On the Forecast Accuracy of Sports Prediction Markets

Stefan Luckner[1], Jan Schröder[1], and Christian Slamka[2]

[1] Institute of Information Systems and Management, Universität Karlsruhe (TH),
76128 Karlsruhe, Germany
{Luckner,Schroeder}@iism.uni-karlsruhe.de
[2] School of Business and Economics, University of Frankfurt,
60054 Frankfurt am Main, Germany
slamka@wiwi.uni-frankfurt.de

Abstract. Accurate forecasts are essential in many areas such as business and sports forecasting. Prediction markets are a promising approach for forecasting future events and are increasingly used to aggregate information on particular future events of interest such as elections, sports events, and Oscar winners. In this paper, we present the results of an empirical study that compares the forecast accuracy of a prediction market for the FIFA World Cup 2006 to predictions derived from the FIFA world ranking and to a random predictor. We find that prediction markets for the FIFA World Cup outperform predictions based on the FIFA world ranking as well as the random predictor in terms of forecast accuracy.

Keywords: Prediction Markets, Forecast Accuracy, Sports Forecasting.

1 Introduction

Accurate forecasts are essential in many areas such as business, sports, and weather forecasting. With the propagation of the Internet, interest in prediction markets as a tool to forecast future events has continuously increased over the last couple of years. Examples for prediction markets include the Iowa Electronic Markets, TradeSports, NewsFutures, the Hollywood Stock Exchange, and STOCCER. Moreover, several major companies such as Hewlett-Packard, Google, or Microsoft are running internal prediction markets for company-specific predictions.

The basic idea of prediction markets is to trade virtual stocks whose payoffs are tied to the outcome of uncertain future events on an electronic market. Although the final payoffs of stocks are unknown during the trading period, rational and risk neutral traders sell stocks if they consider the stocks to be overvalued and buy stocks if they consider the stocks to be undervalued [1]. During this time, the trading price reflects the traders' aggregated beliefs about the likelihood of a future event. Market prices can thus be interpreted as predictions. The theoretical justification of regarding market prices as estimates of the outcome of future events is founded in the Hayek hypothesis [2]. It states that asymmetrically dispersed information is best aggregated using a price mechanism. Moreover, if all the available information is reflected in the market prices, the market is informationally efficient [3, 4].

One of the main reasons for the emergence of prediction markets is that they have shown a high forecasting accuracy [e.g., 5, 6] compared to traditional forecasts such as expert predictions or surveys. Good performance has also been demonstrated in corporate environments [7-9]. In this paper, we study their forecasting accuracy in the field of sports forecasting, more precisely for predicting the outcomes of soccer matches during the FIFA World Cup 2006. We examine whether prediction markets outperform a random predictor and forecasts that are based on historic data about the success of national soccer teams. The FIFA world ranking is based on such historic data and can thus be used as a benchmark for our study.

The remainder of the paper is structured as follows: In the next section, we present work related to the analysis of the forecasting accuracy of prediction markets in general as well as for soccer games and tournaments in particular. Section 3 deals with the design of the event to be analyzed, namely the FIFA World Cup 2006, as well as the design of the STOCCER prediction market platform including its markets and descriptive statistics about participation in the markets. In section 4, the predictive accuracy of the STOCCER markets is analyzed by comparing the predictions to a random predictor and predictions derived from the FIFA world ranking. Section 5 summarizes the main results.

2   Related Work

Information aggregation with prediction markets has a long tradition in economic research. Starting with the 1988 U.S. presidential election market the focus of earlier research in the field was mainly to study the accuracy of predictions derived from the trading actions taken within these markets. Political stock markets beat election polls in many cases [10, 11]. Research in the business forecasting area suggests that prediction markets can perform better than traditional methods such as business meetings [12]. Plott and Chen show that prediction markets on sales forecasting were significantly better than official forecasts in 6 out of 8 cases [7]. Over the last three years, the Hollywood Stock Exchange (HSX) almost perfectly predicted the Oscar award winners [13] and has beaten the individual and average forecasts of five self-selected experts [14].

