An Empirical Investigation of the Forecast Accuracy of Play-Money Prediction Markets and Professional Betting Markets

Christian Slamka
University of Frankfurt, slamka@wiwi.uni-frankfurt.de

Stefan Luckner
Institute IISM, Universität Karlsruhe, luckner@iism.uni-karlsruhe.de

Thomas Seeman
University of Erlangen-Nürnberg, luckner@iism.uni-karlsruhe.de

Jan Schroder
Universität Karlsruhe, Schroeder@iism.uni-karlsruhe.de

Follow this and additional works at: http://aisel.aisnet.org/ecis2008

Recommended Citation
http://aisel.aisnet.org/ecis2008/236

This material is brought to you by the European Conference on Information Systems (ECIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2008 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
AN EMPIRICAL INVESTIGATION OF THE FORECAST ACCURACY OF PLAY-MONEY PREDICTION MARKETS AND PROFESSIONAL BETTING MARKETS

Christian Slamka, School of Business and Economics, University of Frankfurt, 60054 Frankfurt am Main, Germany, slamka@wiwi.uni-frankfurt.de

Stefan Luckner, Institute of Information Systems and Management, Universität Karlsruhe (TH), 76128 Karlsruhe, Germany, luckner@iism.uni-karlsruhe.de

Thomas Seemann, Institute of Business Administration, University of Erlangen-Nürnberg, 90403 Nürnberg, Germany, thomas.seemann@wiso.uni-erlangen.de

Jan Schröder, Institute of Information Systems and Management, Universität Karlsruhe (TH), 76128 Karlsruhe, Germany, schroeder@iism.uni-karlsruhe.de

Abstract

Prediction markets have proven high forecasting performance in many areas such as politics, sports and business-related fields compared to traditional instruments such as polls or expert opinions. The advantage of real-money prediction markets is to provide participants with a clear incentive to reveal their true opinion on the outcome of an event. However, it is to date unclear whether play-money prediction markets, where participants cannot suffer any losses, perform well compared to other, more strongly incentivized instruments. Thus, the goal of this paper is to compare the forecast accuracy of play-money prediction markets with that of instruments with a monetary incentive to make as accurate predictions as possible, namely fixed-odds betting. We present the results of an empirical study that compares the forecast accuracy of a play-money prediction market for the FIFA World Cup 2006 to predictions derived from odds issued by two professional betting companies. Additionally, we compare the prediction market with two more benchmarks, namely the FIFA world ranking which is based on historic data and a random predictor. We find that the play-money prediction markets for the FIFA World Cup are about as accurate as betting markets. Moreover, the prediction markets clearly outperform the FIFA world ranking as well as the random predictions.

Keywords: Prediction Markets, Forecast Accuracy, Sports Forecasting, Betting
1 INTRODUCTION

The emergence of the Internet within the last decade has not only made many services ubiquitously available for companies around the globe, but has also created the possibility for users to share and exchange their knowledge with like-minded people. Besides the simple exchange of knowledge that can be implemented in a rather informal and scarcely structured way, more sophisticated information systems like prediction markets provide mechanisms that let users express their opinions on the outcomes of future events by the exchange of stocks. Modern IS designs which allow systems to scale up in terms of number of users, number of future events or trading frequency (e.g., Soukhoroukova, 2005) have enhanced the accessibility of such knowledge exchanges. Nowadays, the restriction of a physical meeting of participants does not hold any more. Moreover, due to the ubiquity of IS and access to it, new information is quickly dispersed, allowing for an instant reaction to events (e.g., Elberse, 2007).

The basic idea of prediction markets is to trade virtual stocks on an electronic market whose pay-offs are tied to the outcome of uncertain future events. For example, a stock can be tied to the outcome of a soccer match, paying off one (virtual) currency unit if the stock’s underlying team wins. Although the final pay-offs 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 (Glosten & Milgrom, 1985). As a result, the trading price reflects the traders’ aggregated beliefs about the likelihood of the future event. Market prices can thus be interpreted as predictions (Forsythe, Frank, Krishnamurthy & Ross, 1995). The theoretical justification of regarding market prices as estimates of the outcome of future events is founded in the Hayek hypothesis (Hayek, 1945). 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 (Fama, 1970; Fama, 1991).

