

Overconfidence in Tournaments: Evidence from the Field[★]

Young Joon Park^a and Luís Santos-Pinto^{b,†}

^aPeking University, HSBC School of Business

^bFaculty of Business and Economics, University of Lausanne

This version: January 25, 2010

Abstract

This paper uses a field survey to investigate the quality of individuals' beliefs of relative performance in tournaments. We consider two field settings, poker and chess, which differ in the degree to which luck is a factor and also in the information that players have about the ability of the competition. We find that poker players' forecasts of relative performance are random guesses with an overestimation bias. Chess players also overestimate their relative performance but make informed guesses. We find support for the “unskilled and unaware hypothesis” in chess: high skilled chess players make better forecasts than low skilled chess players. Finally, we find that chess players' forecasts of relative performance are not efficient.

JEL Codes: A12; C93; J41.

Keywords: Tournaments; Rationality; Field Experiment.

★For helpful comments and discussion we are grateful to Vincent Crawford, Ilyan Georgiev, José Mata, Joel Sobel, and two anonymous referees. We also thank conference and seminar participants at University of California San Diego, Universidade Nova de Lisboa, and the Tinbergen Institute. We would also like to thank Viejas Casino and Sintra's Chess Club for allowing the field surveys.

†Corresponding author. Luís Santos-Pinto. Faculty of Business and Economics, University of Lausanne, Internef 535, CH-1015, Lausanne, Switzerland. Ph: 41-216923658. Fax: 41-216923435. E-mail address: LuisPedro.SantosPinto@unil.ch.

1 Introduction

A large body of empirical evidence from social psychology indicates that people display a systematic tendency to overestimate relative skill. In settings where relative skill matters for making decisions this may have important implications for behavior. One such setting is a tournament. The decision to participate in a tournament or the choice of how much effort to put in depend on accurate expectations of relative skill.

This paper uses field surveys to test the rationality of players' beliefs about their relative performance in "Texas Hold'em" poker and chess tournaments. The main finding of the paper is that players in real-world poker and chess tournaments overestimate their relative performance. This happens when players are given incentives for correct self-assessments and under two different ways of measuring beliefs of relative performance.

We chose poker and chess because we are interested in the extent to which different degrees of luck, skill and information may lead to different beliefs about outcomes. Skill is the most important factor in both types of tournaments but luck plays a larger role in poker than in chess. Another fundamental difference is that chess players usually have better information about the skills of their competitors than poker players.

The surveys were distributed in two poker tournaments—University of California San Diego's 2004 Winter and Spring Poker Classics both held at Viejas Casino in California—and one chess tournament—Sintra's 2005 Chess Open, held in Sintra, Portugal. Before the start of each tournament we ask participants to provide a point forecast of their relative performance. We observe the actual rank of each player in the tournament. When the tournament is over the forecast error of each player is computed and players are paid according to the precision of their forecasts under a quadratic scoring rule.

We also ask players to choose between receiving a sure payment and nine different bets whose payments are contingent on relative performance being above c percent of the population, with $c \in \{10, 20, \dots, 90\}$. This is a new measure of beliefs of relative performance, based on the observation of choices among alternatives, that can be compared with players' forecasts. The choice of bet question is a more stringent test of bias in beliefs than the point forecast question. In the forecasting problem, a risk neutral player who overestimates or underestimates relative performance by the same amount faces the same loss. By contrast, in the betting problem, a risk neutral player who overestimates relative performance by 10% incurs a larger loss than if he underestimates it by 10%. Thus, the optimal bet of a risk neutral player should be smaller than his optimal point forecast.

We test for bias in player's forecasts using a standard test of unbiasedness of forecasts. We test if forecasts are not random choices taking into account the boundness of the dependent variable. Incomplete information about skill together with the fact that forecasts are restricted to lie in a bounded interval force players near the low end of the scale to overestimate relative performance, on average, and players near the high end to underestimate.

We find that players' forecasts of relative performance are *biased*: on average, a poker player overestimates relative performance by 7 to 10 percentiles and a chess player by 6 to 7 percentiles. Both biases in forecasts are significant at 5% level. Poker players' betting behavior also reveals overestimation of relative performance. In the Spring Poker Classic, 78.6% of players chose bets that pay when performance is above the median, the bias in poker players' bets is significant at 5% level. In contrast, only 63.8% of chess players chose bets that pay when performance is above median. The bias in chess players' bets is only significant at 10% level. This is consistent with the fact that the choice of bet question is a more conservative test of bias in beliefs than the point forecast question.

Additionally, we find that poker players' forecasts and bets are random guesses. By contrast, chess players' forecasts and bets are informed guesses. We also find that the forecast errors of high skilled chess players are smaller than those of low skilled ones. This finding is consistent with the "unskilled and unaware" hypothesis proposed by Kruger and Dunning (1990). This hypothesis states that the low skilled players lack the cognitive skills to evaluate their ability and so make worse self-assessments of skill than the high skilled players. Finally, we find that chess players' forecasts of relative performance are *inefficient*: chess players could have made better forecasts of relative performance if they had used their knowledge about the quality of the competition to make their forecasts.

The tendency that individuals have to overestimate their abilities was discovered in the field of social psychology. According to Myers (1996), a textbook in social psychology, "(...) on nearly any dimension that is both *subjective* and *socially desirable*, most people see themselves as better than average." For example, Svenson (1981) reports that between 81% and 90% of Americans think they are safer drivers than the median driver. Myers (1996) cites a study according to which: "In Australia, 86 percent of people rate their job performance as above average, 1 percent as below average." Baker et al. (1998) cite a survey of General Electric Company employees according to which: "58 percent of a sample of white-collar clerical and technical workers rated their own performance as falling within the top 10 percent of their peers in similar jobs, 81 percent rated themselves as falling in the top 20 percent. Only about 1 percent rated themselves below the median." However, there are limitations with the psychological evidence. One of them is that individuals are not provided with incentives to think carefully about their predictions.¹

This paper is the first to document overconfidence in field tournaments using monetary rewards for accurate forecasts. Our study contributes to a growing literature in economics on the existence and consequences of behavioral biases. Camerer and Lovallo (1999) investigate the impact of overconfidence on entry in markets. They consider a market entry game where subjects' payoffs are based on rank, which is determined either randomly or through a test of skill. They

¹Rabin (1998) and Santos-Pinto and Sobel (2005) discuss the evidence on people's tendency to overestimate their abilities.

find that there is more entry when relative skill determines payoffs, which suggests that individuals overestimated their ability to do well on the test relative to others. They also find that there is more entry when individuals self-select into the experiment knowing that higher skill implies higher earnings. They call this finding reference-group neglect.

Hoelzl and Rustichini (2005), Moore (2002), Moore and Kim (2003) identify a subject’s beliefs about relative performance by asking the subject whether a reward should be based on a skill-based test or the outcome of a random device. There is overestimation of relative performance when more than half of the subjects prefer to be rewarded on the basis of their performance on the test than on the basis of a randomization device that selects a winner with probability one half. The experiments find overestimation on easy tests and underestimation on hard tests. Monetary payments significantly reduced overestimation of relative performance but did not improve subjects’ choices.²

Clark and Friesen (2008) study forecasts of relative performance in two tasks: (1) maximizing a two variable unknown function by moving contiguously from cell to cell on a spreadsheet and (2) decoding five letter words. Forecast precision was rewarded with a quadratic scoring rule in 8 sessions and there were no incentives for precise forecasts in 4 sessions. Clark and Friesen find overestimation of relative performance in 3 out of 12 sessions, underestimation in 2 out of 12 sessions, and lack of bias in 7 out of 12 sessions. The use of a quadratic scoring rule did not reduce either forecast bias or variance over non-incentive forecasts.