During the last decade, prediction markets were also employed in the field of sports forecasting, especially for predicting the outcome soccer matches and tournaments. The first markets to be reported in literature date back to 1994 and 1998 [15]. So far, most of the papers on soccer prediction markets deal with the prediction accuracy of markets e.g. compared to betting odds. Schmidt and Werwatz [15] analyze a 2000 European Championship market to find out whether a prediction market is a better predictor in terms of forecasting accuracy than a random predictor. As a second benchmark they use betting odds for the same event from several betting companies. Volkland and Strobel [16] use quite the same techniques for analyzing their data for a 2002 FIFA World Cup prediction market. With different market parameters (e.g. number of participants and number of markets) the results are different to the ones found by Schmidt and Werwatz. In fact, all results of [15] were reversed in [16]. While in [15] the random predictor performed worse than the markets' predictions, the markets' performance in [16] was hardly better than

predicting by choice. Also, relative to the prediction markets expert bookmakers forecasted more accurately in [16] compared to the ones in [15]. Volkland and Strobel state that not a single explanatory factor could be identified which was significant in explaining predictive success or accuracy by the markets.

Spann and Skiera compared the large number of 837 German premier league soccer games to the corresponding results of so called tipsters as well as betting odds [17]. Their results are quite encouraging. According to their study, prediction markets slightly outperform the betting odds results while the tipsters are strongly dominated by both.

Compared to the markets analyzed in [15], the number of participants in the 2002 FIFA World Cup prediction market was about half of the number in [15] but these smaller number of traders had to deal with about three times the number of markets and double the number of trading days compared to the 2000 European Championship market. Volkland and Strobel state that further research should aim for more participation in the markets to achieve more sophisticated analyses. We thus analyze data from a more liquid prediction market for the FIFA World Cup 2006. Instead of comparing the market prediction to betting odds we analyze whether prediction markets outperform forecasts that are based on historic data about the success of national soccer teams.

3   STOCCER – A 2006 FIFA World Cup Prediction Market

In this section we firstly describe the event we studied, namely the FIFA World Cup 2006. Secondly, we present the prediction markets we operated during the World Cup in order to predict the outcome of the tournament and the 16 matches in the final rounds.

3.1   The FIFA World Cup 2006

The most important soccer tournament worldwide in 2006, the FIFA World Cup, was played from June 9th to July 9th with 32 participating national teams. In the preliminary rounds, in eight groups with four teams each, each team played against the other three teams in the group. The top two teams in each group advanced to the round of 16, where the winning team of a group played the second of one of the remaining groups. The remaining matches of the tournament were played in a sudden death system until the final. Additionally, one game was played for the third place between the losers of the two semi-final games. All in all, 48 matches were played in the preliminary rounds and 16 in the final rounds, resulting in a total of 64 matches. Draws were possible in the preliminary rounds but not in the final rounds. In case of a draw in the final rounds the match went into overtime and where necessary there were penalty shootouts.

3.2   The STOCCER Exchange

In order to study the prediction accuracy of sports prediction markets, we ran a market during the FIFA World Cup 2006. More than 1.500 traders registered for our

