How Do Beliefs about Skill Affect Risky Decisions?*

Adrian Bruhin

University of Lausanne

Luís Santos-Pinto

University of Lausanne

David Staubli

University of Lausanne

October 3, 2016

JEL Codes: C91; D81

Keywords: Individual risk taking behavior; Self-confidence; Laboratory experiment

*Luís Santos-Pinto (Corresponding author), University of Lausanne, Faculty of Business and Economics (HEC Lausanne), Internef 535, CH-1015, Lausanne, Switzerland, Ph: 41-216923658, Fax: 41-216923435, LuisPedro.SantosPinto@unil.ch; Adrian Bruhin, University of Lausanne, Faculty of Business and Economics (HEC Lausanne), Switzerland, Adrian.Bruhin@unil.ch; David Staubli, University of Lausanne, Faculty of Business and Economics (HEC Lausanne), Switzerland, David.Staubli@unil.ch

How do Beliefs about Relative Skill Affect Risky Decisions?

Abstract

Beliefs about relative skill matter for risky decisions such as career choices, market entry, and financial investments. Yet in most laboratory experiments risk is exogenously given and beliefs about relative skill play no role. We use a laboratory experiment free of strategy confounds to isolate the impact of beliefs about relative skill on risky choices. We find that low (high) skill individuals are more (less) willing to take risks on gambles with probabilities depending on relative skill than on gambles with exogenously given probabilities. Additionally, the correlation between stated and revealed beliefs – beliefs estimated from choices – is only moderate, suggesting that relying exclusively on stated beliefs may be misleading.

1 Introduction

In most laboratory experiments risk is exogenously given and individuals' beliefs about relative skill play no role in their decisions. In the real world beliefs about relative skill matter for many decisions. Examples include, following a career path where performance is highly dependent on relative skill (e.g., being a lawyer, or a musician) versus choosing a career path where performance does not depend much on it (e.g., taking a public sector job); entering a market where post-entry payoffs depend on relative skill (e.g., opening a restaurant) versus staying out and earning a certain amount (e.g., working as a waiter); or managing your own financial portfolio versus delegating this to an asset manager depends on one's perception of relative skill at picking financial assets.

This paper uses a laboratory experiment to investigate how beliefs about relative skill affect risky decisions. To answer this question we elicit certainty equivalents (CEs) of luck and skill gambles using a multiple price list format. Both types of gambles are binary as they involve two possible prizes. In a luck gamble the probability of getting the higher of the two possible prizes is given while in a skill gamble it corresponds to the subject's relative skill, measured by relative performance in a cognitive ability test. Consequently, behavior in the luck gambles only depends on preferences towards risk whereas behavior in the skill gambles depends on preferences towards risk as well as beliefs about relative skill.

To isolate the impact of beliefs about relative skill on subjects' choices we use two methods: a model free approach and a structural model approach. The model free approach directly compares the CEs of luck and skill gambles that offer the same prizes and winning probabilities. The advantage of this approach is that it does not rely on a particular model. The structural model approach uses all CEs of luck and skill gambles to estimate the parameters of a Cumulative Prospect Theory model (Kahneman and Tversky (1979), Tversky and Kahneman (1992)). The advantage of this approach is that it not only uses all the data but also allows us to identify revealed beliefs about relative skill based on subjects' choices. However, it relies on the identifying assumption that subjects apply the same utility and probability weighting functions for evaluating the luck and skill gambles.

The model free analysis shows that low (high) skill individuals have higher (lower) CEs of skill gambles than of luck gambles that offer the same prizes and winning probabilities. This indicates that low (high) skill individuals are more (less) willing to take risks on gambles where performance is determined by relative skill. In contrast, we find that intermediate skill individuals have similar CEs of skill gambles and luck gambles that offer the same prizes and winning probabilities.

Next, we investigate how these results relate to subjects' beliefs about relative skill. Subjects have to state a probability distribution over their relative performance in the cognitive ability test. We apply a quadratic scoring rule (QSR) to incentivize truthful disclosure of beliefs. Subjects' stated beliefs about relative skill display three main patterns. First, on average, subjects do not display a strong tendency to overestimate their relative skill. Second, the majority of subjects, however, has biased beliefs about relative skill: for 70.8% of them the relative skill is more than one standard deviation away from their stated mean relative skill; for 31.7% of them the relative skill lies outside the support of the stated belief distribution. Third, the biases in beliefs correlate with relative skill: low skill subjects overestimate their relative skill while high skill subjects underestimate it. Together with the model free analysis this suggests that stated beliefs about relative skill drive differences in behavior in the skill and luck gambles.

The structural model provides us with estimates for revealed beliefs and for risk preference parameters – utility curvature, optimism-pessimism, and likelihood sensitivity. To model the utility function we use power utility (constant relative risk aversion) and to model probability weighing we use the two-parameter probability weighting function in Goldstein and Einhorn (1987). The estimated parameters tell us that subjects display moderate degrees of risk aversion, pessimism, and likelihood insensitivity. These estimates of subjects' risk preferences are plausible given the existing laboratory evidence (see Wakker (2010)). The revealed beliefs confirm the patterns obtained with the stated beliefs: there is no strong overall tendency to overestimate relative skill, the majority of subjects has biased beliefs about relative skill, low skill subjects overestimate their relative skill while high skill subjects underestimate it. However, the correlation between revealed and stated beliefs is at a moderate 46%.

Since we have both stated beliefs as well as revealed beliefs, estimated via the structural model, we can compare their relative power in predicting differences in behavior across the luck and skill gambles. We find that biases in revealed beliefs predict differences between CEs of luck and skill gambles better than biases in stated beliefs. This and the moderate correlation between revealed and stated beliefs indicates that relying exclusively on the latter may be misleading.

Our paper contributes to the understanding of occupational choices in competitive labor markets (Lucas Jr (1978), Jovanovic (2004)), the low returns from self-employment and entrepreneurship (Hamilton (2000), Moskowitz and Vissing-Jorgensen (2002)), and phenomena such as perception of skill at selecting financial assets and excess trading (Odean (1998), Glaser and Weber (2007)), and over-entry into markets (Dunne et al. (1988), Mata and Portugal (1994)). Our results suggest that these behaviors are driven by systematically biased beliefs about relative skill: low (high) skill individuals are more (less) likely to select into competitive occupations, become entrepreneurs, select financial assets on their own, or enter markets because they systematically overestimate (underestimate) their relative skill. Consequently, biased beliefs about relative skill can lead to misallocations of talent: the wrong people might enter professions where performance depends on relative skill whereas the right people might be crowded out.

The paper also contributes to the experimental literature that studies the role of beliefs about skill for individual decisions (e.g., Camerer and Lovallo (1999), Moore and Kim (2003), Hoelzl and Rustichini (2005), Moore and Cain (2007), Clark and Friesen (2009), Bolger et al. (2008), Park and Santos-Pinto (2010), Eil and Rao (2011), Burks et al. (2013), Mobius et al. (2011), Charness et al. (2013), Cain et al. (2015)). Some of these studies involve strategic situations where individuals must make inferences about the rivals' strategies and beliefs (e.g., Camerer and Lovallo (1999), Charness et al. (2013)). In contrast, a major contribution of our paper is to introduce a non-strategic setting which allows us to directly identify the role that beliefs about skill play in explaining risky choices.

Another contribution is that the model free as well as the structural model approaches are both based on revealed preference. This adds to the literature which, so far, has mostly relied on subjects' stated beliefs (e.g., Clark and Friesen (2009), Park and Santos-Pinto (2010), Eil and Rao (2011), Burks et al. (2013), Mobius et al. (2011)). Our results indicate that revealed beliefs are only moderately correlated with stated beliefs and better predict differences in behavior across luck and skill gambles.

2 Experimental Design

The experiment consists of three stages which are explained in detail below. In the first stage, each subject completes a cognitive ability test and is assigned to a skill level according to her relative performance in the test. In the second stage we elicit CEs of luck and skill gambles. Finally, in the last stage we elicit subjects' beliefs about relative skill.

2.1 Cognitive Ability Test and Assignment of Subjects to Skill Levels

Subjects perform a Raven's Matrices test with 20 questions. In each question subjects are asked to identify the missing element that completes a pattern out of six possible choices. The difficulty of the questions varies in order to create heterogeneity in the test score results. Each subject had 20 minutes to complete the test. To incentivize the subjects to do their best, we paid CHF 0.50 per correct answer and told subjects that a good test result could help them earning more money later on during the experiment.[1]

A subject's relative performance in the cognitive ability test determines her relative skill among the 20 participants. The relative skill of each participant is assigned to one of 10 skill levels θ from the set

[1]At the time of the experiment one CHF corresponded to roughly 1.05 USD.

$\{0.05, 0.15, \ldots, 0.85, 0.95\}$ as follows: the two best subjects (the tenth decile) in the room are allocated to the 0.95 skill level, the two next best subjects (the ninth decile) in the room to the 0.85 skill level, etc., and the two worst subjects (the first decile) in the room to the 0.05 skill level.[2] Later on, these skill levels will correspond to the winning probabilities in the skill gambles.

The assignment of subjects to skill levels was explained in detail. After reading the instructions subjects answered comprehension questions about the assignment. They could only go on after answering these questions correctly. While they were making their risky decisions subjects could click on a button to see the list showing which winning probability was implied by which skill level. For further details please refer to the Experimental Material Appendix.

2.2 Risky Decisions

After the cognitive ability test and the assignment of subjects to skill levels we elicited CEs of luck and skill gambles. A luck gamble pays CHF x_h^L with probability p and CHF x_l^L with probability $1 - p$. We vary p from 5% to 95% in 10% increments across three prize combinations (x_h^L, x_l^L) from the set

$$\{(140, 0), (120, 20), (100, 40)\}. \tag{1}$$

The 10 winning probabilities p and the three prize combinations yield a total of 30 luck gambles.

