**Replication: Belief Elicitation with**

**Quadratic and Binarized Scoring Rules**

**Nisvan Erkal,[[1]](#footnote-2)† Lata Gangadharan,[[2]](#footnote-3)‡ and Boon Han Koh[[3]](#footnote-4)§**

# ONLINE APPENDIX

# Appendix A

**Additional Analyses**

|  |  |
| --- | --- |
|  |  |
| (a) Number of risky decisions | (b) Subjects’ risk preferences |
| Figure A1: Decisions in the risk task | |

|  |  |
| --- | --- |
|  | |
|  | |
| (a) Risk-neutral subjects | |
|  |  |
| (b) Risk-averse subjects | (c) Risk-loving subjects |
| Notes:   1. The distributions of reported beliefs are flipped across the horizontal axis for rounds with > 0.5. The solid line in each panel represents a reported belief of 0.5. 2. The distributions of reported beliefs are not statistically significantly different between the two mechanisms irrespective of the subjects’ risk preferences (Kolmogorov-Smirnov tests for: (i) risk-neutral subjects: p-value = 0.539; (ii) risk-averse subjects: p-value = 0.758; and (iii) risk-loving subjects: p-value = 0.978). | |
| Figure A2: Reported beliefs by risk preferences (Part 1, = 0.5) | |

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
| **Variable** | **Treatment Q** | **Treatment B** | **Difference** |
| Age (years) | 21.82 | 22.38 | p = 0.554 |
|  | (0.66) | (0.69) |  |
| Female | 0.60 | 0.53 | p = 0.465 |
|  | (0.06) | (0.06) |  |
| Economics major | 0.07 | 0.08 | p = 0.732 |
|  | (0.03) | (0.04) |  |
| Postgraduate student | 0.20 | 0.37 | p = 0.043 |
|  | (0.05) | (0.06) |  |
| Born in Australia | 0.62 | 0.70 | p = 0.340 |
|  | (0.06) | (0.06) |  |
| # past experiments participated | 0.45 | 0.37 | p = 0.591 |
|  | (0.10) | (0.12) |  |
| Studied probability/statistics in high school and/or university | 0.87 | 0.85 | p = 0.796 |
| (0.04) | (0.05) |  |
| # risky choices in risk task | 3.25 | 3.02 | p = 0.380 |
|  | (0.17) | (0.21) |  |
| Observations | 60 | 60 |  |
| Note: Sample means given. Standard errors of means given in parentheses. | | | |
|  | | | |
| Table A1: Summary statistics of subjects’ characteristics | | | |

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  | ≠ 0.5 | | = 0.5 |
|  | DIST | NAD | DIST |
| Variables | (1) | (2) | (3) |
| Risk-averse () | 0.011 | 0.020 | -0.015 |
|  | (0.024) | (0.023) | (0.023) |
| Risk-loving () | -0.004 | 0.005 | -0.035 |
|  | (0.038) | (0.037) | (0.027) |
| QSR () | 0.012 | 0.010 | -0.034 |
|  | (0.032) | (0.026) | (0.021) |
| QSR × Risk-averse () | -0.069 | -0.070\* | 0.022 |
|  | (0.037) | (0.031) | (0.027) |
| QSR × Risk-loving () | 0.033 | -0.015 | 0.010 |
|  | (0.049) | (0.045) | (0.033) |
| Constant () | 0.185\*\*\* | -0.110\* | 0.096 |
|  | (0.050) | (0.047) | (0.079) |
| Observations | 480 | 480 | 120 |
| R-squared | 0.313 | 0.135 | 0.137 |
| Control for distance of objective probability | Y | Y | - |
| Control for round | Y | Y | Y |
| Individual controls | Y | Y | Y |
| Notes:   1. DIST: absolute difference between reported belief and 0.5; NAD: negative absolute difference between reported belief and induced objective probability. | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | |
| 1. Robust standard errors in parentheses. Standard errors are clustered at the subject level. | | | |
| 1. \* , \*\* , \*\*\* . | | | |
| 1. The regression model is   where is either DIST or NAD. is the vector of controls comprising the distance of the induced objective probabilities from 0.5, the round number within Part 1 of the experiment, as well as subjects’ characteristics, which include age, gender, whether the subject is pursuing a major in economics, whether the subject is a postgraduate student, whether the subject is born in Australia, whether the subject has studied probability and statistics, and previous experience with economic experiments. | | | |
|  | | | |
| Table A2: OLS estimates controlling for risk preferences (Part 1) | | | |