The paper is organized as follows. The next section explains the field survey design. Section 3 summarizes the hypotheses to be tested. Section 4 shows that players’ forecasts are biased. Section 5 discusses the precision of players’ forecasts. Section 6 discusses players’ bets. Section 7 shows that the forecast errors of the high skilled chess players are smaller than those of the low skilled chess players. Section 8 shows that chess players’ forecasts are not efficient. Section 9 discusses the findings and their implications. Section 10 concludes the paper. The Appendix contains the survey used in Sintra’s Chess Open, theoretical results about optimal point forecasts and bets, and the prize structure of each tournament.

2 Field Survey Design

A Texas Hold’em poker tournament is an *elimination* tournament where players are randomly assigned to different tables that can sit up to 10 players. Tables are reshuffled as players who ran out of chips drop out. When a final table with 10 players is reached there is no further reshuffling. In Texas Hold’em poker

²As Hoelzl and Rustichini (2005) point out, the drawback of this measure of beliefs of relative performance is that subjects are facing the choice between a lottery with objective uncertainty–outcome of the random device–and lottery with subjective uncertainty–the outcome of the test of skill. Thus, if subjects suffer from ambiguity aversion, this measure is likely to underestimate the subjective perception that subjects have of their relative performance.

each player gets two cards face-down, to be combined with five community cards dealt face-up in the middle—the first three simultaneously (called the flop), then a fourth (the turn), then a fifth (the river)—to make the best five-card hand.

At the start of a Texas Hold'em poker tournament the role of luck is very important. High ability players with weak hands can be eliminated by low ability players with stronger hands. This happens because at the start of the tournament players' earnings are very similar so a high ability player is sometimes forced to bet against a low ability player that has a stronger hand. As the tournament evolves the role of luck becomes less important since the earnings of the high ability players become increasingly larger than the earnings of the low ability players.

Most chess tournaments are neither elimination nor *round-robin* tournaments (each player playing against every other player). Typically a *Swiss System* is used. According to this system players are initially matched in pairs either drawn at random or seeded according to Elo ratings.³ After the first round, players who win receive a point, those who draw receive half a point and losers receive no points. All players proceed to the next round where winners are pitted against winners, losers are pitted against losers, and so on. In the subsequent rounds players face opponents with the same (or almost the same) score. Modifications are made to ensure that no player is paired against the same opponent twice and that each player plays an equal number of games with white and black.

A Swiss tournament can handle many players without requiring an impractical number of rounds. However, the final rankings of a Swiss tournament are usually more random than those of a round-robin tournament, depending on the tiebreakers used. Even though the correct player usually wins, and the correct player usually ends up in the last place, the players in between are only sorted roughly without a good tiebreaker depth. So, although chess is itself a skill-based task, rankings in a Swiss chess tournament involve some luck.

The previous paragraphs show us that skill and luck play different roles in a Texas Hold'em poker tournament and in chess tournaments. Clearly, skill is the most important factor in determining players' performance in both types of tournaments but luck plays a larger role in a Texas Hold'em poker tournament than in a Swiss chess tournament.

Another fundamental difference between poker and chess tournaments is the amount of information that players have about the quality of the competition. Typically, chess players have better information about the skills of their competitors than poker players. This happens because Elo ratings are a very

³A description of the Swiss System can be found at <http://scichess.org/faq/swiss.html>. The Elo rating system in chess is a means of comparing the relative strengths of chess players, devised by Arpad Elo. Players gain or lose rating points depending on the Elo rating of their opponents. If a player wins a game of chess in a rated tournament, they gain a number of rating points that increases in proportion to the difference between their rating and their opponent's rating. The central statistical assumption of the ELO system is that any player's tournament performances, spread over a long enough career, will follow a normal distribution. A detailed description of the formulae and theory behind the system can be found at <http://home.clear.net.nz/pages/petanche/ratings/descript.htm>.

informative measure of skill in chess and players usually know their own Elo rating and the Elo ratings of their opponents.

We performed the survey at two “Texas Hold’em” poker tournaments held at Viejas Casino in California. The first tournament—“Winter Poker Classic”—was held on March, 7th, 2004. In this tournament there were 155 players each paying a 10 USD entry fee and receiving 1500 USD worth of chips. Once the player used up all chips, he would be eliminated. The total prize pool was 1670 USD. The second tournament—“Spring Poker Classic”—was held on May, 23rd, 2004. In this tournament there were 167 players each paying an entry fee of 20 USD. The total prize pool was 3000 USD. The prize structure of each tournament is depicted in Table A1 in the Appendix.

To obtain poker players’ forecasts of relative performance we asked them the following question: “*Of all the individuals participating in the poker tournament, what percentage do you think will be eliminated before you?*” Players were instructed to answer the question by choosing a whole number between 0 and 99. The survey also informed players that numbers close to zero indicate that they predict that they will be among worst players in the tournament, and that numbers close to 99 indicate that they predict that they will be among the best players in the tournament.

Sintra’s Chess Open was held in July, 17th, 2005 in Sintra, a village near Lisbon. There were 93 chess players in the tournament. The entry fee for members of Sintra’s Chess Club was 3 EUR while non-members had to pay 6 EUR. The total prize pool was 1100 EUR. The prize structure of the tournament is depicted in Table A2 in the Appendix.

Sintra’s Chess Open used the Swiss system. At the start of the tournament players with similar Elo ratings were matched in pairs. The Elo ratings of players were public information since the organizers of the tournament posted them at the entrance of the room where the tournament was held. After the first round, players were placed in groups according to their score (winners in the 1 group, those who drew go in the 1/2 group, and losers go in the 0 group) and then matched in pairs inside each group. Each round the same procedure was used. There were a total of 8 rounds each lasting 20 minutes. The relative performance of each chess player in the tournament was calculated by the organization using the Swiss method.

Like in poker tournaments, we asked players in Sintra’s Chess Open to predict their relative performance. The main novelty is that we asked chess players to report their own Elo rating and the percentage of players in the tournament with a smaller Elo rating. This gives us an idea of players’ information about the quality of the competition.

Based on each player’s forecast of relative performance and his actual performance, we calculated the forecast error of each player, E_i , defined as $E_i = F_i - P_i$, where F_i is player i ’s forecast of relative performance and P_i is player i ’s relative performance, with F_i being an integer between 0 and 99 and P_i being a real number in $[0, 100)$. The monetary reward of player i , R_i , as a function of

player i 's forecast error, was determined by the quadratic scoring rule

$$R_i = \begin{cases} M - [\text{Int} (|E_i|)]^2, & \text{if } \text{Int} (|E_i|) < X \\ 0 & \text{if } \text{Int} (|E_i|) \geq X \end{cases} ,$$

where $\text{Int} (x)$ is the closest integer which is smaller than x . In the Winter Poker Classic $M = 10$ USD and $X = 4$, in the Spring Poker Classic $M = 20$ USD and $X = 5$, and in Sintra's Chess Open $M = 10$ EUR and $X = 4$.

We chose the quadratic scoring rule because of its simplicity and the fact that it allows us to test the rationality of players' forecasts using ordinary least squares regressions. DeGroot (1970) shows that the quadratic scoring rule is incentive compatible for a risk neutral player. Propositions 1 and 2 in the Appendix show that the quadratic scoring rule is also incentive compatible for a player with uniform or an unimodal and symmetric distribution of beliefs, regardless of the player's preferences towards risk.⁴

We did not use a binary lottery payoff scheme to induce risk neutrality from the part of players due to the lack of control associated with performing a field survey. Most players left the room where the tournament was being held immediately after being eliminated so there was no way they could observe the lottery being drawn.

The survey also asked players for demographic characteristics such as age, sex, and academic major. On average, players took 10 minutes to read, answer, and return the survey. Each player filled his survey individually and returned it to us right after he finished it. The reply rate in the Winter Poker Classic was 87.1%, the one in the Spring Poker Classic was 77.2%, and 64.5% in Sintra's Chess Open. We postpone the discussion of self-selection bias to Section 9.

