Available online at www.sciencedirect.com

SCIENCE DIRECT®

International Journal of Forecasting 21 (2005) 551–564

international journal of forecasting

www.elsevier.com/locate/ijforecast

Odds-setters as forecasters: The case of English football

David Forrest[a,*], John Goddard[b], Robert Simmons[c]

[a]Centre for the Study of Gambling, University of Salford, Salford, M5 4WT, UK
[b]School of Business and Regional Development, University of Wales, Bangor, Gwynedd, LL57 2DG, UK
[c]The Management School, Lancaster University, Lancaster, LA1 4YX, UK

Abstract

Sets of odds issued by bookmakers may be interpreted as incorporating implicit probabilistic forecasts of sporting events. Employing a sample of nearly 10 000 English football (soccer) games, we compare the effectiveness of forecasts based on published odds and forecasts made using a benchmark statistical model incorporating a large number of quantifiable variables relevant to match outcomes. The experts' views, represented by the published odds, are shown to be increasingly effective over a 5-year period. Bootstraps performed on the statistical model fail to outperform the expert judges. The trend towards odds-setters displaying greater expertise as forecasters coincided with a period during which intensifying competition is likely to have increased the financial penalties for bookmakers of imprecise odds-setting. In the context of a financially pressured environment, the main findings of this paper challenge the consensus that subjective forecasting by experts will normally be inferior to forecasts from statistical models.

Keywords: Football; Odds; Ordered probit; Comparative forecasting—causal, judgement; Bootstrap-evaluation

1. Introduction

In a paper in the International Journal of Forecasting, Forrest and Simmons (2000) examined the behaviour and performance of newspaper football (soccer) tipsters who pronounced on English league matches. When the experts' performances were measured against that of a simple statistical model employing only very obvious team strength and form indicators, they were found wanting. A series of likelihood-ratio tests implied that the tipsters failed properly to process the information used by the statistical model and, moreover, that only one of them appeared to use successfully (unspecified) information extraneous to the statistical model. The authors concluded that experience in the realm of football forecasting was consistent with the consensus in the general literature on expert forecasts: "in nearly all cases where the data can be quantified, the predictions of the (statistical) models are superior to those of the expert" (Makridakis, Wheelwright, & Hyndman, 1998, p.492).

* Corresponding author. Tel.: +44 (0) 161 295 3674; fax: + 44 (0) 161 295 2130.
E-mail address: d.k.forrest@salford.ac.uk (D. Forrest).

0169-2070/$ - see front matter © 2005 International Institute of Forecasters. Published by Elsevier B.V. All rights reserved.
doi:10.1016/j.ijforecast.2005.03.003

It is possible, however, that the consensus view on the role of judgmental forecasting should be qualified. A reasonable hypothesis to test might propose that accuracy in judgmental forecasting will improve as the payoff to accuracy increases. Perhaps, with minds concentrated by heavy financial consequences from wrong forecasting decisions, the gap between the performance of experts and statistical models would narrow or even disappear.

Professional sport, and football in particular, again provides a convenient forum for testing this hypothesis. The newspaper columnists included in the tipsters study are not the only professionals engaged in assessing prospects in advance of British games. Bookmakers publish odds for each possible result of a match, which may be represented as being derived from subjective probabilistic forecasts of match outcomes. In contrast to newspapers, for which the weekly football column is only a small part of the total product, bookmakers have an obvious and very strong interest in accurate probabilistic forecasting.

Betting volumes in British football are large and growing. Turnover with British bookmakers was claimed by Global Betting and Gaming Consultants (2001) to be close to £2bn in 1998. Mintel Intelligence Report (2001) saw it as the fastest growing sector in British gambling. But the importance to bookmakers of expertise in football odds-setting does not rest only on the high volume of business. Also relevant is the unusual way in which the particular betting market is organised. In conventional bookmaker betting (e.g. horse racing worldwide or team sports betting in Las Vegas casino operations), odds shift during the betting period in response to the weight of bettors' money. If a particular price is overly generous, for example, heavy wagers are likely to be made by the public, and bookmakers will defend themselves by shortening the odds offered. The consequences of the initial error in odds-setting are, then, mitigated to the extent that the 'good value bet' is effectively withdrawn from sale at some point in the betting period.

This is not possible in the case of football betting in Britain (and several European countries). In football, the dominant form of betting is 'fixed odds'. Odds for weekend matches, for example, are determined by odds-setters' meetings at each bookmaker on the previous Monday (Sharpe, 1997), and printed on entry forms available by the Wednesday. Odds are therefore announced simultaneously at the various bookmakers. Although the bookmakers retain the right to change odds up to matchday kick-off time, they rarely make adjustments. These odds then remain available for several days, up to the start of the relevant match, regardless of betting volumes and indeed regardless of new information (e.g. on players' injuries or midweek match results). If bets are mispriced, the financial consequences for bookmakers may be serious because they are committed to continue to sell the mispriced bet even though betting volumes may have alerted them to the fact that, at least in the view of the public, odds-setters have made an inappropriate assessment of outcome probabilities.

Good forecasting therefore matters to bookmakers, certainly more than to newspapers that employ tipsters. Further, the pressure on bookmakers to produce good forecasts to be used in odds-setting has increased over time. From about 1999, the year when UK bookmakers set up their own offshore subsidiaries to serve the domestic and overseas betting markets, international competition intensified to the extent that bettors began to have access to a wide range of betting firms, located around the world but quoting odds on English football.

Up to 1999, British bookmakers usually refused to accept bets on single matches ('singles betting') and would only accept bets on combinations of matches, normally three. One of the consequences of the Internet revolution in the gambling industry was that British bookmakers, one by one, abandoned their traditional restriction against betting on single football matches until, in 2003, all such restrictions were removed. During this period the UK Government was induced to remove the 6.75% betting duty that had applied before October 2001. The ending of the restriction preventing betting on single matches is likely to have increased the competitive pressure on bookmakers. Previously, an error that made one particular price look attractive might not necessarily or even usually, present an opportunity for a positive expected return at the bookmaker's expense, because a combination bet across three or more matches had to be placed. The removal of betting tax was likewise a significant development because, by raising transactions costs substantially, the tax had deterred well-informed professional bettors (capable of identifying mispriced bets) from participating in the market.

Given the substantial and growing importance to bookmakers of accurate assessment of the prospects in each football game played, it is unsurprising that expertise attracts high rewards in the relevant labour market, with odds-setters' positions regularly advertised in the betting press at salaries approximately double those of full professors of statistics in British universities. Bookmakers willing to pay such high salaries might anticipate that the level of expertise procured should be capable of producing subjective probabilistic forecasts at least as good as those generated by available statistical models. Comparing forecasts implicit in match odds with forecasts from an information-rich statistical model therefore offers a more plausible route for drawing general conclusions about the potential of subjective forecasting than exploring the performance of newspaper tipsters, where the level of accountability is likely to be low and the incentive to invest in high-level talent small.

