Forecasts of relative performance in tournaments: evidence from the field

Luís Santos-Pinto and Young-Joon Park

Universidade Nova de Lisboa

5. June 2004

Forecasts of Relative Performance in Tournaments: Evidence from the Field⋆

Young-Joon Park
University of California, San Diego

Luís Santos-Pinto†
Universidade Nova de Lisboa

This version: March 22, 2007

Abstract

This paper uses a field experiment to investigate the quality of individuals' forecasts of relative performance in tournaments. We ask players in luck-based (poker) and skill-based (chess) tournaments to make point forecasts of rank. The main finding of the paper is that players' forecasts in both types of tournaments are biased towards overestimation of relative performance. However, the size of the biases found is not as large as the ones often reported in the psychology literature. We also 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.

⋆We are thankful to Vincent Crawford, Ilyan Georgiev, José Mata, and Joel Sobel and to seminar participants at University of California, San Diego, Universidade Nova de Lisboa, and Tinbergen Institute for many helpful comments and suggestions. We would also like to thank Viejas Casino and Sintra's Chess Club for allowing the field experiments.
†Corresponding author. Luís Santos-Pinto. Universidade Nova de Lisboa, Faculdade de Economia, Campus de Campolide, PT-1099-032, Lisboa, Portugal. Tel.: +351-21-380-1640. Fax: +351-21-387-0933.
E-mail address: lspinto@fe.unl.pt.

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.[1] 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 a field experiment to test the rationality of players' forecasts of relative performance in tournaments. The experiment took place in two poker tournaments–UCSD'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. We chose poker and chess tournaments because a poker tournament is a luck-based competition whereas a chess tournament is a skill-based competition.

Before the start of each tournament we distribute a survey to participants where we ask them, among other things, 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 quality of their forecasts. We use a quadratic scoring rule to reward forecast accuracy.

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 $c \in \{10, 20, \ldots, 90\}$. 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.

We test for bias in players' forecasts and bets. We also test if players' forecasts and bets are significantly different from random choices. To perform these tests we use a parametric approach that takes into account the fact that incomplete information about relative 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.

Our findings are described in detail later in the paper. In summary, 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. Players' betting behavior is consistent with their forecasts. In the Spring Poker Classic, 78.6% of players chose bets that pay when performance is above the median. In Sintra's Chess Open 63.8% of chess players chose bets that pay when performance is above median.

Additionally, we find that poker players' forecasts and bets are not signifi-

[1] Tournaments are commonly used incentive schemes in organizations. For example, salespeople are often paid bonuses that depend on their sales relative to those of the other salespeople in the firm. Most managers are involved in promotion tournaments: vice-presidents compete to be promoted to president and senior executives compete to become CEO. Elections, litigation, auctions, athletic contests, and racing games can also be viewed as tournaments.

cantly different from random guesses with an overestimation bias. By contrast, chess players' forecasts and bets are significantly better than random choices. We also find support for 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 *not efficient:* 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.

This paper is an additional contribution to the literature that documents the existence of behavioral biases in judgment and decision making. The tendency that individuals have to overestimate their relative skill was discovered in the field of social psychology. Two seminal contributions are Dunning et al. (1989) and Kruger and Dunning (1999) who show that overestimation of relative skill varies systematically with several factors.[2] 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.

There is a growing literature in experimental economics on the causes and consequences of overestimation of relative skill. Camerer and Lovallo (1999) investigate the impact of this bias 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 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.[3]

Clark and Friesen (2003) 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 accuracy was rewarded with a quadratic scoring rule in 8 sessions and there were no incentives for accurate forecasts in 4 sessions. Clark and Friesen found 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.

Ferraro (2003) investigates forecasts of relative performance in three intro-

[2] For example, the more ambiguous is the definition of the skill the greater is the overestimation effect, overestimation is higher in tasks that require a greater number of skills, overestimation decreases with task difficulty, and overestimation is higher when individuals think they can control the outcome of a task than when they think that the outcome of a task is mostly determined by chance.