In the Winter and Spring Poker Classics players were asked for their addresses and their earnings from taking the survey were sent by mail. In Sintra's Chess Open players had the option of receiving their earnings by mail or at the end of the tournament. Most players chose to receive them at the end of the tournament.

3 Hypotheses

The main hypothesis that will be tested in this field survey is that players' forecasts of relative performance are biased.

H1a Forecasts are biased.

⁴Alternatively, we could have chosen a scoring rule where the loss is proportional to the absolute value of the forecast error. De Groot (1970) shows that this scoring rule induces risk neutral players to report the median rather than the mean. To test the rationality of players' forecasts under this alternative scoring rule we would need to use least absolute deviations regressions—see Basu and Markov (2003). Camerer (1982) uses a scoring rule where individuals are paid something when they are exactly correct and nothing otherwise. This rule induces risk neutral players to report the mode rather than the mean.

To test H1a we perform a standard test of unbiasedness of forecasts by regressing players' forecast errors on a constant (or intercept). We take as null hypothesis that forecasts are unbiased (the intercept is zero) and as alternative hypothesis that forecasts are biased (the intercept is different from zero). If the estimated coefficient for the intercept is statistically significant, then we find evidence for H1a.

We are also interested in the accuracy of players' forecasts, that is, how good players are at evaluating their relative performance.

H2a Forecasts are not random guesses.

To test H2a we regress players' positions in the tournament on a constant and on their forecasts. We take as null hypothesis that forecasts are random guesses (the slope coefficient is zero) and as alternative hypothesis that forecasts are not random guesses (the slope coefficient is different from zero). If the estimated slope coefficient is statistically significant, then we find evidence for H2a.

We use an alternative measure of beliefs of relative performance. To do that we ask players to choose among different bets with payments that depend on their relative performance in the tournament. Thus, we also test if players bets are biased and if they are not random choices.

H1b Bets are biased.

H2b Bets are not random choices.

To test H1b we take as null hypothesis that players' mean choice of bet is smaller than or equal to their mean position in the tournament. We take as alternative hypothesis that players' mean choice of bet is greater than their mean position in the tournament. To test H2b we take as null hypothesis that players' mean reward for choice of bet is less than or equal to the reward of a random choice of bet. We take as alternative hypothesis that players' mean reward for choice of bet is greater than the reward of a random choice of bet.

Kruger and Dunning (1999) report a series of experiments that show that high skilled individuals make better self-assessments than low skilled individuals. However, their measure of relative skill only relies in a single observation and effects of experience or familiarity with the task are not taken into account. In Sintra's chess tournament we have a very good objective measure of relative skill—the Elo rating—and we know the number of chess tournaments that each player has played before. This allows a more stringent test of Kruger and Dunning's "unskilled and unaware" hypothesis.

H3 The forecast errors of chess players with high Elo ratings are smaller than the forecast errors of chess players with low Elo ratings.

To test H3 we regress the absolute forecast error of chess players on a constant, on their Elo ratings, the number of tournaments they played, and on an interaction term (Elo ratings times the number of tournaments played). We

take as null hypothesis that the forecast errors of chess players with high Elo ratings are greater than or equal to the forecast errors of chess players with low Elo ratings. We take as alternative hypothesis that the forecast errors of chess players with high Elo ratings are smaller than the forecast errors of chess players with low Elo ratings. If the coefficient on Elo rating is negative and statistically significant, then we find evidence for H3.

By definition, rational forecasts must also be efficient, that is, players must make use of all available information to make their forecasts. In the survey we ask chess players to provide their best estimate of the percentage of the population in the tournament with a lower Elo rating. This allows us to test for efficiency in players' forecasts. Since we expect chess players' forecasts to be biased we also test if they are inefficient.

H4 Forecasts of chess players are inefficient.

To test H4 we regress the relative performance of chess players on a constant, on their forecasts, and on their estimates of the percentage of the players in the tournament with a lower Elo rating. We take as null hypothesis that forecasts are efficient (the coefficient of the second predictor variable is zero) and as alternative hypothesis that forecasts are inefficient (the coefficient of the second predictor variable is different from zero). If the estimated coefficient of the second predictor variable is statistically significantly, then we find evidence for H4.

4 Forecast Bias

Table I displays the distribution of forecasts in each tournament divided into intervals of 10 percentiles starting in the interval $[0, 10]$ and ending in $[90, 99]$.

Table I
Distribution of Players' Forecasts in Tournaments

Forecasts	Winter Poker Classic		Spring Poker Classic		Sintra's Chess Open	
	share	cum. share	share	cum. share	share	cum. share
[0, 9]	0.7	0.7	0.8	0.8	10.0	10.0
[10, 19]	4.4	5.3	3.9	4.7	6.7	16.7
[20, 29]	8.2	13.5	14.6	19.3	8.3	25.0
[30, 39]	5.9	19.4	3.9	23.2	5.0	30.0
[40, 49]	6.7	26.1	10.1	33.3	8.3	38.3
[50, 59]	14.1	40.2	13.2	46.5	8.3	46.6
[60, 69]	14.8	55.0	13.2	59.7	16.7	63.3
[70, 79]	10.4	65.4	12.4	72.3	5.0	68.3
[80, 89]	16.2	81.6	12.3	84.6	18.3	86.6
[90, 99]	18.4	100.0	15.4	100.0	13.4	100.0

Inspection of Table I reveals a clear tendency for overestimation of relative performance in all tournaments. Only 26.1% (Winter Poker Classic), 33.3%

(Spring Poker classic) and 38.3% (Sintra’s Chess Open) of players who took the survey forecast to finish below the median.

To test if players’ forecasts of relative performance are biased we run the ordinary least squares (OLS) regression $E_i = \alpha + \varepsilon_i$, where E_i is the forecast error of player i and α is the intercept. The results for each tournament are summarized in Table II.

Table II

OLS Regression Results for Forecast Bias

	Winter Poker Classic	Spring Poker Classic	Sintra’s Chess Open
Intercept	10.02 (0.003)*	7.13 (0.045)*	6.98 (0.025)*
	n=122	n=116	n=60

Dependent variable: Forecast error

p -values in parentheses. * denotes statistical significance at 5% level.

We see that the mean forecast error in the Winter Poker Classic is equal to 10.02 percentiles, the mean forecast error in the Spring Poker Classic is 7.13 percentiles and 6.98 percentiles in Sintra’s Chess Open. The mean forecast errors in all tournaments are greater than zero at 5% significance level. This shows that, on average, players’ forecasts in all tournaments are biased towards overestimation of relative performance.

5 Forecast Precision

To test if players’ forecasts of relative performance are not random guesses we need to have an idea of how well players’ forecasts predict relative performance. One way to do that is to run the OLS regression

$$P_i = \alpha + \beta F_i + \varepsilon_i, \tag{1}$$

where F_i is player i ’s forecast and P_i is player i ’s position in the tournament. If we find that the fit of this regression is good and that the estimate for the slope is significantly greater than zero, then there is evidence that players forecasts are not random guesses. By contrast, if we find that the fit of this regression is bad and that the estimate for the slope is not significantly different from zero, then players’ forecasts are random guesses.

However, the OLS estimates in (1) would be biased.⁵ Incomplete information about relative skill together with the fact that relative performance is

⁵This happens because the dependent variable is bounded by 0 and 100. The nature of the bias can be demonstrated as follows. If $0 \leq P_i \leq 100$, then $-\alpha - \beta F_i \leq \varepsilon_i \leq 100 - \alpha - \beta F_i$. Thus, the fact that we have a limited dependent variable implies that the error term is regulated by an upper and a lower bound that depends on the independent variable. So, the distribution of the error term depends on the value of the independent variable and it is not identically distributed. OLS requires, among other things, that the error term is identically distributed and uncorrelated with the regressor.

restricted to lie in a bounded interval force people near the low end of the scale to overestimate relative performance, on average, and people near the high end to underestimate.

