Eliciting Minimum Acceptable Probabilities

Pre-Analysis Plan

Martin Strobel   Maria Polipciuc*

Very preliminary—please do not cite

April 16, 2021

1 Main page description (public)

Interventions

Intervention(s)

We compare minimum acceptable probabilities (MAPs) for which a participant considers a binary lottery equally desirable to a sure payoff across treatments within individual. Individuals have to decide in randomized order in three scenarios (treatments). The treatments vary the distribution of the winning probability of the lottery.

Trial Start Date Will we have a trial, for instance to test understanding of instructions?

Intervention Start Date

TBD

Intervention End Date

TBD

Primary Outcomes

Primary Outcomes (end points)

MAPs

Primary Outcomes (explanation)

The MAPs are elicited as indifference values.

Experimental Design

Experimental Design

We elicit MAPs in an online experiment in a lottery similar to the *Decision Problem* in Bohnet and Zeckhauser (2004). The aim is to test an assumption re-

quired for the MAP elicitation procedure to be incentive-compatible. Specifically, we test whether there is evidence that participants' declared MAPs are independent of the underlying distribution from which the probability of success of the gamble is drawn. We do this by exogenously manipulating this distribution, and comparing the resulting MAPs across treatments.

Experimental Design Details

Not available

Randomization Method

Computer

Randomization Unit

Individual

Was the treatment clustered?

No

Planned Number of Observations

TBD

Was IRB approval obtained (only for "In Development" and "On-going" trials)? If so, also

IRB Name

IRB Approval Date

IRB Approval Number

2 Introduction

The Minimum Acceptable Probability (MAP) is an elicitation procedure for an indifference value (expressed as the probability or as the number of favorable out-

comes) between a sure payoff and trusting someone. In other words, it is the value which makes an individual indifferent between engaging and not engaging in a trusting interaction.

The concept was introduced by Bohnet and Zeckhauser (2004) and it has been used—with slight variations—to elicit a determinant of trust, betrayal aversion, in a difference-in-difference design. This procedure is incentive compatible if "a principal adheres to the Substitution Axiom of von Neumann-Morgenstern utility [...] A MAP is a cutoff value relating to preferences, and the estimated value of p^* [*NR: the winning probability*] should not affect it." (Bohnet et al., 2008, p. 298). While empirical violations of this axiom—which implies expected utility— have been documented (see footnote 5 on p. 275 in Li et al., 2020, for a list of studies finding empirical violations), to our knowledge it is not yet established whether this violation is empirically relevant in the context of the MAP elicitation procedure.[1]

This paper investigates whether the MAP is influenced by a participant's subjective expectation of the chance of a favorable outcome. Specifically, we remove the social and strategic aspects of a trust decision, and study the MAP for accept-

---

[1]Another important point is the following observation made in footnote 4 on p. 815 in Bohnet et al. (2010): "Note that a principal cannot affect the probability he or she receives in the lottery, because it in no way relates to the answer that he or she provides." This is true by design; however, we believe that there is another dimension of risk exposure than the probability received in the lottery that the participant can control through their MAP. The participant can be viewed as facing a compound risk, which consists of (i) the probability received in the lottery and (ii) the probability that they receive a lottery with a certain winning probability or higher at a certain MAP, which is determined by its cumulative density (or mass, if the distribution is discrete) function. From this, the underlying distribution of the winning probability in the lottery may influence the MAP due to its impact on (ii). An implication is that the difference between treatments might be due to different underlying distributions of the winning probability across treatments, not to the different nature of the lotteries (risky/social/strategic). In other words, the difference might be due to risk aversion—this could potentially explain why several studies which keep subjective probabilities constant across treatments and do not find evidence of betrayal aversion (Fetchenhauer and Dunning, 2012; Polipciuc and Strobel, 2020).

ing a risky lottery. We manipulate participants' subjective expectations about the lottery's winning chances by varying the "world" in which they report a MAP. The experiment consists of three "worlds" (treatments), which participants experience sequentially in randomized order. For each decision, the world is described to participants before the MAP elicitation, and it is characterized by a distribution over winning probabilities. A state of the world is a probability drawn from this distribution prior to decision-making, but about which the participant is only informed after having made his/her decision. Should participants change their MAP depending on the world they face, this is evidence that the MAP elicitation procedure (which uses the strategy method) is not incentive compatible unless one accounts for subjective expectations.[2]

Below we present the sample selection procedure, the experimental design, and the empirical strategy.

3   Research Strategy

This project will collect experimental data on an online platform dedicated to academic research (Prolific) in April 2021. Participants will be exposed to three treatments sequentially, in randomized order. In each of the treatments, participants have to state the MAP (minimum acceptable probability) for which they prefer a lottery over a sure payment. This is a version of Bohnet and Zeckhauser's

---

[2]When uncertainty is resolved—even if this is irrelevant for strategic or informational purposes—has also been shown to matter for decisions (see footnote 11 on p. 29 in Johnson et al., 2019). The experiment jointly tests whether different "worlds" matter, in a context where uncertainty is resolved after making the decision (the actual state of the world is revealed to participants after they state their MAP).

