Sports Forecasting: A Comparison of the Forecast Accuracy of Prediction Markets, Betting Odds and Tipsters

MARTIN SPANN[1]* AND BERND SKIERA[2]
[1] *School of Business and Economics, University of Passau, Germany*
[2] *School of Business and Economics, Johann Wolfgang Goethe-University, Frankfurt am Main, Germany*

ABSTRACT

This article compares the forecast accuracy of different methods, namely prediction markets, tipsters and betting odds, and assesses the ability of prediction markets and tipsters to generate profits systematically in a betting market. We present the results of an empirical study that uses data from 678–837 games of three seasons of the German premier soccer league. Prediction markets and betting odds perform equally well in terms of forecasting accuracy, but both methods strongly outperform tipsters. A weighting-based combination of the forecasts of these methods leads to a slightly higher forecast accuracy, whereas a rule-based combination improves forecast accuracy substantially. However, none of the forecasts leads to systematic monetary gains in betting markets because of the high fees (25%) charged by the state-owned bookmaker in Germany. Lower fees (e.g., approximately 12% or 0%) would provide systematic profits if punters exploited the information from prediction markets and bet only on a selected number of games.   Copyright © 2008 John Wiley & Sons, Ltd.

KEY WORDS   sports forecasting; prediction markets; combined forecasts; rule-based forecasts

INTRODUCTION

Sports forecasting is important for sports fans, team managers, sponsors, the media and the growing number of punters who bet on online platforms (Vlastakis *et al.*, 2007). Widespread demand for professional advice regarding the results of sporting events is met by a variety of expert forecasts, usually in the form of recommendations from tipsters (Forrest and Simmons, 2000). In addition, betting odds offer a type of predictor and source of expert advice regarding sports outcomes. Whereas fixed odds reflect the (expert) predictions of bookmakers (Pope and Peel, 1989), the odds in

*Correspondence to: Martin Spann, School of Business and Economics, University of Passau, Innstr. 27, 94032 Passau, Germany. E-mail: spann@spann.de

pari-mutuel betting markets indicate the combined expectations of all punters, which implies an aggregated expert prediction (Plott *et al.*, 2003).

Prediction markets (PMs), first applied to forecast political election results (Forsythe *et al.*, 1992, 1999) and later business outcomes (Dahan *et al.*, 2006; Elberse, 2007; Elberse and Eliashberg, 2003; Gruca *et al.*, 2003; Jank and Foutz, 2007; Pennock *et al.*, 2001; Spann and Skiera, 2003), increasingly attempt to predict sporting events (Luckner and Weinhardt, 2007; Servan-Schreiber *et al.*, 2004; Wolfers and Zitzewitz, 2006), which suggests they provide an additional method of sports forecasting. In essence, PMs bring a group of participants together via the Internet and let them trade shares of virtual stocks that represent bets on the outcomes of future market situations; the stocks' value depends on the realization of those chosen market situations. When an outcome associated with a specific market situation occurs, each share of virtual stock receives a cash dividend (payoff) (e.g., $1 if the predicted team wins, $0 otherwise). In a PM, each participant contributes his or her knowledge to the market by trading, so the stock prices represent participants' aggregated knowledge and thus the PM's prediction (Hahn and Tetlock, 2006; Spann and Skiera, 2003).

The availability of multiple forecasting methods raises questions about their effective use. Previous studies consider the performance of betting odds and tipsters (for recent summaries, see Forrest *et al.* (2005) or Andersson *et al.* (2005)), respectively betting odds and prediction markets (Servan-Schreiber *et al.*, 2004) for sports forecasting, but knowledge about their comparative performance versus PMs remains scarce, because no studies compare all three forecasting methods (Andersson *et al.*, 2005; Boulier *et al.*, 2006; Chen *et al.*, 2005; Forrest *et al.*, 2005; Goddard and Asimakopoulos, 2004; Paton and Vaughan Williams, 2005). Furthermore, little is known about the potential similarity of forecasts across methods, their performance or their ability to improve forecast accuracy if used in a weighting-based or rule-based combination. However, such knowledge is important because it might allow punters to systematically earn money on those markets. In addition, it provides recommendations for sports and betting companies on how to improve their forecasts.

Therefore, this article empirically compares the forecast accuracy of PMs, tipsters and betting odds, as well as weighting- and rule-based combinations of those forecasts. We present the results of an empirical study that uses data from 837 games across three seasons of the German premier soccer league. In consideration of the vast sums of money at stake in betting markets, we also determine whether the forecasts of the three methods or their combinations enable systematic profits. Thus, we contribute to the sports forecasting literature by providing the first large-scale empirical study of the three forecasting methods and their combinations that compares their forecasting accuracy and ability to enable profits for punters in betting markets.

The remainder of this article is structured as follows. In the next section we describe the three forecasting methods, then use the third section to describe the data, the performance measures and the calculations required for the three expert forecasting methods. In the fourth section we compare the forecast accuracy of the three forecasting methods, as well as of their combinations. The final section concludes our paper.

DESCRIPTION OF THREE FORECASTING METHODS

Prediction markets

The fundamental concept behind PMs suggests that markets can solve information problems (Hayek, 1945). Related to this, the efficient market hypothesis posits that prices always reflect all available information (Fama, 1970). A competitive market achieves market efficiency through the price

mechanism, the most efficient instrument for aggregating asymmetrically dispersed information possessed by market participants (Hayek, 1945; Smith, 1982). Therefore, prices in a competitive market offer an aggregate reflection of all participants' public and private information and thus serve as a good predictor (Spann and Skiera, 2003). As a result, markets possess the positive characteristics of information elicitation and aggregation, immediate reaction to new information and scalability with respect to the number of participants (Dahan and Hauser, 2002; Oliven and Rietz, 2004). These characteristics make them a potentially promising method for solving information problems (Spann *et al.*, 2009; Tziralis and Tatsiopoulos, 2007).

In the various PMs for sporting events (e.g., sports.us.newsfutures.com, www.tradesports.com, www.wsex.com), participants trade virtual stocks related to future market situations, namely, the outcomes of sporting events. The cash dividend (payoff) of these shares of virtual stocks depends on the actual outcome of the event; therefore, the price of one share of a virtual stock should correspond to the PM's aggregate expectation of the event outcome and, in turn, the (discounted) expected cash dividend of a share of stock.

Participants in the PM use their (individual) expectations of the outcome to derive an (individual) expectation of the cash dividend of the related share of virtual stock. Accordingly, they compare their expected cash dividend with the PM's aggregate expectation, which is a function of the stock price, as a means to trade their individual expectations. For example, if a participant anticipates that the L. A. Lakers will score 100 points in a specific game, the cash dividend of the related share of virtual stock would be $100, and each point would correspond to $1. In the case of a current stock price of $95 ($105)—that is, an expectation of 95 (105) points—the stock is undervalued (overvalued), according to the estimates of this participant, who therefore could try to attain an expected profit of $5 by buying (selling). If the potential gains in the virtual portfolio value create a sufficiently high incentive for participants to perform well in the PM, it becomes their best strategy to engage in transactions on the basis of their best individual expectations. Thus, the participants reveal their true expectations of future market situations through their buying and selling activities (Oliven and Rietz, 2004; Spann and Skiera, 2003).

