Online bookmakers' odds as forecasts: The case of European soccer leagues

Article *in* International Journal of Forecasting · July 2010

DOI: 10.1016/j.ijforecast.2009.10.005 · Source: RePEc

Marko Robnik-Sikonja
University of Ljubljana


VOLUME 26, NUMBER 3
JULY–SEPTEMBER 2010
ISSN 0169-2070
ELSEVIER
international journal of forecasting
Special Issue: Sports Forecasting
Guest Editors: Herman O. Stekler and
Leighton Vaughan Williams
IIF
International Institute of Forecasters

Available online at www.sciencedirect.com
ScienceDirect
International Journal of Forecasting 26 (2010) 482–488
international journal of forecasting
www.elsevier.com/locate/ijforecast

Online bookmakers' odds as forecasts: The case of European soccer leagues

E. Štrumbelj*, M. Robnik Šikonja

Faculty of Computer and Information Science, Tržaška 25, 1000 Ljubljana, Slovenia

Abstract

In this paper we examine the effectiveness of using bookmaker odds as forecasts by analyzing 10,699 matches from six major European soccer leagues and the corresponding odds from 10 different online bookmakers. We show that the odds from some bookmakers are better forecasts than those of others, and provide empirical evidence that (a) the effectiveness of using bookmaker odds as forecasts has increased over time, and (b) bookmakers offer more effective forecasts for some soccer leagues for than others.

Keywords: Sports forecasting; Brier score; Statistical tests; Soccer; Betting

1. Introduction

Betting on sports and other types of events attracts the attention of casual bettors, pundits and scientists alike. Betting markets are of particular interest to researchers because there are many similarities between wagering in betting markets and trading in financial markets. These markets consist of different bookmakers, who offer odds for the outcomes of uncertain events, and bettors, who decide whether or not to bet on events. The actual value of the bet is known once the uncertainty has been resolved (i.e., the actual outcome of an event is known). The main goal of bookmakers is to make a profit. This drives them to set the odds high enough to be competitive, but low enough that betting on them is not profitable. Therefore, though the posted odds may not reflect the bookmakers' true probabilistic beliefs, they can still be viewed as probabilistic assessments of a sporting event's outcome, or, in other words, as *forecasts*. So, how effective are these forecasts? Are the odds from some bookmakers better forecasts than those from other bookmakers? Are bookmakers' odds equally effective for different leagues, and does this change over time? In this paper we try to answer these and other similar questions by analyzing soccer matches from six major European soccer leagues, together with the fixed-odds offered for the matches by ten different online bookmakers.

* Corresponding author.
E-mail addresses: erik.strumbelj@fri.uni-lj.si (E. Štrumbelj), marko.robnik@fri.uni-lj.si (M.R. Šikonja).

0169-2070/$ - see front matter
doi:10.1016/j.ijforecast.2009.10.005

1.1. Related work

Most related work on sports betting has focussed on the efficiency of betting markets. Earlier work was predominantly about parimutuel horse-race betting. Unlike fixed-odds betting, where a price is fixed at the moment of betting, parimutuel prices are not known until the market closes. Most authors find evidence that the parimutuel horse-racing market is not efficient and provide profitable strategies, for example Bolton and Chapman (1986), Hausch, Ziemba, and Rubinstein (1981), Hausch and Ziemba (1995), and Lo (1995); although some add that the results should be verified on larger data sets to provide more confident results, for example Asch, Malkiel, and Quandt (1984). Hausch and Ziemba (1990), Ali (1998), and Swindler and Shaw (1995) argue that parimutuel markets are indeed weakly efficient. Also relevant is the work of Dixon and Pope (2004), which is a continuation of the work of Pope and Peel (1988). Dixon and Pope focus on the UK fixed-odds soccer betting market data from 3 bookmakers (years 1993–1996) and find several market inefficiencies, including a reverse favorite-longshot bias. Note that we do not focus on market efficiency in this paper. Instead, we focus on treating bookmaker odds as forecasts and investigating their quality. Related work has provided evidence that bookmaker odds are a very good source of match outcome forecasts. For example, Forrest, Goddard, and Simmons (2005) analyze English soccer games and find that bookmakers' forecasts cannot be outperformed by statistical models. Several related forecasting publications analyze and incorporate knowledge other than betting odds into forecasting match outcomes. For example, Scheibehenne and Broderb (2007) and Pachur and Biele (2007) show that even name recognition by laymen offers some information about the outcome of a sports event, though less than expert knowledge or betting odds. Andersson, Edman, and Ekman (2005), Song, Boulier, and Stekler (2007), and Forrest and Simmons (2000) show that both lay and expert predictions are outperformed by statistical models, which are in turn usually worse than bookmaker odds.

