How accurate do markets predict the outcome of an event? The Euro 2000 soccer championships experiment*

Carsten Schmidt† Axel Werwatz‡

22 November 2001, revised 12 March 2002

Abstract

For the Euro 2000 Soccer Championships an experimental asset market was conducted, with traders buying and selling contracts on the winners of individual matches. Market-generated probabilities are compared to professional bet quotas, and factors that are responsible for the quality of the market prognosis are identified. The comparison shows, that the market is more accurate than the random predictor and slightly better than professional bet quotas, in the sense of mean square error. Moreover, the more certain the market predicts the outcome of an event the more accurate is the prediction.

Keywords: experimental asset markets, prognosis, market efficiency
JEL Classification: C93, D4, G1

*The financial support of the German Science Foundation (Grants nos. SFB 373, A1 and C5) and the International Institute of Infonomics/Heerlen is gratefully acknowledged. We thank Jan Hansen and Rudolf Müller for discussion and their support in conducting the experiment. We greatefully acknowledge insightful conversations on this subject with Werner Güth, Manfred Königstein, Loreto Llorente, Christian Schade, Peter Schmidt, Martin Strobel and Harald Uhlig and comments from audiences at Barcelona, Berlin, Heerlen and Wulkow.

†Max Planck Institute for Research into Economic Systems, Strategic Interaction Group, Kahlaische Str. 10, D-07745 Jena, Germany, Email: cschmidt@mpiew-jena.mpg.de

‡Institute of Statistics, Sonderforschungsbereich 373, Department of Economics, Humboldt-Universität zu Berlin, Spandauer Str. 1, D-10178 Berlin, Germany, Email: axel@wiwi.hu-berlin.de

1 Introduction

This paper deals with the predictive performance of markets drawn on the outcome of uncertain events. Recently, scientist and companies have attempted to predict quality and duration of complex industrial projects (Ortner, 2000) or future sales of investment goods (Chen and Plott, 1998; Plott, 2000), on the basis of experimental asset markets. This kind of market is operated as a futures exchange, where subjects buy and sell money-backed contracts drawn on the outcome of highly uncertain events. As opposed to a prognosis based on expert advice, such a market is able to aggregate the information of a large number of participants.

The market as a combination device has been studied in the information aggregation literature, demonstrating that markets are able to aggregate information fast and efficient (Forsythe and Lundholm, 1990; Plott and Sunder, 1988). Political stock market experiments in the U.S. were able to generate more precise election result forecasts than traditional polls (Forsythe et al., 1999), and these market forecasts proved to be slightly more accurate than polls, also in a European context (Berlemann and Schmidt, 2001). The idea to interpret prices of contracts in terms of an assessment of the event's probability goes back to Hanson (1995).

In this paper, the predictive accuracy of experimental asset markets is investigated in the context of an Internet experiment dealing with a sports event. To be more precise, we are going to evaluate the predictions of 21 markets, made on the winning team of a specific soccer match during the European soccer Championships in 2000. To evaluate the experimental results, we took a benchmark model, the rolling dice, and empirical expert data, professional bookmakers' odds, as a basis for comparison.

A new kind of market game, so-called "fantasy markets", emerged on the Internet. These market games make use of the technological progress that came with the World Wide Web and, unlike the described market experiment, do not involve real money investment. Regularly, traders with extraordinary performance win prices, high-quality consumer goods, while all the other traders do not receive monetary rewards. This implies that traders are not subject to monetary losses. On the one hand, this kind of game attracts a large number of participants. On the other hand, traders often cancel participation when their portfolios face severe losses. Several researchers used data of market based games to find out whether the posted prices correlate with the events observed. For movie box office returns Pennock et al. (2001) discovered a strong correlation to the Hollywood Stock Exchange movie forecast[1] and the predictions of a movie expert[2]. In some cases they found the forecasts generated by fantasy market data more reliable than expert opinions. Furthermore, an empirical study was conducted as to what degree transfer fees of baseball players and market

[1] `http://www.hsx.com`
[2] `http://boxofficemojo.com`

prices in an artificial stock market[3] were indicators for the performance of individual baseball players (Sade, 2001).

Contracts drawn on the outcome of sports events have a historic background in commercial stock markets. In the seventies, at a time, when the German stock market played a minor role, an active non-stock market trading of gold medals (Olympics), World Champion titles (soccer) and results of political parties (federal elections) was observed (Handelsblatt, 1998). At that time private bets were backed with actual deals. When transactions were conducted during regular trading sessions the exchange supervision was called to act (Handelsblatt, 1998). Yet, brokers are not the only target group fascinated with trading sports contracts. In the experimental economics community there were at least two market based experiments for soccer events prior to our analysis. Experimental economists at CREED/Amsterdam conducted an unpublished market for the World Championships in 1994. For the 1998 World Championships we ran a small market at Humboldt University where participants from 10 different countries traded contracts drawn on soccer teams via the Internet. Several "fantasy markets" were also conducted for major soccer events. In the realm of finance, an experiment was set up for the 1998 World Championships to test the influence of market makers on large-scale stock markets (Bochow et al., 1999). The authors are aware of a least three commercial markets that were conducted during the Euro 2000 Championships.[4]

As we proceed, we are going to compare the bookmaker's odds with the market prognosis. Empirical data on sports betting odds is a data source for empirically testing of the market efficiency hypothesis in finance. Unlike stock markets, the fundamental values of bets are revealed after the result of the event is published. Racetrack betting markets odds are thought to be good estimates of win probabilities (Thaler and Ziemba, 1988) and some observers find point spread betting markets to deliver unbiased predictions of the event's outcome (Gandar et al., 1998). Nevertheless, in the empirical literature there are several statistical biases reported, most prominently the favorite-longshot bias in horse track racing (Thaler and Ziemba, 1988), which states that betters overestimate the longshots' probability of winning. Most empirical studies that deal with economic evidence for the market efficiency hypothesis conclude, that betting strategies which take advantage of known biases might be successful, even when betting fees are taken into account (Vergin and Scriabin, 1978; Gandar et al., 1988; Dixon and Coles, 1997), yet these market inefficiencies tend to dissipate over time (Gray and Gray, 1997). Unfortunately, we were not able to test for statistical effects due to the limited scope of the available market data. Similar to an economic testing of the market efficiency hypothesis, we explore the question whether subjects are able to earn profits by using the bookmaker's odds to participate in the market and the other way around.