Therefore, to summarise, the English football betting market is a particularly interesting and useful vehicle to use to compare expert and statistical forecasting systems. This market has been the subject of considerable structural change caused by the growth of Internet betting alternatives to traditional bookmakers. As a result, competitive pressure on odds-setters has increased. This market differs from many other sports betting markets in that odds are fixed, with consequent incentives for odds-makers to generate accurate forecasts of match outcomes. A further difference from American sports betting markets is the existence of three outcomes (home win, away win and draw) rather than two outcomes as in Major League Baseball. Therefore, we need to apply an ordered probit model rather than a standard logit or probit model.

The remainder of the paper proceeds as follows. In Section 2 we describe a data-rich benchmark statistical forecasting model, against which odds-setters' performance will be assessed. The benchmark model uses as inputs very detailed information on the teams contesting each individual football match. A different estimated version of the benchmark model is used to produce probabilistic forecasts for matches played during each of the five seasons for which comparisons between the model's and the odds-setters' forecasting performance are drawn. Each version of the benchmark model is estimated using data from all English Premier League and Football League matches played during the preceding 15 seasons.

In Section 3, we report the results of a series of likelihood-ratio tests which compare the accuracy of the forecasts obtained from the benchmark model and the corresponding forecasts implicit in bookmakers' odds in anticipating the results of nearly 10 000 matches played between 1998 and 2003. We ask whether odds-setters' decisions capture accurately all the information in the benchmark model and whether they appear to benefit from extra information not utilised by the benchmark model. In this way, we test the efficacy of bookmaker odds as a forecasting tool separately for each of five betting firms covered by our electronic archive of betting odds (www.mabels-tables.com). Of the five firms, Coral, Hill, Ladbroke and Stanley are major high-street bookmakers, while SuperSoccer is a specialist agency that supplies odds to most small independent bookmakers in the UK.

In Section 4, we go on to explore an additional though related issue. Boulier and Stekler (2003) compared the performance of the betting market and that of a well-known commentator (from *The New York Times*) in forecasting the outcomes of American football (NFL) matches. They found that spreads in the betting market offered superior guidance to those of the expert. Furthermore, when they estimated a model to 'explain' the forecasts of the expert, the fitted values from that model offered superior forecasts to those of the expert himself. This suggests that, when experts depart from the forecasts suggested by their 'normal' or 'average' method of processing quantitative information, they become less rather than more reliable. By implication, adjustment of forecasts to take account of subjective information may be a bad idea. We test this general proposition in our context, English football (soccer), by comparing the forecasting performance of a model including only information about bookmaker odds with that of a model embodying instead fitted values of bookmaker odds.

Finally, Section 5 summarises our findings concerning the level of expertise demonstrated by odds-setters according to our various criteria, and how this may have changed over time.

Throughout the paper, we assess odds-setters' expertise by the results from models that include odds as an input. We do not draw *direct* comparisons between the probabilistic forecasts of the benchmark statistical model, published odds and match results because published odds are not intended as forecasts. They are prices that are set with commercial objectives in view, and may therefore contain biases that cater to bettor preferences. In many betting markets, positive or negative longshot bias has been identified. For example, Woodland and Woodland (1994, 2001, 2003) have found that odds in baseball and (ice) hockey betting markets are more 'generous' in terms of expected returns in respect of teams with lower win-probabilities. The opposite result is well documented for racetrack (horse and greyhound) betting markets (Vaughan Williams, 1999). Although, in preliminary analysis for this paper, we found scant evidence for longshot bias in the football betting market, we nevertheless allowed for any bias there may be by using (as the basis for our evaluation of their performance) an ordered probit model that generated 'bookmaker forecasts' from information on raw odds. This avoids falsely attributing to bad forecasting any tendency of odds to systematically over- or under-state the chances of the favoured team in a match.

2. A benchmark statistical forecasting model

A number of authors have attempted to model the process by which the results of football matches are determined, notwithstanding that predictions for single matches are notoriously unreliable (Norman, 1998). A strand in the applied statistics literature, derived from Maher (1982), proposes that the scores of the two teams contesting any match may be modelled using independent Poisson distributions, with means reflecting both the goal-scoring record of the team and the goal-conceding record of its opponent. The model was operationalised by Dixon and Coles (1997) and Rue and Salvesen (2000) in attempts to identify profitable betting strategies for football. Dixon and Pope (2004) compare probabilistic forecasts obtained from the Dixon–Coles model with probabilities inferred from UK bookmakers' prices for fixed-odds betting. Other contributors, including Dobson and Goddard (2001), Forrest and Simmons (2000), Goddard and Asimakopoulos (2004) and Kuypers (1999), have adopted the less computationally demanding methodology of ordered probit or logit. These techniques model home win–draw–away win match results directly, rather than indirectly according to the probability distributions of the scores by each team. Regressors typically include measures of team strength and form, such as the difference in league position and the outcomes of recent matches for each team.

Goddard (2005) has compared the performance in forecasting home–draw–away match results of models in which the dependent variable is the number of goals scored and conceded by each team; and models with a discrete (home–draw–away) match result dependent variable. The difference between the forecasting performance of goals- and results-based models is found to be relatively small. Here, we use as our benchmark forecasting model the most parsimonious specification considered by Goddard (2005): an ordered probit model including a match results dependent variable and covariates based on lagged match results and other relevant information. The model produces forecasts based solely on historical information that is publicly available before the start of the match in question. Given the large number of matches (typically sixty) for which odds-setters must make decisions at a single sitting, it is reasonable to suppose that their published odds will be influenced, either explicitly or implicitly, by publicly-available indicators. A fair test of their expertise is therefore whether they process such information more or less efficiently than the benchmark statistical model.