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  | ≠ 0.5 | | = 0.5 |
|  | DIST | NAD | DIST |
| BSR vs. QSR | (1) | (2) | (3) |
|  |  |  |  |
| (i) Risk-averse subjects | -0.057\*\* | -0.060\*\* | -0.011 |
| [] | (0.020) | (0.019) | (0.018) |
|  |  |  |  |
| (ii) Risk-neutral subjects | 0.012 | 0.010 | -0.034 |
| [] | (0.032) | (0.026) | (0.021) |
|  |  |  |  |
| (iii) Risk-loving subjects | 0.045 | -0.005 | -0.023 |
| [] | (0.038) | (0.037) | (0.024) |
|  |  |  |  |
| Notes:   1. DIST: absolute difference between reported belief and 0.5; NAD: negative absolute difference between reported belief and induced objective probability. | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | |
| 1. Differences between mechanisms for each group of individuals in the table are calculated using the estimated coefficients reported in Table A2. Marginal effects are given with robust standard errors in parentheses. Standard errors are clustered at the subject level. | | | |
| 1. The null hypothesis for each test is that there is no difference in the distance or accuracy of reported beliefs between the BSR and QSR for the corresponding group of individuals. The null hypotheses expressed in terms of the coefficients in the regression models are presented in brackets. In all comparisons, marginal effects reflect the difference in behavior under the QSR relative to the BSR. | | | |
| 1. \* , \*\* , \*\*\* . | | | |
|  | | | |
| Table A3: Comparison of BSR and QSR by risk preferences (Part 1) | | | |

|  |  |  |  |
| --- | --- | --- | --- |
|  | | | |
|  | ≠ 0.5 | | = 0.5 |
|  | DIST | NAD | DIST |
|  | (1) | (2) | (3) |
| **(a) BSR** | | | |
| (i) Risk-averse vs. Risk-neutral | 0.011 | 0.020 | -0.015 |
| [] | (0.024) | (0.023) | (0.023) |
|  |  |  |  |
| (ii) Risk-averse vs. Risk-loving | 0.015 | 0.015 | 0.021 |
| [] | (0.034) | (0.034) | (0.023) |
|  |  |  |  |
| (iii) Risk-loving vs. Risk-neutral | -0.004 | 0.005 | -0.035 |
| [] | (0.038) | (0.037) | (0.027) |
|  |  |  |  |
| **(b) QSR** | | | |
| (i) Risk-averse vs. Risk-neutral | -0.058\* | -0.050\* | 0.008 |
| [] | (0.029) | (0.024) | (0.017) |
|  |  |  |  |
| (ii) Risk-averse vs. Risk-loving | -0.087\*\* | -0.040 | 0.033 |
| [] | (0.028) | (0.023) | (0.019) |
|  |  |  |  |
| (iii) Risk-loving vs. Risk-neutral | 0.029 | -0.010 | -0.025 |
| [] | (0.035) | (0.027) | (0.020) |
|  |  |  |  |
| Notes:   1. DIST: absolute difference between reported belief and 0.5; NAD: negative absolute difference between reported belief and induced objective probability. | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | |
| 1. Differences between the two groups of individuals in the table are calculated using the estimated coefficients reported in Table A2. Marginal effects are given with robust standard errors in parentheses. Standard errors are clustered at the subject level. | | | |
| 1. The null hypothesis for each test is that there is no difference in the distance or accuracy of reported beliefs between the two groups of individuals under the BSR (panel a) or the QSR (panel b). The null hypotheses expressed in terms of the coefficients in the regression models are presented in brackets. In all comparisons, marginal effects reflect the difference in behavior of the second group relative to the first group of subjects. | | | |
| 1. \* , \*\* , \*\*\* . | | | |
|  | | | |
| Table A4: Within-mechanism comparisons by risk preferences (Part 1) | | | |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | | | | | |
|  | ≠ 0.5 | | | | = 0.5 |
|  | DIST | NAD | DIST | NAD | DIST |
| Variables | (1) | (2) | (3) | (4) | (5) |
| QSR () | -0.023 | -0.033\* | -0.014 | -0.028 | -0.019 |
|  | (0.016) | (0.014) | (0.022) | (0.020) | (0.013) |
| = 0.25/0.75 () |  |  | -0.100\*\*\* | 0.050\*\*\* |  |
|  |  |  | (0.009) | (0.012) |  |
| QSR × ( = 0.25/0.75) () |  |  | -0.017 | -0.011 |  |
|  |  |  | (0.020) | (0.020) |  |
| Constant () | 0.255\*\*\* | -0.114\* | 0.300\*\*\* | -0.138\*\* | 0.093 |
|  | (0.044) | (0.045) | (0.044) | (0.046) | (0.072) |
| Observations | 480 | 480 | 480 | 480 | 120 |
| R-squared | 0.099 | 0.070 | 0.278 | 0.112 | 0.118 |
| Control for round | Y | Y | Y | Y | Y |
| Individual controls | Y | Y | Y | Y | Y |
| Notes:   1. DIST: absolute difference between reported belief and 0.5; NAD: negative absolute difference between reported belief and induced objective probability. | | | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | | | |
| 1. Robust standard errors in parentheses. Standard errors are clustered at the subject level. | | | | | |
| 1. \* , \*\* , \*\*\* . | | | | | |
| 1. In columns (3) and (4), the regression model is   where is either DIST or NAD. is the vector of controls comprising the round number within Part 1 of the experiment, as well as subjects’ characteristics, which include age, gender, whether the subject is pursuing a major in economics, whether the subject is a postgraduate student, whether the subject is born in Australia, whether the subject has studied probability and statistics, and previous experience with economic experiments. | | | | | |
|  | | | | | |
| Table A5: OLS estimates for DIST and NAD controlling for objective probabilities  (Part 1) | | | | | |