[3] `http://wallstreetsports.com`

[4] The public TV station ZDF together with the online broker Consors organised a market in Germany; The online catalog Yahoo and the game organiser Neopoly also conducted a market for this event with participants from Germany, England and France; Another market was conducted by Phillips corporation in the Netherlands.

The game of soccer itself has been studied by Palomino et al. (2001). They model the probability of a team to score during a game and evaluate empirical data. This approach is relatively novel in that it takes events during a soccer game into account as most of the research on soccer has disregarded this. Our experimental design tries to explicitly exclude events during the game, by closing the individual match markets prior to the start of the game.

The contributions of this paper are as follows. We compare the market-generated prices to a benchmark model and to professional bet quotas, on the basis of statistical hypothesis testing. We identify the variables that are responsible for the quality of the markets' predictive success. We support the hypothesis, that the market yields significantly more accurate results than the random predictor and that it is more accurate than professional bet quotas in the sense of mean square error. Moreover, the more certain the market predicts the outcome of an event, the more accurate is the prediction, while this effect can not be observed for the bookmakers' odds.

In the following section we give an overview on the organization of the soccer event, as well as on the setup of market and bookmaker. Section 3 presents a description of the experimental setup. Section 4 reviews the results. Open research questions for further exploration and concluding remarks are given in section 5.

2 Overview

2.1 The event

The Euro 2000 Soccer Championships were held from June 10th to July 2nd in the Netherlands and Belgium. 16 European national soccer teams qualified for this tournament. On the whole, 31 matches were scheduled, with 24 matches for the group stage and seven matches for the play-off stage. The rules of this tournament are briefly presented in the following. The tournament is organised in two parts. During the group stage four groups are drawn each including four teams. In each group the teams play round robin. For each match the winning team receives three points, while both competing teams receive one point for a draw. Teams are ranked according to their total number of points. In case that more than one team wins the same number of points, the goal difference is used as a tie breaker. Out of each group the two top teams enter the play-off stage. Starting with the play-off stage, the outcome of a match has to be decided. When no decision is made after 90 minutes of regular playing time, the competitors need to play extra time. In the 30 minutes of extra playing time, the first goal recorded - also called the *golden goal* - decides the match. When no decision is reached in extra time, the match is decided on penalties.

Table 1: Outcomes of the matches of the European Soccer Championships 1980-1996

	Result[a]		1980-1996
Group stage		n	72
	team of record wins	1	44.4%
	draw	0	30.6%
	team of record looses	2	25.0%
Play-off stage		n	18
	decision after 90min	1&2	50.0%
	no decision a. 90min	0	50.0%

[a] Source: `http://www.fussballarchiv.de`

One may suspect, that the three possible outcomes of a match in the European Championships are not equally likely. Two reasons, the home advantage and a low likelihood of draws, are regularly mentioned. Unlike compared to sports leagues, there is (about) no home advantage for the competing teams in such a type of tournament. All teams, except the team of the organizing country, play on foreign ground. To answer the question whether a draw occurs in one out of three matches, we evaluated the results of all European Championship matches from 1980 until 1996.[5] The analysis revealed that in the group stage 30.6% of all matches led to a draw. Out of all matches in the play-off stage 50.0% were decided after regular time. Table 1 summarises these historic results.

The empirical data support the hypothesis, that the three possible events for a match are equally likely to occur. Therefore, we set-up the random predictor as a benchmark model. In the following, we refer to the team from whose perspective the result is defined as the *team of record*.

2.2 The bookmaker

The structure of soccer betting, so-called "toto", is quite simple. Recall the three possible outcomes of a match: *team of record wins*, *draw* and *team of record looses*. These outcomes are assigned to values 1, 0 and 2, respectively. The bookmaker offers a fixed quote for every outcome of the match. A quote of 2.5 for a win on the team of record implies that for a one Euro bet the better receives 2.50 Euro in case of win and zero in the other cases. During play-offs the rule is changed and the bet is redefined as "winning in regular playing time". Therefore, matches that are decided after regular time are assigned 0, similar to a draw.

[5] The soccer tournament is conducted every four years and follows the rules laid down since 1980. From 1980 to 1992 eight teams participated in the tournament with the group stage consisting of two groups with four teams, respectively. Since 1996 the number of teams was increased to 16.

For the evaluation of the experimental data in the later sections, we use data of the toto sports betting system "ODDSET" run by the German state-owned lottery. Betters can participate via Internet and at any location in Germany, where state owned lotteries are accepted. The participation is restricted to German residents. In the lottery betters choose combinations of toto bets. Combinations can be selected from up to 90 bets, typically the soccer matches from Tuesday to Monday.[6] The objective of a bet is a combined prediction of 3 to 10 bets. For example, if the better predicts all matches (out of three) correctly, the pay-off amounts to the multiplication of the three toto quotes times the investments, in case of failure the better receives nothing. Participants can invest in bets of 2.50, 5, 10, 15, 20, 25, or 50 Euro.[7] The investment is restricted to 300 Euro per week and the maximum pay-off is limited to 50,000 Euro.

Apart from the decision on how to set the quotes considering the probability of an event's outcome, the bookmaker's objective is to set the quote in a way that the better's investments distribute evenly on all three outcomes. In this case the bookmaker takes no risk. The bookmakers' implicit prognosis, computed from the quotes, may therefore be biased by a demand effect.