1.2. Notation

Let $\mathcal{X}$ and $\mathcal{B}$ be sets of matches and bookmakers, respectively. A regular *soccer match*, $x \in \mathcal{X}$, has three possible *outcomes*. Either the *home* or the *away* team wins, or it ends in a *draw*. The outcome of a match can be described using the triplet $r : \mathcal{X} \to \{0, 1\}^3$, a mapping from the set of all matches to a binary triplet, with the restriction that exactly one of the numbers equals 1 in each triplet. A *bookmaker*, $b \in \mathcal{B}$, may or may not offer odds for a specific match. When odds are offered, they are in the form of a *triplet*, with one number for each possible match outcome. This can be described by $o : \mathcal{X} \times \mathcal{B} \to \{\} \cup (1, \infty)^3$. We use decimal odds. For example, here are the match odds for a match between Chelsea and Manchester United: Chelsea (2.10), Draw (3.2), and Man Utd (3.75). Therefore, if we bet on Chelsea and they win, we get 2.1 times the amount we bet. On the other hand, the odds for betting on Manchester United are higher, which implies that Chelsea is a slight favorite. Let this example be a match x where the odds are offered by bookmaker b. We can write $o(x, b) = \{2.1, 3.2, 3.75\}$; or, if we break down the triplet into the home win, away win, and draw components, $o_H(x, b) = 2.1$, $o_A(x, b) = 3.75$, and $o_D(x, b) = 3.2$. These bookmaker odds can also be viewed as probabilistic forecasts of the match outcome. In simplified form, the odds 2.1 imply that the probability of Chelsea winning is $\frac{1}{2.1} = 0.48$. Similarly, the probability of Manchester United winning is 0.27 and the probability of a draw is 0.31. However, when we sum these probabilities, we get 1.06, which is more than 1.00. The extra 6% is the result of bookmakers lowering the odds in order to ensure a profit, which is known as the *bookmaker margin*. The bookmaker margin of bookmaker b for match x is: $\mathrm{mrg}(x, b) = \frac{1}{o_H(x,b)} + \frac{1}{o_D(x,b)} + \frac{1}{o_A(x,b)} - 1$. We eliminate the margin by normalizing the forecasts so that the probabilities sum up to 1 (i.e., dividing the odds-implied probabilities by the margin). All of the results presented in this paper are for normalized probabilities.

2. Data description

The data used in this paper consist of 10,699 matches across several seasons between the years 2000 and 2006 from the following six major European soccer leagues: English Premier League (*E0*), English Championship (*E1*), Scottish Premiership League (*SC0*), Italian Serie A (*I1*), French Ligue 1 (*F1*), and Spanish La Liga (*SP1*). From this point

Table 1
Number of match odds triplets available for each bookmaker and soccer league.

	b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
E0	1520	760	1988	2088	2044	760	2055	380	380	2096
E1	2202	1102	2884	3015	2953	1086	2990	551	552	3056
SC0	910	455	1194	1246	1143	451	1228	228	228	1257
I1	1241	686	1251	1225	1222	545	1242	349	346	1234
F1	1463	706	1462	1439	1376	118	1464	364	363	1444
SP1	1512	760	1509	1511	1503	744	1513	379	377	1515

Table 2
Mean forecast probabilities, standard deviations, and observed relative frequencies for each of the three possible outcomes across all bookmakers.

		b_0	b_1	b_2	b_3	b_4	b_5	b_6	b_7	b_8	b_9
Home	μ	0.45	0.45	0.45	0.443	0.451	0.449	0.451	0.451	0.448	0.449
	σ	0.135	0.14	0.135	0.124	0.133	0.139	0.135	0.142	0.142	0.134
	Freq.	0.459	0.456	0.459	0.458	0.456	0.456	0.457	0.453	0.452	0.458
Draw	μ	0.271	0.274	0.272	0.277	0.269	0.267	0.271	0.274	0.271	0.272
	σ	0.033	0.036	0.032	0.032	0.029	0.032	0.032	0.035	0.034	0.032
	Freq.	0.272	0.282	0.271	0.27	0.271	0.263	0.27	0.274	0.275	0.269
Away	μ	0.279	0.277	0.278	0.281	0.281	0.284	0.279	0.275	0.281	0.28
	σ	0.122	0.126	0.121	0.109	0.121	0.125	0.122	0.128	0.128	0.121
	Freq.	0.269	0.263	0.27	0.272	0.274	0.281	0.273	0.273	0.273	0.274

