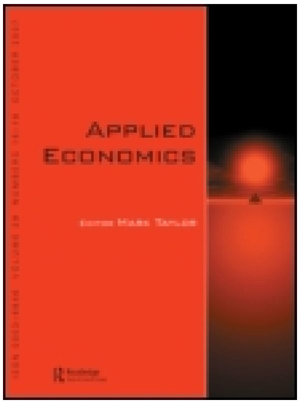


Applied Economics
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/raec20

# Sentiment in the betting market on Spanish football

David Forrest [a] & Robert Simmons [b]
[a] Centre for the Study of Gambling, University of Salford, Salford M5 4WT, UK
[b] Department of Economics, The Management School, Lancaster University, Lancaster, LA1 4YX, UK
Version of record first published: 30 Oct 2009.

To cite this article: David Forrest & Robert Simmons (2008): Sentiment in the betting market on Spanish football, Applied Economics, 40:1, 119-126

To link to this article: http://dx.doi.org/10.1080/00036840701522895

*Applied Economics*, 2008, **40**, 119–126



# Sentiment in the betting market on Spanish football

David Forrest[a] and Robert Simmons[b,*]

[a]*Centre for the Study of Gambling, University of Salford, Salford M5 4WT, UK*
[b]*Department of Economics, The Management School, Lancaster University, Lancaster, LA1 4YX, UK*



We employ a sample of over 3000 bets available on matches from the top tier of Spanish football in an examination of the efficiency of betting odds offered in the on-line betting market. Odds appear to be influenced by the relative number of fans of each club in a match, with supporters of the more popular team offered more favourable terms on their wagers. We report similar findings for a sample of games from Scotland. The results contrast with studies of American sports betting markets but are consistent with competitive behaviour by profit maximizing bookmakers in a market where bettors can choose between several operators.

## I. Introduction

In an early review of betting markets, Thaler and Ziemba (1988) distinguished between weak and strong efficiency. According to their usage of the terms, a set of odds is weak-efficient if odds are sufficiently reflective of objective probabilities that no strategy exists that would give bettors a positive expected return. The more stringent requirement of strong efficiency is that no strategy exists that would improve on the (negative) expected return from betting randomly.

A substantial literature on sports betting reports several violations of strong efficiency. For example, Gray and Gray (1997), in an analysis of the betting market on America's National Football League (NFL), noted a tendency for home teams and underdogs to beat the bookmaker spread disproportionately often. However, bookmaker commissions are high relative to what has to be paid to transact in conventional financial markets[1] and, as a result, it is hard to convert awareness of such biases to consistently profitable betting rules (Sauer, 1998). At least weak efficiency therefore appears to characterize prices in wagering markets.

The search for biases in spreads or odds has, however, been rather narrow in terms of the dimensions of efficiency for which tests have been conducted. The NFL market is organized on the basis of handicaps, so that bets are offered on which team will record a better result in the match than the 'spread' announced by the sports book. In this sort of market, researchers have focused on potential biases in the favourite/underdog and home/away dimensions.

Other sports betting markets are odds based, with wagers offered on a win by either team at odds varying according to their relative strengths. For these markets, published work has concentrated on bias in the short-odds/long-odds dimension (see, for example, Cain *et al.*, 2000, on soccer and Woodland and Woodland, 2001 and 2003, on ice hockey and

*Corresponding author. E-mail: r.simmons@lancaster.ac.uk

[1] In the NFL, each team is quoted at odds of 10/11 to beat the spread. With equal money wagered on each side, bookmakers would earn 10% of amount won by bettors (this is termed 'vigorish') or 4.55% of total stakes.

*Applied Economics* ISSN 0003–6846 print/ISSN 1466–4283 online © 2008 Taylor & Francis
http://www.tandf.co.uk/journals
DOI: 10.1080/00036840701522895

baseball, respectively) or on home team bias (Dixon and Coles, 1997, Dobson and Goddard, 2003, ch. 8.3, both concerned with soccer).