2.3 The market

The design of the market experiment was similar to that of the Iowa Electronic Markets (Forsythe et al., 1992; Forsythe et al., 1999), that host markets on the outcome of political and economic events. In order to participate the trader had to invest real money. The maximum investment of an individual trader was restricted to 50 Euro.[8] With the investment the experimenter provided a trading account on the electronic marketplace. On the marketplace participants were given the option to take part in several individual markets. Each market was organised as follows. On the primary market subjects bought and sold zero-risk unit-portfolios from and to the bank. A unit-portfolio included one of each different contracts on the market. The bank guaranteed to buy and sell this unit-portfolio at a constant price during the entire experiment. On the secondary market the subjects were asked to buy and sell individual contracts. As market institution a continuous double auction with queues was introduced to implement transactions between the participants. As opposed to markets with an initial public offering (IPO), that use a fixed amount of contracts, the institution of the unit-portfolio allows for a dynamic emission of contracts during the experiment. Finally, the zero sum property of the market mechanism guaranteed that the organiser did not make any profits.

[6] There are other kinds of bets offered, like point spreads, and bets on other sports events in non-soccer season. Due to the objective of the paper we concentrate on soccer matches, especially on those conducted during the Euro 2000 Soccer Championships.

[7] We use an exchange rate of 2 DM for 1 Euro for ease of processing, the actual exchange rate is 1.95583 DM.

[8] German participants received an additional five Euro on top of their investment, the total value of the portfolio of a German trader amounted to a range of six to 55 Euro.

The experiment was accessible via the Internet and started five days prior to the soccer tournament. On the whole, it was conducted for 28 days during which time it was accessible around the clock. Participants were recruited from two portal sites, one located in Germany and one in the Netherlands. In Germany the market was made public by the name ribaldo7.de and by the name of voetbalmarkt.nl in the Netherlands. The Dutch organiser only permitted members of the University of Maastricht to participate. To participate at the German portal site subjects were required to have a German bank account.

Altogether, 238 subjects participated in the Euro 2000 market - 188 from Germany and 49 from the Netherlands. The subjects were mostly employees (50%) and students (27%). The average age was 30 years with a standard deviation of 9.6. Not surprisingly, only 10% of the participants were female.

The subjects' commitment was remarkably high. The total investment was 8,933 Euro with an average of 37.53 Euro (18.25 std.dev.). The 238 traders conducted 2,950 trades, on average more than 24 trades per participant. Standard web sites, like search engines and catalogs, on average receive well below 10 page impressions per visit (PI/Visit).[9] The Euro 2000 market made participants and visitors generate a remarkable average of 15 PI/Visit.

3 Experimental setup

The experiment consisted of a set of markets, that the participants were asked to choose from. Each time a trader logged on to the experiment, he was right away connected to the Championship market. This market was open for the entire time of the experiment, which implies that subjects could trade during the matches. The market consisted of 16 contracts representing the teams that qualified for the Euro 2000. The pay-off rule of this market was a very simple one, winner-takes-it-all design. The contract of the Euro 2000 champion yielded a pay-off of 1,000 Centicent[10] all other 15 contracts yielded a zero pay-off.

In addition, participants were asked to participate in markets for individual matches.[11] These markets were organised similar to soccer toto. For matches during the group stage a market

[9]Page impression is a measure used in web site marketing to quantify the number of requested web pages. As opposed to the measure "hits", where the number of transmitted files is aggregated, a request of a web page, also containing several files, is counted only once. A visit is a series of one or more page impressions, served to one user, which ends after a 30 minutes or more break between successive page impressions for that user.

[10]The Centicent was used as denomination during the experiment. 100 Centicent convert into one Euro cent.

[11]We offered four group markets that are not evaluated in this paper. For the sake of completeness we briefly describe the organization of these markets. Every group market contained for each of the four teams a different contract. The contracts of the two teams that participated in the play-off stage payed off 50 Centicent each, the other two zero. The markets were liquidated right after the results of the groups were available.

consisted of three different contracts: *team of record wins*, *draw* and *team of record looses*. In the tournament's play-off stage the matches continued until a winner was determined. For the seven matches of the play-offs, the number of contract types was reduced to two. In both settings, the group and the play-off stage, the winning contract yielded a pay-off of 100 Centicent and whereas the other contracts, two in the group stage, and one in the play-off stage, yielded a zero pay-off. The markets were open between three and five days prior to the matches and closed at the beginning of the corresponding match. In such a setup all information available to the traders should be included in the closing price of a contract. The day after the match all contracts were liquidated and the money was made available to the traders via the online accounts. This way, subjects regained liquidity to invest in upcoming markets.

The Championship market was open all the time, whereas all other markets were open for a limited time only. At the utmost, traders could participate in nine parallel markets. Figure 1 displays a stylised time sequence of the opening times of the markets.

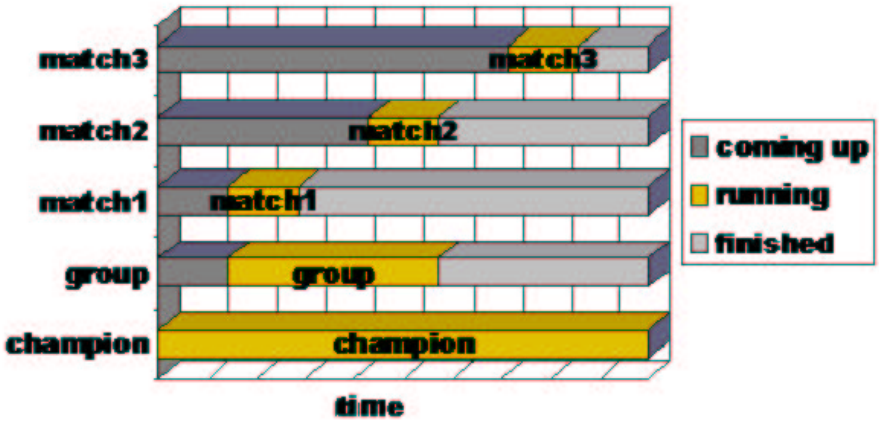

Figure 1: Stylised sequence of opening hours of different markets

The following hypotheses are evaluated in the study: (1) the markets should predict the outcome of matches more precise than a rolling dice, (2) the markets' quotes should be meaningful: the more certain the market the more often the prediction of the market should be right, and (3) the markets should generate more accurate probabilities than the bookmaker's odds. There are several reasons for these hypotheses. First, the market is a zero sum game, with investments payed back to the participants. Second, the markets do not have a demand effect due to a home team bias on the

state of the participants. In case of the experiment German and Dutch participants were active on the same market, while the German bookmaker allows for German betters only. Finally, the bookmaker needs to include the demand for each of the three bets in the quotes. Therefore, the bookmaker's implicit forecast of the matches' outcome is possibly biased.