In BZ, the state of the world is determined before participants make a decision (but they are not told from which "world" i.e. distribution of states it has been drawn). Similarly to our experiment, they are only informed about the state of the world after having made their decision.

(2004) *Decision Problem* (p. 469). The treatments differ in the underlying distribution from which the lottery is drawn, which makes participants' decision one which requires reducing compound objective lotteries. Given the complexity of the tasks, we will recruit participants with completed higher education (also students?), to increase the chances that task comprehension is not an issue.

The pre-analysis plan will be registered at the AEA RCT registry before the start of the data collection.

3.1 Recruitment

Participants are registered users on the online platform Prolific. This platform is tailored for academic research, and gathers demographics about registered users. We will send an invitation to the experiment only to participants who have completed higher education (maybe also to students?), for the reasons mentioned above.

4 Design

The study consists of three parts. The first part describes the task and ask comprehension questions (unincentivized). Only those passing this check are directed to the main part, which is incentivized. After this, those who complete the main part go through a survey.

As mentioned above, participants in the experiment are asked to state their MAP in a *Decision Problem*: what is their MAP for taking a gamble rather than accepting a sure payoff? The experiment combines a within-subject design (with all participants going through all treatments sequentially) with a between-subject

design.

Below we present the design and instructions in more detail.

4.1   Description of the task

See the document 'Explanation.pdf'.

4.2   Comprehension questions

TBD

4.3   MAP elicitation

Here will follow the exact wording of each "world" description.

WORLD 1

WORLD 2

WORLD 3

For how to phrase the MAP question, my suggestion is to take inspiration from Johnson et al. (2019) (forthcoming in the Journal of Decision and Uncertainty), as per their example on page E.2 in their online appendix here: `https://personal.eur.nl/wakker/oldpsabs.htm#prince21.1` (this is a WTA task, so it has to be adapted). I also post it here:

4.4   Survey

Here will follow the exact wording of the questions.

FIGURE 1′. Instructions for WTA with matching and endowment (Question 1)

In each of the 50 envelopes, one option is to keep your mug, and the other option is to give up your mug for a money amount. The note in each envelope is as follows.

> Option 1: Keep your mug
>
> Option 2: Give up your mug for €x

The money amount x varies between €0 and €10 in different envelopes.*
Please give us instructions which option from your envelope to give to you. Do so by specifying an indifference value (in cents).

My indifference value is € ....,....

If x in my envelope falls below (is worse than) my indifference value, then give me Option 1 (mug).
If x in my envelope equals my indifference value, then I am indifferent.**
If x in my envelope exceeds (is better than) my indifference value, then give me Option 2 (€x).

* Five envelopes contain a randomly generated amount between €0 and €1, five contain a randomly generated amount between €1 and €2, ..., and five contain a randomly generated amount between €9 and €10. Thus the amount in your envelope can be any amount, in cents, between €0 and €10.

** Then a coin toss will decide whether Option 1 or 2 is given.

5 Empirical Strategy

The experiment is meant to study how much of betrayal aversion can be explained by subjects not respecting ROCL (reduction of objective compound lotteries). In each treatment, the subjects are faced with a different distribution of the possible states of the world (the fortune wheels). Since these distributions are given, participants are faced with a two stage lottery, in which are probabilities are objective.

whether the world in which one chooses the indifference value for the MAP influences the MAP. If Bohnet et al. (2008) are right and the estimated value

of the true winning probability does not influence MAP, then we should not see significant differences between treatments.

However, there is evidence that individuals change their views on what is a fair payoff/what they can request as an indifference value depending on the world they are in (Bohnet et al., 2008, 2010).[3]

A world A is friendlier than another world B if the cumulative density (mass) function of p^* in A is below the cumulative density (mass) function of p^* in B for p^* in [0,1], where the range of states for worlds A and B is the same. We expect this threshold to be higher in "friendlier" worlds, and lower in "less friendly" worlds, for instance because of (i) a different perception of what is a fair chance or (ii) anchoring on higher values (for instance, a higher mean of p^*) or (iii) a different thrill from gambling or (iv) different aspiration levels or (v) different reference points—leading to different loss aversion effects or (vi) higher order risk attitudes or think of other reasons.[4]

This leads to the following hypotheses.

