

# Political Identity and Trust: Supplementary Appendix

Dylan Minor\*

Pablo Hernández-Lagos<sup>†</sup>

## Abstract

In this supplementary appendix we explain why incentives to hedge are negligible in our setting. We elaborate on three arguments. First, Blanco et al. (2010) do not find evidence of hedging using a design prone to hedging and a game similar to ours. Second, in our setting the maximal deviation from truthful reporting is very small when risk aversion is extreme. Finally, when risk aversion is very high but not extreme truthful reporting is optimal for any belief.

**Keywords:** *Trust, Beliefs, Political Identity*

---

\*Northwestern University

<sup>†</sup>New York University Abu Dhabi.

## Incentives to hedge

In here we elaborate on why hedging motives are negligible in our setting. We base our argument on Blanco et al. (2010) (hereafter BEKN) ideas and on analytical reasoning specific to our setup. BEKN pointed out that in sequential prisoners' dilemma type games, hedging may be a concern when both first-mover behavior and beliefs are incentivized. Using a sequential prisoners' dilemma (SPD) similar to the one we use in the paper, the authors seek to maximize the incentives to hedge. BEKN ask each participant to complete the following sequence of tasks only once. First, each participant makes a decision in the role of second-mover for the case the first-mover cooperates. Then each participant states beliefs regarding second-mover choices of other players and finally each participant makes the first-mover decision. There is no feedback between tasks. Using a between-subject design, BEKN compare the decisions in this SPD across two conditions: Both belief elicitation and action are paid (SPDHedge) and either belief or action is paid (SPDNoHedge). Their design features this sequence aiming to distort both first-mover decisions and belief statements to intensify hedging motives when both belief elicitation and actions are paid. Their data, however, reveals no evidence of hedging. Moreover, in the post-experimental questionnaire, none of their subjects even hinted at hedging (BEKN, p. 425).

Given our participants were drawn from the general US population, we decided to pay every decision to avoid complicating exposition and maximize participation. We also presented decision tasks sequentially and participants were not informed about what decision task will come next. The sequence of decisions in our experiment is also different from BEKN. We first ask participants to make their decisions in the role of first-mover (Player A), then in the role of second-mover (Player B) and finally to report their beliefs.<sup>1</sup> This procedure makes under-reporting the only hedging alternative because participants in the role of first-mover did not know they would have to report their beliefs so the first-mover decision should not respond to hedging motives. In this sense, BEKN task is more likely to lead to hedging than the one in the paper and yet they do not find evidence of hedging in their set up. Nevertheless, it is still worth elaborating on to what extent hedging would be a concern in our game.

---

<sup>1</sup>If studied in isolation, our belief elicitation procedure is incentive compatible even if subjects are risk averse. That is, expected utility from the belief elicitation procedure is maximized when the actual belief is reported (we elaborate on this below).

## Analytical results

In what follows we give analytical arguments as to why hedging incentives in our setting are negligible. Hedging means that a given subject reports a lower probability of reciprocation than what he or she actually believes. Let us denote  $p$  the actual belief the participant one is matched with reciprocates and  $\pi$  the reported number of people who reciprocate. Given that we did not know ex-ante the total number of subjects who would answer the survey, we asked them to report the fraction  $\pi/N$  in percent terms, where  $N$  is the number of participants, instead of  $\pi$  (other studies, such as Costa-Gomes and Weizsacker 2008 also use percent elicitation rather than number of participants). Conditional on  $N$ , reporting  $\pi$  and reporting  $\pi/N$  are equivalent. We therefore state all the results in this supplementary appendix in terms of  $\pi/N$  for a general  $N$  and later on discuss how hedging changes for particular values of  $N$ .

It is useful first to make explicit the relationship between beliefs, which we label  $p$ , and reported number of people who reciprocate,  $\pi$ , for a given  $N$ . This relationship is given by the probability that  $\pi$  out of  $N$  participants reciprocate and it depends on  $p$  as follows:

$$P_\pi = \binom{N}{\pi} p^\pi (1-p)^{N-\pi}.$$

$P_\pi$  is a single-peaked function that attains its maximum when  $\pi/N$  is in  $[p \frac{(N+1)}{N} - \frac{1}{N}, p \frac{(N+1)}{N}]$  which converges to  $p$  as  $N \rightarrow \infty$ . To see this, we look for  $\pi$  such that

$$\frac{\binom{N}{\pi} p^\pi (1-p)^{N-\pi}}{\binom{N}{\pi-1} p^{\pi-1} (1-p)^{N-\pi+1}} \geq 1 \tag{1}$$

and

$$\frac{\binom{N}{\pi} p^\pi (1-p)^{N-\pi}}{\binom{N}{\pi+1} p^{\pi+1} (1-p)^{N-\pi-1}} \geq 1. \tag{2}$$