experimental market called STOCCER[1]. The web trading interface we provided is depicted in Figure 1. Our experimental market started on May 15th 2006 with the championship market and ran until the end of the FIFA World Cup on July 9th 2006. The trading platform was open to the public 24 hours a day, 7 days a week. On average, we had more than 1,600 transactions per day.
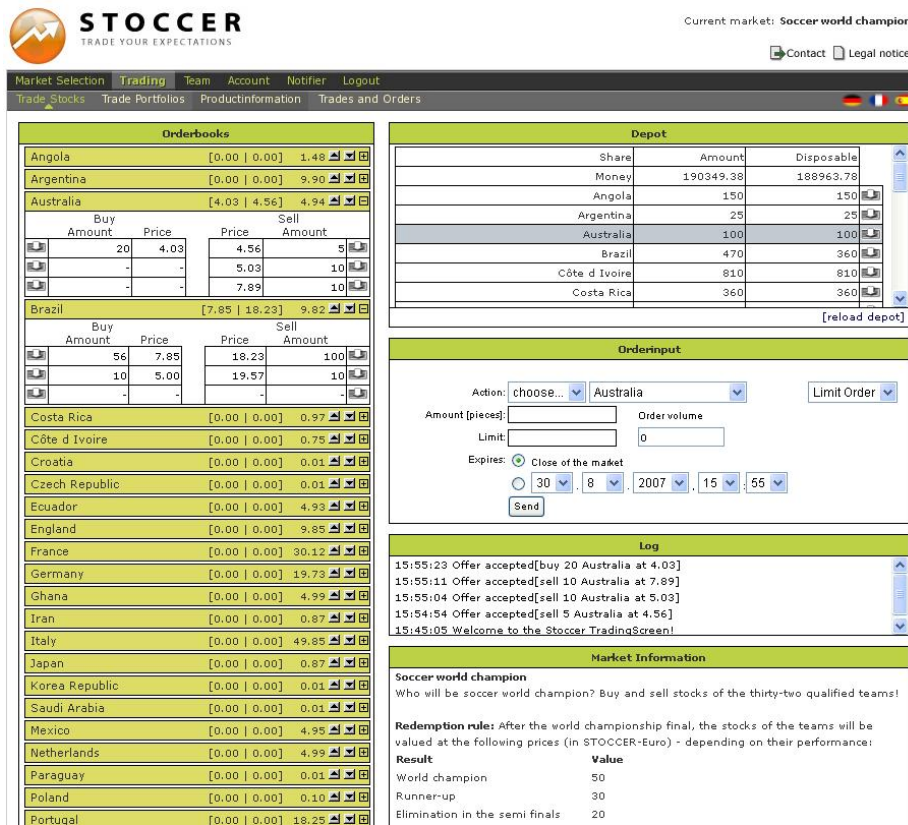

Fig. 1. Trading screen of STOCCER

In total, we ran 19 markets: 16 for the matches in the final rounds starting with the round of sixteen, two goal getter markets and the so called championship market where we traded shares of all the 32 national teams taking part in the FIFA World Cup 2006. The markets we use in this study are also shown in Table 1 with some more information on the number of virtual stocks we traded in each of the markets, market start and end time as well as information on how the shares were valued at the close of the market.

---

[1] www.stoccer.com

In case of the championship market, a virtual stock of the world champion was valued at 50 virtual currency units, a virtual stock of the vice-world champion was valued at 30 virtual currency units etc. In case a team did not reach the round of 16 the team's virtual stocks were worthless at the end. For the 16 matches in the final round we traded three stocks per match since there were three possible outcomes for every match – either team A won or team B won or there was a draw after the second half. We introduced the third virtual stock ("draw") although there were no draws possible in the final rounds of the tournament. The reason for this was that we did not want to consider overtimes and penalty shootouts as we regarded their outcomes as more or less unpredictable. The virtual stock corresponding to the event that actually occurred at the end of a match was valued at 10 virtual currency units after the match; the other two assets were worthless.

Table 1. Markets operated during the FIFA World Cup 2006

Type	Number of virtual stocks	Final payoff	Start	End
Championship	1 per country (32)	World champion: 50 Vice-WC: 30 Semi finals: 20 Quarter finals: 10 Round of 16: 5 Otherwise: 0	May 15th 2006	July 9th 2006
Match	3 per match: team A wins, team B wins, tie after 2nd half	Event occurred: 10 Otherwise: 0	2 days before the matches	At the end of the matches

Concerning the financial market design, all markets used a continuous double auction (CDA) in combination with limit orders. In order to issue shares we decided to make use of so called unit portfolios. A portfolio contains one piece of share of every virtual stock which is traded in the respective market. The portfolio price equals the sum of the payoffs for one share of every virtual stock in a market and was e.g. 10 virtual currency units in the match markets. It thus corresponded to the redemption value for correctly predicting the outcome of a match. Buying and selling of portfolios from and to the market operators is therefore risk free for traders and was possible at all times.