To address this problem we use the transformation of variables technique. One way to map the variable P_i , which is bounded by 0 and 100, to the real line is to use a logit transformation. The logit transformation of player i 's relative performance and forecast are given by $U_i = \ln(P_i/(100 - P_i))$, and $Z_i = \ln(F_i/(100 - F_i))$, respectively. The transformation implies U_i and Z_i are unconstrained variables.⁶ We use the transformed series to run the ordinary least squares regression $U_i = \alpha + \beta Z_i + \varepsilon_i$. Table III displays the results obtained for each tournament.

Table III

OLS Regression Results for Forecast Precision			
	Winter Poker Classic	Spring Poker Classic	Sintra's Chess Open
Constant	0.14 (0.455)	0.01 (0.976)	-0.22 (0.302)
Forecast	0.03 (0.732)	0.11 (0.374)	0.50 (0.000)*
	n=122, R ² =0.001	n=116, R ² =0.007	n=60, R ² =0.39

Dependent variable: Logit transformation of relative performance

p -values in parentheses. * denotes statistical significance at 5% level.

We see from Table III that the fit of Winter and Spring Poker Classic regressions is very bad: the R -squared is equal to 0.1% in the Winter Poker Classic and 0.7% in the Spring Poker Classic. We also see that in both poker tournaments the estimated coefficients are not significantly different from zero.⁷ Thus, we find evidence against H2a for poker tournaments, that is, poker players' forecasts in both tournaments are random guesses. The lack of accuracy of poker players' forecasts in both tournaments implied that the earnings from their forecasts were quite low as it can be seen in Table A3 in the Appendix.

In contrast, the fit of the Sintra's Chess Open regression is 39%. The estimated coefficient for the slope is 0.5 and is significantly different from zero at 5% significance level. Thus, we find evidence for H2a in Sintra's Chess Open, that is, chess players' forecasts of relative performance are not random guesses.

⁶See Zarembka (1974) on the transformation of variables technique. This transformation of variables has also been used by Chen and Giovannini (1992) for testing the rationality of exchange rate forecasts within a band.

⁷Another aspect that needs to be taken into consideration is that players who forecast their performance to be in the bottom of the scale may make larger forecast errors than players who forecast their performance to be in the top of the scale (the transformation of variables may or may not change this pattern of heteroscedasticity). If there is heteroscedasticity in the transformed model, then the OLS estimates are unbiased but inefficient. To address this possibility we run a robust regression using Stata 7.0. We found that the robust standard errors are essentially identical to the OLS standard errors.

6 Betting Behavior

In the Spring Poker Classic and in Sintra’s Chess Open players were also asked to choose among different bets whose payments depended on their relative performance in the tournament. For example, in the Spring Poker Classic each player was offered the choice of getting a sure payment of \$2.00 or betting on his relative performance. There were nine possible bets whose payments were contingent and a player being above c percent of the population, with $c \in \{0, 10, 20, \dots, 90\}$. The bets paid $\$200/(100 - c)$ if a player was eliminated after c percent of the population and zero dollars otherwise.

Proposition 3 in the Appendix shows that for risk neutral players, the choice of bet question is a more stringent test of overestimation of relative performance than the point forecast question. In the forecasting problem, a risk neutral player who overestimates or underestimates relative performance by the same amount faces the same loss. By contrast, in the betting problem, a risk neutral player who overestimates relative performance by 10% incurs a larger loss than if he underestimates it by 10%. Thus, the optimal bet of a risk neutral player should be smaller than his optimal point forecast.

The answers to the choice of bet question in each tournament are summarized in Table IV. The first column of Table IV reports the payoff of each bet, the second column the ratio of the number of players that were paid for that choice of bet to the number of players who chose that bet, the third column the share of players in the Spring Poker Classic who chose each bet, and the fourth column the cumulative share. The remaining four columns provide similar information for Sintra’s Chess Open.

Table IV
Players’ Choices and Earnings from Bets

Spring Poker Classic				Sintra’s Chess Open			
reward	paid	share	cum. share	reward	paid	share	cum. share
\$2.00	14/14	10.0	10.0	1.00 eur	8/8	13.8	13.8
\$2.22	1/1	0.7	10.7	1.11 eur	6/7	12.1	25.9
\$2.50	2/4	2.9	13.6	1.25 eur	3/4	6.9	32.8
\$2.86	3/6	4.3	17.9	1.43 eur	1/1	1.7	34.5
\$3.33	3/5	3.5	21.4	1.67 eur	0/1	1.7	36.2
\$4.00	7/17	12.2	33.6	2.00 eur	7/10	17.2	53.4
\$5.00	8/18	12.9	46.5	2.50 eur	0/4	6.9	60.3
\$6.66	0/3	2.1	48.6	3.33 eur	2/3	5.2	65.5
\$10.00	4/24	17.1	65.7	5.00 eur	5/13	22.4	87.9
\$20.00	6/48	34.3	100.0	10.00 eur	3/7	12.1	100.0
rewards	\$281.79			rewards	95.5 eur		
players	140			players	58		
mean	\$2.01			mean	1.65 eur		

A quick inspection of Table IV shows us that only 21.4% of players who answered the choice of bet question in the Spring Poker Classic and 36.2% of

players who did it in Sintra’s Chess Open chose bets that paid them when their performance was below the median. This suggests that players’ bets also reveal overestimation of relative performance.

To test if poker players’ bets are unbiased, hypothesis H1b, we compare poker players’ bets to their ranks in the Spring Poker Classic. Ranking bets from 5 (the sure thing), 15 (the \$2.22 bet), to 95 (the \$20 bet) we find that the average choice of bet of poker players is 68.33. The average rank is the 51.54th percentile. we take as null hypothesis that the the average choice of bet is less than or equal to 51.54 and as alternative hypothesis that the average choice of bet is greater than 51.54. The p -value for this one-tailed hypothesis test is 0.000 so we can reject H1b for poker players at 5% significance, that is, we find that poker players’ bets are biased towards overestimation of relative performance. Ranking chess players’ bets from 5 (the sure thing), 15 (the \$1.11 bet), to 95 (the \$10 bet) we find that the average choice of bet of chess players is 53.42. The average rank is the 47.61th percentile. The p -value for the one-tailed test is 0.091 so we cannot reject H1b for chess players at 5% significance but we can reject it at 10%. Thus, we find more evidence for overestimation of relative performance in chess players’ forecasts than in their bets.

Table IV also shows us that the average reward for choice of bet of poker players is \$2.01. This value is not different, at 5% significance level, from the expected reward of a random choice of bet in the Spring Poker Classic: \$2.00.⁸ This provides evidence against hypothesis H2b for poker players, that is, poker players’ bets are random choices. By contrast, the average reward for choice of bet of chess players is \$1.65. This value is greater, at 5% significance level, than the average reward of a random choice of bet in Sintra’s Chess Open: 1 euro. This provides evidence in favor of hypothesis H2b for chess players, that is, chess players’ bets are not random choices. Thus, players bets are consistent with their forecasts.⁹

7 Unskilled and Unaware

Kruger and Dunning (1999) report a series of experiments with easy skill-based tasks that support the “unskilled-unaware hypothesis”, that is, that the high skilled individuals are better informed about their skills than low skilled individuals. One possible explanation for this finding is that high skilled players are more experienced than low skilled players and that greater experience implies better information about relative skill.