The notation for the benchmark forecasting model is as follows. $y_{i,j}$ is the result of the match between home team i and away team j, coded as shown below. The latent variable $y^*_{i,j}$ is a linear function of a set of covariates relevant for forecasting match results. The covariates are described briefly below, and are defined in full in Appendix A. $\varepsilon_{i,j}$ is a random disturbance term, assumed to follow the standard Normal distribution. μ_{1s} and μ_{2s}

are the cut-off parameters, which control the overall proportions of home wins, draws and away wins in season s. The structure of the benchmark model is ordered probit:

$$\begin{array}{llll} \text{Home win} & \Rightarrow & y_{i,j}=1 & \text{if} \quad \mu_{2s}<y^*_{i,j}+\varepsilon_{i,j} \\ \text{Draw} & \Rightarrow & y_{i,j}=0.5 & \text{if} \quad \mu_{1s}<y^*_{i,j}+\varepsilon_{i,j}<\mu_{2s} \\ \text{Away win} & \Rightarrow & y_{i,j}=0 & \text{if} \quad y^*_{i,j}+\varepsilon_{i,j}<\mu_{1s} \end{array} \tag{1}$$

Having obtained an estimated version of Eq. (1) over some specific sample period, out-of-sample fitted match result probabilities for the three possible match result outcomes are obtained by rearranging Eq. (1) as follows:

$$\begin{array}{l} \text{Home win probability} = p^{\mathrm{H}}_{i,j} = \mathrm{prob}\left(\varepsilon_{i,j}>\mu_{2s}-y^*_{i,j}\right) = 1-\Phi\left(\mu_{2s}-y^*_{i,j}\right) \\ \text{Draw probability} = p^{D}_{i,j} = \mathrm{prob}\left(\mu_{1s}-y^*_{i,j}<\varepsilon_{i,j}<\mu_{2s}-\hat{y}^*_{i,j}\right) = \Phi\left(\mu_{2s}-y^*_{i,j}\right)-\Phi\left(\mu_{1s}-y^*_{i,j}\right) \\ \text{Away win probability} = p^{\mathrm{A}}_{i,j} = prob\left(\varepsilon_{i,j}<\mu_{1s}-y^*_{i,j}\right) = \Phi\left(\mu_{1s}-y^*_{i,j}\right) \end{array} \tag{2}$$

In Eq. (2), $y^*_{i,j}$ is the fitted value of the latent variable for the match in question, and μ_{1s} and μ_{2s} are the estimated values of these parameters for the final (most recent) season in the estimation period.

In order to generate predictions (in the form of match result probabilities) for each of the five seasons 1998–9 to 2002–3 inclusive, the benchmark model is estimated using data for the preceding 15 seasons in each case. Accordingly, forecasts for season 1998–9 are generated from a version of the model estimated using data for seasons 1983–4 to 1997–8 inclusive; forecasts for 1999–2000 are generated from the model estimated over 1984–5 to 1998–9; and so on. Preliminary experimentation indicated that extending the estimation period up to about 15 seasons produced tangible benefits in terms of improved forecasting accuracy, but there was little or no further gain beyond 15 seasons.

The estimation results for the final version of the benchmark model, estimated over seasons 1987–8 to 2001–2, are reported in Table 1. Overall, the results of a Wald test indicate that the 59 covariates are jointly significant (chi-square statistic=1628.6, critical value at 5% level=77.9). Below, we comment briefly on the interpretation of the estimation results for each set of covariates; see also Goddard (2005) and Goddard and Asimakopoulos (2004).

2.1. Win ratios over the previous 24 months

The win ratio variables $\{P^d_{i,y,s}, P^d_{j,y,s}\}$ are the benchmark model's principal team quality indicators. The indexing of these variables partitions each team's recent win ratio data into components from the present season ($y=0$, $s=0$); from the previous season but within 12 months of the current match ($y=0$, $s=1$); from the preceding season and more than 12 months before the current match ($y=1$, $s=1$); and from two seasons ago but within 24 months of the current match ($y=1$, $s=2$). For teams promoted and relegated within the last two seasons, the win ratio data are further partitioned by division ($d=0$ is the same division, $d=+1$ is one division above; $d=-1$ is one division below; and so on).

2.2. Recent match results

The recent match results variables $\{R^{\mathrm{H}}_{i,m}, R^{\mathrm{A}}_{i,n}\}$ incorporate information about home team i's m most recent home results and n most recent away results, for $m\leq 9$ and $n\leq 4$. The equivalent variables for away team j include the m most recent away results and n most recent home results, again for $m\leq 9$ and $n\leq 4$.

2.3. Importance of match for championship, promotion and relegation issues

If a match is important for championship, promotion or relegation issues for one team but unimportant for the other, the match result is likely to be influenced by the difference between the incentives for the two

Table 1
Ordered probit estimation results: estimation period 1987–1988 to 2001–2002 inclusive

1. Win ratios over previous 24 months ($P^{d}_{i,y,s}$, $P^{d}_{j,y,s}$)

	Home team (i)				Away team (j)			
	0–12 months (y=0)		12–24 months (y=1)		0–12 months (y=0)		12–24 months (y=1)	
Matches played	Current season (s=0)	Last season (s=1)	Last season (s=1)	Two seasons ago (s=2)	Current season (s=0)	Last season (s=1)	Last season (s=1)	Two seasons ago (s=2)
Two divisions higher (d=2)				0.362				0.110
				(0.610)				(0.626)
One division higher (d=1)		1.915***	0.828***	0.493**		−1.470***	−0.776***	−0.625***
		(0.256)	(0.213)	(0.206)		(0.254)	(0.212)	(0.205)
Current division (d=0)	1.769***	1.207***	0.659***	0.464***	−1.330***	−0.935***	−0.590***	−0.403***
	(0.153)	(0.139)	(0.131)	(0.131)	(0.150)	(0.137)	(0.130)	(0.130)
One division lower (d=−1)		0.888***	0.470***	0.421***		−0.590***	−0.322***	−0.200*
		(0.124)	(0.117)	(0.109)		(0.125)	(0.118)	(0.109)
Two divisions lower (d=−2)				−0.051				−0.443**
				(0.191)				(0.197)

2. Most recent match results ($R^{H}_{i,m}$, $R^{A}_{i,n}$, $R^{H}_{j,n}$, $R^{A}_{j,m}$)

Number of matches ago (m,n)	1	2	3	4	5	6	7	8	9
Home team (i)									
Home matches	0.012	0.005	0.027***	−0.001	0.003	−0.009	0.007	−0.000	0.009
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Away matches	0.006	0.020**	0.019**	−0.009					
	(0.009)	(0.008)	(0.008)	(0.008)					
Away team (j)									
Home matches	−0.021**	−0.022**	−0.006	−0.015*					
	(0.009)	(0.009)	(0.009)	(0.008)					
Away matches	−0.015*	−0.011	−0.017**	−0.019**	−0.009	0.002	−0.007	−0.009	−0.029***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)

3. Other explanatory variables

	$SIGH_{i,j}$	$SIGA_{i,j}$	CUP_i	CUP_j	$DIST_{i,j}$	$\Delta AP_{i,1}$	$AP_{i,2}$	$\Delta AP_{j,1}$	$AP_{j,2}$
	0.150***	−0.084***	−0.103***	0.062**	0.054***	0.196***	0.148***	−0.182***	−0.169***
	(0.031)	(0.032)	(0.025)	(0.025)	(0.008)	(0.036)	(0.022)	(0.036)	(0.022)

Obs.=29,480; ln(L)=−30,460.9; ***=coefficient significant at 1% level; **=5% level; *=10% level. Standard errors in parentheses.

teams. For the purposes of the estimation, a match is deemed to be important if it is possible (before the match is played) for the team in question to win the championship or be promoted or relegated, if all other teams currently in contention for the same outcome take one point on average from each of their remaining fixtures.