In contrast, a skill gamble pays CHF x_h^S with probability θ_i and CHF x_l^S with probability $1 - \theta_i$ where θ_i is the skill level of subject i which is

[2] To break ties subjects had to give their best guess about the sum of nine numbers that were displayed on matrix for 10 seconds. If this did not resolve the tie we broke the ties randomly. Note that since there were two subjects for each skill level, the tiebreaker was not relevant in every case. The performance in the estimation questions that breaks the tie between the 3^{rd} and the 4^{th} is inconsequential since rank 3 and rank 4 give rise to the same skill level.

endogenously determined by her relative performance in the cognitive ability test. Since the winning probability θ_i corresponds to the subject's skill level we can only vary the prizes of the skill gambles. We consider nine different prize combinations (x_h^S, x_l^S) from the set

$$\begin{aligned}\{&(180,0),(150,30),(120,60),\\ &\qquad (140,0),(120,20),(100,40),\\ &\qquad\qquad (100,0),(80,20),(60,40)\}\,. \qquad (2)\end{aligned}$$

Note that for the prize combinations $(140, 0)$, $(120, 20)$, and $(100, 40)$ there are corresponding luck gambles. We will exploit this correspondence in the model free approach to directly compare CEs between skill and luck gambles.

To elicit CEs of both luck and skill gambles we use the multiple price list format (see Andersen et al. (2006) for a discussion). We have subjects choosing between a series of certain payoffs and either a luck or a skill gamble. The series of certain payoffs covers the payoff range of the gamble's prizes and decreases in 18 $(= 19 - 1)$ equally sized amounts. Subjects typically start by preferring the first certain payoff to the gamble and then switch to the gamble before the last certain payoff. The arithmetic mean of the last certain payoff preferred to a gamble and the first certain payoff not preferred to a gamble determines the CE of the gamble. In the example depicted in Figure 1 the CE is approximated as $(84 + 91)/2 = 87.50$. If the subject always chooses the gamble the CE is approximated as $(140 + 133)/2 = 136.50$. Likewise, if she always chooses the certain payoff the CE is approximated as $(7 + 0)/2 = 3.50$.

We impose a unique switching point per multiple price list by automatically filling in all choices following the first switch (for details see Andersen et al. (2006) and Tanaka et al. (2010)). This has two advantages. First, it allows to dermine the CE for every gamble. Second, it substantially reduces

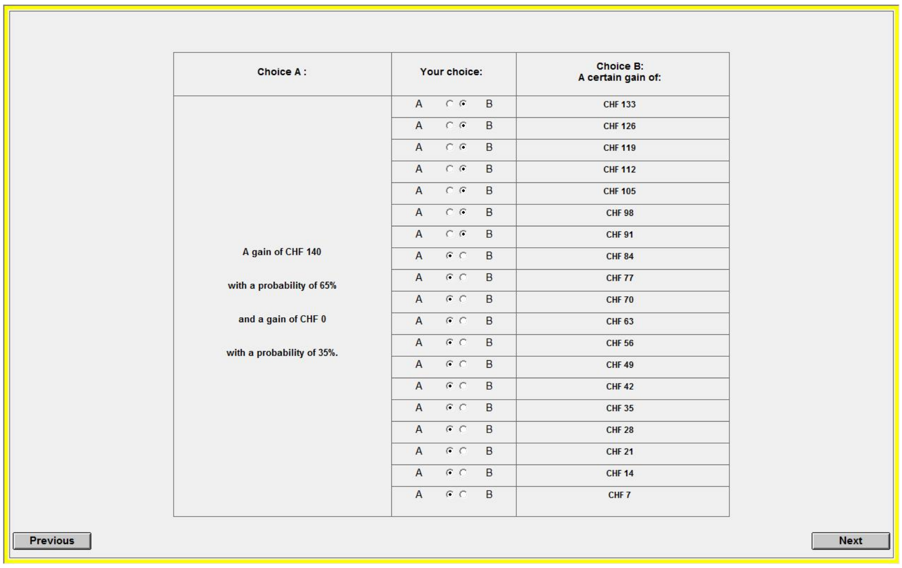


Figure 1: Interface of the luck gamble with x_h = CHF 140 and x_l = CHF 0 and a winning probability of 0.65. In this case, the elicited CE is CHF 87.50 (the arithmetic mean of CHF 84 and CHF 91).

the number of clicks given the high number of gambles.[3]

The luck and skill gambles were presented in blocks. Across sessions the order of the blocks was balanced to cancel out potential order effects (see Charness et al. (2012)). Within a given block the gambles were ordered according to two dimensions: attractiveness (first-order stochastic dominance) and prize spread ($x_h - x_l$). To control for order effects across these two dimensions we assign one fourth of the subjects to each of the four configurations: (i) increasing attractiveness and increasing spread, (ii) increasing attractiveness and decreasing spread, (iii) decreasing attractiveness and increasing spread, and (iv) decreasing attractiveness and decreasing spread.

Since the main focus of the experiment is on risky choices we always

[3]In terms of additional assumptions, a unique switching point imposes weak monotonicity of the utility function, i.e. subjects weakly prefer a higher amount of money to a lower amount of money. Nevertheless, the unique switching points do not prevent choices across gambles that are inconsistent with first-order stochastic dominance.

elicit risky decisions before beliefs.

2.3 Belief Elicitation

After making risky decisions subjects stated their beliefs about their relative skill. We elicit each subject's entire belief distribution across the 10 skill levels. The subject can move sliders to shift probability mass across each skill level which are represented by a histogram in real-time. We incentivize the disclosure of beliefs with a quadratic scoring rule (QSR) where the maximum payment is CHF 10 and the minimum is CHF 0 (Kadane and Winkler, 1988). Our main measure of the subject i's stated belief about relative skill is the mean belief $\mu_i = \sum_{m=1}^{10} f_i(\theta_m)\theta_m$, where $f_i(\theta_m)$ stands for i's stated probability mass of having skill level θ_m, with $m = 1, ..., 10$. For further details see Appendix A.[4]

Note, however, that the QSR assumes that subjects weigh probabilities linearly. If subjects engage in probability weighting their stated beliefs will reflect subjectively distorted probabilities instead of objective ones. In the discussion of the results we will first ignore this issue when we compare the findings of the model free approach to the subjects' stated beliefs. However, we will come back to it in the context of the structural model approach when we compare revealed beliefs to stated beliefs.

2.4 Subjects and Payments

The experiment was performed at the University of Lausanne in April and May 2015 in the computer lab using the software z-Tree (Fischbacher, 2007). All subjects were students of the University of Lausanne and the École Polytechnique Fédérale de Lausanne (EPFL), recruited via ORSEE (Greiner,

[4] We also elicit each subject's mode of her belief distribution across skill levels. We ask each subject to provide a point estimate of her relative skill in the cognitive ability test and pay CHF 5 for a correct guess and CHF 0 for an incorrect guess. This allows us to distinguish rational from irrational beliefs using the method of Burks et al. (2013).

2004). The experiment was conducted in 16 separate sessions with 20 subjects each. One subject left without finishing the experiment. Thus, our data comprises 319 subjects.[5]

In order to provide incentives for truthful revelation of preferences and beliefs, subjects were randomly paid for only one of their risky decisions (Azrieli et al. (2012) show that this is the only incentive-compatible way to use multiple price lists). We use the prior incentive scheme (Prince) proposed by Johnson et al. (2014). At the start of the experiment subjects were told that they would be paid for only one of their decision situations and that decision situation was contained inside a closed envelope that they drew randomly from a box. Inside each envelope there is a decision sheet with the multiple price list corresponding to that decision. Subjects were told that they could not open the envelope before the payment stage. In the payment stage each subject was paid according to the decision situation in the envelope, the subject's choice in that decision situation, the subject's performance on the cognitive ability test (if the decision situation involved a skill gamble and the subject preferred the skill gamble to the certain amount) and the realization of two ten-sided dice (if the subject preferred the luck or skill gamble to the certain amount).

Each subject received a show-up fee of CHF 10. Subjects were paid in private, one-by-one at the end of the experimental session. The information on a subject's earnings from each task and total earnings was printed on a sheet of paper that was shown to each subject. Total earnings per subject varied between CHF 17.50 and CHF 205.00 with a mean of CHF 105.85 and median of CHF 113.50. Each session lasted for approximately two hours, comprising roughly 90 minutes for decisions and 30 minutes for payments. Total earnings paid to subjects across the 16 sessions were CHF

[5]The subject left after finishing the cognitive ability test. The skill levels are thus determined including that subject.

$34,213.00$.[6] Compared to most other experimental work on individual decision making these are high stakes per hour. The instructions can be found in the Experimental Material Appendix.

3 Results

In this section, we first present the results of the model free approach. Second, we report the results on the cognitive ability test and the subjects' stated beliefs. Finally, we show the findings of the structural model approach.

3.1 Model Free Approach

The model free approach compares the CEs of luck and skill gambles that offer the same prizes and winning probabilities. This allows us to directly isolate the effect of beliefs about relative skill on the behavior towards skill gambles. For each subject, there are always three comparable pairs of luck and skill gambles offering prizes from the set $\{(140, 0), (120, 20), (100, 40)\}$. For example, for a subject with skill level $\theta_i = 0.55$, we compare the CEs of the three skill gambles with the CEs of the three luck gambles that have a corresponding winning probability of $p = 0.55$.

On average, the CEs of comparable luck and skill gambles do not differ. The average CE of the skill gambles offering prizes from the set $\{(140, 0), (120, 20), (100, 40)\}$ is CHF 57.105, while the average CE of the corresponding luck gambles is CHF 57.807. The difference amounts to CHF -0.702 and is not significantly different from 0 (p-value=0.576).

However, a more disaggregate look at the data reveals a correlation between relative skill and differences between CEs of skill and luck gambles.

[6] Since we needed exactly 20 subjects per session, we overbooked each session and sent home the supernumerous subjects by compensating them with the show-up fee of CHF 10.