⁸If a player’s performance is worse than the 10th percentile and makes a random choice of bet his expected reward is $(1/10) \times \$2 + (9/10) \times \$0 = \$0.2$. If a player’s performance is better than the 10th percentile and worse than the 20th percentile and makes a random choice of bet is expected reward is $(1/10) \times \$2 + (1/10) \times \$2.22 + (8/10) \times \$0 = \0.422 . Doing the same for players in the other deciles we obtain expected rewards for players in each decile. Thus, if all players made random choices of bet the average reward for choice of bet should be approximately equal to $(\$0.2 + \$0.422 + \$0.672 + \dots + \$5.857)/10 \simeq \$2$.

⁹The Pearson correlation coefficient between poker players’ bets and their forecasts is 0.625 and that between chess players’ bets and their forecasts is 0.77.

In Sintra’s Chess Open we have a very informative objective measure of relative skill—the Elo ratings—and we know players’ previous experience with chess tournaments.¹⁰ This means we can study the impact of relative skill on the precision of chess players’ forecasts while taking into account experience effects. To do that we regress the absolute forecast error of chess players on their Elo ratings, the number of tournaments they played, and an interaction term (Elo ratings times the number of tournaments played). If the coefficient on Elo rating is negative and significant, then we find support for H3, that is, we find evidence that the forecast errors of chess with high Elo ratings are smaller than the forecast errors of chess players with low Elo ratings. Thus, we run the OLS regression

$$|E_i| = \alpha + \beta_1 CExp_i + \beta_2 CElo_i + \beta_3 (CElo_i \times CExp_i) + \varepsilon_i,$$

where $CExp_i = Exp_i - \text{mean}(Exp_i)$, and $CElo_i = Elo_i - \text{mean}(Elo_i)$. The predictor variables are centered to avoid multicollinearity problems and to be able to interpret β_1 and β_2 as average effects of Exp_i and Elo_i on $|E_i|$, respectively. The results obtained for this regression are reported in Table V.

Table V

OLS Regression Results for Forecast Precision,
Experience and Elo in Sintra’s Chess Open

Constant	22.1738 (0.000)*
CExp	0.1073 (0.061)
CElo	-0.0289 (0.001)*
CExp×CElo	-0.0001 (0.076)

n=49, R²=0.23

Dependent variable: Absolute forecast error

p-values in parentheses. * denotes statistical significance at 5%.

Table V shows that coefficient for Elo rating is negative and significant at 5% level. This means that, controlling for experience effects, the higher the Elo rating of a player, the smaller his absolute forecast error. Thus, we find support for H3, that is, we find evidence that the forecast errors of high skilled chess players are smaller than those of low skilled chess players. This finding is consistent with Kruger and Dunning’s (1999) “unskilled-unaware hypothesis” and can not be explained by the fact that high skilled chess players are more experienced in chess tournaments than low skilled chess players.

¹⁰Of the 93 players that took part in Sintra’s Chess Open, 70 had Elo ratings and 23 did not. The range of Elo ratings was from 1090 points to 2441 points. The average Elo rating was 1865 points. Of the 60 players who took our survey only 49 reported the number of chess tournament they had played before. Of these 49 players, 7 had no previous experience with chess tournaments, 15 had participated in 1 to 10 tournaments, 17 had participated in 11 to 100 tournaments, and 10 had participated in 150 to 400 tournaments.

8 Efficiency of Forecasts

We already know that chess players' forecasts are biased. Can we say anything about efficiency? If a chess player makes an efficient forecast of relative performance then he must use all the available information that he has about his relative skill to make that forecast. Since we asked players to provide their best estimate of the percentage of the population in the tournament with a lower Elo rating we can use this variable to test H4, that is, that chess' players forecasts are inefficient.

To do that we run the OLS regression $U_i = \alpha + \beta_1 Z_i + \beta_2 W_i + \varepsilon_i$, where $W_i = \ln(Lelo_i/(100 - Lelo_i))$, with $Lelo_i$ being player i 's assessment of the percentage of the population that has a lower Elo rating. If β_2 is significantly different from zero, then there is evidence that chess players' forecasts are not efficient. The results obtained for this regression are displayed in Table VI.

Table VI

OLS Regression Results for Test of Efficiency of
Players' Forecasts in Sintra's Chess Open

	Regression 1		Regression 2
Constant	-0.07 (0.634)	Constant	0.14 (0.373)
Forecast	0.60 (0.000)*	Forecast	0.35 (0.003)*
		Lower Elo	0.24 (0.003)*
	n=44		n=44
	R ² =0.54		R ² =0.63
	Adjusted R ² = 0.53		Adjusted R ² = 0.62

Dependent variable: Logit transf. of relative performance

p-values in parentheses. * denotes statistical significance at 5%.

The results from the two regressions in Table VI show us that chess players' forecasts are not efficient. The model with the explanatory variable Lower Elo (regression 2) has a better fit than the model without it (regression 1). Thus, we find evidence that supports H4, that is, chess players' forecasts are inefficient. In fact, chess players could have made better forecasts if they had taken into consideration their own subjective assessments of the percentage of the population with a smaller Elo rating.¹¹

9 Discussion

In this section we discuss the findings and implications of the paper.

¹¹The Pearson correlation coefficient between chess players' subjective assessments of the percentage of the population with a smaller Elo rating and the actual percentage with smaller Elo ratings was 0.87.

9.1 Bias and Rationality

The overconfidence bias in chess player's forecasts of relative performance is particularly surprising since chess players have good information about the quality of the competition.¹² This finding can be interpreted from two different perspectives. From the perspective of advocates of rational expectations the bias in chess players' forecasts of relative performance is small and so we should not worry about it. By contrast, from the perspective of advocates of behavioral biases in human judgment, the fact that the bias persists when there is plenty of information about the quality of the competition constitutes strong evidence against rational expectations. Moreover, since in most tasks it is hard to find measures of relative performance as informative as the Elo rating is for chess, advocates of behavioral biases, would argue that biases in judgments of relative performance are widespread.

Why did poker and chess players overestimate their relative performance? There are at least four explanations for this question in the economics' literature. First, the reference group neglect explanation proposed by Camerer and Lovallo (1999) tells us that individuals are not aware of the fact that the people who choose to participate in a tournament are more skilled than a random person. Second, the bias might be due to a positive correlation between risk preferences and skill. If low skilled players are risk averse and high skilled players are risk seeking, then forecasts may be biased towards the positive side even though there is no overestimation of relative performance. Third, the bias might be explained by individuals' tendency to attribute failure to bad luck and success to skill. This explanation is called the self-serving bias in causal attributions and it was first formalized by Gervais and Odean (2001). Van den Steen (2004) shows how the self-serving bias may preclude learning. Finally, Santos-Pinto and Sobel (2005) show that skill investment and egocentric comparisons can lead individuals to become overconfident.

9.2 Difference in Biases in Poker and Chess

According to the psychology literature individuals are more overconfident when they think that they have control over the outcome of the task. Consistent with this view, Camerer and Lovallo (1999) find more entry in the skill-based treatment than in the luck-based treatment of the market entry game.

As we have seen, skill is the most important factor in determining players' performance in poker and chess tournaments but luck plays a larger role in a Texas Hold'em poker tournament than in Swiss chess tournament. However, we found that overconfidence is larger in poker than in chess. Does this finding contradict previous studies? Not necessarily.