Simplifying the first inequality leads to  $\frac{(N-\pi+1)p}{\pi(1-p)} \geq 1$  or  $p(N+1)/N \geq \pi/N$  after re-arranging terms. Similarly, the second inequality yields  $p(N+1)/N - 1/N \leq \pi/N$ . Putting these two expressions together:

$$\frac{p(N+1)}{N} - \frac{1}{N} \leq \frac{\pi}{N} \leq \frac{p(N+1)}{N}.$$

This condition makes explicit what truthful reporting means. In what follows we derive a similar condition that makes explicit what hedging means, in this context.

Having an expression for  $P_\pi$  is useful because we can use it to derive an expression for the expected utility of both the first-mover and belief elicitation tasks together.<sup>2</sup> The expected utility of a Player A who decides to trust as a function of how many people he reports will reciprocate ( $\pi$ ) is given by

$$EU[\pi; p] = P_0 \left[ \frac{0}{N}u(10) + \frac{N-0}{N}u(0) \right] + P_1 \left[ \frac{1}{N}u(10) + \frac{N-1}{N}u(0) \right] + \dots \\ + P_\pi \left[ \frac{\pi}{N}u(13) + \frac{N-\pi}{N}u(3) \right] + \dots + P_{N-1} \left[ \frac{N-1}{N}u(10) + \frac{1}{N}u(0) \right] + P_N \left[ \frac{N}{N}u(10) + \frac{0}{N}u(0) \right]. \quad (3)$$

where  $u(\cdot)$  is a strictly concave utility function. For example, if Player A reports  $\pi > 1$  out of  $N$  participants will reciprocate, the term  $P_1 \left[ \frac{1}{N}u(10) + \frac{N-1}{N}u(0) \right]$  is the probability that only one Player B out of  $N$  participants reciprocates times the expected utility. In that case, Player A obtains 10 units as a result of Player B's cooperation if she happens to be matched with the only one Player B that reciprocates—which occurs with probability  $P_1 \times 1/N$ . She does not receive anything from the belief elicitation task because she stated  $\pi$  out of  $N$  would cooperate, but in reality only 1 out of  $N$  does. Similarly,  $P_\pi \left[ \frac{\pi}{N}u(13) + \frac{N-\pi}{N}u(3) \right]$  is the the probability that exactly  $\pi$  out of  $N$  Players B reciprocate times the expected utility. Note in this case Player A receives 3 additional units in case the Player B she is matched with cooperates (for a total of 13 units) and 3 units if the Player B she is matched with defects (the payoff from the belief elicitation procedure only) because she reported correctly that  $\pi$  out of  $N$  people would reciprocate. And so on.

Our task is to find  $\pi$  that maximizes  $EU[\pi; p]$  and show under what conditions under-reporting is optimal. That is, under what conditions  $\pi/N$  is not in  $[p^{\frac{(N+1)}{N}} - \frac{1}{N}, p^{\frac{(N+1)}{N}}]$  and in particular

---

<sup>2</sup>Writing  $P_\pi$  explicitly also shows that when studied in isolation, our belief elicitation procedure is incentive compatible even when participants are risk averse. To see this, note the expected utility of the belief elicitation procedure viewed in isolation is

$$EU_{Belief}[\pi; p] = P_0u(0) + P_1u(0) + \dots + P_\pi u(3) + \dots + P_Nu(0) \\ = \sum_{k=0}^N P_k u(0) + P_\pi u(3) - P_\pi u(0) \\ = \sum_{k=0}^N P_k u(0) + P_\pi [u(3) - u(0)],$$

which implies that maximizing  $EU_{Belief}[\pi; p]$  is equivalent to maximizing  $P_\pi$  for any well-defined utility function.

falls below  $[p\frac{(N+1)}{N} - \frac{1}{N}]$ . We proceed now to write  $EU[\pi; p]$  in a way in which the hedging motive becomes more apparent.

The first step is to note that the expected utility of trusting can be written as a sum involving  $P_i$ . Precisely,

$$pu(x) + (1-p)u(y) = \sum_{i=0}^N P_i \left[ \frac{i}{N}u(x) + \frac{N-i}{N}u(y) \right]$$

which follows from manipulating the right-hand side and an appropriate change of variables.