In this study, while we control for potential longshot and home field bias in the set of odds we use, our focus is different from that of the bulk of the literature. We search for a species of bias that mainstream finance terms 'sentiment'. Just as it is possible that the pattern in rates of return to stocks is influenced by how well known particular companies are, so it may be possible that there are systematically different returns to betting according to whether one wagers on more or less glamorous teams. The mechanism that could generate such a result is explored in Section II. Subsequently, Section III offers an empirical model and presents results from the betting market on Spanish Premier League football, a league selected for examination because it includes an exceptional range of clubs as measured in terms of resources and fan base: such variation should allow sentiment bias to be identified more precisely. In Section IV, we check whether our results are replicated for another league with very 'big' and very 'small' clubs, the Scottish Premier. Section V offers our concluding remarks.

The study of 'sentiment bias' is not only necessary to make knowledge of the efficiency of wagering markets more complete. We argue also that not taking account of sentiment bias when measuring other biases may lead to false inference. For example, studies on American team sports betting, noted above, report negative longshot bias, i.e. inferior returns accrue to betting on short-odds teams. This could be interpreted as evidence of risk-averse bettor preferences. However, without controlling for possible sentiment bias within the setting of a formal statistical test, it could just as easily be a spurious result. On average, short-odds teams are the teams with the greatest playing resources, which are affordable because they have the largest numbers of supporters. If bookmakers exploit fans' preferences as between teams to worsen the terms available on popular clubs, this will deliver lower returns to short-odds bets but the explanation would not in fact be related to bettor preferences over risk. Our multivariate approach will permit separate evaluation of longshot, home field and sentiment biases in the betting market for the world's biggest spectator sport, where there is often strong emotional attachment to particular teams. Such emotional attachment may extend even beyond national borders as fans in one country commonly select clubs in other national leagues whose fortunes they follow simply to make events in the foreign league more interesting.

## II. How Might Sentiment Bias Odds?

In a pari mutuel betting market, the direction in which sentiment would bias odds is clear. Suppose one team in a match has many more supporters than the other. If followers of sport feel a patriotic urge to bet on 'their' club, there will be an 'excessive' proportion of stakes placed on the popular team, driving down the odds to a level which makes betting on the 'bigger' team a more unfair bet than that on the 'smaller' team. To be sure, well informed independents and professional bettors may recognize at this point that backing the smaller team is the financially superior bet but, given that organizers' take-outs are as high as is typical in gambling, they are unlikely to wager sufficiently heavily for the bias to be eliminated (see models of horse and dog pari mutuel markets with informed and noise traders by, for example, Hurley and McDonough, 1995; Terrell and Farmer, 1996; Vaughan Williams and Paton, 1998).

Avery and Chevalier (1999) were amongst the very few in the literature to consider the issue of sentiment bias in sports betting markets. Their application was to NFL betting offered by Las Vegas sports books. Their prior was that losses from backing 'glamorous' teams would be abnormally high. This prediction relied on an assumption that the market behaved as if pari mutuel because bookmakers adjusted spreads so as to balance their books, i.e. their quotes were designed to induce sufficient betting for each team that they held equal liabilities whichever team in a match beat the spread.

In the event, losses in the NFL market were indeed abnormally high for bets on the group of teams labelled 'glamorous' by Avery and Chevalier. But whether they captured the process that resulted in this outcome is questionable. We are advised by industry sources that Las Vegas sports books do not in fact seek to hold a balanced book on each match. We know from investigation of bookmaker records in the much larger illegal sector in the United States that these bookmakers are accustomed to take large positions, i.e. they often stand to lose heavily if a particular team wins (Strumpf, 2003). And British retail bookmakers regularly report big losses or wins for themselves (sufficient to have a significant impact on annual profit) according to whether the England soccer team achieves or fails to achieve victory in an international fixture, i.e. they do not fully adjust odds to deter an excessive inflow of 'sentimental' money.