An explanation for this finding might be a self-selection effect due to the different nature of the poker and chess tournaments. In any tournament the players who overestimate their skills the most are the ones that are more likely

¹²Previous studies could not address this issue since they lack reliable objective measures of relative skill like the Elo rating.

to enter. Now, suppose that individuals are attracted to tournaments not only for the utility they can get from the money they win net of the entry fee but also from the utility from playing. Participants in the poker tournaments were eliminated when they lost all their chips. If a poker player thinks that he is one of the worst players, then he expects to play only a few games of poker for the \$10 or \$20 entry fee. In this sense it is not rational to participate in poker tournaments with this kind of beliefs. In contrast, a chess player who thinks he is one of the worst could still play interesting games even after having lost some games (against opponents of approximately the same strength). So, playing a chess tournament without winning a prize may be more satisfying than leaving a poker tournament early and without a prize. This could lead to a smaller self-selection effect in chess tournaments.¹³

It might also be that, in the subjective view of players, poker players perceive that the role of skill in a poker tournament is greater than it actually is. If players forecasts are based on this view, then overestimation of relative performance in poker can be greater than overestimation in chess.¹⁴

The smaller bias found in chess might also be due to cultural differences. Most players in Sintra's Chess Open from Portugal whereas most players in the Winter and Spring poker classics are from the US. If individuals from the US have a greater tendency to overestimate their abilities than individuals from Portugal, then the bias in chess should be smaller than the bias in poker.

Niederle and Vesterlund (2007) find evidence of gender difference in overconfidence. They show men are more likely to participate in a tournament than women do because men are more overconfident than women. The proportion of female participants in our tournaments is negligible and so we cannot analyze the impact of gender on overconfidence.¹⁵ The low participation of female players can be considered as evidence consistent with Niederle and Vesterlund (2007).

9.3 Self-Selection and Incentives

We can not rule out the possibility that the overconfidence observed in the data is due to a self-selection bias. We had a reply rate of 87.1% in the Winter Poker Classic, 77.2% in the Spring Poker Classic, and 64.5% in Sintra's Chess Open. If the players who did not answer our survey have a tendency to underestimate

¹³Differences in risk aversion between those who self-select into a poker or a chess tournament might also explain why the bias is larger in poker than in chess.

¹⁴In fact, we found some support for this possibility. We asked players in the survey how they thought their position in the tournament would be determined. Players could chose among seven options that ranged from "*Only by relative skill*" to "*Only by luck*" with 5 other options in between. On average, players in both poker and chess tournaments thought that "skill is more important than luck but that luck plays a large role in determining relative performance." So, on average, there were no significant differences in poker and chess players' perceptions of the role of luck in poker and in chess, respectively.

¹⁵Precisely, the proportion of female participants were 8.2% (10 out of 122) and 6.9% (8 out of 116) for the Winter and Spring Poker tournaments, respectively, and 5% (3 out of 60) for Sintra's chess tournament.

their performance, then forecasts might not be biased.¹⁶

The size of the biases in this paper is modest when compared to those often found in the psychology literature. The use of financial incentives for precision in forecasts could be an explanation for this finding. This is consistent with Hoelzl and Rustichini (2005), who show that monetary incentives can reduce overestimation bias (but do not improve precision).¹⁷ In fact, the main limitation of our paper might be the fact that we only provide modest monetary incentives for precision in forecasts and bets. If monetary incentives are modest, then a player's expected rewards from taking the survey are small and do not depend much on his forecasts or bets. We cannot rule out the possibility that the biases would disappear if players would have been given greater monetary incentives.

9.4 Implications

In settings where skill matters for making decisions overestimation of relative performance may have important implications for behavior. One such setting is a tournament. The decision to participate in a tournament or the choice of how much effort to put in depends on correct beliefs about skill.

If gamblers have a tendency to overestimate their relative performance in tournaments, then their utility will decrease due to their misguided choices (e.g., putting in too much or too little effort). However, managers of casinos can gain if they take this into account when they design tournaments. To do that they should either increase entry fees or reduce tournament prizes.

In the workplace, if workers overestimate their relative skill, they will also overestimate the probability of favorable outcomes like bonuses or promotions. If this is the case, then managers of firms should, on average, prefer incentive schemes featuring payments contingent on relative performance (e.g., rank-order tournaments or incentive schemes composed partly by fixed pay and partly by variable pay dependent on the magnitude of relative performance) to individualistic incentive schemes (e.g., fixed salary plans or piece rates).

10 Conclusion

This paper finds overconfidence in two different field tournaments, chess which is a tournament where skill is very important and subjects have excellent information about their relative skill, and poker where luck is very important and relative ability is likely much more uncertain. In both cases we find evidence of overconfidence both in forecasts and bets. We find that in the more uncertain environment of poker, both forecasts and bets are random guesses. In the chess

¹⁶In future research on this topic we plan to give players a monetary reward just for filling in and returning the survey. This will improve reply rates and reduce self-selection bias.

¹⁷It is not clear if using monetary incentives improves the precision of individuals' forecasts. For a good discussion of this topic see Camerer and Hogarth (1999).

tournament, where players have good information about their ability, both forecasts and bets are informed guesses, players with higher ability make better estimates than those with lower ability, and players forecasts are not efficient.

The paper contributes to a growing literature in economics that studies the existence and consequences of behavioral biases. Closely related studies include Camerer and Lovallo (1999), Hoelzl and Rustichini (2005), Moore (2002), Moore and Kim (2003), and Clark and Friesen (2008). All of these previous studies are laboratory experiments that are opened to the criticism of lack of ecological validity. This paper is the first to document the existence of overconfidence about relative performance in field tournaments using monetary rewards for accurate forecasts.

References

- Baker, G., Jensen, M., and K. Murphy, "Compensation and Incentives: Practice vs. Theory," *Journal of Finance*, Vol. XLIII, no. 3, (1988), 593- 616.
- Basu, S. and S. Markov (2004). "Loss Function Assumptions in Rational Expectations Tests on Financial Analysts Earnings Forecasts," *Journal of Accounting and Economics*, 38: 171-203.
- Camerer, C. (1990). "Do Markets Correct Biases in Probability Judgement? Evidence from Market Experiments" in J. Kagel and L. Green (Eds.), *Advances in Behavioral Economics 2*, Northwood, NJ: Ablex Publ. 125-172.
- Camerer C. and D. Lovallo (1999) "Overconfidence and Excess Entry: An Experimental Approach." *The American Economic Review*, Vol. 89, No. 1, 306-318.
- Camerer, C. and R. Hogarth (1999). "The Effects of Financial Incentives in Experiments: a Review and Capital-labor-production Framework," *Journal of Risk and Uncertainty* 19, 7-42.
- Chen, Z. and A. Gionvanni (1992). "Target Zones And The Distribution Of Exchange Rates: An Estimation Method," *Economics Letters*, Vol. 40(1), 83-89.
- Clark, J. and L. Friesen (2008). "Overconfidence in Forecasts of Own Performance: An Experimental Study," *The Economic Journal*, Vol. 119, Issue 534, 229-251.
- DeGroot, M. (1970). *Optimal Statistical Decisions*, McGraw Hill.
- Gervais, S. and T. Odean (2001). "Learning to be Overconfident," *Review of Financial Studies*, Vol. 14, no. 1, 1-27.
- Hodges, B., G. Regehr, and D. Martin (2001). "Difficulties in Recognizing One's Own Incompetence: Novice Physicians Who Are Unskilled and Unaware of It," *Academic Medicine*, Vol. 76, no. 10, S87-S89.
- Hoelzl, E. and A. Rustichini (2005). "Overconfident: Do You Put Your Money on It?" *The Economic Journal*, Vol. 115, 305-318.
- Holt, C. and S. Laury (2002). "Risk Aversion and Incentive Effects," *The American Economic Review*, Vol. 92, No. 5, 1644-1655.
- Krueger, J., and R. Mueller (2002). "Unskilled, Unaware or Both? The Better-Than-Average Heuristic and Statistical Regression Predict Errors in Estimates of Own Performance," *Journal of Personality and Social Psychology*, vol. 82, no. 2, 180-188.
- Kruger, J. (1999). "Lake Wobegon Be Gone! The Below-Average Effect and the Egocentric Nature of Comparative Ability Judgments," *Journal of Personality and Social Psychology*, 77(2), 221-232.
- Moore, D. (2002). "Egocentric Biases and the Failure of Strategic Prediction." International Association of Conflict Management, 15th Annual Conference. June 9-12, Park City, Utah.