2.4. FA Cup involvement

Early elimination from the FA Cup may have implications for a team's results in subsequent league matches, although the effect could operate in either direction. A team eliminated from the cup may be able to concentrate efforts on the league, suggesting an improvement in league results; or cup elimination may cause a loss of confidence, suggesting a deterioration. The estimated coefficients on CUP_i and CUP_j suggest that the second of these two effects dominates.

2.5. Geographical distance

The covariate $\mathrm{DIST}_{i,j}$, the geographical distance between the home towns of the teams contesting the match, controls for a tendency for home advantage to be less pronounced in matches between teams located close together, and more pronounced in matches between teams from distant cities or towns.

2.6. Attendance relative to league position

The covariates $\mathrm{AP}_{i,k}$ and $\mathrm{AP}_{j,k}$, based on average attendance data relative to league position for $k=1$, 2 seasons prior to the current season, allow for a 'big team' effect on match results: for given values of other controls, large-market teams are more likely (and small-market teams less likely) to win. This effect might reflect the direct influence of the crowd on the match result, or the ability of teams with larger attendances to spend more heavily on acquiring and retaining playing talent.

3. Odds-setters' forecasting performance

In the English football betting market, bookmaker odds are quoted in the form: a-to-b home win; c-to-d draw; and e-to-f away win. If b is staked on a home win, the overall payoffs to the bettor are $+a$ (the bookmaker pays the winnings and returns the stake) if the bet wins, and $-b$ (the bookmaker keeps the stake) if the bet loses. These quoted prices can be converted to the home win, draw and away win 'probabilities': $\theta^{\mathrm{H}}_{i,j}=b/(a+b)$; $\theta^{\mathrm{D}}_{i,j}=d/(c+d)$; $\theta^{\mathrm{A}}_{i,j}=f/(e+f)$. However, the sum $\theta^{\mathrm{H}}_{i,j}+\theta^{\mathrm{D}}_{i,j}+\theta^{\mathrm{A}}_{i,j}$ invariably exceeds one, because the prices contain a margin to cover the bookmaker's costs and profits. Implicit home win, draw and away win probabilities which sum to one are $\theta^{\mathrm{H}}_{i,j}=\theta^{\mathrm{H}}_{i,j}/(\theta^{\mathrm{H}}_{i,j}+\theta^{\mathrm{D}}_{i,j}+\theta^{\mathrm{A}}_{i,j})$, and likewise for $\theta^{\mathrm{D}}_{i,j}$ and $\theta^{\mathrm{A}}_{i,j}$. The bookmaker's 'over-round' is $\lambda_{i,j}=\theta^{\mathrm{H}}_{i,j}+\theta^{\mathrm{D}}_{i,j}+\theta^{\mathrm{A}}_{i,j}-1$, and $\lambda_{i,j}/(1+\lambda_{i,j})$ is the take-out rate if the bookmaker holds equal liabilities in respect of each of the three possible match outcomes.

In Tables 2 and 3, comparisons are drawn between the forecasting performance of the match result probabilities derived from the benchmark model, and the implicit probabilities derived from the bookmakers' odds. These comparisons are based on all matches for which a set of odds and a set of benchmark model probabilities are available. (The model does not provide forecasts for matches involving teams that were admitted to the league within their first two seasons of league membership, because for such teams insufficient lagged match results data are available to define values for all of the benchmark model's covariates. In addition, in the 2002–3 season there was one match that was rescheduled at very short notice, for which bookmakers' odds were not published.)

As a descriptive measure of the accuracy of probability forecasts, Table 2 reports the Brier Score (Boulier & Stekler, 2003; Brier, 1950). For any set of odds-setter's home win probabilities, the Brier Score is

Table 2
Brier Scores for forecasting performance, odds-setters' implicit probabilities and the benchmark model's forecast probabilities

	1998–1999	1999–2000	2000–2001	2001–2002	2002–2003
Home win					
Coral	0.236	0.231	0.233	0.235	0.235
Hill	0.236	0.230	0.232	0.234	0.236
Ladbroke	0.235	0.230	0.232	0.234	0.236
Stanley	0.236	0.230	0.233	0.235	0.235
SuperSoccer	0.236	0.231	0.233	0.235	0.236
Benchmark model	0.234	0.231	0.234	0.234	0.238
Draw					
Coral	0.201	0.199	0.197	0.195	0.196
Hill	0.201	0.198	0.198	0.195	0.196
Ladbroke	0.201	0.199	0.197	0.195	0.196
Stanley	0.201	0.199	0.198	0.195	0.196
SuperSoccer	0.201	0.198	0.198	0.195	0.197
Benchmark model	0.200	0.198	0.197	0.195	0.197
Away win					
Coral	0.189	0.186	0.185	0.185	0.197
Hill	0.189	0.186	0.185	0.184	0.196
Ladbroke	0.189	0.185	0.185	0.184	0.196
Stanley	0.190	0.186	0.185	0.185	0.196
SuperSoccer	0.190	0.186	0.185	0.185	0.197
Benchmark model	0.189	0.185	0.186	0.185	0.198

Table 3
Maximised log-likelihood values and LR tests: various ordered probit regressions