[3] Moore and Cain (2005) use the same experimental design as Camerer and Lovallo (1999) with the added feature that skill-dependent payoffs are based on either an easy or a difficult test of skill. They found more entry when rank was determined by relative performance on the easy test than when rank was determined randomly. They found less entry when rank was determined by relative performance on the difficult test than when rank was determined randomly.

ductory microeconomic classes at Georgia State University. Students in these classes took three non-cumulative multiple-choice exams that made up most of their final grade. Immediately after completing each exam, subjects were asked to forecast their relative performance on the exam.[4] Ferraro found that 80% of the subjects that took the first exam believed they were above the 50th percentile. He also found that overestimation of relative performance was not reduced over time. By the third exam, 83% of all subjects still believed they performed above the 50th percentile.

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.[5]

The main contribution of this paper is to show that overestimation of relative skill is present in luck-based (poker) as well as in skill-based (chess) real world tournaments. The previous studies only considered skill-based tasks. The paper also shows that the bias can exist even when individuals have very good information about the relative skill of their competitors (this was the case in the chess tournament). Previous studies could not address this issue since they lacked reliable measures of relative skill. This finding is at odds with Camerer and Lovallo's reference group neglect explanation for overestimation of relative skill. On a methodological level, this is the first study to use a statistical test on the accuracy of players' forecasts of relative performance that takes into account the boundness of the dependent variable.

The paper is organized as follows. The next section describes the hypotheses to be tested. Section 3 explains the experimental design. Section 4 shows that players' forecasts are biased. Section 5 discusses forecast accuracy. Section 6 looks at players' bets. Section 7 shows that the unskilled are unaware of their skills. Section 8 shows that chess players' forecasts are not efficient. Section 9 discusses the results. 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 prize structures of tournaments.

[4] Forecast accuracy was rewarded with a quadratic scoring rule. In one class, subjects who were closest to predicting their actual percentile received $25 each. In the other two classes, subjects received $5 if their prediction was within one percentile point accurate, $4 if within two percentile points and $1 within three percentile points.

[5] As 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.

2 Hypotheses

The main hypothesis that will be tested in this field experiment is that players forecasts of relative performance are rational. By definition, rational forecasts must be unbiased, that is, there must be no systematic tendency for overestimation or underestimation of relative performance.

H1a Forecasts are unbiased.

We will also use an alternative measure of beliefs of relative performance to check whether players' beliefs are biased or not: we let players bet on their assessments of relative performance. Thus, we also test if players bets are unbiased.

H1b Players' bets are unbiased.

We are also interested in having an idea of how accurate players' forecasts are. Forecasts may be extremely inaccurate, that is, they may not be distinguishable from random guesses. Alternatively, forecasts may be significantly better than random guesses. The same applies to players' bets.

H2a Forecasts are not random guesses.

H2b Players' bets are not random choices.

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 is not very good in that it 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 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 unskilled are unaware.

By definition, rational forecasts must also be efficient, that is, players must make use of all available information to make their forecasts. In the experiment we ask players to provide an assessment of the quality of the competition. This allows us to test for efficiency in players' forecasts.

H4 Forecasts are efficient.

Hypotheses H1 and H2 are tested in poker and chess tournaments. Hypotheses H3 and H4 are only tested in the chess tournament.

3 Field Experiment Design

To study the quality of individuals' forecasts of relative performance in luck-based tasks we performed the field experiment at two "Texas Hold'em" poker tournaments held at Viejas Casino in California.[6]

The first tournament–"Winter Poker Classic"–was held on March, 7th, 2004. In this tournament there were 155 players each paying a $10 entry fee and receiving $1500 worth of chips. Once the player used up all chips, he would be eliminated. The total prize pool was $1670. 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. The total prize pool was $3000. The prize structure of each tournament is depicted in Table A1 in the Appendix.

To obtain 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 euros while non-members had to pay 6 euros. The total prize pool was 1100 euros. 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.[7] 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

[6] In "Texas Hold'em" tournaments luck plays a large role in determining players' positions. In Texas Hold'em poker 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 the tournament 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 and 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.