Accurate forecasts are essential in many areas such as business, sports or weather forecasting, and prediction markets are considered to provide a way to improve forecasting accuracy (Spann & Skiera, 2003; Servan-Schreiber, Wolfers, Pennock & Galebach, 2004) compared to traditional forecasts such as expert predictions, polls or surveys. Good performance has also been demonstrated in corporate environments (Chen & Plott, 2002; Ortner, 2000; Plott, 2000). Thus, it is not surprising that several prominent prediction markets such as the Iowa Electronic Markets, TradeSports, NewsFutures or the Hollywood Stock Exchange have emerged. Furthermore, several major companies such as Hewlett-Packard, Google, or Microsoft are running internal prediction markets for company-specific predictions.

However, the comparison of the forecast accuracy of prediction markets with traditional methods to-date, although reasonable under practical considerations, is associated with some drawbacks. In surveys or polls for example, the forecasting accuracy is strongly dependent on the representative choice of subjects, their veridic answers to questions and the choice of statistics (Berg, Forsythe, Nelson & Rietz, 2001). In prediction markets, on the other hand, participants do not need to be selected representatively and are incentivized to reveal their true expectation about the outcome (Berg et al., 2001). Also, when questioning experts or using business meetings, which include usually at least one expert, the incentive to give correct predictions about the outcome of an event is low as experts are not explicitly rewarded or punished in monetary terms for their given predictions (e.g. Sunstein, 2006). Thus, experts are not incentivized to reveal their private information. However, in real-money prediction markets, the financial success of participants is directly tied to the ability to make accurate predictions, providing them with incentives to reveal their private information.

In this paper we are interested in a comparison of the forecast accuracy of play-money prediction markets with a forecasting instrument where the forecaster is clearly incentivized to reveal his most accurate prediction on the outcome. In fixed-odds betting, which we choose as our primary
benchmark, odds on which participants can bet money are set by the betting company. In order to avoid losses, betting companies are incentivized to make predictions as accurate as possible (Forrest, Goddard & Simmons, 2005). With large sums of money at stake, the monetary incentive to predict well is much stronger than in the described traditional forecasting instruments, but should also be stronger than in prediction markets with no money at stake in play-money markets and little money at stake in real-money markets.

In the field of sports forecasting, we study the forecasting accuracy for predicting the outcomes of soccer matches during the FIFA World Cup 2006. The forecast accuracy of a play-money market where traders could not suffer any losses is compared to two professionally operated fixed-odds betting markets. Additionally, we examine whether the prediction markets outperform a random predictor as well as forecasts that are based on historic data about the success of national soccer teams.

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 and related work about the efficiency of betting markets. Section 3 describes 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. Also, the analyzed betting markets and the other two benchmarks are outlined. In section 4, the results with respect to the comparison of the forecast accuracy are presented and discussed. Section 5 summarizes the main results.

## 2 RELATED WORK

### 2.1 Prediction Markets

Information aggregation with prediction markets does not have a long tradition in economic research. Starting with the 1988 U.S. presidential election market, the focus of the research field was mainly to study the accuracy of predictions derived from the trading activity within these markets. Political stock markets outperformed election polls in many cases as, for example, many results from the Iowa Electronic Market demonstrate (Berg et al., 2001; Forsythe, Nelson, Neumann & Wright, 1992). Moreover, research in the business forecasting area suggests that prediction markets can perform better than traditional methods such as business meetings (Sunstein, 2006). Chen and Plott (2002) show that prediction markets on sales forecasting were significantly better than official company forecasts in 6 out of 8 cases. Over the last three years, the Hollywood Stock Exchange almost perfectly predicted the Oscar award winners (Lamare, 2007) and has beaten the individual and average forecasts of five experts (Pennock, Lawrance, Giles & Nielsen, 2001). Spann and Skiera (2003) found that, regarding the hit-rate, where a “hit” is present when a forecasting instrument has the lowest mean absolute percentage error of all instruments, the Hollywood Stock Exchange is superior to two analyzed experts.