The second step is to add and subtract  $P_\pi \left[ \frac{\pi}{N}u(10) + \frac{N-\pi}{N}u(0) \right]$  to (3). That is, from equation (3) we take out of the summation the term  $P_\pi \left[ \frac{\pi}{N}u(13) + \frac{N-\pi}{N}u(3) \right]$  and we replace it with  $P_\pi \left[ \frac{\pi}{N}u(10) + \frac{N-\pi}{N}u(0) \right]$ . Since this latter term does not belong in equation (3) we must subtract it. Equation (3) can be written as follows:

$$\begin{aligned} EU[\pi; p] &= \sum_{k=0}^N P_k \left[ \frac{k}{N}u(10) + \frac{N-k}{N}u(0) \right] + P_\pi \left[ \frac{\pi}{N}u(13) + \frac{N-\pi}{N}u(3) \right] - P_\pi \left[ \frac{\pi}{N}u(10) + \frac{N-\pi}{N}u(0) \right] \\ &= pu(10) + (1-p)u(0) + P_\pi \left[ \frac{\pi}{N}u(13) + \frac{N-\pi}{N}u(3) \right] - P_\pi \left[ \frac{\pi}{N}u(10) + \frac{N-\pi}{N}u(0) \right] \\ &= \underbrace{pu(10) + (1-p)u(0)}_{\text{Expected utility from trusting}} + \underbrace{P_\pi \left[ \frac{\pi}{N}(u(13) - u(10)) + \frac{N-\pi}{N}(u(3) - u(0)) \right]}_{\text{Expected utility from reporting } \pi, \text{ when } p \text{ is actual belief}} \end{aligned}$$

By writing  $EU[\pi; p]$  this way we are able to flesh out potential hedging motives. The first term, the expected utility of trusting, does not depend on  $\pi$ . The second term therefore reflects the fundamental trade-off. On the one hand,  $P_\pi$  is maximized by truthful reporting. On the other hand,  $\frac{\pi}{N}(u(13) - u(10)) + \frac{N-\pi}{N}(u(3) - u(0))$  is maximized by setting  $\pi = 0$  when  $u(\cdot)$  is (increasing and) concave. Therefore, it may be the case that when  $u$  is “very” concave (i.e.,  $u(3) - u(0) \gg u(13) - u(10)$ ) participants under-report.<sup>3</sup>

In order to make the argument that hedging motives are negligible, it is instructive to consider the case in which risk aversion is extreme. That is, where  $u(3) - u(0) > 0$  and  $u(13) - u(10) = 0$ . In this extreme case, truthful revelation of beliefs is maximally distorted by the incentive to hedge.

---

<sup>3</sup>If participants are risk neutral (e.g.,  $u(x) = x$ ) then  $EU[\pi : p] = pu(10) + (1-p)u(0) + P_\pi \times 3$  so there are no incentives to hedge.

Analytically,  $EU[\pi; p]$  now can be written as follows

$$EU[\pi; p] = pu(10) + (1-p)u(0) + P_\pi \left[ \frac{N-\pi}{N} (u(3) - u(0)) \right].$$

Since  $\pi$  only enters the second term in this expression and  $u(3) - u(0)$  is a constant, we focus on maximizing  $P_\pi \left[ \frac{N-\pi}{N} \right]$  over  $\pi$ .

Note that

$$\begin{aligned} P_\pi \left[ \frac{N-\pi}{N} \right] &= \binom{N}{\pi} p^\pi (1-p)^{N-\pi} \left[ \frac{N-\pi}{N} \right] \\ &= \frac{N!}{(N-\pi)! \pi!} p^\pi (1-p)^{N-\pi} \frac{N-\pi}{N} \\ &= \frac{(N-1)!}{(N-\pi-1)! \pi!} p^\pi (1-p)^{N-\pi} \\ &= \frac{(N-1)!}{(N-\pi-1)! ((N-1) - (N-\pi-1))!} p^{((N-1)-(N-\pi-1))} (1-p)^{N-\pi-1} (1-p) \end{aligned}$$

denoting  $\rho = (1-p)$  it follows that

$$\begin{aligned} P_\pi \left[ \frac{N-\pi}{N} \right] &= \rho \frac{(N-1)!}{(N-\pi-1)! ((N-1) - (N-\pi-1))!} (1-\rho)^{((N-1)-(N-\pi-1))} \rho^{N-\pi-1} \\ &= \rho \binom{N-1}{N-\pi-1} \rho^{N-\pi-1} (1-\rho)^{((N-1)-(N-\pi-1))}. \end{aligned}$$

This is a single peaked function of  $N - \pi - 1$  and its maximum point satisfies (following the same argument in (1) and (2) after re-arranging terms)

$$\rho - \frac{1}{(N-1)+1} \leq \frac{N-\pi-1}{(N-1)+1} \leq \rho.$$

After further manipulating this expression, the opportunity to hedge leads a rational participant to state  $\pi/N$  such that

$$p - \frac{1}{N} \leq \frac{\pi}{N} \leq p.$$

Note that when  $N$  is very large, hedging motives are negligible even for extreme risk aversion in our setting. Table 1 compares the truthful reporting interval with the hedging interval when risk aversion is extreme.