By making individual expectations tradable, a PM creates a market for predictions about future market situations, in which participants compete according to their individual expectations. Thus, the stock prices reflect the participants' aggregated information. Extensive studies using both empirical data and laboratory experiments support the informational efficiency of such markets (see overviews by Fama, 1970, 1991), as do the powerful results of political stock markets (Forsythe *et al.*, 1999).

Tipsters

Expert forecasts of sport outcomes often come from so-called 'tipsters', whose predictions appear in sports journals or daily newspapers. Tipsters are usually independent experts who do not apply a formal model but rather derive their predictions from their experience or intuition (Forrest and Simmons, 2000). They generally provide forecasts for only a specific selection of games, often related to betting. No immediate financial consequences result from the predictions of tipsters.

Empirical evidence regarding the forecast accuracy of tipsters shows that their ability is limited. Forrest and Simmons (2000) show that tipsters perform better than random forecasting methods but worse than a forecasting method that always predicts a home win (three tipsters correctly predicted 41.09%, 42.56% and 42.86%; wins by home teams occurred 47.5% of the time). Andersson *et al.* (2005) also reveal, paradoxically, that soccer experts fail to predict more accurately than people with limited knowledge of the game. These authors suggest their finding indicates the experts' inefficient

use of information, as well as laypersons' effective use of fast, frugal heuristics. Their result also mirrors research that found poor forecasting abilities of stock market experts (e.g., Törngren and Montgomery, 2004) and economists for business trends (e.g., Mills and Pepper, 1999).

Betting odds

Extensive analyses in economics and business literature suggest that betting odds provide an efficient forecasting instrument (Gandar *et al.*, 1998; Pope and Peel, 1989; see also the recent special issue of *Applied Economics* on the Economics of Betting Markets). For a recent summary of the history of sports wagering, see Vlastakis *et al.* (2007). Bookmakers determine fixed betting odds according to their expectations of game outcome probabilities and, once they are published, fixed odds rarely change. These fixed odds therefore represent expert predictions by bookmakers (Pope and Peel, 1989).

Andersson *et al.* (2005) compare the performance of experts and laypeople in predicting the outcomes of the soccer World Cup 2002; the difference in their performance is not statistically significant. Forrest *et al.* (2005) also compare the forecasting performance of several British bookmaking companies for the outcomes of English soccer games over a 5-year period and find that the forecasting performance increases over time. However, Goddard and Asimakopoulos (2004) reveal, in the context of English soccer league matches during the 1999–2000 season, that considering additional information, such as previous outcomes, team quality indicators and geographical distance between the teams, leads to betting strategies with a positive gross return of +8%. Boulier *et al.* (2006) similarly analyze the forecasting performance of betting market spreads for outcomes during 1994–2000 in the American National Football League (NFL) but show that no information beyond the point spread explains outcomes significantly better. Finally, Paton and Vaughan Williams (2005), also in the context of English premier soccer league games, indicate that information about initial bid–offer spreads of four major UK sports spread betting companies improves predictions, making them slightly better than the predictions of individual betting companies. Their results suggest betting odds provided by betting companies have a rather high forecasting accuracy, which is plausible, because betting companies with inefficient odds would not survive. However, despite these various analyses and considerations, no previous studies compare their results with those of PMs or tipsters.

DESCRIPTION OF THE DATA

Dataset

We forecast the outcomes of games in Germany's premier soccer league over three seasons: 1999–2000, 2000–2001 and 2001–2002. Germany's premier soccer league includes 18 teams that each play twice in a season, which equals 34 tournament rounds with nine games each, or 306 games per season, and thus 918 games in all three seasons. In the 1999–2000 and 2000–2001 seasons, a tournament round had the following structure: two games on Friday, five games on Saturday and two games on Sunday. In the 2001–2002 season, a tournament round instead meant seven games on Saturday and two games on Sunday. In contrast to many other sports, especially in the United States, a soccer game has three possible outcomes: win, lose or draw. In the case of a draw, each team receives one league point; in case of a win, the winning team receives three league points and the loser none.

Table I. Data samples (actual outcomes in each sample)

Season	Sample	No. obs.	% home	% draw	% away
1999/2000	Games predicted by PM and betting odds	288	46.18%	28.13%	25.69%
	Games predicted by tipster	203	49.26%	27.59%	23.15%
	Games predicted by PM, betting odds and tipster	203	49.26%	27.59%	23.15%
2000/2001	Games predicted by PM and betting odds	267	52.81%	21.35%	25.84%
	Games predicted by tipster	241	48.96%	23.65%	27.39%
	Games predicted by PM, betting odds and tipster	207	49.28%	22.71%	28.02%
2001/2002	Games predicted by PM and betting odds	282	52.48%	21.99%	25.53%
	Games predicted by tipster	277	53.43%	20.58%	25.99%
	Games predicted by PM, betting odds and tipster	268	53.36%	20.90%	25.75%
All 3 (1999–2002)	Games predicted by PM and betting odds	837	50.42%	23.89%	25.69%
	Games predicted by tipster	721	50.76%	23.58%	25.66%
	Games predicted by PM, betting odds and tipster	678	50.88%	23.45%	25.66%

For each tournament round, we collected game outcomes, stock prices on a PM (www.bundesligaboerse.de) for predictions of those game outcomes, tipster predictions (win, draw, lose) of the most popular German sports journal (*Sport Bild*) and the fixed betting odds of the largest German state-owned bookmaker (*Oddset*). The PM provided predictions for 91.18% of all games (N_{PM} = 837 games), so we collect betting odds for the same sample of games (N_{BET} = 837). However, we have fewer observations of tipsters' predictions, because the sports journal did not publish predictions for the two games on Friday during the 1999–2000 and 2000–2001 seasons. In addition, the journal arbitrarily ignored predictions for games in some weeks, which leaves us with N_{TIP} = 721 predictions by tipsters. Therefore, predictions associated with all three methods are available for 678 games. While the number of observations is smaller than those of studies that analyze the forecast accuracy of betting markets (e.g., Pope and Peel, 1989: 1291 matches; Cain *et al.*, 2000: 2855 matches; Dixon and Pope, 2004: 6629 matches; Vlastakis *et al.*, 2007: 12,841 matches; Graham and Stott, 2008: 11,000 matches), it is substantially larger than those of studies analyzing the forecast accuracy of prediction markets (e.g., Jank and Foutz, 2007: 262 observations; Pennock *et al.*, 2001: 161 observations; Servan-Schreiber *et al.*, 2004: 208 observations; Spann and Skiera, 2003: 152 observations). Table I provides descriptive statistics regarding the number of observations and proportion of actual home victories, draws and away victories in each sample and season. These results roughly match English league soccer outcomes; as Goddard and Asimakopoulos (2004) report, home teams win in 45.3% of games, away wins occur in 28.0% and draws happen in 26.7% of all games.