on, the abbreviations are used when referring to individual leagues. The data cover match odds from 10 online bookmakers: Bet365 (b_0), Betandwin (b_1), Gamebookers (b_2), Interwetten (b_3), Ladbrokes (b_4), Sporting Odds (b_5), Sportingbet (b_6), Stan James (b_7), VC Bet (b_8), and William Hill (b_9). Fewer matches are available for some bookmakers, either due to missing data or because the bookmaker did not cover a particular league or season. Information about the number of matches available for each bookmaker is shown in Table 1, and the details of the mean forecasts, standard deviations, and observed relative frequencies can be found in Table 2.

3. The forecasting quality of bookmaker odds

We use two approaches to evaluate the effectiveness of using bookmaker odds as forecasts. In the first approach we treat each of the three possible outcomes as a separate binary event and each component of the bookmakers' odds triplet as an individual forecast. To measure the quality of these forecasts we use the Brier score (see Brier, 1950). The Brier score is a proper scoring rule, as was shown by Hendrickson and Buehler (1971), so the expected score is minimized by reporting the true subjective probabilities. In other words, it encourages the forecaster both to make careful assessments and to be honest. Therefore, adjusting the odds in order to either maximize the profit or minimize the liability would result in a worse score. The Brier score for the home win forecast of match x by bookmaker b is the squared difference between the forecast probability and the true outcome:

$$B_H(x,b) = \left(\frac{1}{o_H(x,b)} - r_H(x)\right)^2. \tag{1}$$

The Brier scores for the away win and draw are defined in a similar way. The match in our illustrative example ended with a 1 : 1 draw (i.e., $r(x) = \{0, 1, 0\}$). Therefore, the scores for the individual forecasts in our introductory example are 0.201, 0.501, and 0.065 for the home, draw, and away outcomes, respectively. Normalized probabilities, 0.453, 0.292, 0.255 were used.

In the second approach, we use the *ranked probability score* (RPS) (see Epstein, 1969; and Murphy, 1969) to evaluate the three forecasts for a soccer match as a multi-category forecast. The ranked probability score is the sum of the differences between the cumulative forecast probability and the cumulative outcome probability:

$$\mathrm{RPS}(x, b) = \sum_{i=1}^{3} \left(\sum_{k=1}^{i} \frac{1}{o_k(x, b)} - \sum_{k=1}^{i} r_k(x) \right)^2 . \quad (2)$$

For the purposes of using Eq. (2), we order the three possible categories (i.e., match outcomes) in the following way: *home* = 1, *draw* = 2, and *away* = 3. This follows the natural order of events during a soccer match, where every transition of the game result from home win to away win (and vice versa) passes through a draw result. We could also use the reversed order (i.e., away, draw, home); however, it can easily be verified that reversing the order does not change the RPS score. Note that RPS is also a proper scoring rule and that the Brier score is a special case of RPS, where the number of categories is 2. For our illustrative example, $RPS(x, b) = (0.453 - 0)^2 + ((0.453 + 0.292) - 1)^2 + (1 - 1)^2 = 0.27$.

First, we investigate whether bookmaker odds give better forecasts than "naïvely" forecasting using the observed relative frequencies of the outcomes of soccer matches (see Table 2). A non-parametric test is used for this investigation, as the distributions of the Brier scores and RPS violate the assumption of normality. Normality was tested using the tests proposed by D'Agostino (1986) and Shapiro and Wilk (1965). The tests are known as the D'Agostiono-Pearson and Shapiro-Wilk normality tests, respectively. The p-values were less than 0.0001 for both tests, across all bookmakers and outcomes, with the alternative hypothesis being that the bookmakers' Brier scores are not normally distributed. We use the Wilcoxon rank-sum test (see Wilcoxon, 1945) to compare each bookmaker with the naïve forecasts across all matches covered by that bookmaker, with the alternative hypothesis being that the bookmaker's forecasting scores are better. The samples are matched, so we can use a paired test. The first column of Table 3 shows the results of this test for the *draw* (i.e., using B_D to evaluate the forecasts). The results for the remaining two outcomes and for RPS have been omitted, because the differences are convincingly significant for all bookmakers (p-values < 0.0001). The significance of the differences between the odds-based forecasts and the naïve forecasts is much less (or non-existent) when predicting a draw. Therefore, bookmaker draw odds provide only a small improvement over forecasting based on the relative frequency (relative to the improvement for the remaining two outcomes) of a draw. This is consistent with the findings of Pope and Peel (1988) and Dixon and Pope (2004), who suggest that a draw is the most difficult outcome for experts to forecast.