|  |  |  |
| --- | --- | --- |
|  | | |
|  | DIST | NAD |
| BSR vs. QSR | (1) | (2) |
|  |  |  |
| (i) = 0.1/0.9 | -0.014 | -0.028 |
| [ (columns 3 and 4)] | (0.022) | (0.020) |
|  |  |  |
| (ii) = 0.25/0.75 | -0.031\* | -0.039\*\* |
| [ (columns 3 and 4)] | (0.015) | (0.014) |
|  |  |  |
| Notes:   1. DIST: absolute difference between reported belief and 0.5; NAD: negative absolute difference between reported belief and induced objective probability. | | |
| 1. Differences between mechanisms for each set of objective probabilities in the table are calculated using the estimated coefficients reported in columns (3) and (4) of Table A5. Marginal effects are given with robust standard errors in parentheses. Standard errors are clustered at the subject level. | | |
| 1. The null hypothesis for each test is that there is no difference in the distance or accuracy of reported beliefs between the BSR and QSR for the corresponding set of objective probabilities. The null hypotheses expressed in terms of the coefficients in the regression models are presented in brackets. In all comparisons, marginal effects reflect the difference in behavior under the QSR relative to the BSR. | | |
| 1. \* , \*\* , \*\*\* . | | |
|  | | |
| Table A6: Comparison of BSR and QSR by induced probabilities (Part 1) | | |

# Appendix B

**Experimental Instructions**

This appendix includes the instructions for treatment B. Instructions for treatment Q are identical, except that we switched the order in which Parts 1 and 2 are presented to the subjects.

**Overview of Experiment**

Thank you for agreeing to take part in this study which is jointly funded by the Australian Research Council and the Department of Economics at the University of Melbourne. Please read the following instructions carefully. A clear understanding of the instructions will help you make better decisions and increase your earnings from the experiment.

You will participate in three experiments today: Experiment 1, Experiment 2, and Experiment 3. You will receive detailed instructions for each experiment before you participate in them. Note that the experiments are independent from each other. That is, your decisions in one experiment will not affect your earnings from the other experiments.

You will be paid for the decisions you make in either Experiment 1, Experiment 2, or Experiment 3. At the end of today's session, the computer will randomly determine which experiment to pay you for. This implies that you should carefully consider all of the decisions you make in each experiment as they may determine your earnings.

During the experiments, we will be using Experimental Currency Units (ECU). At the end of the session, we will convert the amount you earn into Australian Dollars (AUD) using the following conversion rate: 10 ECU = 1 AUD. Your final payment today will also include a $10 participation fee.

At the end of Experiment 3, you will be asked to fill out a brief questionnaire asking you some general questions. All of the decisions you make in today's session will remain anonymous.

Please do not talk to one another during the experiment. Please also refrain from using your mobile phones, tablets, and/or laptops. If you have any questions, please raise your hand and we will come over to answer your questions privately.

**Experiment 1**

Experiment 1 consists of five rounds of an individual decision-making task. At the end of the experiment, if you are paid for Experiment 1, then the computer will randomly pick one of the five rounds for payment.

In each round, the computer will present you with an urn containing 100 balls. Some of the balls are **red**, and some of them are **blue**. Note that you will be presented with a new urn at the beginning of each round. Hence, the exact composition of the urn will be different in each round.

At the beginning of each round, the computer will reveal to you the exact composition of the urn. That is, the computer will reveal how many red balls and how many blue balls there are in the urn. The computer will then randomly draw a ball from the urn.

Your task is to predict how likely it is that the ball randomly drawn by the computer is red. To make your guess, you will need to choose a number between 0 and 100. A higher number means that you think the ball is more likely to be red.

Your payment in each round is determined as follows. You will receive either 0 ECU or 200 ECU. Your chance of receiving 200 ECU depends on both your guess and the color of the ball randomly drawn by the computer.

Specifically, your chance of receiving 200 ECU is determined by the following formulas:

Suppose you state a high number as your guess that the ball is red. The formulas above imply that your chance of receiving 200 ECU is high if the computer draws a red ball, and your chance of receiving 200 ECU is low if the computer draws a blue ball. Hence, you should carefully consider how likely it is that the computer will draw a red ball or a blue ball.

To illustrate, suppose your guess that the ball drawn by the computer is red is 100. Then, if the computer draws a red ball, your chance of receiving 200 ECU will be . If the computer draws a blue ball, your chance of receiving 200 ECU will be . Hence, your guess should depend on which of these two outcomes you think is more likely.

Here are two more examples explaining how your chance of receiving 200 ECU will be determined based on your guess and the color of the ball drawn by the computer.

**Example 1.** Suppose you guess 70 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that it has randomly drawn a red ball. Then, your chance of receiving 200 ECU will be .