The reason for Avery and Chevalier's empirical findings, and whether the direction of sentiment bias is necessarily in a particular direction, cannot therefore properly be understood by an approach based essentially on treating the market as if it were pari mutuel rather than organized by bookmakers. Levitt (2004) provides a more realistic framework for analysis. According to Levitt, bookmaker markets are fundamentally different from conventional financial markets in terms of the role of the market maker who would in this case naturally be modelled as maximizing expected winnings (i.e. client losses). Levitt shows that if bettors have a preference for backing options with a particular characteristic (favourites in his example or, in our case, popular teams), then bookmakers will adjust to this feature of demand by increasing the price (i.e. worsening the odds) on those particular bets. But there is a limit on the adjustments to odds that will be feasible if bookmaker 'vigorish' is fixed: if odds on the favourite were shifted too much, this would push odds for the outsider to a point where positive expected returns were available for well informed professional bettors.

Kuypers (2000) also offered a theoretical model but one specific to the context we examine: he was the first to note the possibility that bookmakers may adjust odds to take account of the presence of committed supporters in the betting market for a particular match. His theoretical model points to a profit maximizing bookmaker setting odds in the context of a market where there is a mixture of 'neutral' bettors, who share his own (objectively correct) assessment of the probabilities of each possible outcome of the game, and 'committed' bettors, whose views are coloured by their support for a particular team and are over-optimistic about its winning. This appears a useful framework to represent the soccer betting market though we prefer to think of the 'fan' bettors as capable of assessing win probabilities objectively even though behaving as if they had misperceptions: we interpret their willingness to wager as taking account of extra utility obtained by demonstrating their support for the team in the betting market as well as in the stadium. In part, their bets are motivated by a desire to increase the extent to which they are stakeholders in the club.

In his illustrative numerical example, Kuypers has ten bettors who have decided to wager on a particular match and will decide which bet to place according to their perception of the value offered by each bet (hone win, draw, away win). Four of the bettors are neutrals and six are fans of Manchester United, Britain's most supported team. United's opponent in Kuypers' hypothetical match is Liverpool. He has no committed Liverpool fans in the betting market but this is not damagingly unrealistic since the same results could be obtained from assuming a mix of committed fans whose *average* assessment of a Manchester United victory was over-optimistic. In other words, he is really examining the affect of net 'stakeholder' support in favour of one of the two teams.

Kuypers demonstrates in this context that bookmaker's expected profit is increased by moving from 'efficient' odds to 'inefficient' odds (that implicitly overstate the probability that United will win). This has some intuitive appeal since the bookmaker is able to take advantage of the eagerness of United supporters to back their team by offering them a lower pay-out in the event they win. However, the model, like that of Levitt, is based on very restrictive assumptions. Bookmaker commission is fixed and determines how much money is attracted to wager on a particular game. That money then distributes itself between the possible outcomes according to the relative odds.

The weakness of this approach in the present context is that it may miss the point of including fan preferences in the model in the first place. United fans are likely to see a bet on 'their' team rather than a bet on the match as the product being offered. In this case they may bet in greater or lesser volume as the odds are varied. Kuypers' assumption, that their only response is to switch bets (to Liverpool) at some reservation price, appears as unlikely as a proposition that, if United replica shirts became expensive enough, fans would buy Liverpool shirts instead.

If one makes United fans interested only in betting on their team, but with the decision on how much to bet conditional on value, one could conceivably (depending on the elasticity of demand with respect to expected loss to a unit bet) generate a prediction that odds on a United win will be biased in favour of, rather than against, the bettor. Levitt's caveat still applies, of course: odds cannot be distorted 'too far'. Expected bookmaker profit from bets on United will be maximized where demand for United bets is unit-elastic; but, with a fixed over-round, the bookmaker must also take account of the effect on net revenue of lengthening the odds for Liverpool bettors as odds are made more favourable to United bettors.