Season	1998–1999	1999–2000	2000–2001	2001–2002	2002–2003
No. of matches	1944	1946	1946	1946	1945
1. Log-likelihood: ordered probit regression of match results on $\phi^{\delta}_{i,j}$					
Coral	−2022.2	−1994.5	−1992.4	−1990.2	−2029.4
Hill	−2021.4	−1990.4	−1989.6	−1984.3	−2029.2
Ladbroke	−2021.4	−1986.8	−1991.1	−1985.1	−2030.6
Stanley	−2023.0	−1990.8	−1992.8	−1987.9	−2029.1
SuperSoccer	−2023.6	−1993.1	−1991.0	−1990.5	−2031.0
2. Log-likelihood: ordered probit regression of match results on $p^{\delta}_{i,j}$					
Benchmark model	−2015.8	−1988.9	−2000.1	−1988.3	−2038.8
3. Log-likelihood: ordered probit regression of match results on $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$					
Coral	−2014.4	−1986.0	−1990.8	−1984.3	−2028.7
Hill	−2014.2	−1984.1	−1988.6	−1981.2	−2028.7
Ladbroke	−2013.9	−1982.1	−1989.6	−1981.8	−2029.9
Stanley	−2014.6	−1984.4	−1991.2	−1983.3	−2028.6
SuperSoccer	−2014.8	−1985.3	−1989.6	−1984.5	−2029.9
4. LR test for significance of $\phi^{\delta}_{i,j}$ in regression of match results on $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$					
Coral	2.80*	5.80**	18.68***	7.88***	20.16***
Hill	3.30*	9.50***	22.94***	14.20***	20.26***
Ladbroke	3.90**	13.50***	21.06***	12.94***	17.88***
Stanley	2.48	8.94***	17.90***	9.98***	20.44***
SuperSoccer	2.14	7.14***	20.96***	7.62***	17.78***
5. LR test for significance of $p^{\delta}_{i,j}$ in regression of match results on $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$					
Coral	15.56***	16.96***	3.32*	11.78***	1.36
Hill	14.50***	12.54***	1.96	6.30**	1.06
Ladbroke	15.08***	9.36***	3.10*	6.50**	1.44
Stanley	16.80***	12.88***	3.34*	9.18***	0.96
SuperSoccer	17.66***	15.60***	2.70	12.06***	2.24
6. Log-likelihood: ordered probit regression of match results on $\hat{\phi}^{\delta}_{i,j}$					
Coral—fitted	−2021.2	−1991.1	−2003.4	−1990.8	−2036.0
Hill—fitted	−2020.8	−1989.9	−2002.9	−1990.7	−2035.1
Ladbroke—fitted	−2021.9	−1990.5	−2002.6	−1991.6	−2035.2
Stanley—fitted	−2021.2	−1990.5	−2002.7	−1991.4	−2034.8
SuperSoccer—fitted	−2022.0	−1991.2	−2003.0	−1991.1	−2035.2

***=significant at 1% level; **=5% level; *=10% level.

$QR=\sum(H_{i,j}-\phi^{H}_{i,j})^2/N$, where $H_{i,j}=1$ if the match between teams i and j resulted in a home win and 0 otherwise, N is the number of matches, and there are equivalent definitions for draws and away wins. QR can also be evaluated for the probabilities obtained from the benchmark model. QR is analogous to the mean square error of a set of probability forecasts. QR always lies within the scale 0 to 1; the smaller QR is within this scale, the more accurate the probability forecasts are.

Panel 1 of Table 3 shows the maximised values of the log-likelihood functions obtained by fitting a set of five ordered probit regressions using $y_{i,j}$ as the dependent variable and $\phi^{\delta}_{i,j}=\phi^{H}_{i,j}+0.5\phi^{D}_{i,j}$ as the sole covariate, where $y_{i,j}$ is the match result and $\phi^{H}_{i,j}$ and $\phi^{D}_{i,j}$ are the implicit home win and draw odds-setters' probabilities, as defined above. The same model is estimated separately for each of the five sets of probabilities. These maximised log-likelihood values provide an alternative basis for drawing comparisons of forecasting accuracy over a specific set of matches, combining the forecast probabilities for all three match result outcomes in a single measure (rather than three separate measures in the case of QR).

The Brier Scores and the maximised log-likelihood values provide a simple basis for comparing the forecasting performance of the odds-setters' probabilities. According to the maximised log-likelihood values, William Hill was the best performer overall (1st out of 5 in two of the five seasons, and 2nd in the other three seasons), and SuperSoccer was the worst performer (2nd once, 4th once and 5th in the other three seasons). However, it is apparent that the variation in the predictability of match results from season to season is much larger than the variation in forecasting performance between the odds-setters. According to the maximised log-likelihood values, all five odds-setters were noticeably less successful in forecasting match results in 1998–9 and 2002–3 than in the other three seasons.

Panel 2 of Table 3 shows the maximised values of the log-likelihood functions obtained by fitting an ordered probit regression using $y_{i,j}$ as the dependent variable and $p^{\delta}_{i,j}=p^{H}_{i,j}+0.5p^{D}_{i,j}$ as the sole covariate, where $y_{i,j}$ is the match result and $p^{H}_{i,j}$ and $p^{D}_{i,j}$ are the benchmark model's home win and draw probabilities, generated in accordance with the procedure described in Section 2. Comparisons between the benchmark model's and the odds-setters' Brier Scores can be obtained from Table 2, and comparisons between the maximised log-likelihood values can be obtained from Panels 1 and 2 of Table 3.

Over the five seasons, these comparisons appear to produce a clear trend: at the start of the period the benchmark model tended to outperform the odds-setters, but by the end of the period the opposite was true. As was also the case for the odds-setters, the benchmark model's forecasting performance was noticeably worse in 1998–9 and 2002–3 than in the other three seasons when, presumably, there was less 'noise' in the pattern of match results.

The Brier Scores permit comparisons between the odds-setters' and benchmark model's probabilities for the match result outcomes (home win, draw, away win) individually. However, on this basis there is little evidence of systematic differences between the forecasting performance of the odds-setters and the model. For home wins, the model outperformed all 5 odds-setters in 2 seasons, and was outperformed by all 5 in 2 seasons. For draws, the model outperformed all 5 odds-setters in 3 seasons, and was outperformed by all 5 in 1 season. For away wins, the model outperformed all 5 odds-setters in 1 season, and was outperformed by all 5 in 3 seasons.

The comparison between the forecasting performance of the odds-setters and the benchmark model varies over time, but it is possible that at any particular time each summary measure (bookmaker probability, $\phi^{\delta}_{i,j}$, and model probability, $p^{\delta}_{i,j}$) contains relevant information that the other does not contain. If so, the forecasting performance of $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$ combined should be superior to that of either individually. In order to investigate whether this is the case, Panel 3 of Table 3 shows the maximised values of the log-likelihood functions obtained by fitting a set of five ordered probit regressions using match outcome $y_{i,j}$ as the dependent variable, and *both* $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$ as covariates. As before, this model is estimated separately for each of the five sets of odds-setters' probabilities. Panels 4 and 5 show the results of likelihood ratio (LR) tests for the individual significance of $\phi^{\delta}_{i,j}$ and $p^{\delta}_{i,j}$, respectively, in the regressions summarised in Panel 3.

Panel 4 assesses the extent to which the odds-setters' probabilities contain useful information that is not incorporated in the benchmark model. For 1998–9, the value added by the odds-setters seems to have been quite marginal, with bookmaker probability, $\phi^{\delta}_{i,j}$, significant at 5% only in the case of one of the odds-setters according to the LR tests. For the other four seasons, however, $\phi^{\delta}_{i,j}$ is significant in all cases, suggesting that the odds-setters probabilities do contain useful information that is not captured by the benchmark model.