4 Results

4.1 Data

Altogether, 31 matches were played during Euro 2000. For 21 of these matches we ran match markets.[12] The prices of the contracts in these markets just before kick-off can be used to generate predictions of the match's outcome. The converting of contract prices into predictions is straightforward: with the market's predicted outcome being the one with the highest contract price.

The contract prices just before kick-off (i.e., the closing prices, with the individual match markets closing right before the beginning of the respective match) should incorporate all relevant information available to the market participants at this time. Indeed, if the match markets efficiently aggregate information, there should not be any other observable variable for the improvement of the prediction.

A second set of predictions can be derived from the contract prices of the competing teams in the Championship market. Just before the start of the match we take as the Championship market's predicted winner the team with the higher contract price. Obviously, this prediction rule will never predict a draw, except in the unlikely event that both teams have identical contract prices.

Finally, from the ODDSET Internet site[13] we obtained bookmaker quotes for 31 matches. At the bookmaker's the bet with the lowest quote is expected to be the most likely outcome, while on the market, the contract with the highest price is expected to pay-off.

4.2 Is the market as reliable as a random predictor?

The goal of this paper is to study an experimental market's capability to predict uncertain events. A natural reference point for the market's performance is a random mechanism that generates uninformed, unsystematic predictions, like the flipping of a coin. Indeed, some observers of soccer

[12] Due to technical difficulties we were not able to conduct match markets for the first 10 matches of the Euro 2000.
[13] `http://www.oddset.de`

argued that the outcome of games is entirely random.[14] Therefore, we formally test the null hypothesis that the market delivers uninformed, random predictions.

4.2.1 Distribution of the test statistics under the null hypothesis

Let X_n be the number of correct predictions in n trials. To perform the test, we need to derive the distribution of X_n under the null hypothesis. At this point, we introduce notation, that describes the games' outcomes as well as their predictions and probabilities. For reasons of simplicity, we ignore the fact that some of the Euro 2000 games had three possible outcomes in this paragraph (the matches in the group stage) while others had only two (the games in the play-off stage).

Let us suppose that each game has three possible outcomes: *team of record wins* (1), *draw* (0), and *team of record looses* (2), and define the random variable G_i accordingly, where i is used to index the games. Let q_{ij} be the true, objective probability that game i has outcome j. For instance, q_{70} indicates the probability that game 7 ends with a draw (outcome 0). Hence,

$$G_i = \begin{cases} 0 & \text{with probability } q_{i0} \\ 1 & \text{with probability } q_{i1} \\ 2 & \text{with probability } q_{i2}. \end{cases}$$

Let R_i be the prediction of a random predictor of game i's outcome. For the time being, we will assume that the random predictor forecasts the outcome of each game independently and identically according to the following probabilistic rule:

$$R_i = \begin{cases} 0 & \text{with probability } \frac{1}{3} \\ 1 & \text{with probability } \frac{1}{3} \\ 2 & \text{with probability } \frac{1}{3} \end{cases}$$

Note the probability that the random predictor correctly predicts the outcome of game i as $P(R_i = G_i)$. It is obvious to show (see section A of the appendix) that $P(R_i = G_i)$ is 1/3 for each game, even if the objective probabilities q_{ij} vary across matches. Consequently, the probability, that the random predictor will deliver x correct predictions in n trials is given by the binomial distribution with the parameters n and 1/3. That is,

$$P(X_n = x) = \left(\begin{array}{c} n \\ x \end{array} \right) \left(\frac{1}{3}\right)^x \left(\frac{2}{3}\right)^{(n-x)}$$

[14] See, for instance, the article titled "Sammer, Kohler, Nemax & Co. Some similarities between the stock market and soccer are striking - Is it all random?" in the April 17, 2000 issue of "Der Tagesspiegel".

4.2.2 Conducting the test

From the distribution of the test statistics under the null hypothesis we are able to obtain the critical values for the desired significance level. Considering that of the 21 matches seven had two possible outcomes while fourteen had three possible outcomes, we obtain the exact distribution of the test statistics as a sum of two binomial distributions. Specifically, $X_n = Y_{14} + Z_7$ with $Y_{14} \sim BV(n = 14, 1/3)$ and $X_7 \sim BV(n = 7, 1/2)$.

Using a significance level of $\alpha = 0.05$ the critical value for a one-sided test in the positive direction is $x_{cr} = 13$, as $P(X_n > 13) = 0.037 < 0.05$ under the null. It turns out, that match markets predicted the outcome of the game correctly in 15 out of 21 cases. Since 15 exceeds the critical value 13, we can reject the null hypothesis that the market is an uninformed, random predictor.

4.2.3 Distribution of the test statistics under the alternative hypothesis

The alternative hypothesis implicit in the one-sided test is that the predicting of the outcome of matches, by market prices is better than using a coin. But even if predictions derived from market prices are indeed superior, they need not be correct all the time. Stated differently, the number of correct predictions, X_n, remains a random variable until all games have been played.

In the previous section, we were able to derive the exact distribution of X_n *under the null* hypothesis by showing that X_n is the sum of two binomially distributed random variables. However, deriving the precise form of the distribution of X_n *under the alternative* hypothesis is impossible since predictions are no longer Bernoulli trials (independent, with a constant probability of correctly predicting the outcome of a game). Rather, market participants observe the teams' performances during the course of the tournament, with market prices, and hence market probabilities, are conditional on the outcomes and events of previous matches.