It is surprising that Kuypers did not test his hypothesis that odds will vary with net levels of support for teams in a match. We offer a test below by including a measure of supporter numbers in a model constructed for testing market efficiency of odds.

## III. Evidence from Spain

We consider bets available on matches played in the Spanish Premier League in the four seasons from

2001/02 to 2004/05 (season 2005/06 was retained as a hold-out sample). For each match, Internet betting odds on each possible outcome, home win, draw and away win, were listed, for seven different bookmakers, in an archive held at www.football-data.co.uk[2]. However, there were missing observations since not every match had odds listed from all seven operators. We chose to use the odds from Interwetten since this was the firm with the highest coverage. This gave us a sample of 1510 matches or 3020 possible bets on a team winning a match.[3]

None of our results depends on using Interwetten odds rather than those of another of the seven (all internationally known) firms. Correlation between sets of odds for each possible pair of bookmakers was very high, ranging from 0.97 to 0.99.

Implicit transactions costs proved to be high in this market. Overround across the 1510 matches had a mean of 0.138 (with very low variance, SD 0.009). Overround is the sum that a bettor would lose if he wagered on all three outcomes in a match at stakes varying such that he would collect €1 (winnings plus return of stake) from the bookmaker regardless of result. If the bookmaker holds equal liabilities for all results, his earnings relative to total stakes will be given by overround/(1 + overround), so bookmaker commission on stakes here would be 15.8%. An interesting puzzle yet to be resolved is why bookmaker commissions are so much higher in Europe than in Las Vegas NFL betting, notwithstanding the apparently very competitive market. The size of the commissions implies of course that any consistent bias in the odds would have to be very large for awareness of it to permit bettors to enjoy positive expected returns.

Interwetten odds are quoted in the standard European format, known as 'decimal odds'. For example, suppose Sevilla is quoted at 2.20 to win. If a €1 bet were successful, the customer would collect €2.20 after the game (€1.20 profit plus €1 return of stake).

An alternative way of expressing odds is 'probability odds', which is the reciprocal of decimal odds, in our example for Sevilla 0.4545. This equals the level of stake necessary if €1 is to be collected in the event of a successful outcome for the bettor. The sum of the probability odds, of course, exceeds one (by the amount of overround) to permit bookmaker profit. Probability odds themselves are not therefore to be interpreted as probabilities implied by the odds. It is necessary to adjust them by multiplying the probability odds on each outcome of a particular match by a constant such that they do sum to one. These are then the probabilities implied by the set of quotes published by the bookmaker. We term each such probability, 'bookmaker probability' (*BOOKPROB*).[4]

Preliminary analysis of the data indicated some advantage to betting on home teams. Across the 1510 matches, the mean value for *BOOKPROB* in respect of home wins was 0.456. However, the proportion of actual home wins was 0.480[5]. This implies that, over the data period, superior returns were obtained from home bets.

We do not suppose that bookmakers mispriced home bets because they wrongly assessed the value of home advantage. Forrest *et al*. (2005) demonstrated that odds setters were capable of processing into the odds a rich variety of data relevant to the result of a match, with at least as much accuracy as the most sophisticated statistical forecasting model for football that has so far been developed (first presented in Dobson and Goddard, 2001, Ch. 8). Any underpricing of home bets must therefore be viewed as a deliberate commercial decision catering to either the preferences or misperceptions of bettors. But these preferences or misperceptions may not relate to home field advantage *per se*. Because the home team wins more than twice as often as the away team in Spanish matches, home teams are disproportionately often the shorter-odds bet in a match and the apparent bias in odds may in fact represent exploitation of bettors' preferences in the short-odds/long-odds dimension. Multivariate analysis is necessary to identify separately home/away, short-odds/long-odds and indeed sentiment bias.