[7] 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/petanque/ratings/descript.htm.

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.[8]

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, a very informative measure of relative skill at chess. We also asked chess players to report 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 reward of player i, R_i, as a function of player i's forecast error, was determined by the quadratic scoring rule

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

where Int (x) is the closest integer which is smaller than x. In the Winter Poker Classic $M = \$10$ and $X = \$4$, in the Spring Poker Classic $M = \$20$ and $X = \$5$, and in Sintra's Chess Open $M = 10$ and $X = 4$ euro.[9]

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.[10]

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 experiment. 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.[11]

[8] A detailed description of Swiss system can be found at http://scichess.org/faq/swiss.html

[9] It is not clear whether using monetary incentives improves individuals' forecasts. For a good discussion of this topic see Camerer and Hogarth (1999).

[10] Alternatively, 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.

[11] It is not clear that this procedure works in practice. For example, Selten et al. (1999) find that the binary lottery payoff scheme does not induce risk neutrality, but on the contrary, it leads to stronger deviations from risk neutrality than a direct money payoff scheme.

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 79%, the one in the Spring Poker Classic was 70% and the one in Sintra's Chess Open was 65%.

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.

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

	Winter Poker Classic		Spring Poker Classic		Sintra's Chess Open	
Forecasts	Players	Share	Players	Share	Players	Share
$[0, 9]$	1	0.7	1	0.8	6	10.0
$[10, 19]$	6	4.4	5	3.9	4	6.7
$[20, 29]$	11	8.2	19	14.6	5	8.3
$[30, 39]$	8	5.9	5	3.9	3	5.0
$[40, 49]$	9	6.7	13	10.1	5	8.3
$[50, 59]$	19	14.1	17	13.2	5	8.3
$[60, 69]$	20	14.8	17	13.2	10	16.7
$[70, 79]$	14	10.4	16	12.4	3	5.0
$[80, 89]$	22	16.3	16	12.4	11	18.3
$[90, 99]$	25	18.5	20	15.5	8	13.3
Total	135	100.0	129	100.0	60	100.0

Inspection of Table I reveals a clear tendency for overestimation of relative performance in all tournaments. We have that 74.1%, 66.7% (poker) and 61.6% (chess) of players who took the survey forecast to finish at or above the median.

To test if players' forecasts of relative performance are unbiased, hypothesis H1a, we run the ordinary least squares regression $E_i = \alpha + \varepsilon_i$, where E_i is the forecast error of player i and α is the intercept.[12] The results for each tournament are summarized in Table II.

[12] This is a standard test of unbiasedness in forecasts.

Table II

OLS Regression Results for Forecast Bias

	Winter Poker Classic	Spring Poker Classic	Sintra's Chess Open
Intercept	10.02 (3.02)***	7.13 (2.03)**	6.98 (2.31)**
	n=122	n=116	n=60

Dependent variable: Forecast error
t statistics in parentheses. ***, **, * denotes statistical significance at the 1%, 5%, and 10% level respectively.

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.

Are the mean forecast errors obtained for our three samples representative of the populations? We think that the answer to this question is yes. First, we have forecast errors for a large majority of the population in both poker tournaments (79% and 70%). Second, in the Winter Poker Classic the players whose order of elimination was not monitored are among the worst performers in the tournament and these are the ones who overestimate relative performance the most. So, the sample mean forecast error in the Winter Poker Classic is a lower bound of the mean forecast error of the population. Third, the mean performance of 28 players in Spring Poker Classic who did not answer the survey but whose order of elimination was monitored is similar to the mean performance of the 116 players who forecasted their position. This makes us believe that the sample mean forecast error in the Spring Poker Classic is representative of the mean forecast error of the population.

What about the level of bias of the whole population in Sintra's Chess Open? Does it differ significantly from 7 percentiles? In this tournament we have forecast errors for 65% of the population. However, since we now have the relative performance of all players in the tournament we can back out the missing forecasts of the 33 players from their relative performance.[13] Doing that we find that the mean forecast of relative performance of the population would be equal to 55.42 percentiles. Since the mean relative performance of the population is equal to 49.47, the mean forecast error would be equal to 6 percentiles.