Calculations of the prediction market forecasts

The PM we investigate, the Soccer Market, attracted approximately 10,000 total registered users, with an average of 1500 active participants for each tournament round. It usually opened on Thursday at 6:00 p.m., and trading ended 5 hours later, at 11:00 p.m. on Thursday. On Friday, Saturday and Sunday, the Soccer Market remained open for 5 hours each day during the games, then closed each tournament round at the end of the last game on Sunday.

The payoff function of each share of virtual stock depends on the number of league points a soccer team gains in one tournament round. In the 1999–2000 season, the ultimate payoff of a share of

stock of the losing team was $100; of a drawing team $200, and of the winning team $400. The minimum $100 payoff for a loss serves to avoid 'penny stocks':

$$d_{\text{Home (Away)},g,r,s} = \$100 + \$100 \cdot Z_{\text{Home (Away)},g,r,s} = \begin{cases} d_s^{\text{Win}} = \$400 \\ d_s^{\text{Draw}} = \$200 \\ d_s^{\text{Loss}} = \$100 \end{cases} \quad (1)$$

$$(s \in S, r \in R_s, g \in G_{r,s}) \text{ for } S = 1999/2000$$

The payoff rule changed for later seasons. Each share of stock of a losing team was $0, of a drawing team $1 and of the winner $3:

$$d_{\text{Home (Away)},g,r,s} = \$1 \cdot Z_{\text{Home (Away)},g,r,s} = \begin{cases} d_s^{\text{Win}} = \$3 \\ d_s^{\text{Draw}} = \$1 \\ d_s^{\text{Loss}} = \$0 \end{cases} \quad (2)$$

$$(s \in S, r \in R_s, g \in G_{r,s}) \text{ for } S = 2000/2001 \text{ and } 2001/2002$$

where

$d_{\text{Home (Away)},g,r,s}$:   cash dividend of a share of stock that models the number of league points the home (away) team gains in the gth game in the rth tournament round of the sth season;

$Z_{\text{Home (Away)},g,r,s}$:   number of league points the home (away) team gains in the gth game in the rth tournament round of the sth season;

$d_s^{\text{Draw (Win/Loss)}}$:   cash dividend of a share of stock in the case of a draw (win/loss) in the sth season;

$G_{r,s}$:   index set of games in the rth tournament round of sth season;

R_s: index set of tournament rounds of sth season;

S: index set of seasons.

In each tournament round of the 1999–2000 season, all participants of the Soccer Market start with the same assets: 1000 shares of each type of team stock and $500,000 (virtual), with the possibility of a maximum virtual loan of $500,000 at a 1% weekly interest rate. For the 2000–2001 and 2001–2002 seasons, the endowment in each tournament round consisted of 1000 shares of each type of team stock and $5000 (virtual) cash, with no loans possible.

Participants are treated alike, regardless of when they enter the Soccer Market. A participant can trade shares according to his or her estimations of the game outcomes by selling shares of a presumably overvalued team stock or buying shares of a presumably undervalued team stock. Portfolio values from one tournament round do not transfer to the following tournament round; instead, the incentive involves monetary rewards for each round. At the end of each tournament round, the participant with the highest (virtual) portfolio value receives $150 (real), the person with the second highest value receives $100, and the third-ranking participant receives $50. There is no risk of actual financial loss. Table II follows the recommendations of Spann and Skiera (2003) to describe the design of the PM.

To determine outcome predictions from the PM, we use the stock prices of the team stocks at the end of trading on the first day, that is, the earliest possible end-of-trading point *before* the first game of a tournament round to predict *all* games of that round. Equation (3) represents the expected league points of a team in a specific tournament round in the 1999–2000 season; equation (4) describes the 2000–2001 and 2001–2002 seasons according to the current stock price:

Table II. Design of the prediction market

Step	Decisions
Choice of forecasting goal	• Forecasting of soccer game outcomes in the German premier league • Payoff function: gain of league points of home (away) team in a tournament round of the German premier soccer league (equation (1) for 1999–2000 season and equation (2) for 2000–2001 and 2001–2002 seasons) • Public access, possible to join at any time
Design of incentives for information revelation	*Composition of initial portfolios/endowment:* • Endowment of 1000 shares of each type of team stock and $500,000 [$5000] (virtual) (each tournament round for every participant) in 1999–2000 [2000–2001 and 2001–2002] seasons • Provision of loan up to $500,000 (virtual) at a 1% weekly interest rate in 1999–2000 season; no loans in later seasons *Remuneration/incentive mechanism:* • Monetary rewards • Rank-order tournament; participant with highest increase in (virtual) portfolio value receives $150 in cash, second highest $100 and third highest $50 • Time intervals: rank-order tournament for each tournament round • No non-performance-based incentives
Financial market design	• Double auction trading mechanism with competitive market maker and open order book • Each tournament round: 5 hours of daily trading from Thursday to Sunday • No short trading • Order types: limit and market without temporal restriction • No position or price limits • No trade fees

$$\hat{Z}^{\mathrm{PM}}_{\mathrm{Home\,(Away)},g,r,s,t} = \frac{\left(p_{\mathrm{Home\,(Away)},g,r,s,t} - 100\right)}{100}$$
$$(s \in S, r \in R_s, g \in G_{r,s}, t < T) \text{ for } S = 1999/2000 \tag{3}$$

and

$$\hat{Z}^{\mathrm{PM}}_{\mathrm{Home\,(Away)},g,r,s,t} = p_{\mathrm{Home\,(Away)},g,r,s,t}$$
$$(s \in S, r \in R_s, g \in G_{r,s}, t < T) \text{ for } S = 2000/2001 \text{ and } 2001/2002 \tag{4}$$

where

$\hat{Z}^{\mathrm{PM}}_{\mathrm{Home\,(Away)},g,r,s,t}$:   expected gain of league points according to the PM for the home (away) team at the tth point in time in the gth game in the rth tournament round of the sth season;

$p_{\mathrm{Home\,(Away)},g,r,s,t}$:   price of a share of the home (away) team's stock at the tth point in time in the gth game in the rth tournament round of the sth season;

T: point of time at the end of the game of the home (away) team in the gth game in the rth tournament round of the sth season.