**Example 2.** In the above example, suppose the ball randomly drawn by the computer is blue. Then, your chance of receiving 200 ECU will be .

To determine whether you receive 200 ECU, the computer will randomly draw a number between 0 and 100. Each number between 0 and 100 is equally likely to be picked. If the number drawn by the computer is less than or equal to your chance of receiving 200 ECU as determined by the formulas above, then you will receive 200 ECU. Otherwise, you will receive 0 ECU. Hence, in Example 1 above, if the number randomly drawn by the computer is less than or equal to 91, then you will receive 200 ECU. Otherwise, you will receive 0 ECU.

|  |
| --- |
|  |
| C:\Users\bhkoh\Dropbox\Uni - Research\Own Research\Beliefs-BSR_QSR\Experiments\Instructions\Figures\BSR-High.png |
| Figure 1: Decision Screen for Experiment 1 |
|  |

Figure 1 shows the decision screen for this experiment. To help you with your decision, you are provided with an on-screen calculator. For each guess that you state, the calculator will report your chance of receiving 200 ECU if the ball randomly drawn by the computer is red or blue. You can use this calculator as many times as you want.

We will provide you with some time to consider your guess carefully. After 60 seconds have lapsed, you will be allowed to state your guess at the bottom of the screen. You may submit your decision at any time after this.

Are there any questions? If not, please proceed to answer the practice questions on the next page. The purpose of these practice questions is to make sure that you understand the experiment. If you have any questions at any time, please raise your hand and an experimenter will come over to answer your questions privately. We will proceed with Experiment 1 once everyone has completed the practice questions.

**Practice Questions (Experiment 1)**

1. I will be paid for the decisions I made in all of the experiments today. True or False?

**Ans:**

1. I will participate in five rounds of a decision task in Experiment 1. If I am paid for Experiment 1, then I will be paid for my decision in one of the five rounds. True or False?

**Ans:**

1. Suppose you guess 40 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that the randomly drawn ball is blue. What is your chance of receiving 200 ECU?

[Note: It is sufficient to write down the formula.]

**Ans:**

1. Suppose you guess 80 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that the randomly drawn ball is red. What is your chance of receiving 200 ECU?

[Note: It is sufficient to write down the formula.]

**Ans:**

**Once you have completed the practice questions, please raise your hand and an experimenter will come over to check your answers.**

Summary

1. You will participate in five rounds of a decision task in Experiment 1. In each round, you will be presented with an urn containing 100 balls. The exact composition of the urn will be different in each round.
2. At the beginning of each round, the computer will reveal to you how many **red** balls and how many **blue** balls there are in the urn. The computer will then randomly draw a ball from the urn.
3. You will be asked to predict how likely it is that the ball randomly drawn by the computer is **red**. You will need to choose a number between 0 and 100. A higher number means that you think the ball is more likely to be red. Your payoff in each round will be determined by both your guess and the color of the ball randomly drawn by the computer.
4. There will be an onscreen calculator that you can use to calculate how your payoff will be affected by both your guess and the color of the ball randomly drawn by the computer. We will provide you with some time to make your decision. Once you have reached a decision, you will need to submit your guess at the bottom of the screen. You may submit your decision at any time after 60 seconds have lapsed.
5. At the end of the experiment, if you are paid for Experiment 1, then the computer will randomly select one of the five rounds for payment.

Are there any questions? If not, we will proceed with three practice rounds for Experiment 1. The purpose of the practice rounds is to allow you to familiarize yourself with the decision screens. Your decisions in the practice rounds will not affect your payments for today's experiment. We will proceed with Experiment 1 once everyone has completed the practice rounds.

**Experiment 2**

As in Experiment 1, Experiment 2 consists of five rounds of an individual decision-making task. At the end of the experiment, if you are paid for Experiment 2, then the computer will randomly pick one of the five rounds for payment.

The procedure for this task is similar to Experiment 1. First, in each round, the computer will present you with a new urn containing 100 balls. Some of the balls are **red**, and some of them are **blue**. The exact composition of the urn will be different in each round.

At the beginning of each round, the computer will reveal how many red balls and how many blue balls there are in the urn. Then, the computer will randomly draw a ball from the urn. Similar to Experiment 1, your task is to predict how likely it is that the ball randomly drawn by the computer is red. However, your payment in Experiment 2 will be determined in the following way.

You will receive an amount between 0 ECU and 200 ECU. The exact amount you will receive depends on both your guess and the color of the ball randomly drawn by the computer.

Specifically, your payoff is determined by the following formulas:

Suppose you state a high number as your guess that the ball is red. The formulas above imply that your payoff is high if the computer draws a red ball, and your payoff is low if the computer draws a blue ball. Hence, you should carefully consider how likely it is that the computer will draw a red ball or a blue ball.

To illustrate, suppose your guess that the ball drawn by the computer is red is 100. Then, if the computer draws a red ball, your payoff will be ECU. If the computer draws a blue ball, your payoff will be ECU. Hence, your guess should depend on which of these two outcomes you think is more likely.