[13] To do that we run the ordinary least squares regression $Z_i = a + bU_i + \varepsilon_i$, where Z_i and U_i are the logit transformations of player i's relative performance and forecast of relative performance, respectively. We obtain

$$\hat{Z}_i = \underset{(1.49)}{0.29} + \underset{(6.06)}{0.77}\, U_i.$$

We use the estimated coefficients to back up the transformed forecasts of the 33 players from their relative performance. After that we invert the transformation to find the value of the forecasts, that is, we calculate $\hat{F}_i = 100e^{\hat{Z}_i} / \left(1 + e^{\hat{Z}_i}\right)$ for each of the 33 players.

5 Forecast Accuracy

The mean absolute forecast error in the Winter Poker Classic was 29.66 percentiles and 31.34 percentiles in the Spring Poker Classic. Such large mean absolute forecast errors suggest that poker players forecasts of relative performance are very inaccurate. By contrast, the mean absolute forecast error of these 60 chess players is 17.03 percentiles.

To test if players' forecasts of relative performance are not random guesses, hypothesis H2a, we need to have an idea of how well players' forecasts predict relative performance. One way to do that is to run the ordinary least squares (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 not distinguishable from random guesses.

However, the OLS estimates in (1) would be biased.[14] Incomplete information about relative skill together with the fact that relative performance is 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 is given by $U_i = \ln\left(P_i/(100 - P_i)\right)$,and the logit transformation of player i's forecast of relative performance is given by $Z_i = \ln\left(F_i/(100 - F_i)\right)$. The transformation implies U_i and Z_i are unconstrained variables.[15] 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.

[14] This 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.

[15] See 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.

Table III

OLS Regression Results for Forecast Accuracy

	Winter Poker Classic	Spring Poker Classic	Sintra's Chess Open
Constant	0.14 (0.75)	0.01 (0.03)	-0.22 (-1.04)
Forecast	0.03 (0.34)	0.11 (0.89)	0.50 (6.07)***
	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
t statistics in parentheses. ***, **, * denotes statistical
significance at the 1%, 5%, and 10% level respectively.

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.[16] Thus, we find evidence against hypothesis H2a for poker tournaments, that is, poker players' forecasts in both tournaments are random guesses. By contrast, we see from Table III that 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 1% significance level. Thus, we find evidence that supports hypothesis H2a for Sintra's Chess Open: chess players' forecasts of relative performance are not random guesses.[17]

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 $c \in \{0, 10, 20, \ldots, 90\}$. 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

[16] Another 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 heteroescedasticity). If there is heteroescedasticity 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.

[17] 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.

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.[18]

Table IV

Players' Choices and Earnings from Bets

Spring Poker Classic			Sintra's Chess Open		
Reward	Paid	Share	Reward	Paid	Share
\$2.00	14/14	10.0	1.00 euro	8/8	13.8
\$2.22	1/1	0.7	1.11 euro	6/7	12.1
\$2.50	2/4	2.9	1.25 euro	3/4	6.9
\$2.86	3/6	4.3	1.43 euro	1/1	1.7
\$3.33	3/5	3.5	1.67 euro	0/1	1.7
\$4.00	7/17	12.2	2.00 euro	7/10	17.2
\$5.00	8/18	12.9	2.50 euro	0/4	6.9
\$6.66	0/3	2.1	3.33 euro	2/3	5.2
\$10.00	4/24	17.2	5.00 euro	5/13	22.4
\$20.00	6/48	34.4	10.00 euro	3/7	12.1
Rewards	\$281.79		Rewards	95.5 euro	
Players	140		Players	58	
Average	\$2.01		Average	1.65 euro	

A quick inspection of Table IV shows us that 78.6% of players who answered the choice of bet question in the Spring Poker Classic and 63.8% of players who did it in Sintra's Chess Open chose bets that paid them when their performance was above the median. Thus, players choices of bet seem to be consistent with their forecasts of relative performance in that they also reveal overestimation of relative performance.