- Moore, D. and T. Kim (2003). "Myopic Social Prediction and the Solo Comparison Paradox." *Journal of Personality and Social Psychology*, 85, No. 6, 1121-1135.
- Myers, D. (1996). *Social Psychology*, New York: McGraw-Hill.
- Niederle, M. and L. Vesterlund (2007). "Do Women Shy Away from Competition? Do Men Compete too Much?" *Quarterly Journal of Economics*, Vol. 122, No. 3, 1067-1101.
- Rabin, M. (1998). "Psychology and Economics," *Journal of Economic Literature*, Vol. XXXVI, 11-46.
- Santos-Pinto, L. and J. Sobel (2005). "A Model of Positive Self Image in Subjective Assessments." *The American Economic Review*, Vol. 95, No. 5, 1386-1402.
- Svenson, O. (1981). "Are we all Less Risky and More Skillful than our Fellow Drivers," *Acta Psychologica*, 94, 143-148.
- Van den Steen, E. (2004). "Rational Overoptimism (and Other Biases)," *The American Economic Review*, Vol. 94(4), 1141-1151.
- Zarembka, P. (1974). "Transformation of Variables in Econometrics," in: Paul Zarembka, ed., *Frontiers in Econometrics*, Academic Press, New York.

11 Appendix

11.1 Sintra's 2005 Chess Open Survey

You are about to answer a survey that, among other things, asks you to make a prediction of your relative position in this chess tournament. Depending upon how well you make your prediction you may be able to earn up to 10 euros (Question 1). The survey also asks you to choose between different lotteries whose prizes depend on your relative position in the tournament (Question 2). Depending on your choice of lottery and how well you perform in the tournament you may earn up to an additional 10 euros. We will send you your payment by mail if you provide us your name and address. If you prefer, you can provide us only your e-mail address and we will tell you your payment by e-mail and then you can give us your address if you wish to receive it by mail. This survey is confidential.

Name: _____ E-mail: _____
Address: _____
Zip code: _____ Age: _____ Sex: _____

Q1: Please read the following question carefully: Of all the individuals participating in this chess tournament what percentage do you think will be ranked below you?

Before you answer note that, after the tournament is over, we will compare your prediction with the ratio of the actual number of players ranked below you to the total number of players. We will then pay you (in euros) for your prediction as follows:

10 if the prediction is less than 1% away from your position;
9 if the prediction is more than 1% and less than 2% away from your position;
6 if the prediction is more than 2% and less than 3% away from your position;
1 if the prediction is more than 3% and less than 4% away from your position;
0 otherwise.

Now, answer the question by choosing a whole number between 0 and 99 (recall that the number you choose represents your best estimate of what percentage of people will be ranked below you. Numbers close to zero indicate that you predict that you will be among worst players in the tournament, numbers close to 99 indicate that you predict that you will be among the best players in the tournament).

Q2: Consider the 10 lotteries below, whose prizes (in euros) depend on your ranking in the tournament. Choose one of the options:

We pay you 1.00 for sure
We pay you 1.11 if at least 10% of players are ranked below you and 0 otherwise
We pay you 1.25 if at least 20% of players are ranked below you and 0 otherwise
We pay you 1.43 if at least 30% of players are ranked below you and 0 otherwise
We pay you 1.67 if at least 40% of players are ranked below you and 0 otherwise
We pay you 2.00 if at least 50% of players are ranked below you and 0 otherwise
We pay you 2.50 if at least 60% of players are ranked below you and 0 otherwise

We pay you 3.33 if at least 70% of players are ranked below you and 0 otherwise
 We pay you 5.00 if at least 80% of players are ranked below you and 0 otherwise
 We pay you 10.00 if at least 90% of players are ranked below you and 0 otherwise

Q3: What is your Elo rating? If you don't know the answer to this question, then choose between: a) I don't have an Elo rating or b) I have an Elo rating but I can't recall it.

Q4: What is your best estimate of the percentage of players in this tournament who have an Elo rating less than yours?

Q5: How many chess tournaments have you played before? Consider that a chess tournament involves monetary prizes and at least 20 players.

Q6: How do you think your position in this tournament will be determined? Choose one

- Only by your relative skill at playing chess
- More by your relative skill than by luck, and luck plays a small role
- More by your relative skill than by luck, and luck plays a large role
- As much by your relative skill as by luck
- More by luck than by your relative skill, and relative skill plays a large role
- More by luck than by your relative skill, and relative skill plays a small role
- Only by luck

11.2 Forecasting Problem

Suppose that an individual's beliefs of relative performance are a continuous random variable X . Let beliefs have density $g(x)$, continuous and with support in $[a, b]$, with $0 \leq a < b \leq 1$. Suppose this individual has initial wealth \bar{w} and utility of wealth $U(w)$. Let f represent the individual's point forecast, with $f \in [0, 1]$. This individual's wealth—a continuous version of the discrete quadratic scoring rule—is given by $w = \bar{w} + [w_0 - (x - f)^2]$, with $w_0 \geq 1$. The optimal point forecast of this individual is given by

$$\max_{f \in [0,1]} \int_a^b U(\bar{w} + w_0 - (x - f)^2) g(x) dx. \quad (2)$$

We will call (2) the point forecast problem. The first-order condition to (2) is given by

$$\int_a^b U'(\bar{w} + w_0 - (x - f^*)^2) 2(x - f^*) g(x) dx = 0.$$

and the second-order condition by

$$\int_a^b [U''(\bar{w} + w_0 - (x - f^*)^2) 2(x - f^*)^2 - U'(\bar{w} + w_0 - (x - f^*)^2)] g(x) dx < 0.$$

If an individual is risk averse we have $U' > 0$ and $U'' < 0$ and the second-order condition is verified. If an individual is risk neutral we have $U' > 0$ and $U'' = 0$ and the second-order condition is also satisfied. If an individual is risk seeking we have $U' > 0$ and $U'' > 0$ and we can't tell if the second-order condition is satisfied or not.

It is a well known result that optimal point forecast of a risk neutral individual is his mean belief of relative performance—see DeGroot (1970) pp. 228. Proposition 1 shows that the optimal point forecast of an individual with uniform beliefs of relative performance is his mean belief of relative performance, regardless of his preferences towards risk.

Proposition 1 *If an individual's beliefs of relative performance have the uniform distribution with support $[a, b]$, with $0 \leq a < b \leq 1$, then $f^* = E(X)$.*

Proof. Using integration by parts, the first-order condition to the point forecast problem is equivalent to

$$U(\bar{w} + w_0 - (b - f^*)^2)g(b) - U(\bar{w} + w_0 - (a - f^*)^2)g(a) = \int_a^b U(\bar{w} + w_0 - (x - f^*)^2)g'(x)dx.$$

If beliefs have the uniform distribution, then $g'(x) = 0$ for all x and $g(a) = g(b)$, so the above condition reduces to $U(\bar{w} + w_0 - (b - f^*)^2) = U(\bar{w} + w_0 - (a - f^*)^2)$, or $f^* = (a + b)/2 = E(X)$, that is, the optimal point forecast of an individual with uniform beliefs of relative performance is his mean belief of relative performance. *Q.E.D.*

Proposition 2 shows that the optimal point forecast of an individual with unimodal and symmetric beliefs of relative performance is his mean belief of relative performance, regardless of his preferences towards risk.