Predictions for game outcomes reflect the differences in the stock prices of two competing teams; we predict a win for the team with the higher stock price. We predict a draw as the game outcome only when the two competing teams achieve identical stock prices.[1]

After determining the prices of the home and away teams in a specific game[2] and given that all outcome probabilities sum to 1, we can calculate the specific outcome probabilities $\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Draw\,(Home/Away)}}\right)$ for a specific game (for details, see the Appendix):

$$\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Home}}\right) = \frac{\phi}{\left(d_s^{\mathrm{Win}} - d_s^{\mathrm{Loss}}\right)\cdot\left(2\cdot d_s^{\mathrm{Draw}} - d_s^{\mathrm{Win}} - d_s^{\mathrm{Loss}}\right)} \tag{5}$$

with

$$\phi = \left(\mathrm{price}_{\mathrm{Home},g,r,s} - d_s^{\mathrm{Loss}}\right)\cdot\left(d_s^{\mathrm{Draw}} - d_s^{\mathrm{Win}}\right) - \left(\mathrm{price}_{\mathrm{Away},g,r,s} - d_s^{\mathrm{Win}}\right)\cdot\left(d_s^{\mathrm{Draw}} - d_s^{\mathrm{Loss}}\right)$$

$$\left(s \in S, r \in R_s, g \in G_{r,s}, 2\cdot d_s^{\mathrm{Draw}} \neq d_s^{\mathrm{Win}} + d_s^{\mathrm{Loss}}, d_s^{\mathrm{Win}} \neq d_s^{\mathrm{Loss}}\right)$$

$$\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Draw}}\right) = \frac{\mathrm{price}_{\mathrm{Home},g,r,s} - d_s^{\mathrm{Loss}} - \dfrac{\phi}{\left(2\cdot d_s^{\mathrm{Draw}} - d_s^{\mathrm{Win}} - d_s^{\mathrm{Loss}}\right)}}{d_s^{\mathrm{Draw}} - d_s^{\mathrm{Loss}}} \tag{6}$$

$$\left(s \in S, r \in R_s, g \in G_{r,s}, 2\cdot d_s^{\mathrm{Draw}} \neq d_s^{\mathrm{Win}} + d_s^{\mathrm{Loss}}, d_s^{\mathrm{Win}} \neq d_s^{\mathrm{Loss}}\right)$$

$$\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Away}}\right) = 1 - \left(\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Draw}}\right) + \mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Home}}\right)\right) \tag{7}$$

$$\left(s \in S, r \in R_s, g \in G_{r,s}\right)$$

Calculations of the betting market forecasts

We use the fixed betting odds of the largest German state-owned bookmaker (*Oddset*), which employs decimal odds and charges a fee of 25%, included in the odds.[3] That fee is substantially higher than the average margin of approximately 12% of most European (non-state-owned) bookmakers (Vlastakis *et al.*, 2007) or the 5% in person-to-person betting on betting exchanges such as Betfair (Smith *et al.*, 2006). We derive the bookmaker's forecasts from the betting odds by retrieving the implied probability of the different game outcomes and standardizing the probabilities to 1:

$$b_{g,r,s}^{\mathrm{Draw\,(Home/Away)}} = \frac{u_{g,r,s}^{\mathrm{Draw\,(Home/Away)}}}{u_{g,r,s}^{\mathrm{Draw}} + u_{g,r,s}^{\mathrm{Home}} + u_{g,r,s}^{\mathrm{Away}}} = \frac{\dfrac{1}{q_{g,r,s}^{\mathrm{Draw\,(Home/Away)}}}}{\dfrac{1}{q_{g,r,s}^{\mathrm{Draw}}} + \dfrac{1}{q_{g,r,s}^{\mathrm{Home}}} + \dfrac{1}{q_{g,r,s}^{\mathrm{Away}}}} \tag{8}$$

$$\left(s \in S, r \in R_s, g \in G_{r,s}\right)$$

---

[1] We also used a less strict definition for the prediction of a draw by allowing small differences in stock prices as predictons of a draw. Such variations had very little influence on the results.

[2] To equalize the differences in stock prices of competing teams in the 2000–2001 and 2001–2002 seasons with those for the 1999–2000 season, we multiply them by a scaling factor of 100. For example, assume the PM expects a home team to gain 2.5 league points and the away team to earn 1.8 league points. In the 1999–2000 season, the stock price difference would be $(2.5 \times 100 + 100) - (1.8 \times 100 + 100) = 250 - 180 = 70$. However, in the 2000–2001 and 2001–2002 seasons, the same prediction yields $2.5 - 1.8 = 0.7$, which we then multiply by 100 to equal the difference in the 1999–2000 season.

[3] The data showed that the numerator in 8 is always equal to 1.25, which indicates a margin of 25%.

Table III. Outcomes predicted by the three methods (actual outcomes in each sample)

Season	Sample	No. obs.	% home	% draw	% away
1999/2000	Outcomes predicted by PM	288	76.04%	1.74%	22.22%
	Outcomes predicted by betting odds	288	79.86%	0.00%	20.14%
	Outcomes predicted by tipster	203	51.23%	30.54%	18.23%
2000/2001	Outcomes predicted by PM	267	77.90%	1.12%	20.97%
	Outcomes predicted by betting odds	267	84.64%	0.00%	15.36%
	Outcomes predicted by tipster	241	50.62%	33.20%	16.18%
2001/2002	Outcomes predicted by PM	282	75.18%	2.13%	22.70%
	Outcomes predicted by betting odds	282	79.08%	0.00%	20.92%
	Outcomes predicted by tipster	277	54.51%	28.16%	17.33%
All 3 (1999–2002)	Outcomes predicted by PM	837	76.34%	1.67%	21.98%
	Outcomes predicted by betting odds	837	81.12%	0.00%	18.88%
	Outcomes predicted by tipster	721	52.29%	30.51%	17.20%

where

$b_{g,r,s}^{\text{Draw (Home/Away)}}$: standardized probability derived from betting odds of a draw (home team win/away team win) in the gth game in the rth tournament round of the sth season;

$u_{g,r,s}^{\text{Draw (Home/Away)}}$: unstandardized probability derived from betting odds of a draw (home team win/away team win) in the gth game in the rth tournament round of the sth season;

$q_{g,r,s}^{\text{Draw (Home/Away)}}$: betting odds of a draw (home team win/away team win) in the gth game in the rth tournament round of the sth season.

Therefore, if the decimal odds of a home win, draw and away win are, respectively, 1.7, 2.8 and 3.3, the standardized probabilities are 47.1%, 28.6% and 24.3%. The highest probability determines the forecast for the game outcome. Our results show that the bookmaker never assigns a draw with the highest probability.

Furthermore, we calculate the expected gain of league points by the home and away teams in a game on the basis of the standardized probabilities for each of the three possible game outcomes:

$$\hat{Z}_{\text{Home (Away)},g,r,s}^{\text{Odds}} = 3 \cdot b_{g,r,s}^{\text{Home (Away)}} + 1 \cdot b_{g,r,s}^{\text{Draw}} + 0 \cdot b_{g,r,s}^{\text{Away (Home)}}$$
$$(s \in S, r \in R_s, g \in G_{r,s}, t < T)$$

(9)

In Table III, we display the shares of outcomes predicted by each method for each season and all three seasons together.

FORECAST ACCURACY OF THREE METHODS

Evaluation criteria

Our criteria for evaluating and comparing the three forecasting methods are as follows:

1. We calculate the percentage of hits for each method, that is, the number of correctly predicted games relative to the total number of predicted games.