During the last decade, prediction markets were also employed in the field of sports forecasting. For instance, Servan-Schreiber et al. (2004) compare the performance of play-money versus real-money prediction markets and find no significant difference in the forecast accuracy. Moreover, both market types outperformed individual forecasts. However, the markets were very likely not completely identical with regard to number of traders, trading activity or trading fees, which might have played a role in the formation of the prediction (for a discussion see Luckner and Weinhardt (2007)). Regarding the comparison of play- versus real-money markets, Rosenbloom and Notz (2006) find that real-money markets performed better than play-money markets in case of non-sports events. In a paper which deals with the comparisons of prediction markets and fixed-odds betting, Schmidt and Werwatz (2002) analyze a 2000 European Championship market to detect 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. The random predictor performed worse than the markets’ predictions. Also, relative to the prediction markets expert bookmakers
forecasted less accurately. One of the key features of these market was the real-money investment which was required; every participant had to deposit a certain amount of money (up to 50€) and thus could suffer losses. As such, these analyzed markets were similar to the ones of the Iowa Electronic Markets, which were shown to be very accurate in the past. However, these markets differ fundamentally from the play-money market STOCCER since our market did not require any real-money investment. Participants could therefore neither lose nor win any money by revealing their expectations.

In a very recent study, Spann and Skiera (2007) analyze the forecast accuracy of a play-money market in comparison to betting odds and tipsters. Their results show that prediction markets slightly outperform betting odds, while both instruments strongly outperform tipsters. In the study presented in this paper, we also examine different market designs, take historical data into account and provide descriptive statistics about the trading actions in the analyzed markets.

2.2 Fixed-odds Betting

In fixed-odds betting, usually one or several professional experts of a company set odds which are usually not, or rarely, adjusted over time (e.g. Forrest et al., 2005). Bettors then accept or reject those bets at some time before the beginning of the decisive event. Essentially, in fixed-odds betting, outside information from potentially knowledgeable bettors is not accounted for in the odds. Studies have mostly shown that fixed-odds betting markets are efficient (Gandar, Dare, Brown & Zuber, 1998; Pope & Peel, 1989). For instance, Pope and Peel (1989) develop a linear probability model which incorporates the probabilities of the actual occurrences of the outcomes and the probabilities implicitly quoted by the odd-setters, deriving betting strategies and showing that no strategy could lead to expected positive returns. However, some inefficiencies such as the favorite-longshot bias (e.g. Cain, Law & Peel, 2000) were also detected, where favorites are undervalued and “long shots”, i.e. subjects with a very low probability of winning, are overvalued.

Forrest et al. (2005) give the main reason why especially in fixed-odds betting markets efficiency is required: “If bets are mispriced, the financial consequences for bookmakers may be serious (…)”. Although a commission fee of 15-25% is usually charged (Woodland & Woodland, 1994) which can palliate possible losses in the short run, under competition, betting companies setting the odds in consequence have a strong incentive to generate efficient, and therewith accurate, quotes.

3 DATA SETS

In this section we firstly describe the event we studied, namely the FIFA World Cup 2006. Secondly, we present the data sources used to forecast the outcome of the World Cup. This includes our prediction market STOCCER as well as sports betting odds and the FIFA world ranking. Additionally, we also employed a random predictor.

3.1 The FIFA World Cup 2006

The 2006 FIFA World Cup hosted 32 participating national soccer teams from June 9th to July 9th 2006. The preliminary rounds included eight groups with four teams each. The top two teams in each group advanced to the final rounds starting with the round of 16. The final rounds applied a sudden death system until the final game. Additionally, one game was played for the third place between the losers of the two semi-final games. In total, 64 matches were played during the tournament, of which 48 matches were played in the preliminary rounds and another 16 in the final rounds.

In the final rounds, matches that were tie after the official 90 minute match time were followed by an overtime period and, if necessary, by a penalty shootout to determine the team qualifying for the next round. In our prognosis data we count ties after the official 90 minute match time as “draw”. We did
not consider overtimes and penalty shootouts because we regarded their outcomes as more or less unpredictable. This interpretation also holds for the betting odds, where a draw was counted the same way.

3.2 The STOCCER Exchange

In order to study the prediction accuracy of sports prediction markets, we ran an exchange during the FIFA World Cup 2006. More than 1,500 traders registered for our experimental markets called STOCCER\(^1\). The web trading interface we provided is depicted in Figure 1. Our exchange started on May 15th 2006 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, more than 1,600 market transactions were executed per day.

Figure 1. Trading screen of STOCCER.