	Maximal hedging	Truthful reporting
$\frac{\pi}{N} \in$	$[p - \frac{1}{N}, p]$	$[p\frac{(N+1)}{N} - \frac{1}{N}, p\frac{(N+1)}{N}]$

Table 1: Hedging

## Scope of hedging

Given that participants were not informed about the number of other participants in their treatment, our first approach to determine the scope of hedging is to show how the intervals in Table 1 differ for different values of  $N$ . Let us start by assuming that a given participant thinks there are 1000 other participants in his treatment and his actual belief is  $p = 0.5$ . If he is to report truthfully he would state any number within  $[0.4995, 0.5005]$ —this comes from plugging in the numbers in the right column of Table 1. Note, however, that it makes little sense to report  $\pi = 500.2$  other participants will reciprocate, as the smallest unit is 1 person. Reporting truthfully therefore means reporting  $\pi = 500$  out of 1000 people. If he hedges, he would report any number in  $[0.499, 0.5]$  (left column of Table 1). Given that the smallest difference between sensible reporting is one person, he will report either  $\pi/N = 499/N$  or  $\pi/N = 500/N$ . That is, at most he will report only 1 person less than what he actually believes.

Likewise, if a participant who thinks that there are 100 other subjects in his treatment who also hold a belief  $p = 0.5$  is to report truthfully, he would state any number within  $[0.495, 0.505]$ . As before, it makes little sense to report something like “50.1 out of 100 other participants will cooperate” because the smallest unit of analysis is one participant. Hence, he should state  $\pi = 50$ , so  $\pi/N = p = 0.5$ . If he is extremely risk averse, he would report any number in the interval  $[0.49, 0.5]$  instead. That is, the maximal under-reporting would again be 1 person short of his actual belief.

In a treatment in which the participant thinks  $N$  is smaller, as in the treatment with the fewest subjects where  $N = 44$ , the no-hedging interval is  $[0.4886, 0.5114]$  while the maximal-hedging interval is  $[0.4773, 0.5]$ . As before, the only report a truthful agent would make is  $\pi = 22$  out of 44, or  $\pi/N = p = 0.5$  because under-reporting, say  $\pi = 21$ , would lead to  $\pi/N = 21/44 = 0.4773$  which is out of the interval. An extreme risk averse participant may report either  $\pi = 21$  or  $\pi = 22$ . Again, he has an incentive to under-report only by 1 participant. Importantly, in all these cases

incentives to hedge do not rule out truthful reporting as truthful reporting is also included in the interval.

To have a visual representation as to how big the difference between the intervals in Table 1 is, we draw them one on top of the other. The following figures show (at scale) the maximal-hedging (in each case, the one on top) and the truthful-reporting intervals for each  $N$  of the main treatments and for  $p = 0.5$  and  $p = 0.75$ . In a nutshell, even for these relatively small numbers the intervals do not differ too much.

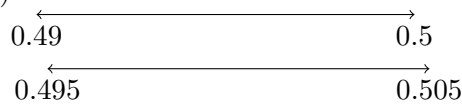
**Each figure shows the maximal hedging interval (top) and the no-hedging interval**

$$p = 0.5$$

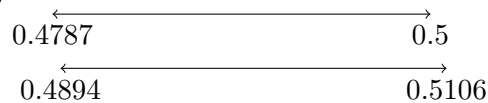
Main Treatment D-NR (N=100)



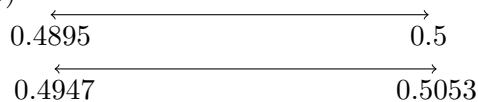
Main Treatment D-D (N=100)



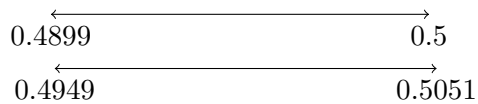
Main Treatment D-R (N=47)



Main Treatment R-NR (N=95)



Main Treatment R-R (N=99)



Main Treatment R-R (N=44)