2. We calculate the root mean squared error (RMSE) for the deviation between the expected and actual gains of league points for each of the two teams in every game (N = total number of games in sample):

$$\text{RMSE} = \sqrt{\frac{\sum_{s \in S} \sum_{r \in R_s} \sum_{g \in G_{r,s}} \left[\left(\hat{Z}_{\text{Home},g,r,s} - Z_{\text{Home},g,r,s} \right)^2 + \left(\hat{Z}_{\text{Away},g,r,s} - Z_{\text{Away},g,r,s} \right)^2 \right]}{N \cdot 2}} \qquad (10)$$

3. We calculate the amount of money the predictions of each forecasting method would have won on the betting market for three possible fee scenarios: (a) with the 25% fee of the (state-owned) betting company; (b) with a fee of 12%, which is common for most European (non-state-owned) betting companies; and (c) with no fee. The calculated profit in all three scenarios indicates the value of each forecasting method. Specifically, the winnings without a fee (0%) show whether forecasting performance is better than the betting odds. The amount after subtracting the betting company's margin denotes whether punters can use the information to make money in a real-world betting market. The 12% fee reveals whether punters could earn money in a (competitive betting market) situation with a fee below the monopolistic fee (25%) of the state-owned betting market in Germany.

In addition, we compare the forecasts of the three methods with those of a naïve model and a pure random draw model. The naïve model always predicts a home win, which is the most frequent game outcome (i.e., the naïve model is not strictly naïve, because it uses this information; Forrest and Simmons, 2000). The pure random draw model randomly predicts one of the three events with overall probabilities, which provides a forecasting accuracy of ($h^2 + d^2 + a^2$), in which h, d and a are the proportions of home victories, away victories and draws in our dataset (Forrest and Simmons, 2000, p. 321).

Forecast accuracy of each method

In Table IV, we compare the hit rates of the PM, the naïve model, random picks and betting odds for the whole sample of 837 games. The PM yields a hit rate of 52.69%, greater than the total number

Table IV. Comparison of forecasting accuracy of prediction markets and betting odds

Instrument	Hit rate	% improve (*p*-value)[a]	RMSE[b]	Profit[c] (25% fee)	Profit[c] (12% fee)	Profit[c] (0% fee)
Prediction market	52.69%		1.29	−9.96%	0.27%	12.30%
Betting odds	52.93%	−0.45% (0.462)	1.22	−10.26%	−0.07%	11.92%
Naïve model: home win	50.42%	4.50% (0.099)	1.71	−10.36%	−0.18%	11.79%
Random draw model	37.73%	39.65% (0.000)	n.a.	n.a.	n.a.	n.a.

No. obs. = 837.
[a] Percentage of improvement of PM over alternative method = [hit rate PM − hit rate of alternative method]/hit rate of alternative method (one-tailed binomial test for difference of hit rate of PM).
[b] Root mean squared error for the deviation between the expected and actual gain of league points for both teams in every game. The naïve model only provides an outcome prediction (home win), from which we derive the expected gain in league points. Thus, the comparability of the RMSE of the naïve model is limited to the RMSE of PM and betting odds predictions, which provide separate predictions for the expected gain of league points for each team.
[c] Profit measured as the (relative) return on betting.

Table V. Comparison of forecast accuracy of different methods

Instrument	Hit rate	% improve (*p*-value)[b]	RMSE[c]	Profit[d] (25% fee)	Profit[d] (12% fee)	Profit[d] (0% fee)
Prediction market	54.28%		1.28	−6.83%	3.75%	16.20%
Betting odds	53.69%	1.10% (0.389)	1.22	−9.01%	1.33%	13.49%
Tipster[a]	42.63%	27.33% (0.000)	1.61	−19.97%	−10.88%	−0.19%
Naïve model: home win	50.88%	6.68% (0.041)	1.70	−9.84%	0.39%	12.44%
Random draw model	37.98%	42.92% (0.000)	n.a.	n.a.	n.a.	n.a.

No. obs. = 678.

[a] The sports journal did not predict all games; therefore, the comparison with the PM and betting odds depends on the same selection of games.

[b] Percentage of improvement of PM over alternative method = [hit rate PM − hit rate of alternative method]/hit rate of alternative method (one-tailed binomial test for difference to hit rate of PM).

[c] Root mean squared error for the deviation between the expected and actual gain of league points for both teams in every game. However, the tipster and naïve model only provide an outcome prediction (home win, draw or away win), from which we derive the expected gain in league points. Thus, the comparability of the RMSE of the tipster and naïve model is limited to the RMSE of PM and betting odds predictions, which provide separate predictions for the expected gain of league points for each team.

[d] Profit measured as the (relative) return on betting.

Table VI. Correlation between forecasts of different forecasting methods

Instrument	PM	Betting odds	Tipster
Prediction market		0.844	0.436
Betting odds	0.844		0.409
Tipster	0.436	0.409	
Home win	0.216	0.196	0.124
Actual outcome	0.251	0.232	0.144

No. obs. = 678; all correlations significant at $p < 0.001$.

The sports journal did not predict all games; the comparison with PM and betting odds is based on the same selection of games.

of home victories (50.42%) and pure random picks (37.73%). Betting odds have a slightly higher hit rate of 52.93% and a slightly lower RMSE, but lead to lower profits. Differences between the PM and betting odds are insignificant, indicating that the forecast accuracy is comparable. Both methods outperform the naïve model of home wins.

Table V displays the hit rates of the PM, betting odds and tipster for the overlapping sample of 678 games. This time, the PM achieves a higher hit rate and profit than the betting odds, but also a higher RMSE. Again, differences between the PM and betting odds are not significant and both methods significantly outperform the tipsters and the naïve model. The tipster's predictions are notably poor; even the naïve model clearly outperforms them. Thus, the forecasting accuracy of the PM and the betting odds is comparable and much better than those of the tipster or the naïve model.

These results fall in line with the correlations of the predictions (Table VI), for which we code the forecast of a home win as '1', a draw as '0' and an away win as '−1'. The correlation between

the predictions of the PM and the tipster is 0.436; that between the PM and naïve model is 0.216. Therefore, the predictions of these methods differ substantially. In contrast, the forecasts of the PM and betting odds correlate at 0.844, indicating their relatively close similarities. However, the forecasts are far from being equal, which indicates that we might be able to exploit these differences by combining the results of the different methods.

Forecast accuracy of combinations of the methods
Several studies show that combining the results of different forecasting methods can improve forecasting accuracy (Armstrong, 2001; Batchelor and Dua, 1995; Blattberg and Hoch, 1990). Therefore, we test the forecast accuracy of a weighting-based combination of forecasts (Blattberg and Hoch, 1990), as well as various rule-based combinations of forecasts.

Accuracy of weighting-based combined forecasts
We follow Blattberg and Hoch (1990), who suggest a 50:50 weighting, and thus a simple averaging of forecasts in a different setting. Therefore, we averaged the predicted number of league points for the home and away teams from the PM and betting odds (see equation (11)). We exclude the tipster, which does not provide a forecast for the expected league points and offers fairly poor predictions:

$$\hat{Z}^{\text{Comb.}}_{\text{Home (Away)},g,r,s} = 0.5 \cdot \hat{Z}^{\text{PM}}_{\text{Home (Away)},g,r,s} + 0.5 \cdot \hat{Z}^{\text{Odds}}_{\text{Home (Away)},g,r,s}$$

$$(s \in S, r \in R_s, g \in G_{r,s})$$

(11)

where

$\hat{Z}^{\text{Comb. (PM/Odds)}}_{\text{Home (Away)},g,r,s}$:   forecast of the weighting-based combination (PM/betting odds) for the expected league points of the home (away) team of the gth game in the rth tournament round of the sth season.