Here are two more examples explaining how your payoff will be determined based on your guess and the color of the ball drawn by the computer.

**Example 1.** Suppose you guess 70 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that it has randomly drawn a red ball. Then, your payoff will be ECU.

**Example 2.** In the above example, suppose the ball randomly drawn by the computer is blue. Then, your payoff will be ECU.

In summary, in Experiment 1, your guess determined your chance of receiving 200 ECU. In Experiment 2, your guess will determine the exact payoff (an amount between 0 and 200 ECU) that you will receive.

The decision screen for this experiment is similar to that in Experiment 1. To help you with your decision, you are provided with an on-screen calculator. For each guess that you state, the calculator will report your payoff if the ball randomly drawn by the computer is red or blue. You can use this calculator as many times as you want.

We will provide you with some time to consider your guess carefully. After 60 seconds have lapsed, you will be allowed to state your guess at the bottom of the screen. You may submit your decision at any time after this.

Are there any questions? If not, please proceed to answer the practice questions on the next page. The purpose of these practice questions is to make sure that you understand the experiment. If you have any questions at any time, please raise your hand and an experimenter will come over to answer your questions privately. We will proceed with Experiment 2 once everyone has completed the practice questions.

**Practice Questions (Experiment 2)**

1. I will participate in five rounds of a decision task in Experiment 2. If I am paid for Experiment 2, then I will be paid for my decision in one of the five rounds. True or False?

**Ans:**

1. Suppose you guess 40 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that the randomly drawn ball is blue. What is your payoff?

[Note: It is sufficient to write down the formula.]

**Ans:**

1. Suppose you guess 80 as the chance that the ball randomly drawn by the computer is red. At the end of the experiment, the computer reveals that the randomly drawn ball is red. What is your payoff?

[Note: It is sufficient to write down the formula.]

**Ans:**

**Once you have completed the practice questions, please raise your hand and an experimenter will come over to check your answers.**

Summary

1. Similar to Experiment 1, you will participate in five rounds of a decision task in Experiment 2. In each round, you will be presented with an urn containing 100 balls. The exact composition of the urn will be different in each round.
2. At the beginning of each round, the computer will reveal to you how many **red** balls and how many **blue** balls there are in the urn. The computer will then randomly draw a ball from the urn.
3. You will be asked to predict how likely it is that the ball randomly drawn by the computer is **red**. You will need to choose a number between 0 and 100. As in Experiment 1, your payoff in each round will be determined by both your guess and the color of the ball randomly drawn by the computer. **However, note that how your payoff will be determined in Experiment 2 is different to that in Experiment 1.**
4. There will be an onscreen calculator that you can use to calculate how your payoff will be affected by both your guess and the color of the ball randomly drawn by the computer. We will provide you with some time to make your decision. Once you have reached a decision, you will need to submit your guess at the bottom of the screen. You may submit your decision at any time after 60 seconds have lapsed.
5. At the end of the experiment, if you are paid for Experiment 2, then the computer will randomly select one of the five rounds for payment.

Are there any questions? If not, we will proceed with three practice rounds for Experiment 2. The purpose of the practice rounds is to allow you to familiarize yourself with the decision screens. Your decisions in the practice rounds will not affect your payments for today's experiment. We will proceed with Experiment 2 once everyone has completed the practice rounds.

**Experiment 3**

In Experiment 3, you will be presented with nine decision tasks. At the end of the experiment, if you are paid for Experiment 3, then the computer will randomly pick one of the nine decision tasks for payment.

In each decision task, you will be presented with two options, labeled Option A and Option B. Option A and Option B differ in how your payment will be determined. Specifically,

* Under Option A, you will receive some fixed amount for sure.
* Under Option B, you will receive some positive amount with some chance, and zero otherwise.

For each decision task, you will need to choose which option you prefer to be paid.

Here are two examples.

**Example 1.** Suppose, in a given decision task, you are presented with the following options.

Option A: You will receive 50 ECU for sure.

Option B: You will receive 100 ECU with a chance of 70%, and 0 ECU otherwise.

Suppose you choose Option A. If this task is chosen for payment, then you will receive 50 ECU from this task.

**Example 2.** In the above example, suppose you choose Option B. To determine your payment, the computer will randomly draw a number between 0 and 100. If the number drawn by the computer is less than or equal to 70, then you will receive 100 ECU. Otherwise, you will receive 0 ECU. For example, suppose the computer draws 57. If this task is chosen for payment, then you will receive 100 ECU for this task.

The nine decision tasks will be presented in a single table. Each row of the table corresponds to an individual decision task. Figure 2 shows an example of the decision screen for this experiment.

|  |
| --- |
|  |
|  |
| Figure 2: Decision Screen for Experiment 3 |
|  |

Are there any questions? If not, we will proceed with Experiment 3.

# Appendix C

**Within-Subject Comparison between BSR and QSR**

In this appendix, we present additional analysis combining data from Part 1 and Part 2 of the experiment (i.e., based on a within-subject comparison between the BSR and QSR). In our analysis, we also test for order effects.