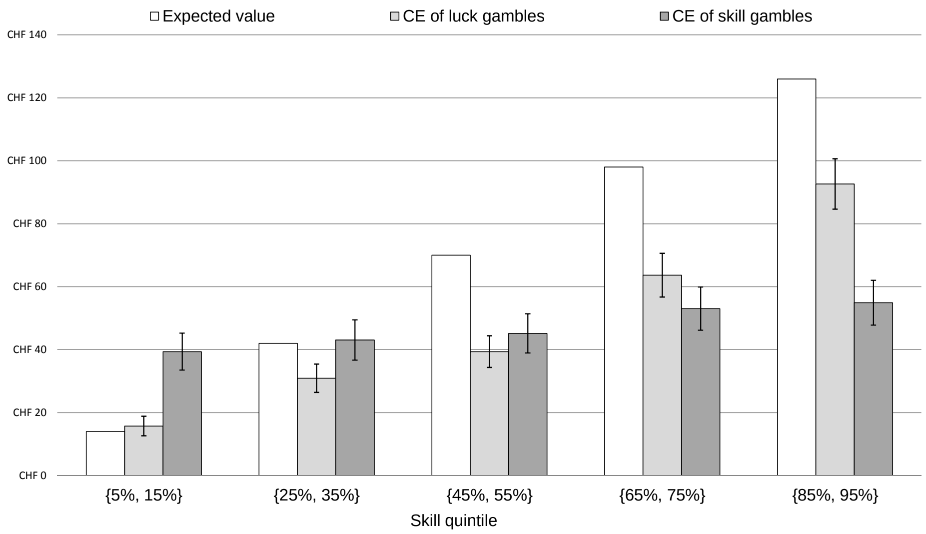


Figure 2: Average expected values of the luck gambles (white bars), average CEs of the luck gambles (light grey bars), and average CEs of the skill gambles (dark grey bars). Depicted is the gamble with winning prize $x_h = 140$ and losing prize $x_l = 0$. The 95% -confidence intervals are based on robust standard errors. Number of observations = 319.

In Figure 2 we divide subjects into quintiles according to their skill levels for luck and skill gambles offering prizes $(140, 0)$.[7] The figure reveals that low skill subjects have higher CEs of skill gambles than luck gambles, while for high skill subjects the reverse happens. The regression shown in Table 1 confirms this correlation over all three comparable pairs of luck and skill gambles. The dependent variable, CE_{ij}^{diff}, represents the difference between subject i's CE of skill gamble j and her CE of luck gamble j. The independent variables are indicators for the skill quintiles. The constant is taken out to avoid perfect multicollinearity. The interpretation of the coefficients is straightforward: subjects in the lowest skill quintile, corresponding to skill levels $\{0.05, 0.15\}$, value the skill gambles on average CHF 16.768 more than

[7]The analogous figures for the luck and skill gambles with the prizes $(120, 20)$ and $(100, 40)$, respectively, are available in Appendix B and display a similar pattern.

Skill {0.85, 0.95}	-25.380***
	(2.044)
Skill {0.65, 0.75}	-9.740***
	(2.379)
Skill {0.45, 0.55}	4.495***
	(1.551)
Skill {0.25, 0.35}	10.709***
	(1.807)
Skill {0.05, 0.15}	16.768***
	(2.372)
Observations	955
R-squared	0.345

Subject cluster-robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1: Regression where the dependent variable is the difference between the CE of skill gamble, i.e., CE_{ij}^{S}, and the CE of the corresponding luck gamble, i.e., CE_{ij}^{L}, and the independent variables are skill quintile dummies. Coefficients can be interpreted in CHF.

the corresponding luck gambles. As skill quintiles increase the difference between the CEs of skill and luck gambles decreases in absolute value down to CHF -25.380 for the top quintile. The numbers in parentheses display the subject cluster-robust standard errors. There are 955 observations, three observations per subject less two observations that were not recorded during the experiment ($955 = 3 \times 319 - 2$).

In sum, these findings yield our first main result:

Result 1 *On average, the CEs of skill gambles are similar to the CEs of corresponding luck gambles. However, subjects in the lower three skill quintiles (from the 5% to the 55% skill levels) have higher CEs of skill gambles than of luck gambles, and subjects in the highest two skill quintiles (65% and 95% skill levels) have lower CEs of skill gambles than of luck gambles.*

3.2 Cognitive Ability Test and Stated Beliefs

The average test score, measured by the number of correct answers, was 12.66 with a standard deviation of 2.80. The distribution of test scores is approximately normal (the p-value of a D'Agostino's Chi-squared test is 0.534). The minimum score was 2 and the maximum score was 19. That is, no subject got all 20 questions correct, so there is no bunching at the top of the score distribution. In a trivially simple test, we would expect (close to) 20 correct answers for each subject. In contrast, in a prohibitively difficult test we would expect, on average, $20/6 = 3.33$ correct answers as subjects would have to guess randomly. Hence, the observed average of 12.66 correct answers is slightly above the middle of those two extremes.[8]

[8] Across the sixteen sessions, there were 87 ties between two or more subjects after answering the cognitive ability test. After the tiebreaker there were 18 ties left that had to be decided by means of the random number generator. Per session there are a total of $\binom{20}{2} = 190$ bilateral comparisons. Across the 16 sessions this amounts to 3040. The 87 ties account for 2.86% and the 18 random tie breaks account for 0.59% of all bilateral

In terms of the subjects' stated beliefs there are three main findings:

Result 2 *On average, subjects do not display a strong tendency to overestimate their relative skill.*

The average stated belief indicates a weak tendency towards overconfidence.[9] On average, subjects state that their individual winning probability is 0.54. This is slightly but significantly different from the true average winning probability of 0.5 (p-value < 0.001).[10]

This weak tendency towards overconfidence needs a comment. At first glance Result 2 seems to be at odds with a significant body of literature where overconfidence is a well-established fact (Lichtenstein et al. (1977), Svenson (1981), Camerer and Lovallo (1999), Van den Steen (2004), Santos-Pinto and Sobel (2005), Burks et al. (2013)). DeBondt and Thaler (1995) names it "perhaps the most robust finding in the psychology of judgement is that people are overconfident." However, overconfidence about relative skill is not an ubiquitous phenomenon. Experimental evidence shows that people overestimate their relative performance on easy tasks (overplacement) but underestimate their performance relative to others on harder tasks (underplacement) (see Kruger (1999)). As pointed out above, our cognitive ability test is a moderate difficulty task and this might be why we don't find a very strong overall tendency for either overplacement or underplacement.

Result 3 *The majority of subjects state biased beliefs about relative skill.*

comparisons.

[9] Moore and Healy (2008) summarize the following three types of overconfidence about skill: (1) overestimation of one's actual performance, (2) overplacement of one's performance relative to others, and (3) excessive precision in one's beliefs.

[10] The distribution of relative skill levels is uniform by construction with 10% of the subjects per disaggregated skill level θ_1, θ_2,..., θ_{10} (and 20% per aggregated skill level). The average winning probability is 0.5 by construction with a standard deviation of 0.288. However, the average winning probability was slightly more than 0.5 since the subject that left without finishing the experiment (but after doing the cognitive ability test) had a winning probability of 0.35.

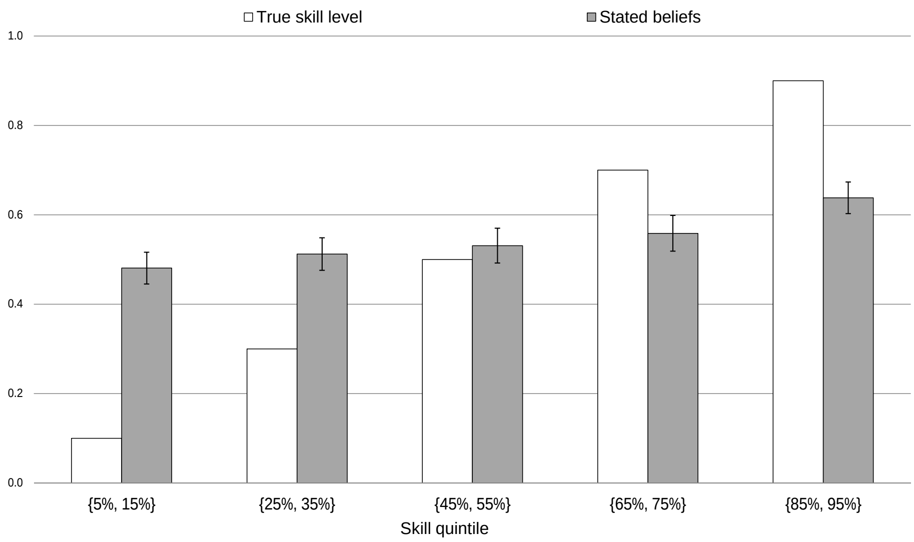


Figure 3: Average true skill levels (white bars) and average stated beliefs (gray bars) per skill quintile. The 95% -confidence intervals are based on robust standard errors. Number of observations = 319.

For 70.8% of subjects the true skill level falls outside the average stated belief plus/minus one standard deviation of the belief distribution. In addition, we find that for 31.7% of subjects the true skill level falls outside the support of their stated belief distribution.

Result 4 *Subjects in the two lowest skill quintiles (from the 5% to the 35% skill level) overestimate their relative skill, and subjects in the two highest skill quintiles (from the 65% to the 95% skill level) underestimate their relative skill.*

Figure 3 exhibits the true skill levels and the average stated belief per skill level. The white bars represent the average of the true skill levels in each quintile. For example, the first skill quintile contains subjects with winning probabilities 0.05 and 0.15 (with equal share) giving rise to the average of 0.1. The gray bars exhibit subjects' average stated beliefs. The error bars display 95%-confidence intervals based on robust standard errors.

We refer to the difference between stated beliefs and true skill levels as the bias in stated beliefs, $b_i^{stated} \equiv \mu_i - \theta_i$. Subjects in the lowest two skill quintiles state, on average, significantly overconfident beliefs. Subjects in the highest two skill quintiles state, on average, significantly underconfident beliefs. Subjects in the middle skill quintile state slightly overconfident beliefs, however, these slight overconfidence is only statistically significant at the 10% level.

