bets. The worst performance by far (46.6%) is for those selections that are least popular with bettors. Thus, while the evidence is mixed, there is at least a suggestion in the data of the possibility that popular bets fare better, at least relative to the least popular choices in the bottom quartile.

Table 5
Win Percentages as a Function of whether a Bet is Popular Among Bettors

| Category of games: | Win percentage as a function of the degree to which bettors prefer this selection (relative to other teams in the same category): | | | | |
	All games, regardless of degree preferred by bettors	Top quartile (i.e. selections most popular with bettors)	Second quartile	Third quartile	Bottom quartile (i.e. selections least popular with bettors)
Home favourites	49.1 [6,741]	46.0 [1,666]	44.6 [1,697]	57.4 [1,696]	48.2 [1,682]
Visiting favourites	47.8 [4,904]	52.0 [1,240]	57.4 [1,210]	43.9 [1,206]	39.2 [1,205]
Home underdogs	57.5 [2,286]	71.0 [568]	52.6 [579]	59.9 [564]	47.5 [575]
Visiting underdogs	50.4 [5,270]	45.5 [1,325]	59.6 [1,288]	45.6 [1,318]	51.0 [1,339]
All favourites	48.6 [11,645]	48.6 [2,906]	49.9 [2,907]	51.8 [2,902]	44.4 [2,887]
All underdogs	52.6 [7,556]	53.2 [1,893]	57.4 [1,867]	49.9 [1,882]	49.9 [1,914]
All bets (favourites and underdogs)	50.1 [19,201]	50.4 [4,799]	52.9 [4,744]	51.0 [4,784]	46.6 [4,801]

Notes: The values reported in the Table are fraction of bets won. The unit of observation is a bet. Games are categorised into quartiles within bet type (e.g. home favourites, visiting favourites, home underdogs, visiting underdogs) according to the fraction of bets on the team in the named category. Top quartile bets correspond to those cases where the greatest fraction of bettors selected the team in the named category. Total number of bets underlying each calculation in brackets.

[30] Interestingly, the win percentage for mock bets placed at http://www.wagerline.com were 49.5% in 2001 – precisely what would be expected if aggregate bettor preferences contained no information. One possible explanation for the difference in win rates between the bettors in my sample and those at wagerline is that my sample contains a more sophisticated set of gamblers who have much more money at risk.

© Royal Economic Society 2004

Table 6 tackles this question more rigorously in a regression framework that estimates the impact of a bet's popularity non-parametrically using a fourth-order polynomial and controlling for other factors such as week of the season dummies and team-fixed effects. In addition, Table 6 reports standard errors corrected for the fact that although there are over 19,000 bets made, there are only a few hundred game outcomes; for a given game, the bet outcomes will be perfectly correlated, so effectively there are only a few hundred degrees of freedom. The results of the Table provide suggestive, but ultimately quite weak, evidence in support of the value of aggregating preferences. In columns 1 and 2, a bet's popularity among bettors enters linearly. The coefficient is 0.11, implying that a ten percentage point increase in the number of bettors favouring a particular gamble is associated with a 1.1 percentage point increase in the likelihood that team wins. While substantively large, the estimate is not statistically significant. When allowing the coefficient on a bet's popularity to vary in a non-linear fashion, the individual coefficients are no longer readily interpretable, so I report average win percents by quartile in the bottom panel of the Table. The patterns observed in columns 3–5 are similar to those in the raw data, with the bets in the top half of the distribution winning more than 50% and bets in the bottom quartile doing especially poorly. It is important to note, however, that the coefficients on bet popularity are in all cases jointly insignificant (as reflected in the p-values reported near the bottom of the Table). Thus, great caution must be exercised in drawing any conclusions from these regressions.