STOCCER hosted two types of markets:

- **In the championship market**, virtual stocks for all the 32 national teams taking part in the FIFA World Cup 2006 were traded. The pay-off of the virtual stocks was valued at 50 virtual currency units for the world champion and at 30 virtual currency units for the vice-world champion. Table 1 shows the pay-off values for teams reaching the semi-finals, quarter finals and the round of 16. In case a team did not reach the round of 16 the team’s virtual stocks rendered worthless at the end of the tournament.

\(^{1}\) www.stoccer.com
In the match markets for the final rounds, we traded three stocks for each of the 16 matches – one for every possible outcome of the matches: either team A wins or team B wins or the match is drawn after the second half. The virtual stock corresponding to the event that actually occurred at the end of a match was valued at 10 virtual currency units; the other two assets were worthless.

<table>
<thead>
<tr>
<th>Type</th>
<th>Number of virtual stocks</th>
<th>Final pay-off</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>Match</td>
<td>3 per match: team A wins, team B wins, tie after 2nd half</td>
<td>Event occurred: 10 Otherwise: 0</td>
<td>2 days before the matches</td>
<td>At the end of the matches</td>
</tr>
</tbody>
</table>

Table 1. **STOCCER's markets operated during the FIFA World Cup 2006.**

Concerning the financial market design these 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 (see Berg & Rietz, 2003). 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 pay-offs 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.

STOCCER operated as a play-money market and provided participants 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 on the main website and a lottery of prizes. The overall Top-100 traders with the highest deposit value after the end of the tournament 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 Euros. Although the total sum of 6,000 seems quite large at first sight it has to be considered that the probability of winning one of the prizes was quite low, taking into account the more than 1,500 traders. In addition, we weekly raffled an iPod among the 20 most active traders of the preceding week. However, we assume that a large proportion of traders were not only driven by financial rewards but also by their enthusiasm and interest in soccer.

### 3.3 Sports Betting Odds

As a benchmark for the prediction market we use the betting odds of two major German sports betting providers: Oddset and Wetten.de. Oddset is Germany's largest betting institution and is owned and controlled by the state. Wetten.de is a popular sports betting provider that is privately held.

Both bookmakers offered fixed odds which bettors could wager against. The fixed odds were set by the bookmaker at the time the odds were placed. For each of the 64 World Cup matches, bets could be placed on a win for the first team (1), a draw (0), and a win for the second team (2). All bets are referring to the score after the official time of 90 minutes. Betting quotes are stated in decimal odds – a bet quoted with 3.5 pays out 3.5 times the wagering amount deployed. A typical betting screen is illustrated in Figure 2. As bookmakers follow a commercial interest and try their best to avoid short-

---

2 [www.oddset.de](http://www.oddset.de)

3 [www.wetten.de](http://www.wetten.de) (international site also available at [www.digibet.com](http://www.digibet.com))
term losses (see section 3.3), the odds include a commission fee. This means that wagering the same amount on all three possible outcomes would lead to a 15-25% loss.

The total sales of Oddset in 2006 were 342.3 million Euros. Although we do not have disaggregated data on single categories, with the main sport in Germany being soccer and the main soccer event in 2006 being the FIFA World Championship, we guess that a considerably large amount of money has been betted on these markets for the World Cup.

3.4 FIFA World Ranking

We use the FIFA world ranking\(^4\) as another benchmark which is based on historic data only. The FIFA world ranking from May 2006 is built on a history of the last eight years and takes into account the following 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 are relevant. For each individual factor, points are assigned which are then aggregated to an index value. For most factors, complex calculations are used to determine the team’s actual state and strength.

3.5 Random Predictor

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 benchmark to evaluate the forecast accuracy of our prediction market. As we can observe three possible outcomes for each individual match, an uninformed, random guess would result in a hit-rate of 33.33% (for details see Schmidt & Werwatz, 2002).

4 EVALUATION OF THE FORECAST ACCURACY

Prediction market prices and thus the forecasts of prediction markets are driven by the information and the expectations of the users trading in the markets (Spann & Skiera, 2003). Beside historic data, traders also consider current information that is available to them as well as ongoing developments during the tournament and even during matches.

\(^4\) http://www.fifa.com/worldfootball/ranking/
To compare the forecasting accuracy of our markets to predictions derived from the random predictor, the FIFA ranking and the betting odds, 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 different instruments for the whole sample of 64 matches.