To test if poker players' bets are unbiased, hypothesis H1b, we need to 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. The t statistic for the hypothesis test that the average choice of bet is not significantly different from 51.54 is equal to 6.56 and the critical value, at 5% significance level, is equal to $t_{5\%}(139) \simeq 1.645$. Thus, we find evidence against hypothesis H1b, that is, we find that poker players' bets are biased towards overestimation of relative performance. Using a similar

[18] 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, and the third column reports the share of players in the Spring Poker Classic who chose each bet. The remaining three colums provide similar information for Sintra's Chess Open.

procedure we also find that chess players' bets are biased towards overestimation of relative performance.[19]

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. Thus, we find evidence against hypothesis H2b, that is, we find that 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. Thus, we find evidence in favor of hypothesis H2b, that is, that chess players' bets are not random choices.

We see that poker players' bets are consistent with their point forecasts in that they also reveal overestimation of relative performance and lack of accuracy. If anything, poker players' betting behavior seems to reveal more overestimation of relative performance than their forecasting behavior. In effect, while 78.6% of players chose bets that paid them when their performance was above the median only 62.8% of players forecasted that their performance would be above median. However, since we did not assess players' preferences towards risk we cannot claim that poker players reveal more overestimation of relative performance in their bets than in their point forecasts.

Chess players' bets are also consistent with their point forecasts in that they also reveal overestimation of relative performance and are better than random choices. However, the tendency towards overestimation of relative performance in chess players' betting behavior is only statistically significant at 10% level whereas chess players' forecasts reveal overestimation of relative performance at 5% significance level.

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.[20]

To test Kruger and Dunning's "unskilled-unaware hypothesis" we use data from Sintra's Chess Open. In this tournament we have a very informative measure of relative skill, a player's Elo rating, and we have asked players for

[19] Ranking 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 t statistic for the hypothesis test that the average choice of bet is not significantly different from 47.61 is equal to 1.35 and the critical value, at 10% significance level, is equal to $t_{10\%}(57) \simeq 1.28$.

[20] However, Burson et al. (2006) show that for difficult skill-based tasks (where there is underestimation of relative performance) the low skilled players are more accurate in their forecasts than the high skilled players.

their previous experience with chess tournaments. This means we can study the impact of relative skill on the accuracy of chess players' forecasts while taking into account experience effects. To do that we run the OLS regression

$$|E_i| = a + b_1 Exp_i + b_2 Elo_i + b_3 (Elo_i \times Exp_i) + \varepsilon_i.$$

The results obtained for this regression are reported in Table V.

Table V

OLS Regression Results for Forecast Accuracy, Experience and Elo in Sintra's Chess Open

Constant	27.05 (6.97)***
Experience	0.21 (1.87)*
Elo	-32.67 (-3.79)***
Exp×Elo	-0.20 (1.77)*
	n=49, R^2=0.28

Dependent variable: Absolute forecast error
t statistics in parentheses. ***, **, * denotes statistical significance at the 1%, 5%, and 10% level respectively.

Table V shows that experience improves forecasts, but the coefficient is only statistically different from zero at 10% significance level . However, the coefficient for relative skill is negative and significant at 1% significance level. That is, the data shows that high skilled chess players make smaller forecast errors than low skilled chess players. Thus we find support for Kruger and Dunning's (1999) "unskilled-unaware hypothesis" and this can not be explained by the fact that high skilled players are more experienced in chess tournaments than low skilled players.

8 Efficiency of Forecasts

For chess players' forecasts of relative performance to be rational they would have to be unbiased and efficient. 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 for efficiency in chess' players forecasts–hypothesis H4.[21] To do that we run the ordinary least squares 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

[21] The Elo ratings of all players were posted at the entrance of the room where the tournament took place.

percentage of the population that has a lower Elo rating. If chess players' forecasts are efficient we expect that β_2 is not significantly different from zero. By contrast, 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 (-0.48)	Constant	0.14 (0.90)
Forecast	0.60 (7.07)***	Forecast	0.35 (3.13)***
		Lower Elo	0.24 (3.15)***
	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
t statistics in parentheses. ***, **, * denotes statistical significance at the 1%, 5%, and 10% level respectively.