In contrast to traditional betting exchanges for sports events, we decided to operate STOCCER as a play-money market. Instead of investing real money every market participant had an initial endowment of 100,000 virtual currency units as well as 100 units of every virtual stock. The only extrinsic incentives for traders to join the market and reveal their expectations were a ranking of their user names and a lottery of prizes. The overall TOP-100 traders took part in a final lottery, where the first prizes were shares of an investment fund with a value of 3,000, 2,000, and 1,000 Euro. In addition, we weekly raffled an iPod among the 20 most active traders of the preceding week.

4   Evaluation of the Forecast Accuracy

Market prices and thus also the forecasts of prediction markets are driven by the information and the expectations of the users trading in the markets. These forecasts are worthless if they do not result in better predictions than randomly drawing one of the possible outcomes. Thus, we use a random predictor as our first benchmark to evaluate the forecast accuracy of our prediction market. Besides historic data traders also consider current information that is available to them as well as ongoing developments during the tournament and even during matches. We examine whether forecasts based on market prices comprising this additional information outperform forecasts that are based on historic data about the success of national soccer teams. We use the FIFA world ranking[2] as our second benchmark since it is based on historic data only. The FIFA world ranking from May 2006 we use for this study was built on a history of the last eight years and takes into account the factors outcome of matches, importance of matches, strength of opponents, regional strength, home and away matches, as well as number of matches and goals. For the index all international "A" matches were relevant and for all factors individual points were given which were aggregated for the index. For most factors complex calculations are used to get an idea of a team's actual state and strength[3].

To compare the forecasting accuracy of our markets to predictions derived from the FIFA ranking and to the random predictor we calculate the hit rate, i.e. the percentage of correctly predicted games, for each forecasting instrument. In Table 2, we compare the hit rate of the three instruments for the whole sample of 64 matches.

Table 2. Comparison of forecasting accuracy (all matches)

Instrument	No. Obs.	Hit rate	% improvement[4]	p-value[5]
Championship market	64	59.38%		
FIFA ranking	64	46.88%	26.66%	0.042
Random draw	64	33,33%	78,14%	< 0,001

In case of the championship market we predict a win for the team with the higher stock price. We predict a draw whenever the stock prices of two teams are the same. In case of the FIFA ranking we predict a win for the team that has the better position in the ranking. For our random predictor we assumed that the three possible outcomes of a match are equally likely to occur. Empirical data supports this hypothesis [15]. Thus, the theoretical hit rate of such a random draw model is 33.33%.

When comparing the hit rates of the championship market, the FIFA ranking and the random predictor for all 64 matches, we find that our championship market indeed

---

[2] http://www.fifa.com/worldfootball/ranking/
[3] The calculation of the ranking is rather complex and we do not go into more details here. Due to its complexity the calculation procedure was changed in the meanwhile.
[4] Percentage of improvement of prediction market over alternative instrument.
[5] Chi-square test for difference to hit rate of prediction market.

yields a higher hit rate than the FIFA ranking and the random draw model. The forecasts can be improved when using a prediction market instead of the other two instruments. Table 2 shows the percentage of improvement of the prediction market over the respective alternative instrument. The difference in the hit rate of the prediction market and the two other instruments is significant in both cases ($p < 0.05$, chi-square test).

As described in Section 3 we also operated separate markets for the 16 matches in the final rounds. To calculate the hit rate in case of the match markets we predict the outcome with the highest stock price out of the three possible outcomes of a match. We compare the forecasts of these 16 match markets to the forecasts of the other three instruments. The results of this comparison are shown in Table 3.

Table 3. Comparison of forecasting accuracy (final rounds)

Instrument	No. Obs.	Hit rate	% improvement[4]	p-value[5]
Match market	16	62,50%		
Championship market	16	43,75%	42,86%	0,121
FIFA ranking	16	25,00%	150,00%	0,002
Random draw	16	33,33%	87,52%	0,010

For the last 16 matches of the tournament, the hit rate of the match markets is significantly higher than the hit rate of the FIFA ranking and of the random draw model. Interestingly, the hit rate is higher in case of the match markets than it is when predicting a win for the team with the higher stock price in the championship market. One reason for this insignificant difference could be the fact that the likelihood of draws is underestimated in the championship market as well as in case of the predictions based on the FIFA ranking. Furthermore, traders in match markets can focus on the outcome of one match at a time.