The pattern of stated beliefs depicted in Figure 3 could be broadly consistent with subjects being uninformed about their skills. To rule out this possibility, we compare each subject's stated belief distribution to the uniform distribution, representing complete skill uncertainty. The corresponding Chi-squared test has the following test statistic:

$$\chi_9^2 = \sum_{m=1}^{10} \frac{(w_m - Nq_m)^2}{Nq_m}$$

where $N = 100$, as there are 100 percentage points to be allocated; $q_m = 0.1$, $\forall m$, as according to the null hypothesis of a uniform distribution there is a 10% probability weight on every skill level; and w_m is the number of percentage points that a subject puts on skill level m (see Kenney and Keeping (1954)). The test statistic is compared against the critical value of a Chi-squared distribution with 9 degrees of freedom, i.e., the number of skill levels minus 1. The null hypothesis that stated beliefs are uniformly distributed can be rejected at both the 5%- and the 1%-significance level for all but 3 subjects. At 10%-significance level the null hypothesis can be rejected for all but one subject.

In conclusion, subjects' stated beliefs indicate that subjects have precise perceptions about their relative skill but these perceptions are biased.

3.3 Structural Model Approach

We now discuss the results of the structural model approach which allows us to infer subjects' revealed beliefs. We present our decision model, explain the estimation strategy, and conclude with the estimation results.

3.3.1 Decision Model

We use Cumulative Prospect Theory (Kahneman and Tversky (1979), Tversky and Kahneman (1992)) as our model of choice under risk and under uncertainty, since this model has proven to be descriptively superior to other models (see Starmer (2000)). According to Cumulative Prospect Theory (CPT), the value of a gamble X that pays x_h with probability p and x_l with probability $1-p$ is

$$V(X) = w(p)u(x_h) + [1 - w(p)]u(x_l),$$

where $u(\cdot)$ is the utility function and $w(\cdot)$ is the probability weighting function. We specify the utility function using the power function that implies constant relative risk aversion (CRRA), the benchmark approach in most empirical work on risk attitudes:

$$u(x;r) \begin{cases} \frac{x^{(1-r)}}{1-r}, & r \neq 1 \\ \ln(x), & r = 1 \end{cases},$$

where x is the monetary outcome of a gamble and the parameter r measures the coefficient of CRRA. The parameter r determines the curvature of the utility function: $r > 0$ corresponds to risk aversion, $r = 0$ to risk neutrality, and $r < 0$ to love for risk.

To specify the probability weighting function, we use the two-parameter function proposed by Goldstein and Einhorn (1987):

$$w(p;\delta,\gamma) = \frac{\delta p^{\gamma}}{\delta p^{\gamma} + (1-p)^{\gamma}},$$

where γ and δ reflect the degree of probability weighting. The parameter γ governs likelihood sensitivity. If $\gamma \in (0,1)$, the function captures the inverse s-shape pattern where low probabilities are upweighted and high probabilities are downweighted. If $\gamma > 1$ we have an s-shape pattern where low probabilities are downweighted and high probabilities are upweighted. The parameter δ reflects the degree of optimism or pessimism. When $\delta \in (0,1)$ the subject is a pessimist, since she downweights the probability of the high prize x_h and upweights the probability of the low prize x_l. When $\delta > 1$ the subject is an optimist, since she upweights the probability of the high prize x_h and downweights the probability of the low prize x_l. When $\gamma = \delta = 1$ there is no probability weighting and $w(p; 1, 1) = p$.

3.3.2 Estimation Procedure

To estimate risk preferences at the individual level we compute the CEs implied by our decision model for luck and skill gambles. Subject i's CE of a luck gamble $j = 1, ..., 30$ that pays x^L_{hj} with probability p_j and x^L_{lj} with probability $1 - p_j$, where p_j is the exogenously given winning probability, is implicitly defined by:

$$u(\hat{CE}^L_{ij}; r^L_i) = w(p_j; \gamma^L_i, \delta^L_i)u(x^L_{hj}; r_i) + [1 - w(p_j; \gamma^L_i, \delta^L_i)]u(x^L_{lj}; r^L_i), \quad (3)$$

where $\hat{CE}^L_{ij}$ is subject i's CE of luck gamble j predicted by the model. Similarly, subject i's CE of a skill gamble $k = 1, ..., 9$ that pays x^S_{hk} with probability θ_i and x^S_{lk} with probability $1 - \theta_i$, where θ_i is the probability of winning for subject i determined by her relative skill is implicitly defined by:

$$u(\hat{CE}^S_{ik}; r^S_i) = w(\theta_i; \gamma^S_i, \delta^S_i)u(x^S_{hk}; r^S_i) + [1 - w(\theta_i; \gamma^S_i, \delta^S_i)]u(x^S_{lk}; r^S_i),$$

where $\hat{CE}^S_{ik}$ is subject i's CE of skill gamble k predicted by the model. The prize combinations (x^L_{hj}, x^L_{lj}) are given by (1), (x^S_{hk}, x^S_{lk}) by (2), and $p_j = .05, .15, ..., .85, .95$.

In the luck gambles both prizes and winning probabilities are known to subjects. The estimation of the parameters is thus straightforward in this case and can be implemented in a similar way as in Bruhin et al. (2010). In skill gambles however, the subjects do not know their winning probability as they do not know their relative skill for sure. If we estimate the parameters of the utility and probability weighting functions using the objective winning probabilities θ_i we would obtain erroneous estimates since biases in beliefs would induce distortions in the estimates of these parameters. To illustrate, consider the case of a subject i that has an objective winning probability $\theta_i = 0.15$ but who believes her winning probability to be 0.45. If this subject evaluates the skill gambles based on her biased belief and we use the objective winning probability in our estimation, we would draw the wrong conclusion that subject i is very optimistic or insensitive to changes in probabilities. Thus, we replace the objective winning probability θ_i by a parameter ξ_i that reveals subject i's belief about her relative skill and needs to be estimated. We refer to the parameter ξ_i as the subject's revealed belief.

We jointly estimate the individual parameters of the utility function, the probability weighting function, and the subject's revealed belief using 30 luck gambles and the 9 skill gambles using the method of maximum likelihood. The main assumption for identifying the subject's revealed belief is that the utility and probability weighting functions remain stable across luck and skill gambles. Formally, this identifying assumption corresponds to:

$$r_i^S = r_i^L = r_i; \quad \gamma_i^S = \gamma_i^L = \gamma_i; \quad \delta_i^S = \delta_i^L = \delta_i; \forall i$$

Using the identifying assumption and rearranging equation (3) we obtain an expression for subject i's predicted CE of luck gamble j:

$$\hat{CE}_{ij}^L = u^{-1}\Big(w(p_j;\gamma_i,\delta_i)u(x_{hj}^L;r_i) + (1 - w(p_j;\gamma_i,\delta_i))u(x_{lj}^L;r_i);r_i\Big).$$

Analogously, the predicted CE of skill gamble k corresponds to:

$$\hat{CE}_{ik}^{S} = u^{-1}\Big(w(\xi_i;\gamma_i,\delta_i)u(x_{hk}^{S};r_i) + (1 - w(\xi_i;\gamma_i,\delta_i))u(x_{lk}^{S};r_i);r_i\Big).$$

The observed CEs can be decomposed into the predicted CEs and a random error: $CE_{ij}^{L} = \hat{CE}_{ij}^{L} + \varepsilon_{ij}^{L}$ for luck gambles, and $CE_{ik}^{S} = \hat{CE}_{ik}^{S} + \varepsilon_{ik}^{S}$ for skill gambles.

There are several reasons why the observed CEs may differ from the predicted ones. The most obvious one is that the multiple price list approach only allows us to approximate the observed CEs by taking an average of the certain amounts just above and below the switching point. Furthermore, Hey and Orme (1994) point out that carelessness, hurry, or inattentiveness can lead to accidentally wrong answers.

Following Hey and Orme (1994) we assume that subject i's random error for luck gamble j follows the distribution $\varepsilon_{ij}^{L} \sim N(0, \sigma_{ij}^{L})$, where the standard deviation σ_{ij}^{L} equals $\kappa_i^{L}(x_{hj}^{L} - x_{lj}^{L})$. Analogously, for skill gamble k: $\varepsilon_{ik}^{S} \sim N(0, \sigma_{ik}^{S})$, where $\sigma_{ik}^{S} = \kappa_i^{S}(x_{hk}^{S} - x_{lk}^{S})$.[11] The variance of the random error is proportional to the corresponding gamble's payoff range, as the subject's CE is elicited with respect to 19 certain amounts that are equally spread out within the gamble's payoff range.[12]

The likelihood function is given by:

$$L(\psi_i; CE_{ij}^{L}, CE_{ik}^{S}, x_{hj}^{L}, x_{lj}^{L}, x_{hk}^{S}, x_{lk}^{S}) = \prod_{j=1}^{30} \frac{1}{\sigma_{ij}^{L}} \phi\Big(\frac{CE_{ij}^{L} - \hat{CE}_{ij}^{L}}{\sigma_{ij}^{L}}\Big) \prod_{k=1}^{9} \frac{1}{\sigma_{ik}^{S}} \phi\Big(\frac{CE_{ik}^{S} - \hat{CE}_{ik}^{S}}{\sigma_{ik}^{S}}\Big),$$

[11] A likelihood-ratio test shows that the data doesn't support the assumption of equal variance of erros in luck and skill gambles. The null hypothesis that $\kappa_i^{S} = \kappa_i^{L}$ is rejected at the 5%-level for 49% of the subjects.

[12] For instance, in the gamble with $x_h = 140$ and $x_l = 0$ a shift of the switching point by one line translates into a change in the CE of 7. On the other hand, if $x_h = 100$ and $x_l = 40$ a shift by one line translates into a change in the CE of 3.