Our main finding is that at the aggregate level, the results are consistent with those reported in Hossain and Okui (2013). However, there is evidence of order effects when we look at individual risk categories.[[4]](#footnote-5) Specifically, for individual risk categories, the results are in line with the theoretical predictions when the subjects are first exposed to the BSR (order-BQ), but not when they are first exposed to the QSR (order-QB).

|  |  |  |  |
| --- | --- | --- | --- |
| **Mechanism** | **Induced Objective Probability** | | |
| **≠ 0.5** | | **= 0.5** |
| **DIST**  **(1)** | **NAD**  **(2)** | **DIST**  **(3)** |
| **BSR** | 0.24 | -0.11 | 0.03 |
|  | (0.01) | (0.00) | (0.01) |
|  | = 480 | | = 120 |
| **QSR** | 0.23 | -0.13 | 0.02 |
|  | (0.01) | (0.01) | (0.01) |
|  | = 480 | | = 120 |
| Notes:   1. DIST = absolute difference between reported belief and 0.5; NAD = negative absolute difference between reported belief and induced objective probability. | | | |
| 1. Sample means given. Standard errors of means given in parentheses. | | | |
| 1. Columns (1) and (2) pool data for rounds with = 0.1, 0.25, 0.75, or 0.9. For = 0.5 (column 3), only DIST is reported since DIST and NAD are equal in absolute terms. | | | |
|  | | | |
| Table C1: Summary statistics of DIST and NAD at aggregate level (Parts 1 & 2) | | | |

Table C1 presents the average DIST and NAD for each elicitation mechanism pooled across all subjects. For rounds with p ≠ 0.5, there are no statistically significant differences in DIST between the BSR and QSR. (randomization test, p-value = 0.679). However, reported beliefs are on average less accurate under the QSR than the BSR (p-value = 0.036). For rounds with p = 0.5, there are no statistically significant difference in DIST between the BSR and QSR (randomization test, p-value = 0.237).

Table C2 presents coefficient estimates of DIST and NAD. Columns (2), (4), and (6) include interaction terms between the order of the mechanisms and the belief elicitation mechanism. The estimates in the table are consistent with the conclusions of the non-parametric tests above, as well as those from HO (column 1 of Table 2 in their paper). However, the estimates in column (4) reveal that the difference in NAD at the aggregate level is not statistically significant when we consider order-BQ and order-QB separately.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  |  | | | |  | |
|  | DIST | | NAD | | DIST | |
| Variables | (1) | (2) | (3) | (4) | (5) | (6) |
| QSR () | -0.005 | -0.006 | -0.024\*\* | -0.030 | -0.011 | -0.020 |
|  | (0.010) | (0.017) | (0.009) | (0.015) | (0.008) | (0.014) |
| Order-QB () | -0.020 | -0.021 | -0.012 | -0.017 | -0.009 | -0.018 |
|  | (0.015) | (0.016) | (0.013) | (0.014) | (0.009) | (0.013) |
| QSR × Order-QB () |  | 0.001 |  | 0.011 |  | 0.017 |
|  |  | (0.020) |  | (0.018) |  | (0.015) |
| Constant () | 0.244\*\*\* | 0.245\*\*\* | -0.074 | -0.071 | -0.015 | -0.012 |
|  | (0.050) | (0.051) | (0.050) | (0.052) | (0.062) | (0.064) |
|  |  |  |  |  |  |  |
| QSR relative to BSR |  |  |  |  |  |  |
| Order-QB: |  | -0.006 |  | -0.030 |  | -0.020 |
| [] |  | (0.017) |  | (0.015) |  | (0.014) |
| Order-BQ: |  | -0.005 |  | -0.019 |  | -0.003 |
| [] |  | (0.011) |  | (0.010) |  | (0.006) |
|  |  |  |  |  |  |  |
| Observations | 960 | 960 | 960 | 960 | 240 | 240 |
| R-squared | 0.228 | 0.228 | 0.070 | 0.071 | 0.107 | 0.110 |
| Control for dist. of obj. prob. | Y | Y | Y | Y | - | - |
| Control for round | Y | Y | Y | Y | Y | Y |
| Individual controls | Y | Y | Y | Y | Y | Y |
| Notes:   1. DIST = absolute difference between reported belief and 0.5; NAD = negative absolute difference between reported belief and induced objective probability. | | | | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | | | | |
| 1. Robust standard errors in parentheses. Standard errors are clustered at the subject level. | | | | | | |
| 1. \* , \*\* , \*\*\* . | | | | | | |
| 1. The regression model is   where is either DIST or NAD. is the vector of controls comprising the distance of the induced objective probabilities from 0.5, the round number within each part of the experiment, as well as subjects’ characteristics, which include age, gender, whether the subject is pursuing a major in economics, whether the subject is a postgraduate student, whether the subject is born in Australia, whether the subject has studied probability and statistics, and previous experience with economic experiments. | | | | | | |
|  | | | | | | |
| Table C2: OLS estimates for DIST and NAD at aggregate level controlling for order (Parts 1 & 2) | | | | | | |