Mean RPS scores for each bookmaker across all matches covered by that bookmaker can be found in Table 3. The third column contains the relative mean RPS scores, which were obtained using the following procedure: for each bookmaker the matches were ordered according to their date and split into two groups of equal size—older matches and newer matches. RPS was then calculated separately for each group, and the column contains the mean RPS of the newer matches, relative to the RPS of the older matches. These results suggest that the effectiveness of using bookmaker odds as forecasts has improved over time across all bookmakers. Note that this procedure was only performed for bookmakers for whom we had data from at least two full seasons available.

4. Are all bookmakers equally effective sources of forecasts?

In the previous section we established that a non-parametric test should be used to compare the Brier scores (and RPS values) of soccer match forecasts. In his critique of related work, Batchelor (1990) suggests the use of the Friedman test for a comparison of forecasters (see Friedman, 1939). However, this test is designed for complete block designs (i.e., the sample Brier scores should be from the same matches), so only matches where odds are available from all bookmakers should be used. This would greatly reduce the number of matches we have available. Therefore, for the purpose of this test, bookmakers b_1, b_5, b_7 and b_8, for whom the fewest odds are available, are omitted. With this omission we have 1504 (*E0*), 2146 (*E1*), 891 (*SC0*), 1191 (*I1*), 1344 (*F1*), and 1490 (*SP1*) matches available with complete match odds for all remaining bookmakers. We use the Friedman test separately on these matches for each outcome and league, to check whether some bookmakers forecast that outcome for that league better than others. The

Table 3
The first column holds the p-values for the hypothesis that draw forecasts ($\overline{B_D}$) are just as effective as forecasting using the relative frequency of a draw. The remaining two columns contain the mean RPS scores and the mean RPS scores of newer matches relative to older matches, for each bookmaker.

Bookmaker	$\overline{B_D}$	$\overline{RPS}$	(Newer matches/older matches)
b_0	< 0.001	0.402	0.959
b_1	< 0.001	0.3971	0.991
b_2	0.001	0.4038	0.958
b_3	0.015	0.4066	0.941
b_4	< 0.001	0.4064	0.953
b_5	0.124	0.4085	–
b_6	< 0.001	0.4051	0.944
b_7	0.002	0.3968	–
b_8	< 0.001	0.3962	–
b_9	0.191	0.4055	0.955

differences in forecasting quality are significant, both for individual outcomes and for all three outcomes combined (i.e., when RPS is used to evaluate the forecasts). The p-values are less than 0.005 for all outcome/league pairs, with the exception of $\overline{B_H}/F1$, where $p = 0.14$.

Note that the Kruskal-Wallis test (see Kruskal & Wallis, 1952) could be used as an alternative to the Friedman test. The Kruskal-Wallis test, which can be applied to a non-block design (thus enabling us to use a larger sample of matches), is the non-parametric equivalent of a one-way analysis of the variance. The Kruskal-Wallis test does not assume that populations follow normal distributions, but it does assume that the distributions are all identical shapes. The null-hypothesis is that the positions of all of the distributions are the same. The shapes are very similar across all bookmakers for each league and match outcome. However, the pairing of bookmaker odds (i.e., the complete block design) is efficient in controlling for the variability between matches, which makes the Friedman test more powerful than the Kruskal-Wallis test. Indeed, the results change if we use the Kruskal-Wallis test instead of the Friedman test. While the differences between bookmakers' forecasts remain significant for draw odds, the significance of the home outcome forecasting differences cannot be confirmed ($p > 0.30$). For the home and away odds, in most cases we can accept the null-hypothesis that the forecasts are not different ($p > 0.9$).

5. Forecasting quality across different leagues

For each bookmaker we have the mean Brier scores and RPS for each league, so we use the Friedman test to check whether the odds for certain leagues are better forecasts than those for other leagues (i.e., whether the mean RPS or Brier scores are significantly different). Each outcome was tested separately, as was the combination of the three outcomes. The p-values of the tests for each outcome can be found in Table 4. We also include the mean rankings of the leagues across all bookmakers. The lower the rank, the better the forecasts, relative to the same outcome for other leagues. For example, the mean rank of 1 for the home win in *SC0* means that (as far as home win forecasting is concerned) the *SC0* home odds receive the best score across all bookmakers. The results reveal that the odds from certain leagues are significantly better forecasts than those from other leagues. Using the post-hoc multiple comparisons test proposed by Nemenyi (1963) (at the 95% significance level), we get the critical distance of 2.51. This test is more conservative than the Friedman test, and only the difference in ranks between *SC0* and *F1* is larger than the critical distance, and therefore significant.