In our multivariate model, the unit of observation is the individual bet that a team will win. 3020 such bets were available. We estimate a probit model as follows:

$$\text{prob (bet } i \text{ wins)} = f(BOOKPROB_i, HOME_i, DIFFATTEND_i)$$

[2] An access fee is payable.

[3] To give sharper focus to the analysis, we include only home or away win bets, ignoring wagers on a match being drawn. Our principal hypothesis is concerned with the influence of fan betting. The draw has no fans.

[4] Let decimal odds for home win, draw and away win be $d_H$, $d_D$ and $d_L$, respectively. *BOOKPROB*, for a home win for example, is then $(1/d_H)/(1/d_H + 1/d_D + 1/d_L)$.

[5] Away wins occurred in 0.197 of the matches while 0.323 were drawn. The proportions reflect greater home advantage than in, for example, the English Premier League.

where *BOOKPROB* is the probability of a win implied by published odds[6] and *HOME* is a dummy variable set equal to one if the bet relates to a team playing at home. *DIFFATTEND* is our measure to test the influence of sentiment (or net stakeholder support). Numbers of passive fans for clubs are unknown but are taken to be proxied by numbers of active supporters and so *DIFFATTEND* is the difference in mean home attendance in the previous season between the subject team and its opponent. For example, Barcelona and Real Madrid are the two most heavily supported clubs in Spain and bets for these teams will commonly attract a large positive value for *DIFFATTEND*. The largest value of *DIFFATTEND* in our data set was 64.04. The series was calculated from attendance data held at www.european-football-statistics.co.uk.

The 3020 observations correspond to pairs of bets (home win and away win) available across 1510 matches. In any pair of bets, if one wins, the other cannot. In estimation by probit, error terms within each pair of observations would therefore be correlated. We allowed for this by employing clustered probit which applies an appropriate correction to generate robust SEs on the coefficients. Each match comprises a cluster.

The null hypothesis of efficiency stipulates that the odds quoted by the bookmaker should reflect all available information relevant to the outcome of the match. Such information includes which team is playing at home and which team is 'bigger'. The coefficients on *HOME* and *DIFFATTEND* should therefore be zero.

Results in the form of marginal effects (measured from a point where the two continuous variables are set equal to their means and *HOME* is set equal to zero) are displayed as Table 1. Efficiency is rejected. The marginal effect of *BOOKPROB* is significantly less than one (at the 95% level) and this implies negative longshot bias such as has been reported in studies cited above for baseball and ice hockey. *HOME* attracted a significantly positive coefficient so that betting on a home team will (*ceteris paribus*) be advantageous to the bettor. Of course, the shorter odds team is more often than not also the home team, so that these two biases will tend to cancel each other out in the majority of matches.

**Table 1. Clustered probit regression results, Spain**

| | Marginal effect | *t*-statistic |
|---|---|---|
| *BOOKPROB* | 0.649 | 4.75 |
| *HOME* | 0.175 | 5.64 |
| *DIFFATTEND* | 0.0021 | 3.20 |
| Number of observations | 3020 | |
| Number of clusters | 1510 | |
| Pseudo- $R^2$ | 0.115 | |
| Log-likelihood | −1711.1 | |

Our focus is on the variable *DIFFATTEND*. The null hypothesis of lack of bias with respect to size of club is decisively rejected. The effect is signed as positive, indicating that, *ceteris paribus*, more favourable terms are extended to bettors on clubs with larger numbers of supporters. This could be directly the result of 'sentiment bias', with pricing taking account of bettor preferences, or it could result from odds setters giving too little weight to size of club as a determinant of match results.[7] We are minded to the former explanation in view of the strength of the evidence that football odds setters behave as if they can process information at least as well as a sophisticated statistical model (Forrest *et al.*, 2005).