Conversely, Panel 5 of Table 3 assesses the extent to which the benchmark model contains useful information that is not also captured by the odds-setters. For 1998–9, 1999–2000 and 2001–2, model probability, $p^{\delta}_{i,j}$, is significant in all cases. However, for 2000–1 $p^{\delta}_{i,j}$ is only borderline significant, and for 2002–3 $p^{\delta}_{i,j}$ is insignificant. Overall these results appear consistent with the notion that relative to the benchmark model, the odds-setters' performance has improved over time. By 2002–3, the odds-setters appear to have been incorporating most of the information that is contained in the benchmark model into their prices.

Within each football season and for each match outcome, in absolute terms the Brier Scores reported in Table 2 for the five bookmakers and for the benchmark model are very similar. This is also true of the maximised log-likelihood values reported in Table 3. It is therefore relevant to assess whether the small absolute differences that do exist are economically important. For example, is the significance of model probability, $p^{\delta}_{i,j}$, in the regression of match results on $p^{\delta}_{i,j}$ and bookmaker probability, $\phi^{\delta}_{i,j}$, in three of the five seasons (Table 3, Panel 5) of any practical importance for a bettor seeking to use the benchmark model to devise a profitable betting strategy? Or are the match result probabilities produced by the benchmark model so similar to those implicit in the bookmakers' odds that we must infer that the model contains little or no useful additional information?

Panel 1 of Table 4 reports the percentage return that would be earned if bets (with identical stakes) were placed indiscriminately on each of the three possible outcomes for every match. These returns are calculated for each of the five bookmakers individually, and using the best available odds (across the five bookmakers) for each match outcome. For each of the bookmakers, negative returns predominantly in the range −10% to −12% merely reflect the magnitude of the bookmakers' over-round. Using the best available odds, this negative percentage return is substantially reduced (but not eliminated): the average return over all five seasons is −6.6%. This reduction in the bookmakers' take-out, achievable solely through arbitrage, appears non-negligible, suggesting that relatively small differences between the Brier Scores and maximised log-likelihood values reflect differences between the bookmakers' prices that are important economically.

Panel 2 of Table 4 reports the percentage return that would be earned if a bettor placed one bet on every match, on the match outcome with the highest expected return according to the benchmark model. The expected returns are calculated by comparing the benchmark model's probabilities with each of the five sets of bookmakers' odds individually, and with the best available odds. In all cases, there are substantial reductions in the bookmakers' take-out. Using the best available odds, the take-out is virtually eliminated: the average return over all five seasons is −0.2%. Again, the reductions in the bookmakers' take-out achievable through arbitrage and by exploiting the benchmark model to select profitable bets, reflect differences between the odds-setters' probabilities and the benchmark model's probabilities that appear to be important economically.

Table 4
Percentage returns from indiscriminate and selective betting strategies

Season	1998–1999	1999–2000	2000–2001	2001–2002	2002–2003
No. of matches	1944	1946	1946	1946	1945
1. Percentage return from all available bets					
Coral	−10.02	−10.79	−11.41	−11.62	−10.38
Hill	−10.01	−10.99	−11.34	−11.96	−11.18
Ladbroke	−10.16	−11.24	−11.55	−12.16	−11.04
Stanley	−9.25	−10.78	−11.07	−11.57	−10.39
SuperSoccer	−9.93	−10.84	−11.07	−11.44	−10.13
Best available odds	−5.31	−6.37	−7.09	−7.65	−6.48
2. Percentage return if bets are placed on the match outcome with the highest expected return, according to the probabilities obtained from the benchmark model					
Coral	−0.94	0.60	−6.35	−4.15	−5.20
Hill	−1.09	−5.62	−8.21	−9.79	−6.27
Ladbroke	−1.52	−4.58	−8.35	−12.14	−8.13
Stanley	1.86	−2.84	−8.76	−4.53	−3.84
SuperSoccer	−1.82	−4.44	−6.15	−8.16	−1.12
Best available odds	2.94	−0.22	−1.77	−2.81	0.82

One possible limitation of the analysis in this section arises from the fact that the same benchmark model is used to generate probabilities for all matches within a single season whereas odds-setters' decisions could in principle be based on versions of their (implicit) model updated much more frequently. For example, benchmark model forecasts for all matches played in the 2002–3 season are generated from the model that is reported in Table 1 (estimated using data for seasons 1987–8 to 2001–2 inclusive). Although 2002–3 data prior to the current match are used to calculate the covariate values that generate the probability forecasts for the current match, these data are not used to update the estimates of the model itself as the 2002–3 season progresses. In principle, however, the information required to do this would be available to bettors during the course of each season. It is therefore relevant to examine whether updating the model in this way might deliver improved forecasting performance.

In order to investigate this issue, Table 5 reports Brier Scores and maximised log likelihood values from ordered probit regressions of match results on benchmark model forecast probabilities, calculated on a similar basis to those reported in Table 2 and Panel 2 of Table 3, for the following:

(i) Benchmark model probabilities for matches played in August to December of season s, obtained from a version estimated using data for seasons s-15 to s-1 (as before);
(ii) Benchmark model probabilities for matches played in January to May of season s, obtained

Table 5
Brier Scores and maximised log-likelihood values for the benchmark model's forecast probabilities: August–December and January–May

	1998–1999	1999–2000	2000–2001	2001–2002	2002–2003
No. matches					
(i)	1025	1002	1011	1057	1044
(ii)/(iii)	919	944	935	888	902
Brier Score, home win					
(i)	0.2389	0.2320	0.2322	0.2371	0.2386
(ii)	0.2288	0.2290	0.2362	0.2301	0.2376
(iii)	0.2286	0.2301	0.2369	0.2300	0.2368
Brier Score, draw					
(i)	0.1991	0.1976	0.1920	0.1957	0.1934
(ii)	0.2015	0.1989	0.2025	0.1943	0.2002
(iii)	0.2016	0.1991	0.2030	0.1942	0.2001
Brier Score, away win					
(i)	0.1851	0.1816	0.1877	0.1905	0.2012
(ii)	0.1939	0.1885	0.1834	0.1790	0.1938
(iii)	0.1936	0.1893	0.1834	0.1791	0.1934
Log-likelihood: ordered probit regression of match results on $p_{i,j}^{\delta}$					
(i)	−1063.4	−1018.2	−1032.5	−1095.5	−1096.7
(ii)	−950.7	−968.4	−966.1	−892.6	−941.7
(iii)	−951.1	−968.7	−966.2	−892.8	−941.7

(i) Denotes matches played in August to December of season s, forecast probabilities obtained from model estimated using data for seasons s-15 to s-1.
(ii) Denotes matches played in January to May of season s, forecast probabilities obtained from model estimated using data for seasons s-15 to s-1.
(iii) Denotes matches played in January to May of season s, forecast probabilities obtained from model estimated using data for January to May of season s-15, seasons s-14 to s-1, and August to December of season s.

from a version estimated using data for seasons s-15 to s-1 (as before);

(iii) Benchmark model probabilities for matches played in January to May of season s, obtained from a version estimated using data for January to May of season s-15, all of seasons s-14 to s-1, and August to December of season s.

