Internal Uncertainty in Belief Formation and Choice Under Risk

Benjamin Enke and Thomas Graeber
Harvard University
enke@fas.harvard.edu
graeber@fas.harvard.edu
This version: July 26, 2019

Internet experiments on Amazon Mechanical Turk (AMT) to study the relationship between internal uncertainty and belief formation / choice under risk.

This pre-registration consists of three distinct sets of experiments that we describe in turn:

1. Choice under risk
2. Belief updating tasks
3. Survey expectations

In addition, in part 4, we pre-register correlational add-on analyses.

PART 1: CHOICE UNDER RISK

I. Experimental setup and measurement of internal uncertainty

On AMT, subjects complete six standard multiple price list tasks to elicit certainty equivalents for lotteries. Each lottery has two states, one of which has a payout of zero and one of which has a strictly positive or strictly negative payout. The non-zero payout and the payout probability are randomized. Parameters in the randomization set are:

- Payouts: $15, $20, $25, -$15, -$20, -$25
- Payout probabilities: 0%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, 100%

(Whenever the non-zero payout is negative, subjects receive a budget of the same amount so that they cannot lose money.)

Subjects make six choices. One of these choices gets randomly selected for payment if the subject is selected for payment. The probability of being selected for payment is 1/3 for each subject.

After each price list, the subsequent screen elicits a measure of internal uncertainty. Here, we ask subjects to indicate a range of certainty equivalents such that they are completely certain that their true valuation for the lottery is contained in that range.

II. Treatments

1. Treatment *Baseline*

In *Baseline*, subjects complete tasks of the type described above. Out of six tasks, two randomly selected ones will feature a compound lottery (within-subject treatment variation). An example of a compound lottery is:

We randomly draw an integer number between 0 and 20, with equal probability. Call this number x.
With probability x%: receive $20
With probability 100%-x%: receive $0

2. Treatment *High Default*

In *High Default,* we present the same lotteries as in *Baseline,* except that (i) no compound lotteries will be implemented and (ii) the probabilities will be presented in terms of the number of balls (out of 100) that are of a given color. An example is:
10 balls are red. If a red ball gets selected: Receive $20
90 balls are blue. If a blue ball gets selected: Receive $0

3. Treatment *Low Default*

In *Low Default,* we present the same lotteries as in *High Default,* except that the zero-payout state will be split up into nine distinct states. An example is:

10 balls are red. If a red ball gets selected: Receive $20
10 balls are blue. If a blue ball gets selected: Receive $0
10 balls are green. If a green ball gets selected: Receive $0
10 balls are black. If a black ball gets selected: Receive $0
10 balls are white. If a white ball gets selected: Receive $0
Etc.

III. Hypotheses

1. In treatment *Baseline*, subjects with higher internal uncertainty exhibit more pronounced (more compressed) probability weighting functions in both the gain and the loss domain.
2. In treatment *Baseline*, under compound lotteries, subjects exhibit more pronounced (more compressed) probability weighting functions in both the gain and the loss domain.
3. In treatment *Low Default*, the probability weighting function is shifted towards zero relative to treatment *High Default*.

IV. Exclusion Criteria

1. We will implement a set of control questions after subjects have read the experimental instructions. Any subject that gets a control question wrong on the first attempt will be excluded from the study immediately, i.e., will proceed to the final payment screen of the study after the control questions.
2. We also implement an attention check. Any subject that gets the attention check wrong will be excluded from the analysis, and not count towards the number of completes.
3. We implement the following analyses:
    a. Analyses that include all data points (subject to the restrictions above).
    b. Analyses in which we drop all observations that have the following characteristic: Divide the lottery's non-zero payout x by the probability p of receiving that payout. Refer to this number as "probability weight" w. We exclude a data point if it exhibits extreme probability weighting:
        i. $x>0$, $p<25\%$ and $w>95\%$
        ii. $x>0$, $p>75\%$ and $w<5\%$
        iii. $x<0$, $p<25\%$ and $w>-95\%$
        iv. $x<0$, $p>75\%$ and $w<-5\%$
    c. Analyses in which we exclude respondents in the bottom decile of the response time distribution.

V. Randomization and Sample size

Treatments *Low Default* and *High Default* will be randomized within session.

We will recruit the following number of completes:

- Treatment *Baseline*: 700
- Treatments *Low Default* and *High Default*: 150 each

PART 2: BELIEF UPDATING

I. Experimental setup and measurement of internal uncertainty

On AMT, subjects complete six balls-and-urns belief updating tasks. In each task, the computer randomly selects one out of two urns according to a known base rate p. One urn contains q>50 red balls and (100-q) blue balls, while the other urn contains q blue balls and (100-q) red balls. The computer randomly selects one or more of these 100 colored balls from the selected urn (with replacement if more than one ball is drawn). The balls are shown to the subject, who then needs to provide a probability estimate (0-100) that either urn was selected. The experiment randomizes the problem parameters p, q and the number of draws N across trials. The set of problem parameters is given by:

- Base rate p: 10, 30, 50, 70, 90
- Diagnosticity q: 70, 90, 100
- Number of draws N: 1, 3

Subjects' estimates are incentivized using a binarized scoring rule. Subjects make six decisions. One of these decisions gets randomly selected for payment if the subject is selected for payment. The probability of being selected for payment is 1/3.

After each belief statement, the subsequent screen elicits a measure of internal uncertainty. Here, we ask subjects to indicate a range of values such that they are certain that the Bayesian posterior for the belief updating task is contained in that range.