An alternative approach (to clustering) to account for nonindependence of observations within the same match is to estimate the probit model with a sample comprising a single bet, selected home or away by random process, on each of the 1510 matches. Of course, exact estimation results will then vary according to which bets are randomly selected. It is therefore appropriate to execute the process a number of times to check that results are robust. We carried out 50 trials. In every single trial, the estimated probit equation included significantly positive coefficients on *HOME* and *DIFFATTEND* and a coefficient on *BOOKPROB* for which the corresponding marginal effect was significantly below one.[8] The outcomes from this approach were therefore the same as from estimation with clustering.

In respect of sentiment, our findings differ from those of Avery and Chevalier who examined NFL betting and found that less generous spreads were offered for 'glamorous teams'. They also stand in contrast to those of Strumpf (2003) who examined the records of illegal New York bookmakers and found

[6] As is customary in European soccer betting, and in contrast to the Las Vegas market on American sports, odds are available about three days before a game and are generally held fixed until the start of the match. There is therefore no distinction to be made between opening and closing values of *BOOKPROB*.

[7] For example, for two teams with similarly poor recent form, it may be more probable that it will be the better supported- and therefore likely better resourced club- that will recover by winning its next game because reversion to mean effects would apply in its case. It is conceivable that odds setters focus on form and not on this mean reversion effect.

[8] All references to significance relate to the 5% level.

that terms of bets were made harsher (relative to Las Vegas spreads or odds) for clients betting on local New York teams which would have strong supporter following in the neighbourhoods served by these bookmakers[9]. Both sets of American results therefore are consistent with bookmakers penalizing bettors on teams with large numbers of supporters. In the Internet market on Spanish football, the bias also exists but, intriguingly, works in the opposite direction.

In our discussion of the models of Kuypers and Levitt, we showed that the prediction of an unambiguously adverse impact on odds for bettors on 'big' teams was dependent on the assumption that relative odds within a match affected only which bet a client would place, not the decision whether to bet at all. We questioned the realism of this assumption. Those betting for a particular team partly because of emotional attachment may indeed be sensitive to odds; but, if the price goes above their reservation price, they will perhaps abandon the bet rather than back their team's opponent. If this behaviour captures the nature of demand, bookmakers would find it worthwhile to shift odds in favour of supporters of the bigger team if they accounted for a sufficiently large part of the market for that match and if their demand were sufficiently price-elastic. We note that the American markets studied by Avery and Chevalier and Strumpf were monopolistic. All but one Las Vegas sports book employs the same odds setting agency, so that there is limited opportunity to shop around for better odds. In the US illegal sector, bookmakers prefer to take on clients for a long-term relationship since trust between the parties is essential when the service rendered is illegal. Again, bettors lack an option to search for better value across a range of bookmakers. The situation is quite different in the Internet betting market on European soccer (and indeed in the retail betting sector of some European countries, notably Britain and Ireland) where a range of odds from several bookmakers, each with a strong reputation for probity, may be viewed on comparison websites. In this context, it is natural that, if 'committed' bettors comprise a significant part of the market, individual bookmakers' prices for bets on popular teams, should be driven down to levels consistent with superior, rather than inferior, value for money.

Biases in the odds are then a feature of the market on Spanish football. If they are persistent such that they may be observed in one period and then guide the bettor to superior (or even positive) returns in the following period, then the situation would be characterized as violating strong (or even weak) efficiency.

Since 'bigger' teams attracted a price advantage (*ceteris paribus*) over 2001/02 to 2004/05, we examined the returns in the 2005/06 season from a simple strategy of always placing a unit bet where *DIFFATTEND* exceeded some threshold level. Table 2 shows the returns, with the threshold set variously at 10, 20 and 30 (thousands). In each case the result was a loss for the bettor but one that was much less severe than the return of $\sim -16\%$ associated with random betting.[10] This confirms that the phenomenon that superior value attached to bets on 'big' teams was not transient. However, the bias was not sufficiently large to overcome the high overround, so that the market remains weak-efficient even though it is not strong-efficient.