To obtain (iii), the benchmark model is updated as if it were re-estimated on 31 December (when roughly half of the season's matches had been completed). The comparison between (ii) and (iii) is the most important feature of Table 5. This comparison suggests however that contrary to the hypothesis articulated above, updating the benchmark model at the season's mid-point has little or no effect on its forecasting performance. According to the Brier Scores, the updated model performs sometimes marginally better, and sometimes marginally worse, than the model estimated with data to the end of the previous season. The maximised log-likelihood values suggest that overall, the performance of the updated model is marginally worse. From inspection of the coefficients of the benchmark model estimated over different sample periods, it is apparent that any systematic variation in the coefficients of this model occurs very slowly, over periods of several years' rather than just a few months' duration. Therefore the benefit gained by updating the model at the season's mid-point is small. Furthermore, updating imposes a small cost, because (iii) requires the estimation of two additional coefficients, with μ_{1s} and μ_{2s} for seasons s-15, and s estimated using data on only half of the matches played in these two seasons. Table 5 suggests that this cost may marginally outweigh the benefit gained by updating the estimated model in this way.

So far, we have found that the performance of the odds-setters improved over our sample period relative to the statistical model. We find evidence of statistically significant differences between the forecast probabilities of the odds-setters and those of the statistical model in three out of five seasons. Our simulations of the returns to indiscriminate and selective betting strategies shown in Table 4 suggest that these differences in forecasting performance do translate into differences in the financial returns from betting that are important economically. Updating the forecasts from the statistical model within seasons does not appear to make any substantial difference to our results.

4. Odds-setters' use of subjective information

In the final stage of the empirical analysis, an attempt is made to distinguish between the contribution to the odds-setters' forecasting performance of the publicly available historical information that is contained in the covariates of the forecasting model, and the contribution of other more subjective information that is not contained in the forecasting model covariates, but which may nevertheless be relevant to an assessment of the match result probabilities. Webby and O'Connor (1996) use the term 'broken leg cue' to refer to unusual information, extra to the normal flow, that may influence a forecaster. This might include information on individual player injuries (literally broken legs in some cases) or impending suspensions, or inferences concerning future match results from subjective analysis of recent team performances (rather than objective analysis of match results). It seems highly likely that odds-setters' prices are influenced by information of this kind. If the subjective information is used effectively, it may contribute positively to the odds-setters' forecasting performance. However, it is also conceivable that the odds-setters' use of 'broken leg cues' adds 'noise' that might have a deleterious effect on their forecasting performance.

Panel 6 of Table 3 shows the maximised values of the log-likelihood functions obtained by fitting a set of six ordered probit regressions using match outcome, $y_{i,j}$, as the dependent variable and predicted bookmaker probability $\hat{\phi}^{\delta}_{i,j}$ as the sole covariate. The predicted bookmaker probabilities, $\phi^{\delta}_{i,j}$, are the fitted values for $\phi^{\delta}_{i,j}$ obtained from five OLS regressions in which the dependent variable is log odds, $\ln\{\phi^{\delta}_{i,j}/(1-\phi^{\delta}_{i,j})\}$, for each of the five sets of odds-setters' implicit probabilities and the independent variables are the full set of forecasting model covariates defined in Appendix A. (For the OLS regressions, the logit transformation is applied because $0 \leq \phi^{\delta}_{i,j} \leq 1$, while the logit is unconstrained). $\phi^{\delta}_{i,j}$ captures the relationship between the publicly available historical information and the odds-setters' implicit probabilities. Each OLS regression can be interpreted as the implicit model employed by the odds-setter to convert historical information into prices. Therefore the comparison between the ordered probit models for match results in which actual and fitted bookmaker probabilities, $\phi^{\delta}_{i,j}$ and $\hat{\phi}^{\delta}_{i,j}$ respectively, are used as the sole covariates provides a measure of the contribution made by the odds-setters' use of subjective information to their forecasting performance.

The comparison between the maximised log-likelihood values in Panels 1 and 6 of Table 3 suggests that at the start of the period, the forecasting performance of the probabilities obtained from the odds-setters' implicit models was superior to that of the probabilities derived from the odds-setters' actual prices. Therefore the odds-setters' use of subjective information appears to have been detrimental to their forecasting performance. This finding is consistent with the results for spread betting on match outcomes in the NFL reported by Boulier and Stekler (2003). By the end of the period, however, the opposite was true: the forecasting performance of the probabilities derived from the odds-setters' actual prices was superior to that of the probabilities derived from their implicit models. Therefore the odds-setters' use of subjective information appears to have had a positive effect on their forecasting performance. Again, these results appear consistent with the notion that the odds-setters' use of information (both objective and subjective) has improved over time.

5. Conclusions

We have assessed the use of information by football odds-setters in British bookmaking firms during five seasons, from 1998–9 to 2002–3. Our key findings are as follows.

a) Early in the period, a data-rich benchmark statistical forecasting model tended to produce better forecasts than probabilities based on published bookmakers' odds; but by the end of the period, the opposite was true.

b) According to a series of likelihood-ratio tests, adding probabilities based on bookmakers' odds improved the forecasting performance of the benchmark model, indicating that odds-setters are privy to, and make effective use of, information not included in the benchmark model. This was a test failed (Forrest and Simmons, 2000) by newspaper tipsters even though their performance was evaluated relative to a much cruder statistical model.
c) In another series of likelihood ratio tests, adding the probability estimates from the benchmark model to a forecasting framework utilising only bookmaker odds was demonstrated to improve forecasting performance, but only for the first three seasons of the study period. By 2001–2 and 2002–3, odds-setters' decisions appeared already to capture fully and accurately the extensive data represented in the benchmark model.
d) We estimated an equation designed to account for the determination of the odds themselves. Early in the period, fitted values from this equation provided more effective guidance on the outcomes of matches than actual odds, indicating that subjective adjustment to results from their implicit models lowered the odds-setters' performance. However, by the end of the study period, the opposite was true and incorporation of subjective judgement appeared to improve performance: odds-setters appear latterly to have learned how properly to assess and use 'broken leg cues'.