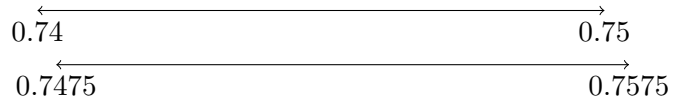


**Each figure shows the maximal hedging interval (top) and the no-hedging interval**

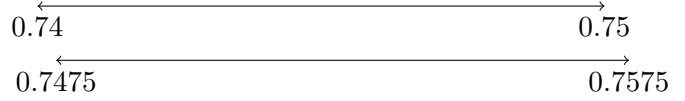


$$p = 0.75$$

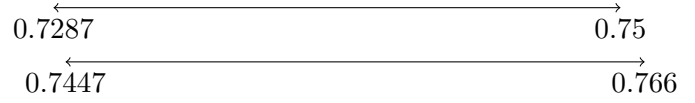
Main Treatment D-NR (N=100)



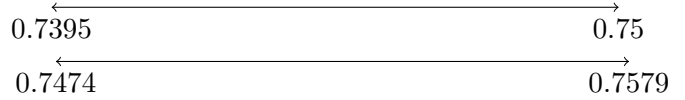
Main Treatment D-D (N=100)



Main Treatment D-R (N=47)



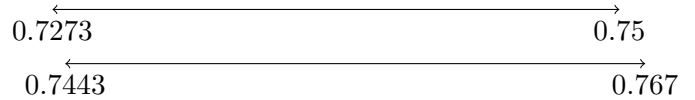
Main Treatment R-NR (N=95)



Main Treatment R-R (N=99)



Main Treatment R-R (N=44)



A second approach is to show whether hedging motives would make a participant to under-report when risk aversion is high, but not extreme—similarly to BEKN. To fix ideas let us assume as in BEKN  $u(x) = x^{1-r}$  and let us focus on a CRRA coefficient  $r = 0.9$  as it entails the highest incentives to under-report in BEKN (see Table 3, 4 and 5 in their appendix). As BEKN (p. 425) points out,  $r = 0.9$  is a large coefficient compared to the range commonly observed in experiments eliciting risk preferences ( $r$  usually ranges between 0.3 and 0.5, see e.g., Holt and Laury, 2002). We show that hedging should not be an issue using such a large coefficient, so it should also not be an issue with smaller coefficients.

As it turns out, truthful reporting is optimal for any belief  $p$  when  $N = 100$ . Figure 1 shows the expected utility (the value of the second term in (4)) as a function of  $\pi$  for different values of  $p \in \{0.01, 0.05, 0.1, \dots, 1\}$ . At 5% intervals it is optimal to report truthfully. Truthful reporting

		Actual belief (p)																				
		0.01	0.05	0.1	0.15	0.2	0.25	0.3	0.35	0.4	0.45	0.5	0.55	0.6	0.65	0.7	0.75	0.8	0.85	0.9	0.95	1
Report ( $\pi/N$ )	0.01	0.4087	0.0344	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.05	0.0031	0.1912	0.0360	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.1	0.0000	0.0168	0.1322	0.0447	0.0034	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.15	0.0000	0.0001	0.0312	0.1068	0.0458	0.0054	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.2	0.0000	0.0000	0.0011	0.0362	0.0893	0.0444	0.0068	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.25	0.0000	0.0000	0.0000	0.0026	0.0371	0.0778	0.0419	0.0076	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.3	0.0000	0.0000	0.0000	0.0001	0.0041	0.0362	0.0687	0.0391	0.0079	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.35	0.0000	0.0000	0.0000	0.0000	0.0001	0.0052	0.0345	0.0615	0.0362	0.0078	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0058	0.0324	0.0555	0.0333	0.0074	0.0006	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.45	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0003	0.0060	0.0301	0.0503	0.0305	0.0068	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0059	0.0277	0.0457	0.0277	0.0059	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.55	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0252	0.0418	0.0249	0.0050	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0051	0.0228	0.0378	0.0221	0.0040	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
	0.65	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0056	0.0252	0.0418	0.0249	0.0050	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000
	0.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0004	0.0044	0.0203	0.0344	0.0193	0.0029	0.0001	0.0000	0.0000	0.0000
	0.75	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0027	0.0151	0.0278	0.0133	0.0009	0.0000	0.0000	0.0000
0.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0019	0.0123	0.0248	0.0101	0.0003	0.0000	0.0000	
0.85	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0094	0.0218	0.0064	0.0000	0.0000	0.0000	
0.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0063	0.0187	0.0024	0.0000	
0.95	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0030	0.0158	0.0000	
1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0035	

Figure 1: CRRA utility  $u(x) = x^{1-r}$  and  $r = 0.9$ .  $N = 100$ .

remains optimal at 1 person (1%) intervals for each  $p \in \{0.01, 0.02, \dots, 0.99, 1\}$  (not shown).