5.1 Hypotheses

5.1.1 Main hypotheses

Hypothesis 1 *The MAP in the world with left skew (more mass on high values of p^*) is higher than the MAP in the world with right skew (more mass on low*

---

[3]Bohnet et al. (2010) discuss culturally different reference points for trustworthiness which can influence the MAP. The claim we test is different: that (some of) the observed difference between treatments in their paper and many others is not due to the social or strategic dimensions of a trusting interaction, but to a difference in underlying distributions of the probability of winning—which could influence the MAP through changing reference points for what is a fair deal.

[4]The post-experimental survey could shed some light on whether (some of) the factors matter. We leave it to future research to study the mechanisms at play, should we find that results in line with our expectations.

values of p^*).

Hypothesis 2   *The MAP in the world with left skew (more mass on high values of p^*) is higher than the MAP in the world with a uniform distribution over p^*.*

Hypothesis 3   *The MAP in the world with right skew (more mass on low values of p^*) is lower than the MAP in the world with a uniform distribution over p^*.*

5.1.2 Heterogeneity

Since the MAP is a way to gauge risk aversion, we expect that in the same world, females state higher MAPs than males on average.

Hypothesis 4   *Within each world, females require higher MAPs on average than males.*

Our treatments vary the objective distribution of the winning probability. How subjects process these probabilities might depend on things like (i) their optimism/pessimism etc. These heterogeneity analyses will be based on subsamples resulting from answers to the post-experimental survey.[5]

Hypothesis 5   *Within each world, more optimistic individuals require lower MAPs on average than pessimistic individuals.*

Maybe a question on taking risks? Concerns with fairness? External reference point for earnings e.g. how much should a 45-minute survey pay for you to be willing to take it (if it pays a fixed amount)?–¿ if international sample, allow them to select currency

---

[5]Gender is among the demographics which we can obtain from Prolific.

9

5.2 Specifications and Analysis

We present the OLS regressions which will be used to test the hypotheses. Additionally, we will also run non-parametric Mann-Whitney U tests.

The main hypotheses (1–3) will be tested using the following regression:

$$MAP_i = \beta + \beta_L L + \beta_R R + \epsilon_i \tag{1}$$

where MAP_i is the MAP chosen by participant i, L is an indicator which takes the value of 1 if the decision was made in the left skew world, R is an indicator which is 1 if the decision was made in the right skew world and ϵ_i is a random error term. Standard errors in the estimation will be clustered at the individual level.

For heterogeneity analyses, we will interact all terms in equation (1) with an indicator variable corresponding to each specific hypothesis. For instance, for Hypothesis 5, all terms will be interacted with indicator variable F_i, which takes the value 1 if the participant is female:

$$MAP_i = \beta + \beta^F F_i + \beta_L L + \beta_L^F LF_i + \beta_R R + \beta_R^F RF_i + \epsilon_i \tag{2}$$

The formal statements of the hypotheses are in the Appendix.

References

Bohnet, I., Greig, F., Herrmann, B., and Zeckhauser, R. (2008). Betrayal aversion: Evidence from Brazil, China, Oman, Switzerland, Turkey, and the United States. *American Economic Review*, 98(1):294–310.

Bohnet, I., Herrmann, B., and Zeckhauser, R. (2010). Trust and the reference points for trustworthiness in Gulf and Western countries. *Quarterly Journal of Economics*, 125(2):811–828.

Bohnet, I. and Zeckhauser, R. (2004). Trust, risk and betrayal. *Journal of Economic Behavior & Organization*, 55(4):467–484.

Fetchenhauer, D. and Dunning, D. (2012). Betrayal aversion versus principled trustfulness—how to explain risk avoidance and risky choices in trust games. *Journal of Economic Behavior & Organization*, 81(2):534–541.

Johnson, C., Baillon, A., Bleichrodt, H., Li, Z., van Dolder, D., and Wakker, P. P. (2019). Prince: An improved method for measuring incentivized preferences. Working paper.

Li, C., Turmunkh, U., and Wakker, P. P. (2020). Social and strategic ambiguity versus betrayal aversion. *Games and Economic Behavior*, 123:272–287.

Polipciuc, M. and Strobel, M. (2020). Betrayal aversion with and without a motive. Working paper.

Appendix A   Hypothesis Testing

A.1   Hypothesis 1

$H0 : \beta_L - \beta_R = 0$

$H1 : \beta_L - \beta_R > 0$

A.2   Hypothesis 2

$H0 : \beta_L = 0$

$H1 : \beta_L > 0$

A.3   Hypothesis 3

$H0 : \beta_R = 0$

$H1 : \beta_R > 0$

A.4   Hypothesis 4

Within each world:

$H0 : \beta^F = 0$

$H1 : \beta^F > 0$

or

$H0 : \beta^F + \beta_L^F = 0$

$H1 : \beta^F + \beta_L^F > 0$

or

$H0 : \beta^F + \beta_R^F = 0$

$H1 : \beta^F + \beta_R^F > 0$

A.5   Hypothesis 5

Analogous to Hypothesis 4.