The Scottish Premier League (*SC0*) stands out as the league where the odds offered by bookmakers are the best forecasts, and the French Ligue 1 (*F1*) as the league where they are the worst (using RPS to compare the leagues). We hypothesize that this may be due to the large gap in team quality between the

Table 4
p-values for the Friedman test with the null hypothesis that the odds from all leagues are equally good forecasts. Mean ranks across all ten real-world bookmakers are also included. The higher the mean rank, the less effective the odds are, relative to other leagues.

	E0	*E1*	*SC0*	*I1*	*F1*	*SP1*	p-value
$\overline{B_H}$	3.0	4.0	1.0	6.0	2.3	4.7	<0.0001
$\overline{B_D}$	2.6	1.4	5.2	4.0	3.5	4.3	0.0002
$\overline{B_A}$	4.1	4.9	1.2	3.1	4.9	2.8	>0.0001
$\overline{RPS}$	3.9	3.1	1.8	4.2	5.2	2.8	0.0011

top two Scottish clubs (Rangers and Celtic) and the remaining ten teams in SC0, meaning that there are more matches which are easier to predict. On the other hand, the *F1* teams are more evenly matched, meaning that the match outcomes are, on average, more difficult to predict. To provide some empirical evidence in support of this hypothesis, we obtained final league standings and point totals for each team and for every season covered by our data. We use the relative standard deviation (RSD) of the point totals as a measure of how evenly matched the teams within each league and season were. For the seasons of *SC0* covered by our data, the RSDs range from 0.39 to 0.44 (with a mean of 0.41). On the other hand, for *F1* the RSDs range from 0.23 to 0.27 (with a mean of 0.24). The mean RSDs of the point totals for other leagues are lower than that of *SC0* and higher than that of *F1*, which is consistent with our hypothesis. If the top two teams in *SC0* (Rangers and Celtic) and all of their matches are removed, then the relative standard deviations of the point totals range from 0.23 to 0.26. This shows that the remaining teams are more evenly matched, and, consistent with our hypothesis, the rank of *SC0* using RPS rises from 1.8 to 5.2, because the odds for the remaining *SC0* teams are worse forecasts. However, even with the omission of these two clubs, the difference between leagues remains significant ($p = 0.011$).

In their study of sentiment bias in Spanish soccer, Forrest and Simmons (2008) also investigated the Scottish Premier League in a set of matches from a time period which partially coincides with the time period of our data. The authors present empirical evidence that *bookmakers take the size of a club into account when pricing the bet, and that competitive forces are capable of moving the odds in favour of large groups of bettors*. For the Scottish Premier League the authors do not find enough evidence that this bias can be exploited in a profitable betting strategy, but the bias is both large and consistent, thus improving the "fairness" and quality of the odds on larger teams. It is safe to assume that a larger club size implies better resources and a greater team strength. Therefore, the findings of Forrest and Simmons can be interpreted as showing that bookmakers offer more accurate odds on better teams. This is consistent with our own findings that the bookmaker odds for the Scottish Premier League can be used effectively as forecasts, because of the large differences in quality between the teams.

6. Conclusion

In this paper we have investigated the use of online bookmaker odds as probabilistic forecasts of soccer match outcomes. Based on the results, we conclude that not all bookmakers' odds are equally good for forecasting soccer match outcomes. This holds across all of the leagues examined, and for each outcome separately and the three outcomes combined. Furthermore, we have shown that the offered odds are significantly better forecasts for some leagues than others, and that the effectiveness of using odds as forecasts has increased over time. As part of future work it would be worth investigating whether more effective bookmaker odds can be used to produce a profitable betting strategy or whether a consensus among bookmakers provides additional information. It would also be interesting to further investigate the relationships between how evenly matched the teams are, the effectiveness of the forecasts, and the sentiment bias (see Forrest & Simmons, 2008). In particular, it would be interesting to investigate whether a sentiment bias exists in the French Ligue 1, where we found the teams to be the most evenly matched and where using bookmaker odds as forecasts was the least effective.