4. Conclusions

This paper utilises a unique data set that includes information on both the prices and quantities of bets placed to analyse the workings of the market for gambling on NFL football. The results provide an answer to the apparent puzzle of why these markets do not look much like markets at all, characterised by small and infrequent price changes. The answer is that the bookmakers are better at predicting game outcomes than the typical bettor. As a consequence, the bookmakers are able to set prices in order to exploit their greater talent, and apparently yielding greater profits than could be obtained if the bookmakers acted like traditional market makers and attempted to equilibrate supply and demand, avoiding taking large stakes in the outcomes of games. In the presence of some bettors who are as skillful in picking games as are the bookmakers, there are limits to how much prices can be distorted without creating profit opportunities. It appears that the bookmakers have distorted prices to the point where they bump up against this limit.

In order to support this mode of price-setting, it cannot be the case that a significant fraction of bettors (or more precisely dollars bet) have better information than the bookmaker. In that case, it would be the bookmaker who would be exploited by this system. At least within my sample of bettors, I see little evidence of heterogeneity in ability to pick winners, suggesting that the bookmaker might be as good or better at picking game outcomes than any of the bettors.

© Royal Economic Society 2004

Table 6

Does Pooling Information Across Bettors Help in Picking Winners?

Variable	Dependent variable = bettor wins game				
	(1)	(2)	(3)	(4)	(5)
Rank of this bet's popularity (relative to other games of this type) (0 = least popular; 1 = most popular)	0.11 (0.11)	0.11 (0.11)	−1.02 (1.82)	−1.16 (1.84)	−1.04 (1.85)
Rank of bet's popularity 2	—	—	4.81 (7.61)	5.48 (7.63)	5.18 (7.55)
Rank of bet's popularity 3	—	—	−6.77 (11.20)	−7.92 (11.23)	−7.57 (10.94)
Rank of bet's popularity 4	—	—	3.01 (5.41)	3.64 (5.44)	3.46 (5.21)
Bet on home favourite	—	−0.013 (0.084)	−0.013 (0.083)	−0.011 (0.083)	−0.013 (0.082)
Bet on home underdog	—	0.073 (0.072)	0.073 (0.072)	0.080 (0.074)	0.055 (0.078)
Bet on visiting favourite	—	−0.026 (0.073)	−0.026 (0.073)	−0.022 (0.073)	−0.054 (0.070)
Constant	0.447 (0.048)	0.450 (0.062)	0.499 (0.119)	—	—
R-squared	0.004	0.008	0.010	0.022	0.048
Include week of season dummies?	No	No	No	Yes	Yes
Include team-fixed effects?	No	No	No	No	Yes
p-value of joint significance of bettor preference variables	—	—	0.81	0.80	0.78
Predicted win percentage by quartile of popularity of this bet:					
Top quartile (i.e. selections most popular with bettors)	54.2	54.1	52.3	52.3	52.1
Second quartile	51.5	51.4	54.2	54.1	54.0
Third quartile	48.8	48.8	49.4	49.5	49.8
Bottom quartile	46.1	44.5	44.6	44.6	46.1

Notes: The dependent variable in all regressions is equal to one if the bettor wins the bet, zero otherwise. The unit of observation is a bet. The rank of a bet's popularity is calculated relative to other games of this type (e.g. home underdog). The method of estimation is weighted least squares, with the weights proportional to the total number of bets placed on the game. Standard errors (in parentheses) are clustered to take into account correlation in the rank of bet popularity variables for different bets placed on the same game.

Given the incentives for the bookmaker to get the spread right, it is hardly surprising that the most talented individuals would be employed as the odds makers.

Perhaps, then, a fundamental difference between gambling and financial markets is that it is possible in gambling to find and hire a small set of individuals (the odds makers) who can systematically do better in predicting game outcomes than can bettors overall. In financial markets, on the other hand, the flow of inside information or the inherent complexity in valuing companies may make it impossible for one individual to do better than the market, meaning that a market maker who acted like a bookmaker would do worse than one who simply equilibrated supply and demand and took advantage of the bid-ask spread. The weight of the evidence regarding the inability of fund managers to systematically beat the market indexes is consistent with this conjecture.

University of Chicago and American Bar Foundation

References

Ali, M. (1977). 'Probability and utility estimates for racetrack bettors', *Journal of Political Economy*, vol. 85, (August), pp. 803–15.

Asch, P., Malkiel, B. and Quandt, R. (1984). 'Market efficiency in racetrack betting', *Journal of Business*, vol. 57, pp. 165–75.