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). In other words, 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.[22]

9 Discussion of Findings

9.1 Bias and Rationality

The paper finds that both poker and chess players' forecasts of relative performance fail the rationality test: they are biased towards overestimation of relative performance. The fact that luck plays almost no role in chess and that most players in Sintra's Chess Open have plenty of information about their relative skill made us expect that chess players' forecasts would be unbiased. However, that was not the case.

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 motivational and cognitive biases

[22] Players subjective assessments of the percentage of the population with a smaller Elo rating where very good when compared to actual Elo ratings.

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 motivational and cognitive biases, would argue that biases in judgments of relative performance are likely to be widespread.

Why did poker and chess players overestimate their relative performance? There are at least four alternative explanations for this bias in the economics' literature. According to the reference group neglect explanation proposed by Camerer and Lovallo (1999) individuals overestimate their relative skill because they are not aware that the people who chose to participate in the tournament are more skilled than a random person. The bias may also 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. Another possibility is that the bias results from individuals' tendency to attribute failure to bad luck and success to skill.[23] Finally, Santos-Pinto and Sobel (2005) show that skill investment and egocentric comparisons can lead individuals to overestimate their relative skill.

We could not test these different explanations with our field experiment. However, the finding that chess players' forecasts are inefficient is at odds with Camerer and Lovallo's (1999) reference group neglect explanation. In Sintra's Chess Open players had very good information about the quality of their competitors but their forecasts did not incorporate fully that information.

Another question that might be raised is: why is the overestimation bias larger in poker than in chess? One explanation for this finding may be a selection effect due to the different nature of the poker and chess tournaments in this experiment. 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. In both cases, players who overestimate their skills the most are the one that are more likely to enter. However, playing in a chess tournament without winning a prize may be more satisfying than leaving a poker tournament early without prize. This could lead to a smaller selection effect in chess tournaments.

9.2 Skill versus Luck

The paper shows that overestimation of relative performance is present in both poker and chess tournaments. Clearly, luck plays a large role in the game of poker and only a small role in the game of chess. According to the psychology literature individuals are more overconfident when they think that they have control over the outcome of the task. However, the overestimation bias that we found in poker tournaments was larger than the one that we found in chess. Do

[23]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.

the findings in this paper stand in contradiction with the psychology evidence? Not necessarily.

It could be that poker players perceive the poker tournament as being more of a skill-based task than a luck-based task. In 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 thought that skill is more important than luck but that luck plays a large role in determining relative performance.

9.3 Accuracy and Information about Competitors

One of the main differences between the poker and chess tournaments in this experiment is that chess players have information about the distribution of skills of their opponents whereas poker players do not. The difference in information sets together with the fact that relative performance is much more random at poker than at chess are likely to explain the fact that poker players' forecast errors are so much larger than those of chess players.

9.4 Monetary Incentives and Bias

There is at least one serious limitation to our measures of beliefs of relative performance. As Hoelzl and Rustichini (2005) show, monetary incentives can reduce overestimation of relative performance (but they do not improve accuracy). We provided modest monetary incentives to players for making accurate forecasts and bets. If a player's beliefs of relative performance are very spread out, then his expected rewards from taking the survey are small and do not depend much on his forecasts or bets. On the other hand, if the distribution of a player's beliefs of relative performance is tight, then the impact of monetary incentives is larger. We cannot rule out the possibility that the overestimation bias would disappear if players would have been given larger monetary incentives.