Still, carrying out a right-sided test, i.e. rejecting the null hypothesis if (too) many games are correctly predicted, can be formally justified (see section B of the appendix). This can be done by showing that if the market's assessment of the probabilities of a game's outcome, as revealed by the market prices, is equal to the true, objective probability then the expected number of correct predictions is higher for the market than for the random predictor.

4.3 Does the market beat the odds?

In this section we are going to examine the question whether market prices are more accurate predictors than the quotes, posted by Germany's nationwide bookmaker, ODDSET.

4.3.1 Total number of correct predictions

For the 21 matches, for which market predictions were available, we find that the bookmaker's quotes yield (about) the same number of correct predictions. There was only one instance in which the market's prediction differed from the bookmaker's: the first-round match Portugal vs. Germany, which the market predicted correctly. Moreover, there were two matches for which the bookmaker quoted equal odds for the events *team of record wins* and *team of record looses*, therefore not yielding a definite prediction of the game's outcome. This leaves us with a total number of fourteen correct predictions derived from the bookmaker's odds, compared to the 15 correct predictions of the market experiment.

4.3.2 Squared Prediction Loss

In the previous sections we focused on the qualitative predictions derived from prices and odds and did not take full advantage of the quantitative information inherent in their actual magnitude. After all, we can assume that market participants are much more confident that the team of record will win if its contract price is 99 rather than say, 51, while the qualitative prediction, *team of record wins*, is identical.

If the team of record turns indeed out to win, then a prediction based on a price of 99 should be regarded superior to a qualitatively equally correct prediction based on a price of 51. On the other hand, if the team of record loses, then the prediction based on 99 is to be regarded inferior to the qualitatively equally incorrect prediction based on 51.

The squared prediction loss is a criterion that takes these quantitative errors into account. In the case of matches with only two outcomes its definition is straightforward (the extension of the criterion to three outcomes is presented in appendix C). Taking the market's predictions as an example we get:

$$SPL_m = \sum_{i=1}^{n} \{I(G_i = M_i) - P_i^*\}^2 = \sum_{i=1}^{n} SPE_i$$

That is, for game i the squared prediction error SPE_i can take on values from 0 (game is correctly predicted ($I(G_i = M_i) = 1$ with probability $P_i^* = 1$) to 1 (game is incorrectly predicted ($I(G_i = M_i) = 0$) with probability $P_i^* = 1$).

For the 21 matches, with market prices available, the SPL is between 0 and 21. We find that the match markets have the smallest squared prediction loss ($SPL_m = 3.25$), when compared to the Championship market ($SPL_C = 4.52$) and the bookmaker ($SPL_O = 5.71$).

Table 2: Logit fit for individual games markets

Model[a]		Coef.	Std. Err.	t	P-value
1	$P^*_{m,i}$	11.97	5.64	2.123	0.034
	constant	-6.32	3.25	-1.947	0.052

[a] Dependent variable: $Y_{m,i}$; $n = 21$.

4.4 What makes the market predict accurately?

Further, we are concerned with the variables that make the market predict accurately. The following logistic regression investigates the dependence of the market's predictive success on variables, measuring market activity and market confidence. The dependent variable presents the correct prediction of an event's outcome.

$$Y_{m,i} = I(M_i = G_i) = \begin{cases} 1 & \text{if } M_i = G_i \\ 0 & \text{otherwise} \end{cases}$$

The explanatory variables include market confidence, calculated by $P^*_{m,i} = \max_j P_m(G_i = j)$ and a set of variables that are commonly used in the information aggregation literature. Berg et al. (1997) state, that the following variables influence the quality of the prediction: (1) number of contract types traded in a market (complexity of the market), (2) pre-election market volumes, and (3) differences in weighted market bid and ask queues on election eve. Berlemann and Schmidt (2001) find evidence for German data, that the absolute error rate of market prices, defined by the absolute difference between price of unit-portfolio at the bank and on the market, influences the market's prediction. In addition, we also tested whether number of traders, total volume and price volatility can account for the quality of the prediction.

It turned out, that only the market confidence $P^*_{m,i}$ has a significant effect. The effect is positive, therefore we conclude that the more confident the market, the more accurate the predictive performance. When we include further variables in the regression, including the absolute error rate of market prices citedecade and price volatility measured by the standard error of the price time series, we observe that the effect takes the expected direction, though the corresponding coefficients are not significant. For the following variables we observed a non-intuitive sign of the coefficient: market volume and number of contract types. The results of the logit regression are summarised in table 2.

Finally, looking at the data of the bookmaker, we asked whether the confidence in the event's outcome is a significant factor. Therefore, we reran the logit regression for the highest probability

of the ODDSET prediction $P^*_{m,i} = \max_j P_m(G_i = j)$. We observed, that the effect of confidence noticed at the bookmaker's is not significant.

4.5 How to bet?

In the last step we ask which strategy a participant should choose to participate in the different institutions. To explore this question we will provide data on the profit of hypothetical one Euro bets at the bookmaker and hypothetical one Euro investments in the market using the last traded price. We use the 21 games which have been covered by the match markets. We want to remark, that the absolute figures have to be considered in the light of the actual results of the Euro 2000. When compared to the historic results presented in table 1 relatively few draws did occur during the Euro 2000.[15]

One idea would be to use the information of the bookmaker and trade on the market, the other way around betters could use the information of the market to engage in ODDSET. With regard to the winner of the match the qualitative predictions of both institutions differed only for one match. Therefore, using the likely winner of one institution to engage in the other seems not to generate additional information. For completeness, we calculated the return for a one Euro bet on each of the 21 games. When using the ODDSET quotes to buy the favorite on the market the profit would amount to 1.97 Euro, the other way around the profit at ODDSET would be 2.80 Euro.