$\psi_i = (r_i, \delta_i, \gamma_i, \xi_i, \kappa_i^S, \kappa_i^L)'$ is the vector of parameters that are estimated by numerically maximizing L. To increase numerical precision and ensure convergence of the individual estimations, we apply the following restrictions on the parameters: $-10 \leq 1 - r_i \leq 10$; $0 \leq \gamma_i \leq 10$; $0 \leq \delta_i \leq 10$; $0.05 \leq \xi_i \leq 0.95$.[13] The reason for the restriction on ξ_i is that individual winning probabilities are 0.95 in the best case and 0.05 in the worst case. In order to account for the possibility that a subject's choices are serially correlated, we estimate cluster-robust standard errors at the subject-level.

The individual estimations allow us to directly compare the revealed beliefs ξ_i to the true skill level, θ_i, for each subject. We define the bias in revealed beliefs as $b_i^{revealed} = \xi_i - \theta_i$. If subject i's revealed belief is correct and she uses it for evaluating the skill gambles, then $b_i^{revealed} = 0$. However, if she evaluates the skill gambles based on a too high (low) revealed belief, then $b_i^{revealed} > 0$ ($b_i^{revealed} < 0$).

3.3.3 Estimation Results

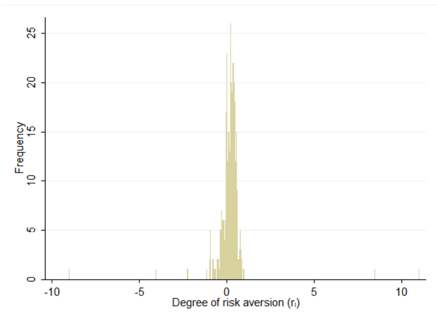
Table 2 shows the summary statistics of individual parameter estimates. The average of ξ_i shows that the revealed beliefs are slighly higher than the true average relative skill of 0.5. Furthermore, the average estimates indicate a moderate degree of CRRA ($r_i > 0$) and pessimism ($\delta_i < 1$) as well as pronounced likelihood insensitivity ($\gamma_i < 1$). Figure 4 exhibits the distributions of the individual parameter estimates of (a) r_i, (b) γ_i, (c) δ_i, and (d) ξ_i.[14]

[13] The bounds for the risk aversion and the probability weighting parameters are hit 11 times (nine subjects hit one bound and one subject hits two bounds). The lower bound for γ is hit six times, the upper bound for δ is hit twice, and the remaining bounds are hit once.

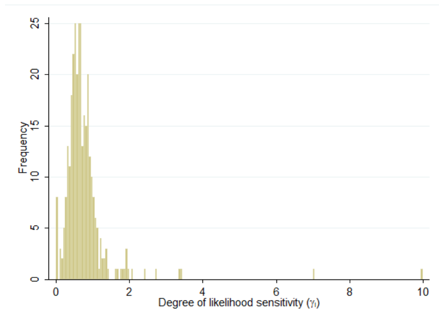
[14] We also tested several alternative specifications for the utility and probability weighting functions, namely constant absolute risk aversion and the two-parameter-version of Prelec's (1998) probability weighting function. They all yield qualitatively almost identical results. In addition, we tested an alternative measure of relative skill. Instead of ranking within each session we constructed ranking within the full sample of 320 subjects.

	Mean	Median	Standard Deviation	Min	Max
Risk aversion (r_i)	0.20	0.24	1.03	-9.00	11.00
Likelihood sensitivity (γ_i)	0.77	0.65	0.76	0.00	10.00
Optimism-pessimism (δ_i)	0.90	0.71	0.90	0.00	10.00
Revealed beliefs (ξ_i)	0.54	0.55	0.21	0.05	0.95
Scaling of error term variance (luck) (κ_i^L)	0.08	0.07	0.05	0.00	0.32
Scaling of error term variance (skill) (κ_i^S)	0.11	0.10	0.06	0.00	0.47

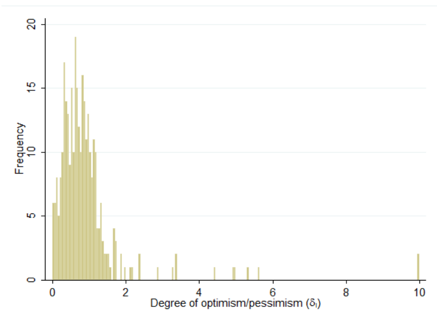
Table 2: Summary statistics of the individual parameter estimates. Number of observations = 319, number of females = 114



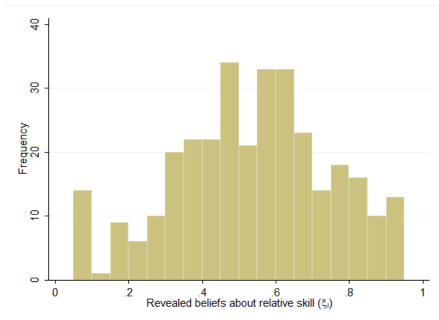
(a)



(b)



(c)



(d)

Figure 4: Histogram of estimated (a) degrees of risk aversion (r_i), (b) degrees of likelihood sensitivity (γ_i), (c) degrees of optimism-pessimism (δ_i), and (d) revealed beliefs (ξ_i). The bar width is 0.05. Number of observations = 319.

Skill {0.85, 0.95}	-0.270***
	(0.025)
Skill {0.65, 0.75}	-0.175***
	(0.029)
Skill {0.45, 0.55}	0.063
	(0.025)
Skill {0.25, 0.35}	0.190***
	(0.025)
Skill {0.05, 0.15}	0.374***
	(0.026)
Observations	319
R-squared	0.570

Note: Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 3: Regression where the dependent variable is the bias in revealed beliefs, i.e., $b_i^{revealed} \equiv \xi_i - \theta_i$ and the independent variables are skill quintile dummies.

Result 5 *The revealed beliefs confirm the pattern found in the stated beliefs: Subjects in the two lowest skill quintiles (from the 5% to the 35% skill level) overestimate their relative skill, and subjects in the two highest skill quintiles (from the 65% to the 95% skill level) underestimate their relative skill.*

Table 3 exhibits the average bias in revealed beliefs per skill quintile. The revealed beliefs of subjects in the two lowest skill quintiles, $\{0.05, 0.15\}$ and $\{0.25, 0.35\}$, indicate overestimation of relative skill. These subjects are attracted to skill gambles because they overestimate their winning probability. In contrast, the revealed beliefs of subjects in the two highest skill quintiles, $\{0.65, 0.75\}$ and $\{0.85, 0.95\}$, indicate underestimation of relative skill. These subjects are unattracted by skill gambles because they underestimate their winning probability. The revealed beliefs of subjects in the

For further details, please refer to Appendix C.

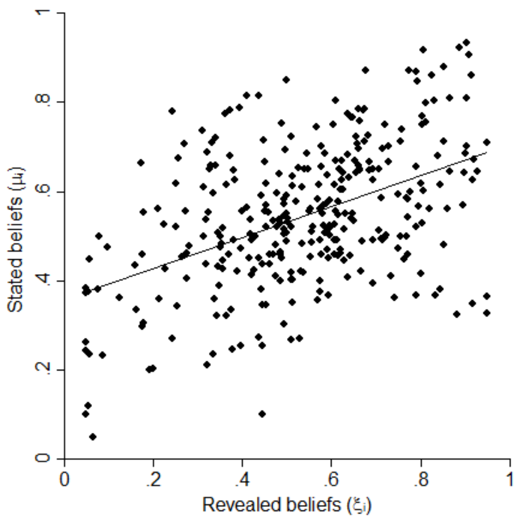


Figure 5: Relation between stated beliefs (μ_i) and revealed beliefs (ξ_i). Scatterplot and regression line. Number of observations = 319.

intermediate skill quintile, $\{0.45, 0.55\}$, indicate a slight overestimation of relative skill which is, however, not significant.

Result 6 *There is, however, only a moderate correlation between stated and revealed beliefs.*

Figure 5 exhibits a scatterplot and a regression line illustrating the correlation between the stated beliefs, μ_i, and the revealed beliefs, ξ_i. If subjects consistently use their stated beliefs to evaluate the winning probabilities in skill gambles, then we would expect a regression line with a slope of one and a strong positive correlation between stated and revealed beliefs.

However, the large dispersion of observations in the scatterplot suggests that the two measures of the subjects' beliefs are only moderately correlated. In fact, the correlation coefficient between μ_i and ξ_i amounts to just 0.461. Moreover, the estimated slope of the regression line (robust standard errors in parentheses),

$$\mu_i = \underset{(0.024)}{0.358} + \underset{(0.043)}{0.348}\xi_i,$$

amounts to only 0.348 which is significantly below one (p-value < 0.001).

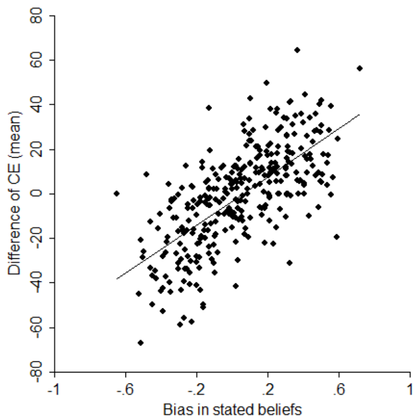
The moderate correlation between stated and revealed beliefs indicates that subjects make choices that are inconsistent with their stated beliefs. One reason for this inconsistency could be that subjects, on average, exhibit a strong degree of probability weighting which invalidates the incentive mechanism for eliciting the stated beliefs. Consequently, relying exclusively on stated beliefs may be misleading.

Moreover, the bias in revealed beliefs is a better predictor for the differences in the subjects' CEs in the skill and luck gambles than the bias in stated beliefs. Figure 6 shows the scatterplots and regression lines illustrating the relationship between the difference of the CEs in the skill and luck gambles and the biases in stated (left-hand side) and revealed (right-hand side) beliefs. The correlation of the differences of the CEs in the skill and luck gambles with the bias in revealed beliefs is 0.796, significantly higher (p-value < 0.001) than the corresponding correlation of 0.672 with the bias in stated beliefs.