We use the differences in the expected league points to predict a win for the team with more expected league points; we predict a draw when the teams have identical expected league points. This weighting-based combination establishes a forecast accuracy of 52.69% ($N = 837$), equal to the hit rate of the forecasts of the PM for all 837 games. It also yields profits on betting markets with 25%, 12% and 0% fees of −13.12%, −0.59% and 11.47%, respectively. Neither the hit rate (one-tailed binomial test, $p > 0.5$) nor the profits (two-tailed paired t-tests, $p > 0.6$) differ significantly from the forecast of the PM or the betting odds for the same sample of 837 games (compare Table IV with Table VII). However, the RMSE of the weighting-based combination is lower than that of the PM, though higher than that of the betting odds. The results for the sample of the 678 games are very similar (compare Table V, second row, with Table VII, last column): neither the hit rate (one-tailed binomial test, $p > 0.4$) nor the profits (two-tailed paired t-test, $p > 0.3$) lead to significantly different results. Therefore, we conclude that our weighting-based combination of forecasts does improve the forecasts of the PM or the betting odds notably.

Accuracy of rule-based combined forecasts
Thus far, we have forecast all games, but we might improve forecast accuracy by concentrating on selected games. This situation more accurately reflects the real world; punters can usually

Table VII. Rule-based and weighting-based combined forecasts

Instrument	Rule-based forecasting				Weighting-based forecasting	
	(i) PM and betting odds agree	(ii) PM and tipster agree	(iii) Betting odds and tipster agree	(iv) PM, betting odds and tipster agree	50:50 forecasts ($N = 837$)	50:50 forecasts ($N = 678$)
Sample overlap[a]	837	678	678	678	837	678
Number of forecasts[a]	778	394	391	380	837	678
Forecast: home win[b]	631	322	330	321	652	528
Forecast: away win[b]	147	67	61	59	182	147
Forecast: Draw[b]	0	5	0	0	3	3
Hit rate (%)	53.98%	56.85%	56.52%	57.11%	52.69%	53.98%
RMSE[c]	1.62	1.54	1.54	1.53	1.25	1.24
Profit[d] (25% fee)	−9.86%	−7.64%	−9.08%	−8.72%	−13.12%	−8.15%
Profit[d] (12% fee)	0.39%	2.87%	1.26%	1.66%	−0.59%	2.28%
Profit[d] (0% fee)	12.44%	15.21%	13.42%	13.86%	11.47%	14.55%
Total profit (0%)[e]	$9678	$5993	$5247	$5267	$10,013	$9865

[a] Number of games.

[b] In case instruments agree, number of games that predict that outcome.

[c] Root mean squared error for the deviation between the expected and actual gain of league points for both teams in every game. The combination methods only provide an outcome prediction (home win, draw or away win), from which we derive the expected gain in league points. Thus, the comparability of the RMSE of the combination methods is limited to the RMSE of direct PM and betting odds predictions, which provide separate predictions for the expected gain of league points for each team.

[d] Profit measured as the (relative) return on betting.

[e] Total profit (0%) equals the profit realized by betting $100 on each selected game.

deliberately bet on only a selected number of games. Therefore, we analyze the quality of the following rules to select the games that we want to forecast:

1. Only forecast if the forecasts of PM and betting odds are the same.
2. Only forecast if the forecasts of PM and the tipster are the same.
3. Only forecast if the forecasts of betting odds and the tipster are the same.
4. Only forecast if the forecasts of PM, betting odds and the tipster are the same.

Table VII shows the results. The rule-based forecasts select between 380 (56.0%) and 778 (93.0%) games in each sample and increase the hit rate to 53.98% (rule 1), 56.85% (rule 2), 56.52% (rule 3) and 57.11% (rule 4). Thus, rule-based combined forecasts increase the hit rate, but none of the improved hit rates is significantly different (one-tailed binomial test) from the hit rate of the PM, that is, 54.28% for the sample of 678 games (Table V). Rule 4 achieves the highest hit rate (57.11%) but selects the fewest games. Betting $100 on each game would lead to winnings of $5267 (13.42%) if the betting companies do not charge fees. This amount is much less than those realized for the rules that select more games and significantly less than the total profit of the weighting-based combined forecasts. Total profits are highest for the PM forecasts ($10,295 for all 837 games, $10,984 for the overlapping 678 games). This total profit is greater than that achieved through weighting-based combined forecasts or relying on betting companies ($9977 for all 837 games, $9146 for the overlapping 678 games). Hence, this result supports the high forecast accuracy of PMs and betting odds.

SUMMARY AND CONCLUSIONS

We compare the forecast accuracy of different methods, namely, prediction markets, tipsters and betting odds, as well as weighting-based and rule-based combinations of those forecasts. The results indicate that PMs and betting odds yield a comparable and good forecast accuracy. PMs would allow punters to make more money on the betting market if the betting company does not charge fees or at least does not charge monopolistic fees. In contrast, tipsters' forecasts are poor, in support of the results of previous studies (Forrest and Simmons, 2000). Our findings also support results cited by, among others, Forrest *et al.* (2005), Boulier *et al.* (2006) and Spann and Skiera (2003), who show that betting odds and prediction markets provide very good forecasts. These previous studies do not, however, compare the forecasting accuracy of those methods and tipsters, whereas we show that PMs and betting odds forecast equally well and both clearly outperform tipsters. Interestingly, PM forecasts could yield profits from betting if the betting market charged moderate fees. A weighting-based combination of the forecasts of PMs and betting odds leads to a slightly higher forecast accuracy, whereas rule-based combined forecasts improve forecast accuracy substantially. However, the latter comes at a cost: it predicts the results of fewer games. Still, our results show that PMs can enhance the accuracy of sports forecasting.

APPENDIX

The price of a share of stock depends on the probability of the outcome of the game, as follows:

$$\text{price}_{\text{Home},g,r,s} = \left[1 - \text{PR}\left(Z_{g,r,s}^{\text{Draw}}\right) - \text{PR}\left(Z_{g,r,s}^{\text{Home}}\right)\right] \cdot d_s^{\text{Loss}} + \text{PR}\left(Z_{g,r,s}^{\text{Draw}}\right) \cdot d_s^{\text{Draw}} + \text{PR}\left(Z_{g,r,s}^{\text{Home}}\right) \cdot d_s^{\text{Win}} \tag{12}$$
$$(s \in S, r \in R_s, g \in G_{r,s})$$

and

$$\text{price}_{\text{Away},g,r,s} = \left[1 - \text{PR}\left(Z_{g,r,s}^{\text{Draw}}\right) - \text{PR}\left(Z_{g,r,s}^{\text{Home}}\right)\right] \cdot d_s^{\text{Win}} + \text{PR}\left(Z_{g,r,s}^{\text{Draw}}\right) \cdot d_s^{\text{Draw}} + \text{PR}\left(Z_{g,r,s}^{\text{Home}}\right) \cdot d_s^{\text{Loss}} \tag{13}$$
$$(s \in S, r \in R_s, g \in G_{r,s})$$

where

$\text{price}_{\text{Home (Away)},g,r,s}$:	price of a share of stock of the home (away) team in the *g*th game in the *r*th tournament round of the *s*th season;
$d_s^{\text{Draw (Win/Loss)}}$:	cash dividend of a share of stock in case of a draw (win/loss) in the *s*th season;
$G_{r,s}$:	index set of games in the *r*th tournament round of *s*th season;
$\text{PR}(Z_{g,r,s}^{\text{Draw (Home/Away)}})$:	probability of the outcome of the *g*th game in the *r*th tournament round of the *s*th season (i.e., draw, home win or away win);
R_s:	index set of tournament rounds of *s*th season;
S:	index set of seasons.