We next examine the behavior of subjects by risk preferences. Panel (a) of Table C3 presents estimates of the differences in behavior between the BSR and QSR for risk-averse, risk-neutral, and risk-loving subjects. Panels (b) and (c) present similar comparisons separately for each treatment (order-BQ or order-QB). These estimates are based on coefficient estimates from OLS regressions of DIST and NAD, which control for subjects’ risk preferences and include interaction terms between the order of the mechanisms and both the belief elicitation mechanisms and subjects’ risk preferences.[[5]](#footnote-6)

We first consider the rounds with ≠ 0.5 (columns 1 and 2 of Table C3). Panel (a) of Table C3 reveals that risk-averse subjects report beliefs that are closer to 0.5 on average under the QSR than the BSR, while risk-loving subjects report beliefs that are further from 0.5 on average (p-values = 0.008 and 0.003, respectively). There is no statistically significant difference in DIST between the mechanisms for risk-neutral subjects (p-value = 0.113). Moreover, reported beliefs of risk-averse subjects are statistically significantly less accurate under the QSR than the BSR (p-value = 0.001), but there are no statistically significant differences in NAD under the two mechanisms for both risk-neutral and risk-loving subjects (p-values = 0.816 and 0.980, respectively). These results are also consistent with those from HO (columns 2-4 of Table 2 in their paper).

Panels (b) and (c) reveal that the differences between the BSR and QSR that we observe under the within-subject design are largely driven by subjects’ behavior when they are first exposed to the BSR (order-BQ). In this treatment, risk-averse subjects report beliefs that are statistically significantly closer to 0.5 and less accurate under the QSR than the BSR (p-values = 0.024 and 0.007, respectively). Moreover, risk-loving subjects report beliefs that are statistically significantly further from 0.5 under the QSR than the BSR, although the difference in NAD is not statistically significant (p-values = 0.001 and 0.856, respectively).

When subjects are first exposed to the QSR (order-QB), the differences in DIST between the two mechanisms are not statistically significant for the risk-averse and the risk-loving subjects (p-values = 0.141 and 0.165, respectively). However, risk-averse subjects continue to report beliefs that are significantly less accurate under the QSR than the BSR (p-value = 0.048). Interestingly, we also observe in this treatment that, contrary to our theoretical predictions, risk-neutral subjects report beliefs that are further from 0.5 under the QSR than the BSR (p-value = 0.045).

|  |  |  |  |
| --- | --- | --- | --- |
|  | ≠ 0.5 | | = 0.5 |
|  | DIST | NAD | DIST |
| BSR vs. QSR | (1) | (2) | (3) |
| (a) Pooled |  |  |  |
| Risk-averse subjects | -0.035\*\* | -0.039\*\* | -0.020\* |
| [] | (0.013) | (0.011) | (0.009) |
| Risk-neutral subjects | 0.030 | -0.003 | -0.008 |
| [] | (0.019) | (0.014) | (0.011) |
| Risk-loving subjects | 0.057\*\* | -0.001 | 0.020 |
| [] | (0.018) | (0.033) | (0.032) |
| (b) Order-BQ |  |  |  |
| Risk-averse subjects | -0.044\* | -0.045\*\* | -0.033\* |
| [] | (0.019) | (0.016) | (0.016) |
| Risk-neutral subjects | 0.039 | 0.000 | -0.028 |
| [] | (0.037) | (0.027) | (0.017) |
| Risk-loving subjects | 0.104\*\* | -0.014 | 0.072 |
| [] | (0.029) | (0.076) | (0.071) |
| (c) Order-QB |  |  |  |
| Risk-averse subjects | -0.025 | -0.032\* | -0.005 |
| [] | (0.017) | (0.016) | (0.009) |
| Risk-neutral subjects | 0.021\* | -0.006 | 0.012 |
| [] | (0.010) | (0.010) | (0.012) |
| Risk-loving subjects | 0.023 | 0.008 | -0.018 |
| [] | (0.016) | (0.017) | (0.013) |
| Notes:   1. DIST = absolute difference between reported belief and 0.5; NAD = negative absolute difference between reported belief and induced objective probability. | | | |
| 1. For = 0.5, only the estimates for DIST are reported since DIST and NAD are equal in absolute terms. | | | |
| 1. Differences between mechanisms for each group of individuals in the table are calculated using the estimated coefficients given in the models below. Marginal effects are given with robust standard errors in parentheses. Standard errors are clustered at the subject level.   Model for panel (a):  Model for panels (b) and (c):  is either DIST or NAD. is the vector of controls comprising the distance of the induced objective probabilities from 0.5, the round number within each part of the experiment, as well as subjects’ characteristics, which include age, gender, whether the subject is pursuing a major in economics, whether the subject is a postgraduate student, whether the subject is born in Australia, whether the subject has studied probability and statistics, and previous experience with economic experiments. | | | |
| 1. The null hypothesis for each test is that there is no difference in the distance or accuracy of reported beliefs between the BSR and QSR for the corresponding group of individuals. The null hypotheses expressed in terms of the coefficients in the regression models are presented in brackets. In all comparisons, marginal effects reflect the difference in behavior under the QSR relative to the BSR. | | | |
| 1. \* , \*\* , \*\*\* . | | | |
|  | | | |
| Table C3: Comparisons of DIST and NAD between mechanisms (Parts 1 & 2) | | | |

# Appendix D

**Theoretical Framework**

In this section, we provide a simple theoretical framework that allows us to form predictions for the subjects’ behavior under each of the two scoring rules.