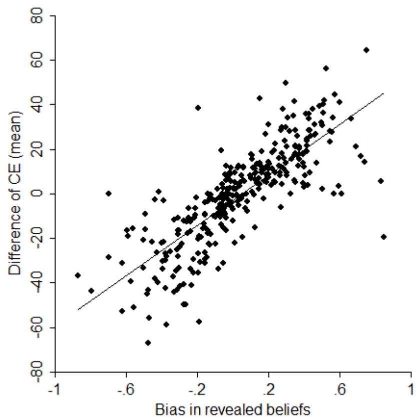
In sum, the comparison between stated and revealed beliefs provides two main insights: First, given the moderate correlation between stated and revealed beliefs, the former should be interpreted with caution. Second, biases in revealed beliefs are a better predictor of the difference between the CEs of skill and luck gambles than biases in stated beliefs.

4 Conclusion

We use a laboratory experiment to study the role played by beliefs about relative skill on risky decisions. To answer this question we compare individuals' behavior in luck and skill gambles using a model free approach and a structural model approach. Both approaches show that low (high) skill individuals are more (less) willing to take risks on gambles where the probabilities depend on relative skill than on gambles with exogenously given



(a)



(b)

Figure 6: (a) Relation between the mean of the differences of CEs between the three comparable pairs of skill and luck gambles and the bias in stated beliefs (b_i^{stated}). (b) Relation between the mean of the differences of CEs between the three comparable pairs of skill and luck gambles and the bias in revealed beliefs ($b_i^{revealed}$). Both panels show the scatterplot and the regression line. Number of observations = 319.

probabilities.

This finding has important implications for career choices, market entry decisions, and financial choices. Biased perceptions of relative skill can lead to misallocations of talent: the wrong people might enter professions where performance depends on relative skill and the right people might be crowded out. More precisely, overconfident low skill individuals might crowd out unbiased intermediate skill individuals and underconfident high skill individuals.

The structural model approach allows us to obtain estimates for revealed beliefs and for risk preference parameters–utility curvature, likelihood sensitivity, and optimism or pessimism. Revealed, like stated beliefs, show that low (high) skill subjects overestimate (underestimate) their relative performance. However, we find a moderate correlation between revealed and stated beliefs. Hence, relying only on stated beliefs can be misleading.

There is scope for further research on this topic. For example, it would be valuable to pursue the extent and implications of the finding that revealed beliefs might be a better predictor of behavior than stated beliefs. This could be done by comparing the out-of-sample predictive power of revealed and stated beliefs.

5 Acknowledgments

We are grateful for helpful comments by Mohammed Abdellaoui, Thomas Astebro, José Mata, Christian Thöni, Marie Claire Villeval, Peter Wakker, and Christian Zehnder. Furthermore, we'd like to acknowledge helpful comments at seminars and conferences at the University of Lausanne, University of Lyon, the ExperiMetrix Workshop at the University of Alicante, and the Second Research in Behavioral Finance Conference in Amsterdam. The funding for this experiment was from the Swiss National Science Foundation grant 135602.

References

Andersen, Steffen, Glenn W Harrison, Morten Igel Lau, and E Elisabet Rutström (2006) 'Elicitation using multiple price list formats.' *Experimental Economics* 9(4), 383–405

Azrieli, Yaron, Christopher P Chambers, and Paul J Healy (2012) 'Incentives in experiments: A theoretical analysis.' In 'Working Paper'

Bolger, Fergus, Briony D Pulford, and Andrew M Colman (2008) 'Market entry decisions: Effects of absolute and relative confidence.' *Experimental Psychology* 55(2), 113–120

Bruhin, Adrian, Helga Fehr-Duda, and Thomas Epper (2010) 'Risk and rationality: Uncovering heterogeneity in probability distortion.' *Econometrica* pp. 1375–1412

Burks, Stephen V, Jeffrey P Carpenter, Lorenz Goette, and Aldo Rustichini (2013) 'Overconfidence and social signalling.' *The Review of Economic Studies* 80(3), 949–983

Cain, Daylian M, Don A Moore, and Uriel Haran (2015) 'Making sense of overconfidence in market entry.' *Strategic Management Journal* 36(1), 1–18

Camerer, Colin, and Dan Lovallo (1999) 'Overconfidence and excess entry: An experimental approach.' *American economic review* pp. 306–318

Charness, Gary, Aldo Rustichini, and Jeroen Van de Ven (2013) 'Self-confidence and strategic behavior'

Charness, Gary, Uri Gneezy, and Michael A Kuhn (2012) 'Experimental methods: Between-subject and within-subject design.' *Journal of Economic Behavior and Organization* 81(1), 1–8

Clark, Jeremy, and Lana Friesen (2009) 'Overconfidence in forecasts of own performance: An experimental study.' *The Economic Journal* 119(534), 229–251

DeBondt, WFM, and R Thaler (1995) 'Financial decision making in markets and firm finance.' *Series of Handbooks in Operational Research and Management Science*

Dunne, Timothy, Mark J Roberts, and Larry Samuelson (1988) 'Patterns of firm entry and exit in us manufacturing industries.' *The RAND journal of Economics* pp. 495–515

Eil, David, and Justin M Rao (2011) 'The good news-bad news effect: asymmetric processing of objective information about yourself.' *American Economic Journal: Microeconomics* 3(2), 114–138

Fischbacher, Urs (2007) 'z-tree: Zurich toolbox for ready-made economic experiments.' *Experimental economics* 10(2), 171–178

Glaser, Markus, and Martin Weber (2007) 'Overconfidence and trading volume.' *The Geneva Risk and Insurance Review* 32(1), 1–36

Goldstein, William M, and Hillel J Einhorn (1987) 'Expression theory and the preference reversal phenomena.' *Psychological review* 94(2), 236

Greiner, Ben (2004) 'The online recruitment system orsee 2.0-a guide for the organization of experiments in economics.' *University of Cologne, Working paper series in economics* 10(23), 63–104

Hamilton, Barton H (2000) 'Does entrepreneurship pay? an empirical analysis of the returns to self-employment.' *Journal of Political economy* 108(3), 604–631

Hey, John D, and Chris Orme (1994) 'Investigating generalizations of expected utility theory using experimental data.' *Econometrica: Journal of the Econometric Society* pp. 1291–1326

Hoelzl, Erik, and Aldo Rustichini (2005) 'Overconfident: Do you put your money on it?' *The Economic Journal* 115(503), 305–318

Johnson, Cathleen A, Aurélien Baillon, Han Bleichrodt, Zhihua Li, Dennie Van Dolder, and Peter P Wakker (2014) 'Prince: An improved method for measuring incentivized preferences.' *Available at SSRN 2504745*

Jovanovic, Boyan (2004) 'The pre-producers.' Technical Report, National Bureau of Economic Research

Kadane, Joseph B, and Robert L Winkler (1988) 'Separating probability elicitation from utilities.' *Journal of the American Statistical Association* 83(402), 357–363

Kahneman, Daniel, and Amos Tversky (1979) 'Prospect theory: An analysis of decision under risk.' *Econometrica: Journal of the Econometric Society* pp. 263–291

Kenney, John Francis, and Ernest Sydney Keeping (1954) 'Mathematics of statistics-part one'

Kruger, Justin (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

Lichtenstein, Sarah, Baruch Fischhoff, and Lawrence D Phillips (1977) *Calibration of probabilities: The state of the art* (Springer)

Lucas Jr, Robert E (1978) 'On the size distribution of business firms.' *The Bell Journal of Economics* pp. 508–523

Mata, Jose, and Pedro Portugal (1994) 'Life duration of new firms.' *The Journal of Industrial Economics* pp. 227–245

Mobius, Markus M, Muriel Niederle, Paul Niehaus, and Tanya S Rosenblat (2011) 'Managing self-confidence: Theory and experimental evidence.' Technical Report, National Bureau of Economic Research

Moore, Don A, and Daylian M Cain (2007) 'Overconfidence and underconfidence: When and why people underestimate (and overestimate) the competition.' *Organizational Behavior and Human Decision Processes* 103(2), 197–213

Moore, Don A, and Paul J Healy (2008) 'The trouble with overconfidence.' *Psychological review* 115(2), 502

Moore, Don A, and Tai Gyu Kim (2003) 'Myopic social prediction and the solo comparison effect.' *Journal of personality and social psychology* 85(6), 1121

Moskowitz, Tobias J, and Annette Vissing-Jorgensen (2002) 'The returns to entrepreneurial investment: A private equity premium puzzle?' Technical Report, National Bureau of Economic Research

Odean, Terrance (1998) 'Do investors trade too much?' *Available at SSRN 94143*

Park, Young Joon, and Luís Santos-Pinto (2010) 'Overconfidence in tournaments: evidence from the field.' *Theory and Decision* 69(1), 143–166

Prelec, Drazen (1998) 'The probability weighting function.' *Econometrica* pp. 497–527

Santos-Pinto, Luis, and Joel Sobel (2005) 'A model of positive self-image in subjective assessments.' *American Economic Review* pp. 1386–1402

Starmer, Chris (2000) 'Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk.' *Journal of economic literature* pp. 332–382

Svenson, Ola (1981) 'Are we all less risky and more skillful than our fellow drivers?' *Acta Psychologica* 47(2), 143–148

Tanaka, Tomomi, Colin F Camerer, and Quang Nguyen (2010) 'Risk and time preferences: linking experimental and household survey data from vietnam.' *The American Economic Review* 100(1), 557–571

Tversky, Amos, and Daniel Kahneman (1992) 'Advances in prospect theory: Cumulative representation of uncertainty.' *Journal of Risk and uncertainty* 5(4), 297–323

Van den Steen, Eric (2004) 'Rational overoptimism (and other biases).' *American Economic Review* pp. 1141–1151

Wakker, Peter P (2010) *Prospect theory: For risk and ambiguity* (Cambridge University Press)

Appendices

A Elicitation of Beliefs

In this Appendix we provide further details on the elicitation of stated beliefs. In addition, we report the beliefs elicited by asking each subject for a point estimate of her relative skill in the cognitive ability test.