At a more disaggregated level,  $N = 1000$ , Table 2 panel a. shows the utility (the value of the second term in (4)) when beliefs are around 500 out of 1000 participants cooperating, so  $p \in \{0.49, 0.491, 0.492, \dots, 0.510\}$ . For each  $p$  (column) in Table 2 panel a. truthful reporting is also optimal. The same result holds when we compute beliefs around  $p = 0.75$ . Table 2 panel b. shows this.

It is important to notice that since our scoring rule pays only when the participant guesses right within a 10% interval, under-reporting may be an issue if  $N = 10$ . This is because even under-reporting by 1 person would make hedging a problem. This is a valid logic as the intervals in Table 1 differ more as  $N$  goes down. The answer, however, is in the negative: There is no hedging even for  $N = 10$  when  $r = 0.9$  (see Figure 3).

All in all, we could make three arguments against hedging. First, BEKN shows that hedging is unlikely to occur in a game very similar to the one in the paper. Second, even for extreme risk aversion hedging may only occur when a given participant has a very precise belief about the percent in the population that will reciprocate trust. If participants do hold beliefs that precise, however, then maximum under-reporting is only by 1 person. Finally, when the risk aversion is high but not extreme, under-reporting is not optimal when  $N = 10$ ,  $N = 100$  and  $N = 1000$ . This differs a little bit from the prescriptions in the appendix in BEKN where some hedging is possible (although the experimental results ended up showing no hedging). The reason behind this difference is the elicitation beliefs. We paid only for truthful revelation (in intervals of 10%) whereas BEKN paid based on a quadratic-scoring rule that also paid for inaccurate predictions.

		Actual belief (p)																				
		0.4	0.41	0.42	0.43	0.44	0.45	0.46	0.47	0.48	0.49	0.5	0.51	0.52	0.53	0.54	0.55	0.56	0.57	0.58	0.59	0.6
Report ( $\pi/N$ ):	0.4	0.014775	0.014746	0.014658	0.014512	0.014310	0.014055	0.013749	0.013396	0.013000	0.012565	0.012097	0.011599	0.011078	0.010537	0.009984	0.009421	0.008854	0.008289	0.007728	0.007177	0.006638
	0.41	0.014718	0.014747	0.014718	0.014630	0.014484	0.014283	0.014028	0.013723	0.013371	0.012975	0.012542	0.012074	0.011577	0.011057	0.010518	0.009965	0.009403	0.008838	0.008273	0.007714	0.007163
	0.42	0.014602	0.014690	0.014728	0.014690	0.014602	0.014457	0.014256	0.014002	0.013697	0.013345	0.012951	0.012518	0.012051	0.011556	0.011036	0.010498	0.009946	0.009385	0.008821	0.008257	0.007699
	0.43	0.014430	0.014457	0.014662	0.014692	0.014662	0.014575	0.014430	0.014229	0.013975	0.013671	0.013320	0.012927	0.012494	0.012029	0.011534	0.011015	0.010478	0.009927	0.009366	0.008804	0.008242
	0.44	0.014202	0.014402	0.014547	0.014635	0.014626	0.014635	0.014547	0.014403	0.014202	0.013949	0.013645	0.013295	0.012902	0.012471	0.012006	0.011512	0.010994	0.010458	0.009908	0.009350	0.008788