In the previous sections it turned out that the market probabilities differ from the bookmaker's probabilities. In the next step we ask how this effect can be used for the engagement in the different institutions. Overall, the market assigns relatively higher win probabilities to the favorites, whereas the bookmaker assigns relatively higher probabilities to the under-dog. Therefore, a strategy to engage in the favorites is relatively higher rewarded at the bookmaker, and a strategy to engage in the under-dog pays-off at the market game. For the particular experiment, the under-dog strategy would have payed-off -12.1 Euro at ODDSET and -7.93 Euro in the market. The favorite strategy has the result described in the above paragraph.

5 Conclusions

In this paper the predictive results of an experimental asset market drawn on a sports event were presented. The evaluation of the experiment's results was done by comparing to a benchmark model, the rolling dice, and the implicit prognosis of bookmaker's odds. The predictions were compared on the basis of statistical hypothesis testing.

[15] The Euro 2000 observed 16.7% draws out of all matches in the group stage.

It turned out, that the market predicts more accurately than the random predictor. The qualitative predictions of the bookmaker's odds are not different from the market's prediction, yet the market was more accurate in terms of a quantitative mean square error. Moreover, the more certain the market predicts the outcome of an event, the more accurate is the prediction. A similar effect could not be observed with regards to the bookmaker's odds.

Further research should aim for a larger data sample, e.g. the complete season of a soccer league. This would provide the opportunity to test the market's prediction for systematic biases. With regards to our work, the well documented biases in sports betting markets, which are described in the market efficiency literature, might be evaluated by means of a double auction market institution. Finally, the relation of market prices and events during a sports game might be evaluated.

References

Berg, J., Forsythe, R. and Rietz, T. (1997). What Makes Markets Predict Well? Evidence from the Iowa Electronic Markets, *in* W. Albers, W. Güth, P. Hammerstein, B. Moldovanu and E. van Damme (eds), *Essays in Honor of Reinhard Selten*, Springer-Verlag, Berlin, pp. 444–463.

Berlemann, M. and Schmidt, C. (2001). Predictive Accuracy of Political Stock Markets - Empirical Evidence from a European Perspective, *Discussion paper*, Sonderforschungsbereich 373, Humboldt-Universität zu Berlin.

Bochow, J., Nguyen, D., Raupach, P. and Wahrenburg, M. (1999). What do Market Makers achieve? Evidence from a large scale experimental stock market, *Discussion paper*, University of Frankfurt.

Chen, K.-Y. and Plott, C. (1998). Prediction markets and information aggregation mechanism: Experiments and application, *Technical report*, California Institute of Technology.

Dixon, M. and Coles, S. (1997). Modelling Association Football Scores and Inefficiencies in the Football Betting Market, *Applied Statistics* **46**: 265–280.

Forsythe, R. and Lundholm, R. (1990). Information aggregation in an experimental asset market, *Econometrica* **58**(2): 309–347.

Forsythe, R., Nelson, F., Neumann, G. and Wright, J. (1992). Anatomy of an Experimental Political Stock Market, *American Economic Review* **82**: 1142–1161.

Forsythe, R., Rietz, T. and Ross, T. (1999). Wishes, Expectations and Actions: A Survey on Price Formation in Election Stock Markets, *Journal of Economic Behavior and Organization* **39**: 83–110.

Gandar, J., Dare, W., Brown, C. and Zuber, R. (1998). Informed traders and price variations in the betting market for professional basketball games, *Journal of Finance* **53**(1): 385–401.

Gandar, J., Zuber, R., O'Brien, T. and Russo, B. (1988). Testing Rationality in the Point Spread Betting Market, *Journal of Finance* **43**(4): 995–1008.

Gray, P. and Gray, S. (1997). Testing Market Efficiency: Evidence From The NFL Sports Betting Market, *Journal of Finance* **52**(4): 1725–1737.

Handelsblatt (1998). Vor der Bundestagswahl. Der neue Handelsblatt-Wahl-Dax. Kohl oder Schröder: Auf wen wetten die Börsianer?, *Newspaper Article 61*, Handelsblatt 3/27/1998.

Hanson, R. (1995). Could gambling save science? Encouraging an honest consensus, *Social Epistemology* **9**(1): 3–33.

Ortner, G. (2000). Aktienmärkte als industrielles Vorhersagemodell, *Zeitschrift für Betriebswirtschaft, ZfB-Ergänzungsheft* **1/2000**: 115–125.

Palomino, F., Rigotti, L. and Rustichini, A. (2001). Skill, Strategy, and Passion: an Empirical Analysis of Soccer, *Technical report 129*, Tilburg University, Center for Economic Research.

Pennock, D., Lawrence, S., Nielsen, F. and Giles, C. (2001). Extracting Collective Probabilistic Forecasts from Web Games, *Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2001)*.

Plott, C. (2000). Markets as Information Gathering Tools, *Southern Economic Journal* **67**: 2–15.

Plott, C. and Sunder, S. (1988). Rational expectations and the aggregation of diverse information in laboratory securities markets, *Econometrica* **56**(5): 1085–1118.

Sade, O. (2001). Who Better Predicts Ex-Ante the Ex-Post Value of a Factor of Production - Managers or the Market? An Experimental Approach Using Baseball Data, *Discussion paper*, University of Utah.

Thaler, R. and Ziemba, W. (1988). Parimutuel Betting Markets: Racetracs and Lotteries, *Journal of Economic Perspectives* **2**(2): 161–174.

Vergin, R. and Scriabin, M. (1978). Winning Strategies for Wagering on National Football League Games, *Management Science* **24**(8): 806–818.