References

Ali, M. M. (1998). Probability models on horse-race outcomes. *Journal of Applied Statistics*, *25*(2), 221–229.

Andersson, P., Edman, J., & Ekman, M. (2005). Predicting the World Cup 2002 in soccer: Performance and confidence of experts and non-experts. *International Journal of Forecasting*, *21*(3), 565–576.

Asch, P., Malkiel, B. G., & Quandt, R. E. (1984). Market efficiency in racetrack betting. *The Journal of Business*, *57*(2), 165–175.

Batchelor, R. A. (1990). All forecasters are equal. *Journal of Business and Economic Statistics*, *8*, 143–144.

Bolton, R. N., & Chapman, R. G. (1986). Searching for positive returns at the track: A multinomial logic model for handicapping horse races. *Management Science*, *32*(8), 1040–1060.

Brier, G. W. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *75*, 1–3.

D'Agostino, R. B. (1986). Tests for normal distribution. In R. B. D'Agostino, & M. A. Stepenes (Eds.), *Goodness-of-fit techniques*. Macel Decker.

Dixon, M. J., & Pope, P. F. (2004). The value of statistical forecasts in the UK association football betting market. *International Journal of Forecasting*, *20*, 697–711.

Epstein, E. S. (1969). A scoring system for probability forecasts of ranked categories. *Journal of Applied Meteorology*, *8*, 985–987.

Forrest, D., Goddard, J., & Simmons, R. (2005). Odds-setters as forecasters: The case of English football. *International Journal of Forecasting*, *21*, 551–564.

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

Forrest, D., & Simmons, R. (2008). Sentiment in the betting market on Spanish football. *Applied Economics*, *40*, 119–126.

Friedman, M. (1939). A correction: The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, *34*, 109.

Hausch, D., Ziemba, W., & Rubinstein, M. (1981). Efficiency of the market for racetrack betting. *Management Science*, *27*(12), 1435–1452.

Hausch, D. B., & Ziemba, W. T. (1990). Strategies for cross-track betting on major horse races. *Journal of Business*, *61*, 61–78.

Hausch, D. B., & Ziemba, W. T. (1995). Efficiency of sports and lottery betting markets. In *Handbooks in operations research and management science, Vol. 9* (Ch. 18).

Hendrickson, A. D., & Buehler, R. J. (1971). Proper scores for probability forecasters. *The Annals of Mathematical Statistics*, *42*(6), 1916–1921.

Kruskal, W. H., & Wallis, W. A. (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, *47*, 583–621.

Lo, S. Y. (1995). The application of ranking probability models to racetrack betting. *Management Science*, *41*(6), 1048–1059.

Murphy, A. H. (1969). On the ranked probability skill score. *Journal of Applied Meteorology*, *8*, 988–989.

Nemenyi, P. B. (1963). *Distribution-free multiple comparisons*. Ph.D. thesis, Princeton University.

Pachur, T., & Biele, G. (2007). Forecasting from ignorance: The use and usefulness of recognition in lay predictions of sports events. *Acta Psychologica*, *125*, 99–116.

Pope, P. F., & Peel, D. A. (1988). Information, prices and efficiency in a fixed-odds betting market. *Economica*, *56*, 323–341.

Scheibehenne, B., & Broderb, A. (2007). Predicting Wimbledon 2005 tennis results by mere player name recognition. *International Journal of Forecasting*, *23*(3), 415–426.

Shapiro, S. S., & Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, *3*, 591–611.

Song, C., Boulier, B. L., & Stekler, H. O. (2007). The comparative accuracy of judgmental and model forecasts of American football games. *International Journal of Forecasting*, *23*(3), 405–413.

Swindler, S., & Shaw, R. (1995). Racetrack wagering and the uninformed bettor: A study of market efficiency. *The Quarterly Review of Economics and Finance*, *35*(3), 305–314.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, *1*(6), 80–83.

Erik Štrumbelj recieved his B.Sc. in computer science from the University of Ljubljana, Slovenia, in 2007. He is currently a Ph.D. student and a Junior Researcher at the University of Ljubljana, Faculty of Computer and Information Science. His research interests include machine learning, data mining, statistics, and forecasting.

Marko Robnik Šikonja recieved his Ph.D. in computer science from the University of Ljubljana, Slovenia, in 2001. He is currently an Assistant Professor at the University of Ljubljana, Faculty of Computer and Information Science. His research interests include machine learning, data mining, knowledge discovery, cognitive modeling, and practical applications. He is the (co)author of more than 40 journal and international conference publications.