	0.45	0.013922	0.014175	0.014375	0.014520	0.014607	0.014635	0.014607	0.014520	0.014375	0.014176	0.013923	0.013620	0.013270	0.012878	0.012447	0.011983	0.011490	0.010974	0.010438	0.009889	0.009332
	0.46	0.013594	0.013896	0.014149	0.014348	0.014493	0.014580	0.014626	0.014580	0.014493	0.014348	0.014149	0.013896	0.013594	0.013245	0.012854	0.012424	0.011961	0.011469	0.010953	0.010419	0.009871
	0.47	0.013220	0.013568	0.013870	0.014122	0.014321	0.014465	0.014552	0.014581	0.014552	0.014465	0.014321	0.014122	0.013870	0.013568	0.013220	0.012829	0.012401	0.011938	0.011447	0.010932	0.010399
	0.48	0.012805	0.013195	0.013543	0.013844	0.014096	0.014294	0.014438	0.014525	0.014554	0.014525	0.014438	0.014294	0.014096	0.013844	0.013543	0.013195	0.012805	0.012405	0.011961	0.011525	0.011091
	0.49	0.012354	0.012781	0.013170	0.013517	0.013818	0.014069	0.014267	0.014411	0.014498	0.014527	0.014498	0.014411	0.014267	0.014069	0.013818	0.013517	0.013170	0.012781	0.012354	0.011893	0.011404
	0.5	0.011870	0.012330	0.012757	0.013146	0.013492	0.013792	0.014043	0.014241	0.014384	0.014470	0.014498	0.014470	0.014384	0.014241	0.014043	0.013792	0.013492	0.013146	0.012757	0.012330	0.011870
	0.51	0.011361	0.011848	0.012307	0.012733	0.013121	0.013467	0.013766	0.014016	0.014214	0.014357	0.014443	0.014470	0.014443	0.014357	0.014214	0.014016	0.013766	0.013466	0.013121	0.012733	0.012307
	0.52	0.010830	0.011340	0.011826	0.012284	0.012709	0.013096	0.013441	0.013740	0.013990	0.014187	0.014330	0.014416	0.014443	0.014416	0.014330	0.014187	0.013990	0.013740	0.013441	0.013096	0.012709
	0.53	0.010282	0.010809	0.011318	0.011804	0.012261	0.012685	0.013072	0.013416	0.013714	0.013963	0.014160	0.014303	0.014389	0.014416	0.014389	0.014303	0.014160	0.013963	0.013714	0.013416	0.013072
	0.54	0.009723	0.010263	0.010789	0.011297	0.011782	0.012238	0.012661	0.013047	0.013391	0.013689	0.013937	0.014134	0.014276	0.014362	0.014390	0.014362	0.014276	0.014134	0.013937	0.013689	0.013390
	0.55	0.009158	0.009705	0.010244	0.010769	0.011276	0.011760	0.012215	0.012638	0.013023	0.013366	0.013663	0.013911	0.014107	0.014249	0.014335	0.014362	0.014335	0.014249	0.014107	0.013911	0.013663
	0.56	0.008591	0.009141	0.009687	0.010224	0.010749	0.011255	0.011738	0.012192	0.012614	0.012998	0.013340	0.013637	0.013885	0.014081	0.014222	0.014308	0.014335	0.014222	0.014081	0.013885	0.013637
	0.57	0.008027	0.008575	0.009124	0.009669	0.010205	0.010729	0.011234	0.011716	0.012169	0.012590	0.012974	0.013315	0.013612	0.013859	0.014054	0.014195	0.014281	0.014335	0.014281	0.014195	0.014054
	0.58	0.007470	0.008012	0.008559	0.009107	0.009651	0.010186	0.010708	0.011213	0.011694	0.012146	0.012566	0.012949	0.013290	0.013586	0.013833	0.014028	0.014169	0.014254	0.014326	0.014254	0.014169
	0.59	0.006924	0.007457	0.007997	0.008543	0.009090	0.009632	0.010167	0.010688	0.011191	0.011672	0.012124	0.012543	0.012925	0.013265	0.013560	0.013807	0.014001	0.014142	0.014227	0.014326	0.014227
0.6	0.006393	0.006912	0.007443	0.007982	0.008527	0.009073	0.009614	0.010148	0.010668	0.011170	0.011650	0.012101	0.012519	0.012901	0.013240	0.013535	0.013781	0.013975	0.014116	0.014201	0.014326	