The paper on newspaper football tipsters by Forrest and Simmons (2000) claimed to validate the consensual view in forecasting that statistical models outperform experts. In the present paper, a much more detailed benchmark statistical model proves to be far from dominant over the views of a group of experts, whose assessments may be inferred from the odds they set for football matches. Furthermore, the performance of these experts has improved in a number of dimensions through a period when an intensification of competitive pressure in bookmaking has made the consequences of poor forecasting performance increasingly costly. Structural changes in the UK betting market have resulted in improved expert forecasts relative to bootstrapped predictions from a sophisticated statistical model. Our news from the laboratory of sport is that subjective probabilistic forecasting may outperform forecasting based on quantification when the financial stakes become sufficiently high. When money is at risk, forecasters employed by British bookmakers appear to make good use of available information.

Acknowledgement

We acknowledge the helpful comments of an anonymous referee in facilitating improvements on an earlier draft.

Appendix A

$P^{d}_{i,y,s}=p^{d}_{i,y,s}/n_{i,y}$, where $p^{d}_{i,y,s}$ =home team i's total 'points' score, on a scale of 1=win, 0.5=draw, 0=loss in matches played 0–12 months (y=0) or 12–24 months (y=1) before current match; within the current season (s=0), the previous season (s=1) or two seasons ago (s=2); in the team's current division (d=0) or one ($d=\pm 1$) or two ($d=\pm 2$) divisions above or below the current division; and $n_{i,y}$=i's total matches played 0–12 months (y=0) or 12–24 months (y=1) before current match.

$R^{\mathrm{H}}_{i,m}$ Result (1=win, 0.5=draw, 0=loss) of i's mth most recent home match.

$R^{\mathrm{A}}_{i,n}$ Result of i's nth most recent away match.

$\mathrm{SIGH}_{i,j}$ 1 if match has championship, promotion or relegation significance for i but not for away team j; 0 otherwise.

$\mathrm{SIGA}_{i,j}$ 1 if match has significance for j but not for i; 0 otherwise.

CUP_{i} 1 if i is eliminated from the FA Cup; 0 otherwise.

$\mathrm{DIST}_{i,j}$ Natural logarithm of the geographical distance between the grounds of i and j.

$\mathrm{AP}_{i,h}$ Residual for i from a cross-sectional regression of the log of average home attendance on final league position, defined on a scale of 92 for the PL winner to 1 for the bottom team in Division 3 of the Football League (FLD3), h seasons before the present season, for h=1,2.

References

Boulier B., & Stekler H. (2003). Predicting the outcomes of National Football League games. *International Journal of Forecasting*, *19*, 257–270.

Brier G. W. (1950). Verification of weather forecasts expressed in terms of probability. *Monthly Weather Review*, *78*, 1–3.

Dixon M. J., & Coles S. G. (1997). Modelling association football scores and inefficiencies in the UK football betting market. *Journal of the Royal Statistical Society. Series C, Applied Statistics*, *46*, 265–280.

Dixon M. J., & Pope P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, *20*, 697–712.

Dobson S. M., & Goddard J. A. (2001). *The economics of football*. Cambridge: Cambridge University Press.

Forrest D. K., & Simmons R. (2000). Forecasting sport: The behaviour and performance of football tipsters. *International Journal of Forecasting*, *16*, 317–331.

Global Betting and Gaming Consultants (2001). 1st annual review of the global betting and gaming market, 2001. West Bromwich: Global Betting and Gaming Consultants.

Goddard, J. A. (2005). Regression models for forecasting goals and match results in association football. *International Journal of Forecasting 21*, 331–340.

Goddard J. A., & Asimakopoulos I. (2004). Forecasting football results and the efficiency of fixed-odds betting. *Journal of Forecasting*, *23*, 51–66.

Kuypers T. (1999). Information and efficiency: An empirical study of a fixed odds betting market. *Applied Economics*, *32*, 1353–1363.

Maher M. (1982). Modelling association football scores. *Statistica Neerlandica*, *36*, 109–118.

Makridakis S., Wheelwright S., & Hyndman R. (1998). *Forecasting methods and applications* (3rd edn.) New York: Wiley.

Mintel Intelligence Report (2001). Online betting. London: Mintel International Group Ltd.

Norman J. (1998). 'Soccer'. In J. Bennett (Ed.), *Statistics in sport*. London: Arnold.

Rue H., & Salvesen O. (2000). Prediction and retrospective analysis of soccer matches in a league. *Journal of the Royal Statistical Society. Series D, The Statistician*, *49*, 399–418.

Sharpe G. (1997). *Gambling on goals: A century of football betting*. Edinburgh: Mainstream.

Vaughan Williams L. (1999). Information efficiency in betting markets: A survey. *Bulletin of Economic Research*, *51*, 1–30.

Webby R., & O'Connor M. (1996). Judgemental and statistical time series forecasting: A review of the literature. *International Journal of Forecasting*, *12*, 91–118.

Woodland L., & Woodland B. (1994). Market efficiency and the favourite-longshot bias: The baseball betting market. *Journal of Finance*, *49*, 269–279.

Woodland L., & Woodland B. (2001). Market efficiency and profitable wagering in the National Hockey League: Can bettors score on longshots. *Southern Economic Journal*, *67*, 983–995.

Woodland L., & Woodland B. (2003). The reverse favourite-longshot bias and market efficiency in Major League Baseball: An update. *Bulletin of Economic Research*, *55*, 113–123.

David Forrest is Reader in Economics in the University of Salford. His research interests are most heavily concentrated in the gambling and sports areas. Recent outlets include *National Tax Journal*, *Economic Inquiry* and *Journal of Management Mathematics*. He is a member of the Board of Editors of *International Gambling Studies*.

John Goddard is Professor of Financial Economics at the School of Business and Regional Development, University of Wales, Bangor. His research interests include the economics of professional team sports, industrial organisation, and the economics of the banking sector. He has recently published articles in these fields in *Journal of Forecasting*, *European Journal of Operational Research* and *Journal of Banking and Finance*. He is co-author with Stephen Dobson of the monograph *The Economics of Football* (Cambridge University Press, 2001).

Robert Simmons is Senior Lecturer in Economics, Lancaster University Management School. His research interests are in labour economics, economics of sports and the economics of gambling. He has recently published articles in these fields in Oxford Economic Papers, Scottish Journal of Political Economy and Oxford Review of Economic Policy. He is a member of the Board of Editors of both Journal of Sports Economics and International Journal of Sport Finance.