As a second measure of internal uncertainty, we elicit subjects' willingness-to-pay to replace their own estimate with the Bayesian posterior. The willingness-to-pay is elicited before stating a guess. We use the Becker-DeGroot-Marschak methodology to determine whether a guess is replaced. A potential replacement is only implemented with a probability of 10% however, and subjects learn whether replacement takes place before stating their guess.

II. Treatments

1. Treatment *Baseline*

In *Baseline*, subjects complete six tasks of the type described above. In each task, we provide subjects with the prior probability and the signal diagnosticity (the composition of ball colors in the urns). An example is:

Probability that urn A gets selected: 60%
Probability that urn B gets selected: 40%

Number of red balls in urn A: 70 out of 100
Number of blue balls in urn B: 70 out of 100

Out of the six tasks, one has a compound structure. This means that (i) the prior probability is 50:50 and (ii) the urn composition is communicated in a compound form. An example is:

Probability that urn A gets selected: 50%
Probability that urn B gets selected: 50%

We randomly draw a number from 60-80, with equal probability. Call this number x.
Number of red balls in urn A: x out of 100
Number of blue balls in urn B: x out of 100

2. Treatment ***Low Default***

In *Low Default,* we transform the belief updating tasks in *Baseline* by increasing the number of urns, holding the objective information structure with respect to urn A constant. There are now 10 urns overall.

We present the same tasks as in *Baseline,* except that urn B is split into nine separate urns with identical color compositions. An example is:

Probability that urn A gets selected: 50%
Probability that urn B gets selected: 4%
Probability that urn C gets selected: 4%
…
Probability that urn I gets selected: 4%
Probability that urn J gets selected: 8%

Number of red balls in urn A: 70 out of 100
Number of blue balls in urn B: 70 out of 100
Number of blue balls in urn C: 70 out of 100
…
Number of blue balls in urn J: 70 out of 100

III. Hypotheses

We define a "belief weighting function" by plotting stated posteriors against Bayesian posteriors.

1. In Baseline, the belief weighting function is inverse S-shaped: stated posteriors are too high for Bayesian posteriors below 50% and too low for Bayesian posteriors above 50%.
2. In treatment *Baseline*, subjects with higher internal uncertainty exhibit more pronounced (more compressed to 50%) belief weighting functions.

3. In treatment *Compound*, subjects exhibit more pronounced (more compressed to 50%) belief weighting functions than in *Baseline*.
4. In *Low Default*, the belief weighting function is shifted downward (toward 0%) relative to treatment *Baseline*.

IV. Exclusion Criteria

1. We will implement a set of control questions after subjects have read the experimental instructions. Any subject that gets a control question wrong on the first attempt will be excluded from the study immediately, i.e., will proceed to the final payment screen of the study after the control questions.
2. We also implement an attention check. Any subject that gets the attention check wrong will be excluded from the analysis, and not count towards the number of completes.
4. We implement the following analyses:
    a. Analyses that include all data points (subject to the restrictions above).
    b. Analyses in which we drop all observations that have the following characteristic: Let the Bayesian posterior be b and the response r. We exclude a data point if
        i. B<25% and r>75%
        ii. B>75% and r<25%
    c. Analyses in which we exclude respondents in the bottom decile of the response time distribution.

V. Randomization and Sample size

Treatments *Baseline, Compound,* and *Low Default* will be randomized within session.

We will recruit the following number of completes:

- Treatment *Baseline*: 700
- Treatment *Low Default*: 300

PART 3: SURVEY EXPECTATIONS

I. Experimental setup and measurement of internal uncertainty

The survey tasks will be added at the end of each of the treatments discussed above. In each task, the subject has to state a probability estimate for a random event.

Following each task, we elicit a measure of internal uncertainty. Here, we ask subjects to indicate a range of probabilities such that they are completely certain that the objective probability of the event in question happening is contained in that range.

Each subject completes three tasks, explained below. Subjects' estimates are incentivized using a binarized scoring rule. Subjects make three decisions. One of these decisions gets randomly selected for payment if the subject is selected for payment. The probability of being selected for payment is 1/3.

II. Questions

1. Inflation beliefs

Subjects are asked to assess the probability that, in a randomly selected year, the inflation rate was less than x%. Across respondents, x varies.

2. Stock market beliefs

Subjects are asked to assess the probability that, in a randomly selected year, the increase in the S&P 500 was less than x%. Across respondents, x varies.

3. Income distribution beliefs

Subjects are asked to assess the probability that a randomly selected household in the US earns less than $x. Across respondents, x varies.

III. Hypotheses

We define a "survey belief weighting function" by plotting stated probabilities against correct probabilities.

1. All belief weighting function are inverse S-shaped: stated posteriors are too high for low probabilities and too low for high probabilities.

2. Subjects with higher internal uncertainty exhibit more pronounced (more compressed to 50%) belief weighting functions.

IV. Exclusion Criteria

1. The same exclusion criteria regarding control questions and attention checks as above for Parts 1 and 2 apply here since survey expectations are elicited from the same sets of respondents.
2. We implement the following analyses:
   a. Analyses that include all data points (subject to the restrictions above).
   b. Analyses in which we drop all observations that have the following characteristic: Let the true probability be b and the response r. We exclude a data point if
      i. B<25% and r>75%
      ii. B>75% and r<25%
   c. Analyses in which we drop the bottom decile of the response time distribution.

V. Randomization and Sample size

Because these questions are added to the treatments described in Parts 1 and 2, the sample size follows automatically and is given by 2000.

PART 4: CORRELATES OF INTERNAL UNCERTAINTY

Here, we take the internal uncertainty measures from Parts 1-3. We correlate them with the respondent's:

- Gender
- Age
- Score on a Raven matrices test
- Educational attainment