For the analysis of the data we prefer the mean of the probability distribution, μ_i, to the point estimates mainly because it doesn't oblige subjects to pick one single skill level. It therefore contains more information on beliefs about relative skill. For instance, if a subject is not sure whether she is

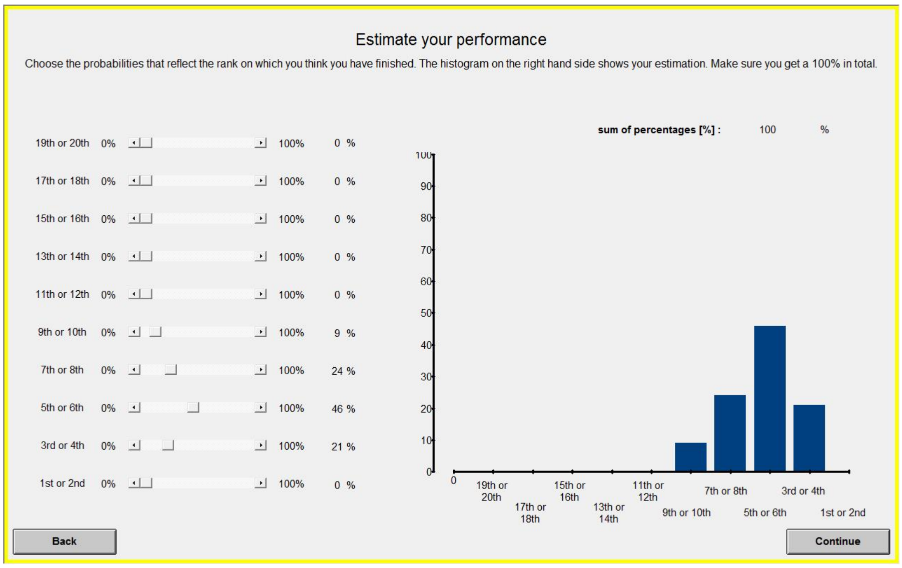


Figure 7: Interface of the elicitation of stated beliefs (full probability distribution)

in skill level θ_8 or θ_9 she can report this by putting some probability weight on either skill level. We apply the following QSR to incentivize truthful disclosure of beliefs. Subjects report their estimated probabilities of belonging to each skill level $q = \{q_1, q_2, ..., q_{10}\}$. If a subject belongs to skill level m, the QSR offers her a payment equal to:

$$QSR_m(q) = 5 + 10q_m - 5\sum_{v=1}^{10} q_v^2$$

The maximum payoff is CHF 10 and the minimum payoff is CHF 0. The maximum payment applies only if the subject puts the entire probability weight on the true skill level. In contrast, CHF 0 is paid out if the entire probability weight is on some other skill level. In any other case, the payment is somewhere between CHF 0 and CHF 10. Figure 7 exhibits the interface that was presented to subjects.

The stated beliefs about relative skill by point estimates allow us to distinguish rational from irrational beliefs. The incentive scheme implied that subjects reveal the mode of their skill beliefs when they are asked for a point estimate on their skill level.[15] Table 4 displays the empirical allocation function for the cognitive ability test. The items on the main diagonal denote the number of individuals who hold correct point estimate beliefs about their relative skill. If all individuals had correct beliefs the matrix would have 64 in all of the main diagonal cells and zeros in all other cells [16]

		Stated beliefs (point estimate)					
		$\{0.05, 0.15\}$	$\{0.25, 0.35\}$	$\{0.45, 0.55\}$	$\{0.65, 0.75\}$	$\{0.85, 0.95\}$	sum
True skill level	$\{0.85, 0.95\}$	0	7	8	38	**11**	64
	$\{0.65, 0.75\}$	0	14	16	**31**	3	64
	$\{0.45, 0.55\}$	0	15	**27**	15	7	64
	$\{0.25, 0.35\}$	0	**17**	23	22	1	63
	$\{0.05, 0.15\}$	**5**	20	25	13	1	64
	sum	5	73	99	119	23	319

Table 4: Absolute frequencies of beliefs about relative skill (points estimate)

Burks et al. (2013) show that the Bayesian model imposes testable implications on how the distribution of relative skill judgements should be related to true skill. Because individuals pick the most likely skill level, the largest (modal) group of individuals thinking they are in a given skill level must belong to that skill level. This is called the diagonal condition (see Burks et al. (2013) for a more detailed description of the properties of Bayesian updating). As Table 4 shows, the diagonal condition is violated for skill levels $\{0.25, 0.35\}$ and $\{0.65, 0.75\}$ and it is not violated for skill levels $\{0.05, 0.15\}$,

[15]According to the data 74% of the subjects give a point estimate that is consistent with the mode of the probability distribution.

[16]Except for the diagonal element in the forth line which would be 63 due to the subject that didn't complete the experiment.

$\{0.45, 0.55\}$, and $\{0.85, 0.95\}$.[17] Hence, the distribution of stated beliefs fails the test of rationality in Burks et al. (2013).

B Figures

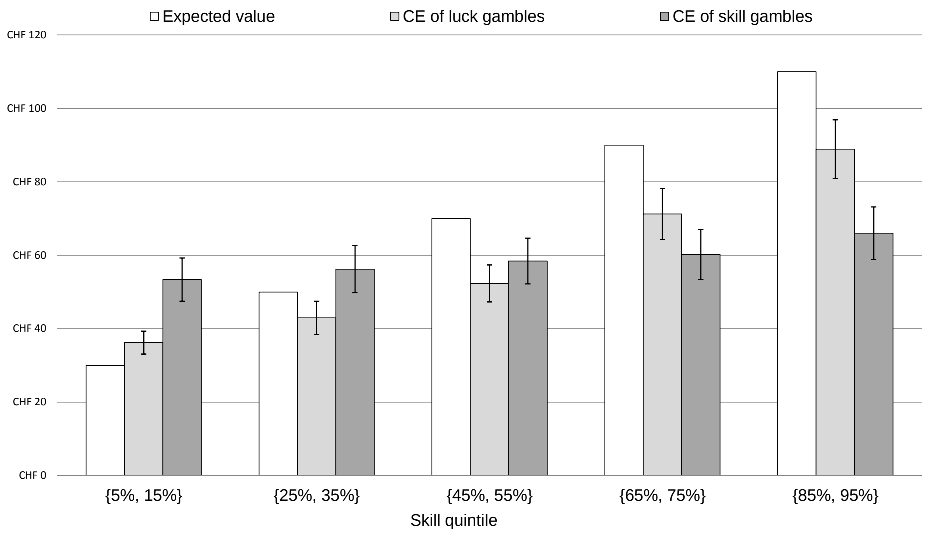


Figure 8: Average expected values of the luck gambles (white bars), average CEs of luck gambles (light grey bars), and average CEs of skill gambles (dark grey bars). Depicted is the gamble with winining prize $x_h = 120$ and losing prize $x_l = 20$. The 95% -confidence intervals are based on robust standard errors. Number of observations = 319.

[17]Of the individuals who put themselves in the skill level $\{0.25, 0.35\}$ or $\{0.65, 0.75\}$, the largest (modal) group of individuals belongs to some other skill level. To illustrate, of the 73 individuals who believe to be in skill level $\{0.25, 0.35\}$ only 17 individuals truly belong to that skill level. This is less than 20 individuals who belong to skill level $\{0.05, 0.15\}$.

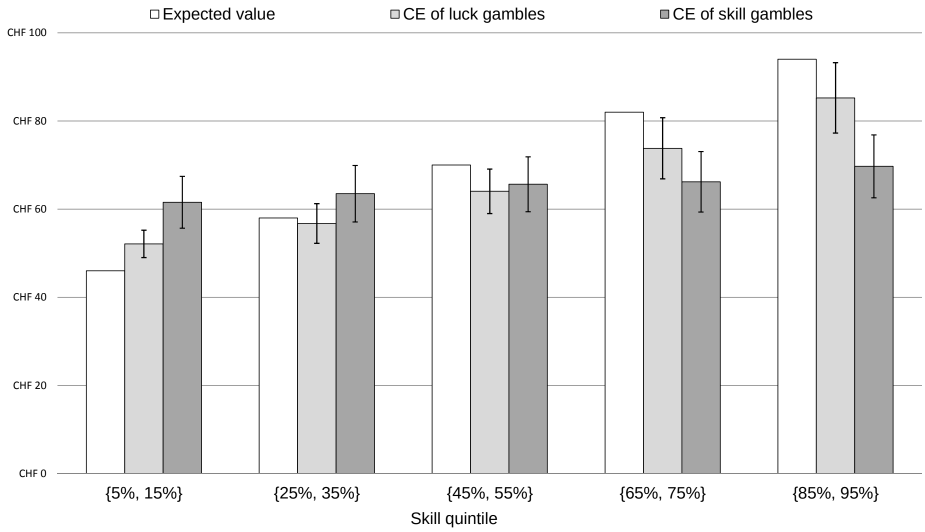


Figure 9: Average expected values of the luck gambles (white bars), average CEs of luck gambles (light grey bars), and average CEs of skill gambles (dark grey bars). Depicted is the gamble with winning prize $x_h = 100$ and losing prize $x_l = 40$. The 95% -confidence intervals are based on robust standard errors. Number of observations = 319.

C Robustness Checks

C.A Specifications of Utility and Probability Weighting Functions

We tested the robustness of the results using alternative functional forms for the utility and the probability weighting function. We estimated alternatively a utility function with constant absolute risk aversion (CARA):

$$u(x; r) = -e^{-xr}$$

where x is the monetary outcome of a gamble and the parameter r measures the coefficient of CARA. $r > 0$ corresponds to risk aversion, $r = 0$ to risk neutrality, and $r < 0$ to love for risk.