However, it is somewhat surprising that the hit rate for the championship market and the FIFA ranking is on average lower for the last 16 matches than it is when taking into account all 64 matches. We think this is plausible since it should be easier to predict the outcome of matches at the beginning of the tournament than at the end. At the beginning, there are numerous underdogs and clear favorites whereas the performance of teams will not differ that much at the end of the tournament. Thus, it is presumably much harder to predict the outcome of matches taking place in the last rounds compared to earlier matches.

5   Summary

In this paper, we presented a sports prediction market we were running during the FIFA World Cup 2006. The goal of our paper is to study the forecasting accuracy of soccer prediction markets compared to forecasts that are based on historic data about the success of national soccer teams and a random predictor. We use a random predictor as our first benchmark since forecasts are worthless if they do not result in better predictions than random draws. We use the FIFA world ranking as a second benchmark

for our study since it is calculated based on historic data. Our results provide evidence that prediction markets outperform predictions derived from the FIFA world ranking, i.e. historic data, as well as a random predictor in terms of forecast accuracy.

References

1. Glosten, L.R., Milgrom, P.R.: Bid, Ask and Transaction Prices in a Specialist Market With Heterogeneously Informed Traders. Journal of Financial Economics 14, 71–100 (1985)
2. Hayek, F.A.v.: The Use of Knowledge in Society. American Economic Review 35(4), 519–530 (1945)
3. Fama, E.F.: Efficient Capital Markets: A Review of Theory and Empirical Work. Journal of Finance 25, 383–417 (1970)
4. Fama, E.F.: Efficient Capital Markets: II. Journal of Finance 46(5), 1575–1617 (1991)
5. Servan-Schreiber, E., et al.: Prediction Markets: Does Money Matter? Electronic Markets 14(3), 243–251 (2004)
6. Spann, M.: Internet-Based Virtual Stock Markets for Business Forecasting. Management Science 49(10), 1310–1326 (2003)
7. Plott, C.R., Chen, K.-Y.: Information Aggregation Mechanisms: Concept, Design and Implementation for a Sales Forecasting Problem, California Institute of Technology (2002)
8. Ortner, G.: Aktienmärkte als Industrielles Vorhersagemodell. Zeitschrift für Betriebswirtschaft (ZfB) - Ergänzungsheft 70, 115–125 (2000)
9. Plott, C.R.: Markets as Information Gathering Tools. Southern Economic Journal 67, 1–15 (2000)
10. Forsythe, R., et al.: Anatomy of an Experimental Political Stock Market. American Economic Review 82, 1142–1161 (1992)
11. Berg, J.E., et al.: Results from a Dozen Years of Election Futures Markets Research. In: Plott, C., Smith, V.L. (eds.) Handbook of Experimental Economic Results (2001)
12. Sunstein, C.R.: Deliberation and Information Markets, in Information Markets: A New Way of Making Decisions. In: Hahn, W., Tetlock, P.C. (eds.) pp. 67–100. AEI-Brookings Press, Washington D.C (2006)
13. Lamare, A.: Hollywood Stock Exchange (HSX.com) Traders correctly picked 7 out of 8 Top Category Oscar Winners to continue its stellar record, cited (2007) Available from: http://www.hsx.com/about/press/070226.htm
14. Pennock, D.M., et al.: The Power of Play: Efficiency and Forecast Accuracy of Web Market Games, in Technical Report, 2000-168. 2001, NEC Research Institute
15. Schmidt, C., Werwatz, A.: How well do markets predict the outcome of an event? The Euro, soccer championships experiment, in Discussion Papers on Strategic Interaction. 2002, Max Planck Institute for Research into Economic Systems: Jena, Germany (2000)
16. Volkland, H.O., Strobel, M.: Surprise, surprise - On the accuracy of the 2002 FIFA Wold Cup prediction markets (2007)
17. Spann, M., Skiera, B.: Sports Forecasting: A Comparison of the Forecast Accuracy of Virtual Stock Markets, Betting Odds and Tipsters, Working Paper, University of Passau (2007)