The strategy of simply wagering for 'big' teams does not however take into account all of the information available from a study of seasons 2001/02 to 2004/05, which suggest additional longshot and home field biases. We next employed the results from our clustered probit estimation to generate forecast probabilities for win bets available from Interwetten in season 2005/06. We examined strategies of placing a unit bet whenever the forecast probability exceeded *BOOKPROB* by some threshold amount. So long as this threshold was set sufficiently high, returns were again favourable (relative to the benchmark) or even positive (Table 3). But it should be noted that the number of bets permitted by the application of a strong filter is low. This is primarily because of a tendency for the large sentiment bias favouring bets on the bigger team to be offset in many matches by the negative bias that works against bets with short-odds.

It is striking, even with this qualification, that positive returns could have been earned in 2005/06 from use of our model so long as the threshold for the gap between forecast and bookmaker probabilities was set above 0.09. However, where the number of bets is small, there is a lot of noise in returns from portfolios of bets. And when we looked back at 2001/02 to 2004/05, it was found that a set of bets corresponding to observations where the fitted value exceeded *BOOKPROB* by 0.09 (or 0.10) would have

[9] Strumpf found that bookmakers also price discriminated by further shifting the odds against clients (making telephone bets) who had a record of regularly backing those teams.
[10] Similar (small negative) returns would typically have accrued to each of these strategies in each of the seasons 2001/02 to 2004/05.

**Table 2. Rates of return to strategies based on values of *DIFFATTEND***

| | Number of bets | Rate of return |
|---|---|---|
| *DIFFATTEND* > 10 | 112 | −0.089 |
| *DIFFATTEND* > 20 | 70 | −0.023 |
| *DIFFATTEND* > 30 | 37 | −0.057 |

*Note*: One unit bets in season 2005/06.

**Table 3. Rates of return to strategies based on forecast probabilities**

| | Number of bets | Rate of return |
|---|---|---|
| *gap* > 0.07 | 71 | −0.106 |
| *gap* > 0.08 | 50 | −0.127 |
| *gap* > 0.09 | 28 | +0.077 |
| *gap* > 0.10 | 18 | +0.128 |

*Note*: One unit bets according to whether '*gap*' exceeds a threshold level; *gap* = forecast probability minus bookmaker probability.

**Table 4. Clustered probit regression results, Scotland**

| | Marginal effect | *t*-statistic |
|---|---|---|
| *BOOKPROB* | 0.827 | 3.44 |
| *HOME* | 0.062 | 1.45 |
| *DIFFATTEND* | 0.0049 | 2.56 |
| Number of observations | 1814 | |
| Number of clusters | 907 | |
| Pseudo- $R^2$ | 0.190 | |
| Log-likelihood | −981.9 | |

delivered a negative rather than a positive return. The profitable strategies in the hold-out period could not therefore have been predicted from modelling for the main sample period. Again, while strong-efficiency appears to be violated, weak-efficiency cannot be rejected: we uncover no betting rules that reliably generate positive returns. This is unsurprising given the size of the bookmaker take. Nevertheless it is interesting that biases in odds are large and particularly that sentiment bias works consistently in favour of supporters of the highest profile clubs. Awareness of this would appear to allow independent bettors at least to reduce their losses by (literally) following the crowds.

## IV. A Comparison: Scottish Football

To test whether our findings might be particular to Spanish football, we gathered data, from the same source as before, for 907 matches played in the Scottish Premier League between 2001/02 and 2004/05. This time the odds used were from the bookmaker William Hill, selected because it had the highest number of matches for which odds were quoted in the archive.

Preliminary analysis of the data again showed an apparent benefit to betting on home teams (mean *BOOKPROB* = 0.433, actual frequency = 0.456). However, home field bias proved not to be statistically significant in multivariate analysis.