10 Conclusion

This paper shows that players in three real world tournaments tend to overestimate relative performance. The bias is present in both luck-based (poker) as well as in skill-based (chess) tournaments. This happens in the presence of financial incentives for accurate forecasts and even when players have very good information about their relative skill. We also find that players are willing to bet on their overly favorable views of relative performance. However, the degree of overestimation that we find is not as large as the ones often reported in the social psychology literature. In poker tournaments the average player overestimates his relative performance by 7 to 10 percentiles. In chess tournaments the average player overestimates relative performance by 6 to 7 percentiles.

References

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.

Burson, K, Larrick, R., and J. Klayman (2006). "Skilled or Unskilled, but Still Unaware of It: How Perceptions of Difficulty Drive Miscalibration in Relative Comparisons," *Journal of Personality and Social Psychology*, 90, 60-77.

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*, v40(1), 83-89.

Clark, J. and L. Friesen (2003). "Rational Expectations of Own Performance: An Experimental Study." Mimeo, University of Canterbury and Lincoln University.

DeGroot, M. (1970). Optimal Statistical Decisions, McGraw Hill.

Dunning, D., J. A. Meyerowitz, and A. D. Holzberg (1989). "Ambiguity and Self-Examination: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability," *Journal of Personality and Social Psychology*, 57(6), 1082-1090.

Ferraro, P. (2003). "Know Thyself: Incompetence and Overconfidence," Experimental Laboratory Working Paper Series, Georgia State University.

Gervais, S. and T. Odean (2001). "Learning to be Overconfident," *Review of Financial Studies*, Vol. 14, N^{o}1, 1-27.

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.

Kruger, J., and D. Dunning (1999). "Unskilled and Unaware of It: How Difficulties in Recognizing Ones Own Incompetence Lead to Inflated Self-Assessments," *Journal of Personality and Social Psychology*, 77, 1121-1134.

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.

Moore, D. and D. Cain (2005). "Overconfidence and Underconfidence," Working Paper, Carnegie Mellon University.

Hoelzl, E. and A. Rustichini (2005). "Overconfident: Do You Put Your Money on It?" *The Economic Journal*, Vol. 115, 305-318.

Isaac R. and J. Duncan (2000). "Just Who You Are Calling Risk Averse?" *Journal of Risk and Uncertainty*, 20, 177-187.

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.

Selten, R., Sadrieh, A., and K. Abbink (1999). "Money Does Not Induce Risk Neutral Behavior, but Binary Lotteries Do Even Worse," *Theory and Decision* 46, 211-49.

Smith V, and J. Walker (1993). "Rewards, Experience and Decision Costs in First Price Auctions," *Economic Inquiry*, 31, 237-244.

Sonnemans, J. and T. Offerman (2001). "Is the Quadratic Scoring Rule really incentive compatible?" Working Paper, University of Amesterdam.

Van den Steen, E. (2004). "Rational Overoptimism (and Other Biases)," *American Economic Review*, 94(4), 1141-1151.

Zarembka, P. (1974). "Transformation of Variables in Econometrics," in: Paul Zarembka, ed., Frontiers in Econometrics, Academic Press, New York.

11 Appendices

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 (Question 1). The survey also asks you to choose between different lotteries whose prizes depend on your relative position in the tournament (Question 3) or on the draw of a random number (Question 10). Depending on your choice of lottery, how well you perform in the tournament, and the draw of the random number you may earn up to an additional 20. 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 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 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 $w = \bar{w} + \left[w_0 - (x - f)^2\right]$, with $w_0 \geq 1$. The optimal point forecast of this individual is given by

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

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

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

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.[24] 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

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

[24] See DeGroot (1970) pp. 228.

or,

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

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

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

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

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

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

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 $c^* < Mode(X)$ since $Mode(X) = \max g(x)$. 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 $c^* < Mode(X)$. So, for a unimodal and symmetric density of beliefs we have that $c^* < Mode(X) = E(X)$. 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

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

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 $c^* < Mode(X) = E(X)$ 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 $Mode(X) < E(X)$. Note that when $g'(a) > 0$ the second-order condition is satisfied and $a < c^* < Mode(X) < E(X)$. When $g'(a) < 0$ we have a corner solution: $c^* = Mode(X) = a$. *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