Avery, C. and Chevalier, J. (1999). 'Identifying investor sentiment from price paths: the case of football betting', *Journal of Business*, vol. 72, (October), pp. 493–21.

Clemen, R. and Winkler, R. (1986). 'Combining economic forecasts', *Journal of Business and Economic Statistics*, vol. 4, (January), pp. 369–91.

DellaVigna, S. and Malmendier, U. (2003). 'Contract design and self-control: theory and evidence', unpublished manuscript, University of California-Berkeley Department of Economics.

Fomby, T. and Samanta, S. (1991). 'Application of stein rules to combination forecasting', *Journal of Business and Economic Statistics*, vol. 9, (October), pp. 391–407.

Gambling Magazine (1999). 'Bookies still counting the costs', published on the internet at http://www.gamblingmagazine.com/articles/31/31-57.htm.

Golec, J. and Tamarkin, M. (1991). 'The degree of price inefficiency in the football betting markets', *Journal of Financial Economics*, vol. 30, pp. 311–23.

Golec, J. and Tamarkin, M. (1998). 'Bettors love skewness, not risk, at the horse track', *Journal of Political Economy*, vol. 106, (February), pp. 205–25.

Gray, P. and Gray, S. (1997). 'Testing market efficiency: evidence from the NFL sports betting market', *Journal of Finance*, vol. 52, (September), pp. 1725–37.

Jullien, B. and Salanie, B. (2000). 'Estimating preferences under risk: the case of racetrack bettors', *Journal of Political Economy*, vol. 108, (June), pp. 503–30.

Kuypers, T. (2000). 'Information and efficiency: an empirical study of a fixed odds betting market', *Applied Economics*, vol. 32, pp. 1353–63.

Lee, M. and Smith, G. (forthcoming). 'Regression to the mean and football wagers', *Journal of Behavioral Decisionmaking*.

National Gambling Impact Study Commission (1999). *Final Report*, Washington, D.C.

Nevada Gaming Control Board (2002). 'Nevada gaming revenues calendar year 2001 analysis', press release, February 12.

Sauer, R., Brajer, V., Ferris, S. and Marr, M. (1988). 'Hold your bets: another look at the efficiency of the gambling market for national football league games', *Journal of Political Economy*, vol. 96, (February), pp. 206–13.

Shin, H. (1991). 'Optimal betting odds against insider traders', ECONOMIC JOURNAL, vol. 101, (September), pp. 1179–85.

Shin, H. (1992). 'Prices of state contingent claims with insider traders, and the favourite-longshot bias', ECONOMIC JOURNAL, vol. 102, (March), pp. 426–35.

Shin, H. (1993). 'Measuring the incidence of insider trading in a market for state-contingent claims', ECONOMIC JOURNAL, vol. 103, (September), pp. 1141–53.

Smith, J. (2002). 'Bookies who lose', posted at http://www.betasia.com and reprinted from *Business Life* (March 2001).

Strumpf, K. (2002). 'Illegal sports bookmakers', Mimeo, University of North Carolina. Department of Economics.

Thaler, R. and Ziemba, W. (1988). 'Parimutuel betting markets: racetracks and lotteries', *Journal of Economic Perspectives*, vol. 2, pp. 161–74.

Vaughn Williams, L. and Paton, D. (1997). 'Why is there a favourite-longshot bias in British racetrack betting markets?', ECONOMIC JOURNAL, vol. 107, (January), pp. 150–8.

Woodland, B. and Woodland, L. (1991). 'The effects of risk aversion on wagering: point spreads versus odds', *Journal of Political Economy*, vol. 99, (June), pp. 638–53.

Woodland, B. and Woodland, L. (1994). 'Market efficiency and the favourite-longshot bias: the baseball betting market', *Journal of Finance*, vol. 49, (March), pp. 269–79.

Zuber, R., Gandar, J. and Bowers, B. (1985). 'Beating the spread: testing the efficiency of the gambling market for national football league games', *Journal of Political Economy*, vol. 93, (August), pp. 800–6.