Knowing the prices $\text{price}_{\text{Home},g,r,s}$ and $\text{price}_{\text{Away},g,r,s}$ enables us to calculate the corresponding probabilities of the PM for the outcomes $\text{PR}(Z_{g,r,s}^{\text{Draw (Home/Away)}})$. Thus, we assume $\text{PR}(Z_{g,r,s}^{\text{Draw}}) + \text{PR}(Z_{g,r,s}^{\text{Home}}) + \text{PR}(Z_{g,r,s}^{\text{Away}}) = 1$, and solving equations (12) and (13) for the two probabilities $\text{PR}(Z_{g,r,s}^{\text{Draw}})$ and $\text{PR}(Z_{g,r,s}^{\text{Home}})$ leads to

$$\text{price}_{\text{Home},g,r,s} - d_s^{\text{Loss}} = \text{PR}\big(Z_{g,r,s}^{\text{Draw}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Loss}}\big) + \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Win}} - d_s^{\text{Loss}}\big) \tag{14}$$
$$(s \in S, r \in R_s, g \in G_{r,s})$$

and

$$\text{price}_{\text{Away},g,r,s} - d_s^{\text{Win}} = \text{PR}\big(Z_{g,r,s}^{\text{Draw}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Win}}\big) + \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Loss}} - d_s^{\text{Win}}\big) \tag{15}$$
$$(s \in S, r \in R_s, g \in G_{r,s})$$

Solving both equations for the probability $\text{PR}(Z_{g,r,s}^{\text{Draw}})$ yields

$$\text{PR}\big(Z_{g,r,s}^{\text{Draw}}\big) = \frac{\text{price}_{\text{Home},g,r,s} - d_s^{\text{Loss}} - \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Win}} - d_s^{\text{Loss}}\big)}{d_s^{\text{Draw}} - d_s^{\text{Loss}}} \tag{16}$$
$$\big(s \in S, r \in R_s, g \in G_{r,s}, d_s^{\text{Draw}} \neq d_s^{\text{Loss}}\big)$$

and

$$\text{PR}\big(Z_{g,r,s}^{\text{Draw}}\big) = \frac{\text{price}_{\text{Away},g,r,s} - d_s^{\text{Win}} - \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Loss}} - d_s^{\text{Win}}\big)}{d_s^{\text{Draw}} - d_s^{\text{Loss}}} \tag{17}$$
$$\big(s \in S, r \in R_s, g \in G_{r,s}, d_s^{\text{Draw}} \neq d_s^{\text{Win}}\big)$$

Then, if we subtract equation (17) from equation (16), we obtain

$$\frac{\text{price}_{\text{Home},g,r,s} - d_s^{\text{Loss}} - \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Win}} - d_s^{\text{Loss}}\big)}{d_s^{\text{Draw}} - d_s^{\text{Loss}}}$$
$$= \frac{\text{price}_{\text{Away},g,r,s} - d_s^{\text{Win}} - \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Loss}} - d_s^{\text{Win}}\big)}{d_s^{\text{Draw}} - d_s^{\text{Win}}} \tag{18}$$
$$\big(s \in S, r \in R_s, g \in G_{r,s}, d_s^{\text{Draw}} \neq d_s^{\text{Loss}}, d_s^{\text{Draw}} \neq d_s^{\text{Win}}\big)$$

which, when rearranged, equals

$$\big(\text{price}_{\text{Home},g,r,s} - d_s^{\text{Loss}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Win}}\big) - \big(\text{price}_{\text{Away},g,r,s} - d_s^{\text{Win}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Loss}}\big)$$
$$= \text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) \cdot \big(d_s^{\text{Win}} - d_s^{\text{Loss}}\big) \cdot \big(2 \cdot d_s^{\text{Draw}} - d_s^{\text{Win}} - d_s^{\text{Loss}}\big) \tag{19}$$
$$(s \in S, r \in R_s, g \in G_{r,s})$$

Solving for the probability $\text{PR}(Z_{g,r,s}^{\text{Home}})$ yields

$$\text{PR}\big(Z_{g,r,s}^{\text{Home}}\big) = \frac{\big(\text{price}_{\text{Home},g,r,s} - d_s^{\text{Loss}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Win}}\big) - \big(\text{price}_{\text{Away},g,r,s} - d_s^{\text{Win}}\big) \cdot \big(d_s^{\text{Draw}} - d_s^{\text{Loss}}\big)}{\big(d_s^{\text{Win}} - d_s^{\text{Loss}}\big) \cdot \big(2 \cdot d_s^{\text{Draw}} - d_s^{\text{Win}} - d_s^{\text{Loss}}\big)} \tag{20}$$
$$\big(s \in S, r \in R_s, g \in G_{r,s}, 2 \cdot d_s^{\text{Draw}} \neq d_s^{\text{Win}} + d_s^{\text{Loss}}, d_s^{\text{Win}} \neq d_s^{\text{Loss}}\big)$$

Then, if we substitute equation (20) into equation (16), we achieve the probability $\mathrm{PR}(Z_{g,r,s}^{\mathrm{Draw}})$:

$$\mathrm{PR}\left(Z_{g,r,s}^{\mathrm{Draw}}\right) = \frac{\mathrm{price}_{\mathrm{Home},g,r,s} - d_s^{\mathrm{Loss}} - \dfrac{\left(\mathrm{price}_{\mathrm{Home},g,r,s} - d_s^{\mathrm{Loss}}\right)\cdot\left(d_s^{\mathrm{Draw}} - d_s^{\mathrm{Win}}\right) - \left(\mathrm{price}_{\mathrm{Away},g,r,s} - d_s^{\mathrm{Win}}\right)\cdot\left(d_s^{\mathrm{Draw}} - d_s^{\mathrm{Loss}}\right)}{\left(2\cdot d_s^{\mathrm{Draw}} - d_s^{\mathrm{Win}} - d_s^{\mathrm{Loss}}\right)}}{d_s^{\mathrm{Draw}} - d_s^{\mathrm{Loss}}} \tag{21}$$

$$\left(s \in S, r \in R_s, g \in G_{r,s}, 2\cdot d_s^{\mathrm{Draw}} \neq d_s^{\mathrm{Win}} + d_s^{\mathrm{Loss}}, d_s^{\mathrm{Win}} \neq d_s^{\mathrm{Loss}}\right)$$

ACKNOWLEDGEMENTS

REFERENCES

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**: 565–576.