Proposition 2 *If an individual's beliefs of relative performance are unimodal and symmetric, then $f^* = E(X)$*

Proof. Let the distribution of beliefs have support in $[a, b]$. Using integration by parts, the first-order condition to the point forecast problem is equivalent to

$$U(\bar{w} + w_0 - (b - f^*)^2)g(b) - U(\bar{w} + w_0 - (a - f^*)^2)g(a) = \int_a^b U(\bar{w} + w_0 - (x - f^*)^2)g'(x)dx.$$

or,

$$U(\bar{w} + w_0 - (a - f^*)^2)g(a) + \int_a^{E(X)} U(\bar{w} + w_0 - (x - f^*)^2)g'(x)dx = \\ U(\bar{w} + w_0 - (b - f^*)^2)g(b) + \int_{E(X)}^b U(\bar{w} + w_0 - (x - f^*)^2)(-g'(x))dx.$$

If $a \leq f^* < E(X)$ and g is symmetric and unimodal, then the first term in the LHS is greater than the first term in the RHS and the value of the integral in the LHS is greater than the value of integral in the RHS. But then the value of the LHS is greater than the value of the RHS, a contradiction. If $E(X) < f^* \leq b$ and g is symmetric and unimodal, the first term in the LHS is smaller than the first term on the RHS and the value of the integral in the LHS is smaller than the value of integral in the RHS. But then the value of the LHS is smaller than the value of the RHS, a contradiction. Thus, it must be that $f^* = E(X)$. *Q.E.D.*

11.3 Betting Problem

Suppose that an individual has beliefs of relative performance given by the density $g(x)$, with support in $[a, b]$, with $0 \leq a < b \leq 1$. Suppose this individual has initial wealth \bar{w} and utility of wealth given by $U(w)$. Let c represent the choice of bet, with $c \in [0, 1]$. This individual's wealth—a continuous version of our discrete bets choice—is given by

$$w = \begin{cases} \bar{w} + \frac{w_0}{1-c}, & x \geq c \\ \bar{w}, & x < c \end{cases},$$

with $w_0 \geq 1$. The optimal bet of this individual is the solution to

$$\max_{c \in [0,1]} G(c)U(\bar{w}) + [1 - G(c)]U\left(\bar{w} + \frac{w_0}{1-c}\right). \quad (3)$$

We will call (3) the betting problem. We can state the following result.

Proposition 3 *If an individual is risk neutral and his beliefs of relative performance are*

- (i) *uniform with support $[a, 1]$, with $0 \leq a$, then his optimal bet is any $c^* \in [a, 1]$;*
- (ii) *uniform with support $[a, b]$ with $0 \leq a < b < 1$ then $c^* = a < E(X) = f^*$;*
- (iii) *unimodal and symmetric, then $a < c^* < Mode(X) = E(X) = f^*$;*
- (iv) *unimodal and positively skewed, then $a \leq c^* \leq Mode(X) < E(X) = f^*$.*

Proof: Let start by proving (i). If an individual is risk neutral and has uniform beliefs with support in $[a, 1]$ then the objective function of the betting problem is $\bar{w} + w_0/(1 - a)$. Since this individual's utility does not depend on his choice of bet he must be indifferent between any bet in $[a, 1]$.

Let us show (ii). If an individual is risk neutral and has uniform beliefs with support in $[a, b]$ with $0 \leq a < b < 1$, then the objective function problem of the betting problem is $\bar{w} + \frac{b-c}{b-a} \frac{w_0}{1-c}$. It is clear that for this case the optimal bet is $c^* = a$.

Let us show (iii). If an individual is risk neutral and has unimodal and symmetric beliefs, then the first-order condition to the betting problem becomes $-g(c^*) \frac{w_0}{1-c^*} + [1 - G(c^*)] \frac{w_0}{(1-c^*)^2} = 0$ or $1 - G(c^*) = g(c^*)(1 - c^*)$. This is equivalent to

$$\int_{c^*}^1 g(x) dx = \int_{c^*}^1 g(c^*) dx. \quad (4)$$

If we can show there exists an x_0 strictly greater than c^* such that $g(c^*) < g(x_0)$ then it must be that $c^* < \text{Mode}(X)$ since $\text{Mode}(X) = \max g(x)$. Suppose, by contradiction that: (1) for all $x > c^*$ we have $g(x) \leq g(c^*)$ and (2) that there exists an $x_0 > c^*$ such that $g(x_0) \leq g(c^*)$. By the well know result that one can integrate inequalities, assumptions (1) and (2) imply that $\int_{c^*}^1 g(x) dx < \int_{c^*}^1 g(c^*) dx$, which contradicts (4). Thus, we must either have that (a) $g(x) = g(c^*)$ for $x \geq c^*$, or (b) there exists an $x_0 > c^*$ such that $g(c^*) < g(x_0)$. Case (a) is a degenerate case. If case (b) holds then we know that $c^* < \text{Mode}(X)$. So, for a unimodal and symmetric density of beliefs we have that $c^* < \text{Mode}(X) = E(X)$. To finish the proof we still need to show that the second-order condition to the betting problem is satisfied. This condition is given by

$$-g'(c^*) \frac{w_0}{1-c^*} - 2g(c^*) \frac{w_0}{(1-c^*)^2} + 2[1 - G(c^*)] \frac{w_0}{(1-c^*)^3},$$

which simplifies to $-g'(c^*) \frac{w_0}{1-c^*}$. We see that the second-order condition is satisfied whenever $g'(c^*) > 0$. But, if $c^* < \text{Mode}(X) = E(X)$ and the distribution is unimodal and symmetric, then it must be that $g'(c^*) > 0$.

Finally, let us show (iv). When $g'(a) > 0$ the proof is similar to that of (iii) with the exception that for a unimodal and positively skewed density of beliefs we have that $\text{Mode}(X) < E(X)$. Note that when $g'(a) > 0$ the second-order condition is satisfied and $a < c^* < \text{Mode}(X) < E(X)$. When $g'(a) < 0$ we have a corner solution: $c^* = \text{Mode}(X) = a$. Q.E.D.

11.4 Prize Structures and Earnings from Forecasts

Table A1: UCSD's 2004 Poker Classic Prizes

Winter Poker Classic - \$10 entry fee		Spring Poker Classic - \$20 entry fee	
Rank	Prize	Rank	Prize
1st place	\$447 (27%)	1st place	\$792 (27%)
2nd place	\$209 (12%)	2nd place	\$370 (12%)
3rd place	\$164 (10%)	3rd place	\$290 (10%)
4th place	\$149 (9%)	4th place	\$264 (9%)
5th place	\$134 (8%)	5th place	\$238 (8%)
6th place	\$119 (7%)	6th place	\$211 (7%)
7th place	\$104 (6%)	7th place	\$185 (6%)
8th place	\$89 (5%)	8th place	\$158 (5%)
9th place	\$75 (4%)	9th place	\$132 (4%)
10th-18th places	\$20 (12%)	10th-18th places	\$40 (12%)
Sum	\$1670 (100%)	Sum	\$3000 (100%)

Table A2: Sintra's Chess Open Prizes

Rank	Monetary Prize	Symbolic Prize
1st place	300 euro (27%)	Trophy
2nd place	180 euro (16%)	Trophy
3rd place	120 euro (11%)	Trophy
4th place	75 euro (7%)	Medal
5th place	50 euro (5%)	Medal
6th-10th places	30 euro (14%)	Medal
11th-15th places	25 euro (11%)	Medal
16th-20th places	20 euro (9%)	Medal
Sum	1100 euro (100%)	-

Table A3: Players' Earnings from Forecasts

Winter Poker Classic		Spring Poker Classic		Sintra's Chess Open	
Reward	Paid	Reward	Paid	Reward	Paid
\$ 0	107	\$ 0	105	0 euro	48
\$ 1	4	\$ 4	3	1 euro	0
\$ 6	6	\$11	2	6 euro	0
\$ 9	4	\$16	1	9 euro	10
\$10	1	\$19	1	10 euro	2
		\$20	4		
Rewards	\$86	Rewards	\$149	Rewards	110 euro
Players	122	Players	116	Players	60
Average	\$.70	Average	\$1.28	Average	1.83 euro