In terms of the probability weighting function we estimated alternatively the the two-parameter function proposed by Prelec (1998) (PLC):

$$w(p; \alpha, \beta) = e^{-\beta(-lnp)^{\alpha}}$$

where α and β reflect the degree of probability weighting. The parameter α governs likelihood sensitivity. If $\alpha \in (0, 1)$, the function captures the inverse s-shape pattern where low probabilities are upweighted and high probabilities are downweighted. If $\alpha > 1$ an s-shape pattern where low probabilities are downweighted and high probabilities are upweighted. The parameter β reflects the degree of optimism or pessimism. When $\beta > 1$ the subject is a pessimist, since she downweights the probability of the high prize x_h and upweights the probability of the low prize x_l. When $\beta \in (0, 1)$ the subject is an optimist, since she upweights the probability of the high prize x_h and downweights the probability of the low prize x_l. When $\alpha = \beta = 1$ there is no probability weighting and $w(p; 1, 1) = p$

The estimation procedure remains the same as described in section 3.3.2, including the parameter restrictions. The alternative functional forms give rise to three additional specifications of the structural model: CRRA & PLC, CARA & Goldstein and Einhorn (1987) (GE), and CARA & PLC.

C.A.1 Summary Statistics on Structural Estimations

Table 5 reports the summary statistics of the individual parameter estimates of the alternative specifications of the structural model.

C.A.2 Correlations between Estimated Revealed Beliefs

Table 6 reports the correlation coefficients between the revealed beliefs obtained under the four combinations of utility function and probability weighting function specifications. The high correlations provide evidence for robustness of our baseline specification.

	Mean	Median	Stand. Dev.	Min	Max
CRRA & PLC					
Risk aversion (r_i)	0.27	0.25	1.00	-2.00	11.00
Likelihood sensitivity (α_i)	0.68	0.62	0.47	0.00	4.69
Optimism-pessimism (β_i)	1.34	1.10	1.04	0.00	10.00
Revealed beliefs (ξ_i)	0.54	0.56	0.21	0.05	0.95
Scaling of error term variance (luck) (κ_i^L)	0.08	0.07	0.04	0.00	0.32
Scaling of error term variance (skill) (κ_i^S)	0.11	0.10	0.06	0.00	0.46
CARA & GE					
Risk aversion (r_i)	0.02	0.01	0.04	-0.04	0.51
Likelihood sensitivity (γ_i)	0.86	0.69	0.95	0.00	10.00
Optimism-pessimism (δ_i)	1.59	0.98	2.1	0.00	10.00
Revealed beliefs (ξ_i)	0.52	0.54	0.21	0.05	0.95
Scaling of error term variance (luck) (κ_i^L)	0.08	0.07	0.04	0.00	0.32
Scaling of error term variance (skill) (κ_i^S)	0.11	0.09	0.06	0.00	0.50
CARA & PLC					
Risk aversion (r_i)	0.04	0.01	0.18	-0.04	2.11
Likelihood sensitivity (α_i)	1.00	0.70	1.40	0.00	10.00
Optimism-pessimism (β_i)	1.18	0.93	1.20	0.00	10.00
Revealed beliefs (ξ_i)	0.52	0.54	0.21	0.05	0.95
Scaling of error term variance (luck) (κ_i^L)	0.08	0.06	0.04	0.00	0.32
Scaling of error term variance (skill) (κ_i^S)	0.11	0.09	0.07	0.00	0.50

Table 5: Summary statistics of the individual parameter estimates for the alternative specifications of the structural model. Number of observations = 319, number of females = 114.

	CRRA & GE	CRRA & PLC	CARA & GE	CARA & PLC
CRRA & GE	1			
CRRA & PLC	0.976	1		
CARA & GE	0.942	0.938	1	
CARA & PLC	0.935	0.963	0.957	1

Table 6: Correlation coefficients between the revealed beliefs obtained under the four different combinations of utility and probability weighting function specifications. Number of observations = 319.

C.A.3 Revealed Beliefs per Skill Quintile

Table 7 reports the bias in revealed beliefs per quintile level. Depicted are the three additional specifications of the utility and the probability weighting function. The estimates of under- and overestimation of relative skill are robust to alternative specifications of the utility and the probability weighting function. The coefficients change only little compared to the baseline estimates (exhibited in Table 3 and described in Result 5).

	CRRA & PLC	CARA & GE	CARA & PLC
Skill {0.85, 0.95}	-0.268***	-0.296***	-0.290***
	(0.024)	(0.026)	(0.025)
Skill {0.65, 0.75}	-0.171***	-0.193***	-0.197***
	(0.029)	(0.029)	(0.030)
Skill {0.45, 0.55}	0.065***	0.048*	0.042
	(0.024)	(0.025)	(0.026)
Skill {0.25, 0.35}	0.208***	0.180***	0.194***
	(0.025)	(0.023)	(0.025)
Skill {0.05, 0.15}	0.383***	0.342***	0.359***
	(0.026)	(0.025)	(0.026)
Observations	319	319	319
R-squared	0.585	0.567	0.574

Note: Robust standard errors in parentheses.

*** p<0.01, ** p<0.05, * p<0.1

Table 7: Regression where the dependent variable is the bias in revealed beliefs, i.e., $b_i^{revealed} \equiv \xi_i - \theta_i$, and the independent variables are skill quintile dummies. Depicted are the alternative specifications of the structural model.

C.B Ranking within the Full Sample

In this section we report the results obtained by determining true skill levels based on ranking within the full sample of 320 subject rather than within each session.[18] We denote this measure of the true skill level of subject i by $\tilde{\theta}_i$. The corresponding bias in stated and revealed beliefs are denoted by $\tilde{b}_i^{stated} \equiv \mu_i - \tilde{\theta}_i$ and by $\tilde{b}_i^{revealed} \equiv \xi_i - \tilde{\theta}_i$, respectively. The correlation between $\tilde{\theta}_i$ and the test score (= 0.963) is significantly higher (p-value $<$ 0.001) than the correlation between θ_i and the test score (= 0.903) because $\tilde{\theta}_i$ does not depend on the random sampling of subjects in each session. However, $\tilde{\theta}_i$ is still not a strictly monotonic transformation of the test score

[18]Note that the subject that left without finishing the experiment is ranked as well.

	CRRA & GE	CRRA & PLC	CARA & GE	CARA & PLC
Skill {0.85, 0.95}	-0.242***	-0.240***	-0.267***	-0.262***
	(0.024)	(0.024)	(0.026)	(0.025)
Skill {0.65, 0.75}	-0.164***	-0.160***	-0.182***	-0.186***
	(0.031)	(0.031)	(0.031)	(0.032)
Skill {0.45, 0.55}	0.080***	0.082***	0.065**	0.059*
	(0.030)	(0.030)	(0.031)	(0.032)
Skill {0.25, 0.35}	0.177***	0.195***	0.168***	0.182***
	(0.026)	(0.027)	(0.025)	(0.028)
Skill {0.05, 0.15}	0.335***	0.344***	0.303***	0.320***
	(0.029)	(0.029)	(0.028)	(0.029)
Observations	319	319	319	319
R-squared	0.487	0.498	0.475	0.478

Note: Robust standard errors in parentheses.

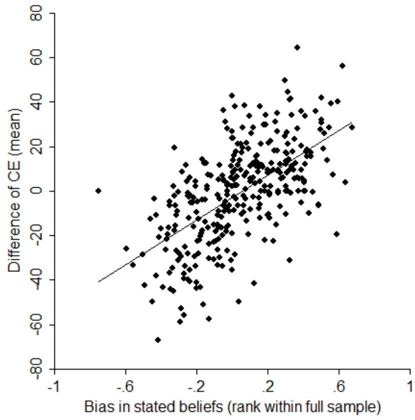
*** p<0.01, ** p<0.05, * p<0.1

Table 8: Regression where the dependent variable is the bias in revealed beliefs, i.e., $\tilde{b}_i^{revealed} \equiv \xi_i - \tilde{\theta}_i$, and the independent variables are skill quintile dummies. Depicted are the four specifications of the structural model. True skill levels determined by ranking within the full sample.

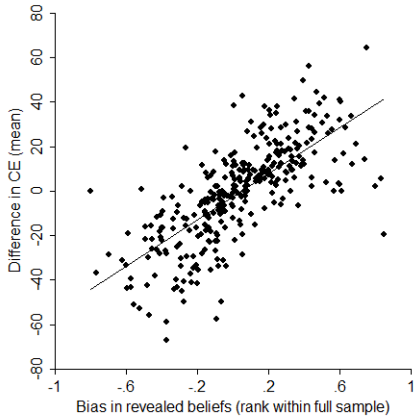
because of the tiebreaker and the random number generator that apply to break ties. It is thus possible that two subjects have the same test score but not the same $\tilde{\theta}$.

Table 8 reports the bias in revealed beliefs per quintile level. Depicted are the four combinations of utility and the probability weighting function specifications. The coefficients change only little compared to the baseline estimates (see Tables 3 and 7). In contrast to the baseline estimates, however, the weak overestimation of relative skill in the intermediate quintile is statistically significant in all four specifications.

Finally, we exhibit in Figure 10 the scatterplots and regression lines illustrating the correlation of the difference in CEs with both the bias in stated beliefs ($\tilde{b}_i^{stated}$) and with the bias in revealed beliefs ($\tilde{b}_i^{revealed}$). Depicted are the estimates obtained by our baseline specification with CRRA utility functions and GE probability weighting functions.



(a)



(b)

Figure 10: (a) Relation between the mean of the differences of CEs between the three comparable pairs of skill and luck gambles and the bias in stated beliefs ($\tilde{b}_i^{stated}$). (b) Relation between the mean of the differences of CEs between the three comparable pairs of skill and luck gambles and the bias in revealed beliefs ($\tilde{b}_i^{revealed}$). True skill levels determined by ranking within the full sample. Both panels show the scatterplot and the regression line. Number of observations = 319.

As in the baseline estimate where subjects are ranked within each session the correlation of the differences of the CEs with the bias in revealed beliefs (= 0.722) is significantly higher (p-value = 0.006) than the correlation of the differences in CEs with the bias in stated beliefs (= 0.601).