Armstrong JS (ed.). 2001. *Principles of Forecasting*. Kluwer: Dordrecht.

Batchelor R, Dua P. 1995. Forecaster diversity and the benefits of combining forecasts. *Management Science* **41**: 68–75.

Blattberg RC, Hoch S. 1990. Database models and managerial intuition: 50% model and 50% managers. *Management Science* **36**: 887–899.

Boulier BL, Stekler HO, Amundson S. 2006. Testing the efficiency of the National Football League betting market. *Applied Economics* **38**: 279–284.

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

Chen Y, Chu C-H, Mullen T, Pennock DM. 2005. Information markets vs. opinion polls. In *Research Paper Proceedings of the 6th ACM conference on Electronic Commerce*, Vancouver, Canada.

Dahan E, Hauser JR. 2002. The virtual customer. *Journal of Product Innovation Management* **19**: 332–353.

Dahan E, Lo AW, Poggio T, Chan NT, Kim A. 2006. Securities trading of concepts (STOC). Working paper, University of California, Los Angeles.

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

Elberse A. 2007. The power of stars: do star actors drive the success of movies? *Journal of Marketing* **71**: 102–120.

Elberse A, Eliashberg J. 2003. Demand and supply dynamics for sequentially released products in international markets: the case of motion pictures. *Marketing Science* **22**: 329–354.

Fama EF. 1970. Efficient capital markets: a review of theory and empirical work. *Journal of Finance* **25**: 383–417.

Fama EF. 1991. Efficient capital markets: II. *Journal of Finance* **46**: 1575–1617.

Forrest D, Simmons R. 2000. Forecasting sport: the bahaviour and performance of football tipsters. *International Journal of Forecasting* **16**: 317–331.

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

Forsythe R, Nelson F, Neumann GR, Wright J. 1992. Anatomy of an experimental political stock market. *American Economic Review* **82**: 1142–1161.

Forsythe R, Rietz TA, Ross TW. 1999. Wishes, expectations and actions: a survey on price formation in election stock markets. *Journal of Economic Behavior and Organization* **39**: 83–110.

Gandar JM, Dare WH, Brown CR, Zuber RA. 1998. Informed traders and price variations in the betting market for professional basketball games. *Journal of Finance* **53**: 385–401.

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

Graham I, Stott H. 2008. Predicting bookmaker odds and efficiency for UK football. *Applied Economics* **40**: 99–109.

Gruca TS, Berg JE, Cipriano M. 2003. The effect of electronic markets on forecasts of new product success. *Information Systems Frontiers* **5**: 95–105.

Hahn B, Tetlock P (eds). 2006. *Information Markets: A New Way of Making Decisions*. AEI–Brookings Press: Washington, D.C.

Hayek FA. 1945. The use of knowledge in society. *American Economic Review* **35**: 519–530.

Jank W, Foutz N. 2007. Using virtual stock exchanges to forecast box-office revenue via functional shape analysis. In *Second Workshop on Prediction Markets*, San Diego, CA, 2007.

Luckner S, Weinhardt C. 2007. How to pay traders in information markets: results from a field experiment. *Journal of Prediction Markets* **1**: 147–156.

Mills TC, Pepper GT. 1999. Assessing the forecasts: an analysis of forecasting records of the Treasury, the London Business School and the National Institute. *International Journal of Forecasting* **15**: 247–257.

Oliven K, Rietz TA. 2004. Suckers are born but markets are made: individual rationality, arbitrage, and market efficiency on an electronic futures market. *Management Science* **50**: 336–351.

Paton D, Vaughan Williams L. 2005. Forecasting outcomes in spread betting markets: can bettors use 'quarbs' to beat the book? *Journal of Forecasting* **24**: 139–154.

Pennock DM, Lawrence S, Giles LC, Nielsen FA. 2001. The real power of artificial markets. *Science* **291**: 987–988.

Plott CR, Wit J, Yang WC. 2003. Parimutuel betting markets as information aggregation devices: experimental results. *Economic Theory* **22**: 311–351.

Pope PF, Peel DA. 1989. Information, prices and efficiency in a fixed-odds betting market. *Economica* **56**: 323–341.

Servan-Schreiber E, Pennock DM, Wolfers J, Galebach B. 2004. Prediction markets: does money matter? *Electronic Markets* **14**: 1–10.

Smith MA, Paton D, Vaughan Williams L. 2006. Market efficiency in person-to-person betting. *Economica* **73**: 673–689.

Smith VL. 1982. Microeconomic systems as an experimental science. *American Economic Review* **72**: 923–955.

Spann M, Skiera B. 2003. Internet-based virtual stock markets for business forecasting. *Management Science* **49**: 1310–1326.

Spann M, Ernst H, Skiera B, Soll JH. 2009. Identification of lead users for consumer products via virtual stock markets. *Journal of Product Innovation Management* (forthcoming).

Törngren G, Montgomery H. 2004. Worse than chance? Performance and confidence among professionals and laypeople in the stock market. *Journal of Behavioral Finance* **5**: 246–251.

Tziralis G, Tatsiopoulos I. 2007. Prediction markets: an extended literature review. *Journal of Prediction Markets* **1**: 75–91.

Vlastakis N, Dotsis G, Markellos RN. 2007. How efficient is the European football betting market? Evidence from arbitrage and trading strategies. Working paper, Athens University of Economics and Business.

Wolfers J, Zitzewitz E. 2006. *Interpreting Prediction Market Prices as Probabilities*. Wharton School, University of Pennsylvania: Philadelphia, PA.

Authors' biographies:

Martin Spann (spann@spann.de) is a Chaired Professor of Marketing and Innovation at the University of Passau, Germany. He has been a Visiting Professor at the Marshall School of Business at the University of Southern California, Los Angeles, and at Bocconi University in Milan, Italy. Martin's research interests are prediction markets, pricing, electronic commerce, and innovation & new product management. His work has been published in Management Science, Information Systems Research, Journal of Marketing, Journal of

Interactive Marketing, Journal of Product Innovation Management, and European Journal of Operational Research.

Bernd Skiera (skiera@skiera.de) is a Chaired Professor of Electronic Commerce at the University of Frankfurt, Germany, and director of the E-Finance Lab and the Retail Banking Competence Center. His research focuses on prediction markets, internet marketing, pricing and customer management. His work has been published in Management Science, Marketing Science, Journal of Marketing Research, Journal of Marketing, Journal of Interactive Marketing, and European Journal of Operational Research.

Authors' addresses:
Martin Spann, School of Business and Economics, University of Passau, Innstr. 27, 94032 Passau, Germany.

Bernd Skiera, School of Business and Economics, Johann Wolfgang Goethe-University, Mertonstr. 17, 60054 Frankfurt am Main, Germany.