Assume that individuals are subjective expected utility maximizers with utility given by , where . denotes the subject’s *true* belief that an event occurs. In our experiment, this is the probability that the ball randomly drawn by the computer is red, and this is objectively known to the subjects. The subject states a report given his belief and the scoring rule. If a scoring rule incentivizes truth-telling, then this implies that it is an optimal strategy for the subject to report under the scoring rule.

## Reported beliefs under the QSR

Under the QSR, the subject receives given his reported belief and the color of the ball drawn by the computer, with . Normalizing , the subject solves

which gives the following first-order condition (FOC)

For a risk-neutral subject, , implying that is the same on both sides of . Hence, the FOC reduces to , giving . A risk-neutral subject truthfully reports his belief under the QSR.

For a risk-averse subject, . Since is increasing in , on the left-hand side (LHS) of is decreasing in . As is also decreasing in , the entire LHS of is decreasing in . Conversely, since is decreasing in , is increasing in . Consequently, the entire right-hand side (RHS) of is increasing in .

To evaluate the optimal strategy for a risk-averse subject under the QSR, we first assume that and check if the FOC holds. Note that this premise implies , and therefore the equality in depends on the terms in on both sides of the FOC. We consider three cases. First, suppose . Then, and the equality in holds. The subject truthfully reports his belief if he thinks that the computer is equally likely to draw a red ball and a blue ball.

Next, suppose while still maintaining the assumption that . Hence, the subject believes that the computer is more likely to draw a red ball than a blue ball. Note that for . Hence, since . Consequently, the LHS of is less than the RHS when . Since the LHS is decreasing in and the RHS is increasing in , will need to decrease for the FOC to hold. As a result, decreases towards 0.5, and the FOC holds for some .[[6]](#footnote-7)

By a similar argument, now suppose while still maintaining the assumption that . Hence, the subject believes that the computer is more likely to draw a blue ball than a red ball. Note that for . Hence, . Consequently, the LHS of is now greater than the RHS when . This implies that will need to increase for the FOC to hold. As a result, increases towards 0.5, and the FOC holds for some .

Overall, under the QSR, a risk-neutral subject will report . However, a risk-averse subject will only do so if 0.5. A risk-averse subject who holds a belief will have an incentive to distort away from and towards 0.5. Intuitively, this distortion occurs because the QSR penalizes the subject more on the margin for an incorrect prediction of the color of the ball as his reported belief becomes more extreme (i.e., when gets further away from 0.5). Consequently, subjects who are risk averse would prefer to smooth out their payoffs by distorting their reports towards the midpoint.

What about the behavior of risk-loving subjects under the QSR? Hossain and Okui (2013) explain that the direction of behavior under the QSR depends on the specific utility function. Nonetheless, subjects generally tend to report relatively extreme probabilities.

## Reported beliefs under the BSR

Under the BSR, the subject receives with probability given his report and the color of the ball drawn by the computer. The subject solves

giving the following FOC

which implies . Hence, under the BSR, a subjective expected utility maximizer will truthfully report his belief independent of his risk preferences.[[7]](#footnote-8)

1. † Department of Economics, University of Melbourne, VIC 3010, Australia. n.erkal@unimelb.edu.au. [↑](#footnote-ref-2)
2. ‡ Department of Economics, Monash University, VIC 3800, Australia. lata.gangadharan@unimelb.edu.au. [↑](#footnote-ref-3)
3. § School of Economics, University of East Anglia, NR4 7TJ, United Kingdom. b.koh@uea.ac.uk. [↑](#footnote-ref-4)
4. HO do not test for order effects. [↑](#footnote-ref-5)
5. In the regression models, we also control for subjects’ characteristics, which include age, gender, whether the subject is pursuing a major in economics, whether the subject is a postgraduate student, whether the subject is born in Australia, whether the subject has studied probability and statistics, and previous experience with economic experiments. We also include as controls the round number within each part of the experiment and the distance of the induced objective probabilities from 0.5 (for 𝑝 ≠ 0.5). [↑](#footnote-ref-6)
6. Note that will not be distorted to a value below 0.5 because both and when and . This implies that the LHS of is greater than the RHS when , and therefore the inequality will now be in the opposite direction. [↑](#footnote-ref-7)
7. Note that the framework we present here assumes that subjects are subjective expected utility maximizers. Hossain and Okui (2013) show that the BSR is incentive compatible even for non-expected utility maximizers. Harrison et al. (2014) also show this to be true when subjects follow a rank-dependent utility (RDU) model. [↑](#footnote-ref-8)