A Random predictor

Suppose that game i has three possible outcomes, denoted by 0, 1, and 2, respectively and let q_{ij} be the objective probability that game i has outcome j ($j \in \{0,1,2\}$. Let G_i be the discrete random variable that records game i's outcome and R_i be the prediction of game i's outcome obtained by randomly choosing between the three possibilities. The probability that this random prediction will be equal to the actual outcome of the game is given by

$$\begin{array}{rcllllll} P(R_i = G_i) & = & P((R_i = 0 \cap G_i = 0) & \cup & (R_i = 1 \cap G_i = 1) & \cup & (R_i = 2 \cap G_i = 2)) \\ & = & P(R_i = 0 \cap G_i = 0) & + & P(R_i = 1 \cap G_i = 1) & + & P(R_i = 2 \cap G_i = 2) \\ & = & P(R_i = 0)P(G_i = 0) & + & P(R_i = 1)P(G_i = 1) & + & P(R_i = 2)P(G_i = 2) \\ & = & 1/3\, q_{i0} & + & 1/3\, q_{i1} & + & 1/3\, q_{i2} \\ & = & 1/3\,(\, q_{i0} \,+\, q_{i1} \,+\, q_{i2}) & & & & \\ & = & 1/3 & & & & \end{array}$$

B Alternative hypothesis

Let M_i be the market's prediction of the outcome of game i and p_{ij} the market's assessment of the probability that game i has outcome j. We define

$$M_i \;=\; k \qquad \text{if} \quad p_{ik} \;=\; max_j\, p_{ij}$$

Suppose that the market's probabilities are equal to the objective probabilities, i.e. $p_{ij} = q_{ij}$ for all i and j. In this case, the probability that the market will correctly predict the outcome of game i is given by:

$$P(M_i = G_i) \;=\; P(\text{“outcome with highest prob. will occur”}) \;=\; max_j\, q_{ij} \;=\; q_i^*$$

Let's define the random variable Y_i^m that measures the market's predictive success of game i as follows

$$Y_i^m \;=\; \begin{cases} 1 & \text{if } M_i = G_i \\ 0 & \text{otherwise} \end{cases}$$

Obviously, $E[Y_i^m] = P(Y_i^m = 1) = q_i^*$. The number of correct predictions in n trials by the market, X_n^m, has expected value

$$E[X_i^m] = \sum_{i=1}^{n} E[Y_i^m] \;=\; \sum_{i=1}^{n} q_i^* \qquad (1)$$

Recall that under the null hypothesis (random predictor) and under the current assumptions (n games with three outcomes each) the number of correct predictions X_n is binomially distributed with mean $E[X_n|H_0] = n \times 1/3 = \sum_{i=1}^{n} 1/3$ which is smaller than (1) since $q_i^* \geq 1/3$.

C Squared prediction loss

For games with three possible outcomes, we computed the squared prediction error as follows. Using the market's prediction of game i to illustrate:

$$SPL_i = I(G_i = M_i)\,[\frac{3}{4}\,(1 - P_i^*)]^2 + (1 - I(G_i = M_i))\,[\frac{3}{2}\,P_i^* \;+\; \frac{3}{4}(\frac{1}{3} - P_i^*)]^2.$$

D Data

Table 3: Prices and probabilities of the match markets

			Last price			Probabilities		
Date	Match	Result	1	0	2	1	0	2
16/6 18:00	D:Czech Rep.-France	2	15	20	89	12.10%	16.13%	71.77%
16/6 20:45	D:Denmark-Netherlands	2	2	4	85	2.20%	4.40%	93.41%
17/6 18:00	A:Romania-Portugal	2	20	4	80	19.23%	3.85%	76.92%
17/6 20:45	A:England-Germany	1	40	30	50	33.33%	25.00%	41.67%
18/6 18:00	C:Slovenia-Spain	2	19	10	84	16.81%	8.85%	74.34%
18/6 20:45	C:Norway-Yugoslavia	2	68	24	45	49.64%	17.52%	32.85%
19/6 20:45	B:Turkey-Belgium	1	26	22	75	21.14%	17.89%	60.98%
19/6 20:45	B:Italy-Sweden	1	90	22	25	65.69%	16.06%	18.25%
20/6 20:45	A:England-Romania	2	70	30	36	51.47%	22.06%	26.47%
20/6 20:45	A:Portugal-Germany	1	49[a]	12	49	44.55%	10.91%	44.55%
21/6 18:00	C:Yugoslavia-Spain	2	50	19	59	39.06%	14.84%	46.09%
21/6 18:00	C:Slovenia-Norway	0	25	16	60	24.75%	15.84%	59.41%
21/6 20:45	D:Denmark-Czech Rep.	2	-	-	62	-	-	100.00%
21/6 20:45	D:France-Netherlands	2	45	-	49	47.87%	-	52.13%
24/6 18:00	1/4:Turkey-Portugal	2	22	n.a.	82	21.15%	n.a.	78.85%
24/6 20:45	1/4:Italy-Romania	1	80	n.a.	31	72.07%	n.a.	27.93%
25/6 18:00	1/4:Yugoslavia-Netherlands	2	19	n.a.	94	16.81%	n.a.	83.19%
25/6 20:45	1/4:Spain-France	2	20	n.a.	78	20.41%	n.a.	79.59%
28/6 20:45	1/2:France-Portugal	1	77	n.a.	35	68.75%	n.a.	31.25%
29/6 18:00	1/2:Italy-Netherlands	1	58	n.a.	59	49.57%	n.a.	50.43%
2/7 20:00	Final:France-Italy	1	62	n.a.	39	61.39%	n.a.	38.61%

[a] The qualitative prediction has been assigned to Portugal because of the higher market volume at this price.