Panel a.

		Actual belief (p)																					
		0.74	0.741	0.742	0.743	0.744	0.745	0.746	0.747	0.748	0.749	0.75	0.751	0.752	0.753	0.754	0.755	0.756	0.757	0.759	0.76		
Report ( $\pi/N$ ):	0.74	0.009938	0.009932	0.009961	0.009845	0.009884	0.009841	0.009823	0.009796	0.009750	0.009713	0.009692	0.009657	0.009618	0.009579	0.009537	0.009499	0.009459	0.009418	0.009378	0.009333	0.009286	
	0.741	0.009912	0.009938	0.009912	0.009842	0.009825	0.009864	0.009841	0.009821	0.009794	0.009763	0.009723	0.009692	0.009653	0.009616	0.009579	0.009535	0.009490	0.009451	0.009410	0.009371	0.009347	
	0.742	0.009822	0.009892	0.009918	0.009892	0.009822	0.009805	0.009844	0.009841	0.009820	0.009792	0.009761	0.009723	0.009692	0.009653	0.009614	0.009574	0.009535	0.009492	0.009452	0.009412	0.009374	
	0.743	0.009787	0.009893	0.009972	0.009998	0.009972	0.009902	0.009875	0.009824	0.009821	0.009799	0.009790	0.009750	0.009725	0.009689	0.009653	0.009611	0.009570	0.009536	0.009493	0.009451	0.009414	
	0.744	0.009609	0.009768	0.009883	0.009953	0.009979	0.009953	0.009882	0.009875	0.009864	0.009840	0.009819	0.009780	0.009750	0.009722	0.009687	0.009649	0.009610	0.009571	0.009529	0.009484	0.009446	
	0.745	0.009391	0.009589	0.009748	0.009863	0.009933	0.009959	0.009933	0.009862	0.009846	0.009824	0.009803	0.009782	0.009760	0.009735	0.009712	0.009685	0.009647	0.009608	0.009568	0.009527	0.009487	
	0.746	0.009135	0.009371	0.009570	0.009728	0.009843	0.009913	0.009937	0.009913	0.009842	0.009826	0.009805	0.009786	0.009765	0.009740	0.009719	0.009691	0.009653	0.009615	0.009576	0.009536	0.009495	
	0.747	0.008845	0.009115	0.009351	0.009550	0.009708	0.009823	0.009893	0.009917	0.009893	0.009822	0.009806	0.009785	0.009764	0.009741	0.009719	0.009690	0.009651	0.009612	0.009572	0.009532	0.009491	
	0.748	0.008525	0.008825	0.009095	0.009331	0.009530	0.009688	0.009803	0.009873	0.009897	0.009873	0.009802	0.009786	0.009765	0.009742	0.009721	0.009692	0.009653	0.009614	0.009574	0.009534	0.009493	
	0.749	0.008181	0.008505	0.008785	0.009075	0.009311	0.009510	0.009668	0.009784	0.009854	0.009879	0.009854	0.009834	0.009813	0.009792	0.009771	0.009750	0.009729	0.009708	0.009687	0.009666	0.009645	
	0.75	0.007815	0.008161	0.008485	0.008785	0.009085	0.009321	0.009540	0.009697	0.009784	0.009834	0.009834	0.009814	0.009793	0.009772	0.009751	0.009730	0.009709	0.009688	0.009667	0.009646	0.009625	
	0.751	0.007434	0.007796	0.008141	0.008466	0.008765	0.009035	0.009271	0.009470	0.009629	0.009744	0.009814	0.009834	0.009813	0.009792	0.009771	0.009750	0.009729	0.009708	0.009687	0.009666	0.009645	
	0.752	0.007042	0.007423	0.007771	0.008121	0.008446	0.008745	0.009015	0.009251	0.009450	0.009609	0.009724	0.009794	0.009814	0.009793	0.009772	0.009751	0.009730	0.009709	0.009688	0.009667	0.009646	
	0.753	0.006643	0.007043	0.007392	0.007742	0.008067	0.008366	0.008636	0.008871	0.009070	0.009229	0.009344	0.009414	0.009434	0.009413	0.009392	0.009371	0.009350	0.009329	0.009308	0.009287	0.009266	
	0.754	0.006243	0.006624	0.006973	0.007323	0.007648	0.007947	0.008217	0.008452	0.008651	0.008810	0.008925	0.009005	0.009045	0.009045	0.009045	0.009045	0.009045	0.009045	0.009045	0.009045	0.009045	0.009045
	0.755	0.005844	0.006224	0.006573	0.006923	0.007248	0.007547	0.007817	0.008052	0.008251	0.008410	0.008525	0.008595	0.008635	0.008635	0.008635	0.008635	0.008635	0.008635	0.008635	0.008635	0.008635	0.008635
	0.756	0.005445	0.005824	0.006173	0.006523	0.006848	0.007147	0.007417	0.007652	0.007851	0.008010	0.008125	0.008195	0.008235	0.008235	0.008235	0.008235	0.008235	0.008235	0.008235	0.008235	0.008235	0.008235
	0.757	0.005046	0.005424	0.005773	0.006123	0.006448	0.006747	0.007017	0.007252	0.007451	0.007610	0.007725	0.007795	0.007835	0.007835	0.007835	0.007835	0.007835	0.007835	0.007835	0.007835	0.007835	0.007835
	0.758	0.004647	0.005024	0.005373	0.005723	0.006048	0.006347	0.006617	0.006852	0.007051	0.007210	0.007325	0.007395	0.007435	0.007435	0.007435	0.007435	0.007435	0.007435	0.007435	0.007435	0.007435	0.007435
	0.759	0.004248	0.004624	0.004973	0.005323	0.005648	0.005947	0.006217	0.006452	0.006651	0.006810	0.006925	0.006995	0.007									

## References

Blanco, M., Engelmann, D., Koch, A. K., and Normann, H. T. (2010). Belief elicitation in experiments: is there a hedging problem? *Experimental Economics*, 13(4), 412-438.

Costa-Gomes, M. A., and Weizsacker, G. (2008). Stated beliefs and play in normal-form games. *The Review of Economic Studies*, 75(3), 729-762.

Holt, C. A., and Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review*, 92(5), 1644-1655.