In case of the STOCCER championship market, we predict a win for the team with the higher stock price prior to the kick-off of the match. We predict a draw whenever the stock prices of two teams are equal. For the betting odds, we consider the outcome with the lowest odds as the prediction. In case of the FIFA ranking we predict a win for the team that has the better position in the ranking.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>No. obs.</th>
<th>Hit rate</th>
<th>% improvement</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Championship market</td>
<td>64</td>
<td>59.38%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oddset odds</td>
<td>64</td>
<td>57.81%</td>
<td>2.72%</td>
<td>0.799</td>
</tr>
<tr>
<td>Wetten.de odds</td>
<td>64</td>
<td>67.19%</td>
<td>-11.62%</td>
<td>0.203</td>
</tr>
<tr>
<td>FIFA ranking</td>
<td>64</td>
<td>46.88%</td>
<td>26.66%</td>
<td>0.042</td>
</tr>
<tr>
<td>Random draw</td>
<td>64</td>
<td>33.33%</td>
<td>78.14%</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

Table 2. Comparison of forecast accuracy (all matches).

When comparing the hit rates of the championship market, the betting odds, the FIFA ranking and the random predictor for all 64 matches, we find that championship market indeed yields a higher hit rate than the FIFA ranking and the random draw model. The difference in the hit rate of the prediction market and the two other instruments is significant in both cases (p < 0.05). The forecasts can be improved when using a prediction market instead of these two instruments. Table 2 shows the percentage of improvement when one replaces the respective alternative instrument with a prediction market. When comparing the championship market with the two sports betting odds, we find the Stoccer championship market to outperform Oddset on the one hand, while Wetten.de is superior to the Stoccer championship market on the other hand. In both cases, however, the difference in the hit rate is not significant. This can be considered as a clear success for the prediction market because the forecast accuracy is similar as in case of betting odds, even though the market was a play-money market and the likelihood of draws is underestimated in the championship market. Based on market prices in the championship market we would only predict a draw if the prices were exactly the same – which is rather unlikely. This also hold for the FIFA ranking where we would predict a draw if two teams were ranked equally.

As described in section 3.2, 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 instruments. The results of this comparison are shown in Table 3.

---

5 Percentage of improvement of championship market over alternative instrument
6 Chi-square test for difference to hit rate of championship market
For the last 16 matches of the tournament, the hit rate of the match markets is again 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 again be the fact that the likelihood of draws is underestimated in the championship market. Furthermore, traders in match markets can focus on the outcome of one match at a time instead of trying to predict the course of the entire tournament. In the final rounds, the match markets also outperform predictions based on the championship market as well as the betting odds – although the difference is not statistically significant. Moreover, the forecast accuracy of the championship market is the same as the forecast accuracy of the betting odds.

At first sight, it is somewhat surprising that the hit rate for the championship market, the betting odds and the FIFA ranking is on average lower for the last 16 matches than it is when taking into account all 64 matches. However, 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 and analyzed a sports prediction market we were running during the FIFA World Cup 2006. The goal of our paper was to study the forecasting accuracy of a play-money prediction market where traders could not suffer any losses compared to forecasts that are based on odds which were set by professional bookmakers. These bookmakers’ odds are strongly incentivized to give predictions as accurate as possible due to the high monetary investments that are at stake. We could not find any statistically significant difference in the forecasts with respect to hit rates. Both methods provide a comparable forecasting accuracy. As another benchmark, we used the FIFA world ranking for our study, which is calculated based on the historic performance of the national soccer teams. We could observe a significantly better forecasting accuracy for our prediction markets as well as betting odds. Moreover, the random predictor is clearly inferior to the forecasts of our prediction markets.

By demonstrating the competitiveness of play-money prediction markets compared to sports betting forecasts, our results align with those attained by Schmidt and Werwatz (2002). However, the markets in their research used real money as an incentive and participants hence were directly punished financially in case of a poor forecasting performance. Play-money prediction markets as the one we analyzed in this paper, however, are much easier to set up and to operate than real-money prediction markets.

---

7 Percentage of improvement of match market over alternative instrument
8 Chi-square test for difference to hit rate of match market
markets due to legal and technical reasons, which also facilitates the development of corresponding IS platforms. In the research line of play-money markets, our results align with the work of Spann and Skiera (2007). When keeping in mind that betting odds have shown to be extremely good predictors and that similar instruments are mostly non-existent in other fields of application beyond sports forecasting, real-money as well as play-money prediction markets indeed seem to be a very promising forecasting instrument.

References