Table 4: Prices and probabilities of the Championship market

			Last price		Probabilities	
Date	Match	Result	1	2	1	2
10/6 20:45	B:Belgium-Sweden	1	57	42	57.58%	42.42%
11/6 14:30	B:Turkey-Italy	2	37	90	29.13%	70.87%
11/6 18:00	D:France-Denmark	1	115	25	82.14%	17.86%
11/6 20:45	D:Netherlands-Czech Rep	1	233	30	88.59%	11.41%
12/6 18:00	A:Romania-Germany	0	40	139	22.35%	77.65%
12/6 20:45	A:Portugal-England	1	49	123	28.49%	71.51%
13/6 18:00	C:Spain-Norway	2	110	36	75.34%	24.66%
13/6 20:45	C:Yugoslavia-Slovenia	0	55	24	69.62%	30.38%
14/6 20:45	B:Italy-Belgium	1	104	76	57.78%	42.22%
15/6 20:45	B:Sweden-Turkey	0	23	29	44.23%	55.77%
16/6 18:00	D:Czech Rep-France	2	15	161	8.52%	91.48%
16/6 20:45	D:Denmark-Netherlands	2	2	202	0.98%	99.02%
17/6 18:00	A:Romania-Portugal	2	24	133	15.29%	84.71%
17/6 20:45	A:England-Germany	1	70	86	44.87%	55.13%
18/6 18:00	C:Slovenia-Spain	2	20	63	24.10%	75.90%
18/6 20:45	C:Norway-Yugoslavia	2	69	32	68.32%	31.68%
19/6 20:45	B:Turkey-Belgium	1	11	55	16.67%	83.33%
19/6 20:45	B:Italy-Sweden	1	140	3	97.90%	2.10%
20/6 20:45	A:England-Romania	2	50	24	67.57%	32.43%
20/6 20:45	A:Portugal-Germany	1	133	29	82.10%	17.90%
21/6 18:00	C:Yugoslavia-Spain	2	34	53	39.08%	60.92%
21/6 18:00	C:Slovenia-Norway	0	11	10	52.38%	47.62%
21/6 20:45	D:Denmark-Czech Rep	2	1	1	50.00%	50.00%
21/6 20:45	D:France-Netherlands	2	262	221	54.24%	45.76%
24/6 18:00	1/4:Turkey-Portugal	2	28	156	15.22%	84.78%
24/6 20:45	1/4:Italy-Romania	1	148	89	62.45%	37.55%
25/6 18:00	1/4:Yugoslavia-Netherlands	2	27	221	10.89%	89.11%
25/6 20:45	1/4:Spain-France	2	79	285	21.70%	78.30%
28/6 20:45	1/2:France-Portugal	1	300	220	57.69%	42.31%
29/6 18:00	1/2:Italy-Netherlands	1	159	380	29.50%	70.50%
2/7 20:00	Final:France-Italy	1	610	347	63.74%	36.26%

Table 5: Odds and probabilities of the bookmaker

			Odds[a]			Probabilities		
Date	Match	Result	1	0	2	1	0	2
10/6 20:45	B:Belgium-Sweden	1	1.85	2.75	2.90	43.28%	29.11%	27.61%
11/6 14:30	B:Turkey-Italy	2	3.45	2.70	1.70	23.22%	29.67%	47.12%
11/6 18:00	D:France-Denmark	1	1.40	3.10	4.70	57.16%	25.81%	17.03%
11/6 20:45	D:Netherlands-Czech Rep	1	1.50	3.00	4.00	53.33%	26.67%	20.00%
12/6 18:00	A:Romania-Germany	0	4.45	3.00	1.45	18.01%	26.72%	55.27%
12/6 20:45	A:Portugal-England	1	3.00	2.80	1.80	26.75%	28.66%	44.59%
13/6 18:00	C:Spain-Norway	2	1.65	2.90	3.35	48.51%	27.60%	23.89%
13/6 20:45	C:Yugoslavia-Slovenia	0	1.50	3.00	4.00	53.33%	26.67%	20.00%
14/6 20:45	B:Italy-Belgium	1	1.95	2.85	2.60	41.08%	28.11%	30.81%
15/6 20:45	B:Sweden-Turkey	0	2.15	2.90	2.30	37.37%	27.70%	34.93%
16/6 18:00	D:Czech Rep-France	2	2.90	2.90	1.80	27.69%	27.69%	44.62%
16/6 20:45	D:Denmark-Netherlands	2	4.05	3.20	1.45	19.77%	25.02%	55.21%
17/6 18:00	A:Romania-Portugal	2	2.40	2.85	2.10	33.50%	28.21%	38.29%
17/6 20:45	A:England-Germany	1	2.30	2.65	2.30	34.87%	30.26%	34.87%
18/6 18:00	C:Slovenia-Spain	2	4.70	3.40	1.35	17.05%	23.57%	59.37%
18/6 20:45	C:Norway-Yugoslavia	2	2.25	2.80	2.25	35.67%	28.66%	35.67%
19/6 20:45	B:Turkey-Belgium	1	3.15	2.80	1.75	25.48%	28.66%	45.86%
19/6 20:45	B:Italy-Sweden	1	1.65	2.70	3.70	48.61%	29.71%	21.68%
20/6 20:45	A:England-Romania	2	1.60	2.80	3.75	50.05%	28.60%	21.35%
20/6 20:45	A:Portugal-Germany	1	2.55	2.95	1.95	31.52%	27.25%	41.22%
21/6 18:00	C:Yugoslavia-Spain	2	2.55	2.65	2.10	31.48%	30.29%	38.23%
21/6 18:00	C:Slovenia-Norway	0	3.60	2.90	1.60	22.26%	27.64%	50.10%
21/6 20:45	D:Denmark-Czech Rep	2	3.00	2.80	1.80	26.75%	28.66%	44.59%
21/6 20:45	D:France-Netherlands	2	2.45	2.75	2.10	32.71%	29.14%	38.16%
24/6 18:00	1/4:Turkey-Portugal	2	2.90	2.90	1.80	27.69%	27.69%	44.62%
24/6 20:45	1/4:Italy-Romania	1	1.60	2.90	3.60	50.10%	27.64%	22.26%
25/6 18:00	1/4:Yugoslavia-Netherlands	2	3.20	2.90	1.70	25.09%	27.68%	47.23%
25/6 20:45	1/4:Spain-France	2	2.90	2.90	1.80	27.69%	27.69%	44.62%
28/6 20:45	1/2:France-Portugal	0	1.75	2.80	3.15	45.86%	28.66%	25.48%
29/6 18:00	1/2:Italy-Netherlands	0	2.95	2.85	1.80	27.22%	28.17%	44.61%
2/7 20:00	Final:France-Italy	0	1.90	2.75	2.80	42.20%	29.16%	28.64%

[a]Source: ODDSET, `http://www.oddset.de`.