Table 4 displays results from multivariate analysis. Again we employed a clustered probit model, estimated over 1814 home- or away-win bets. The estimated coefficient on the home dummy is again positive but proves insufficiently strong to be statistically significant. The marginal effect on *BOOKPROB* is again below one but this time one lies within the 95% confidence interval, so an absence of longshot bias cannot be rejected. However, *DIFFATTEND* is again both positive and statistically significant. This is further evidence that bookmakers take into account the size of a club when pricing the bet and that competitive forces are capable of moving odds in favour of large groups of bettors with a nonfinancial preference across the set of wagers on offer.

## V. Concluding Remarks

The article has identified biases in odds available on football matches in an Internet betting market. Findings challenge Kuypers' prediction that less favourable odds will be offered for bets on more heavily supported teams. More generally, they are inconsistent with Levitt's proposition that bookmakers will exploit bettors' preferences adjusting odds such that more popular bets are more highly priced.

The reason for the inconsistency between the theoretical predictions and the empirical evidence is likely to be that the models developed consider bettor behaviour after they have decided to bet on a particular event and assume that bettors may then decide to wager on either side depending on their preferences and on relative odds. Where bettors' preferences are nonfinancial, this is likely not to capture the decision process of bookmaker clients who, if faced with odds which are financially unattractive on the one side and clash with their nonfinancial preferences on the other, may determine not to bet at all or else to seek out bets with rival bookmakers. Competitive forces may then generate more, rather than less, favourable prices for those interested in betting in accordance with non-financial preference. These more favourable prices are revealed by multivariate analysis of odds available on Spanish (and Scottish) football when a proxy is included to account for relative numbers of supporters of the two teams.

Our recommendations for future research are that, in theoretical work, further refinement in the modelling of bookmaker behaviour would be appropriate and that in empirical work allowance should be made for potential sentiment bias. Otherwise, analysis may incorrectly associate apparent violations of efficiency with other specific biases that have been the focus of most of the previous literature on wagering markets.

## References


Avery, C. and Chevalier, J. (1999) Identifying investor sentiment from price paths: the case of football betting, *Journal of Business*, **72**, 493–521.

Cain, M., Law, D. and Peel, D. (2000) The favourite-longshot bias and market efficiency in UK football betting, *Scottish Journal of Political Economy*, **47**, 25–36.

Dixon, M. and Coles, S. (1997) Modelling association football scores and inefficiencies in the UK football betting market, *Journal of the Royal Statistical Society, Series C, Applied Statistics*, **46**, 265–80.

Dobson, S. and Goddard, J. (2001) *The Economics of Football*, Cambridge University Press, Cambridge.

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

Hurley, W. and McDonough, L. (1995) A note on the Hayek hypothesis and the favorite long shot bias in pari mutuel betting, *American Economic Review*, **85**, 949–55.

Kuypers, T. (2000) Information efficiency: an empirical study of a fixed odds betting market, *Applied Economics*, **32**, 1353–63.

Levitt, S. (2004) Why are gambling markets organised so differently from financial markets?, *Economic Journal*, **114**, 223–46.

Sauer, R. (1998) The economics of wagering markets, *Journal of Economic Literature*, **36**, 2021–64.

Strumpf, K. (2003) Illegal sports bookmakers, University of North Carolina at Chapel Hill, Working Paper.

Terrell, D. and Farmer, A. (1996) Optimal betting and efficiency in parimutuel betting markets, *Economic Journal*, **106**, 846–68.

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

Vaughan Williams, L. and Paton, D. (1998) Why are some favourite-longshot biases positive and some negative?, *Applied Economics*, **30**, 1505–10.

Woodland, L. and Woodland, B. (2001) Market efficiency and profitable wagering in the National Hockey League: can bettors score on longshots?, *Southern Economic Journal*, **67**, 983–95.

Woodland, L. and Woodland, B. (2003) The reverse favourite-longshot bias and market efficiency in Major League Baseball: an update, *Bulletin of Economic Research*, **55**, 113–23.