

## Pre-Analysis Plan (Part 2)

06/04/2021

### General Remarks

This pre-analysis plan describes additional treatments as part of the research project “Paternalism and Incentives” and the study “Why High Incentives Cause Repugnance: A Framed Field Experiment” (please find the discussion paper on which this plan is based on in the attachment and available for download at <https://ssrn.com/abstract=3850618>). In the first part of the project (as preregistered on February, 05, 2020 and described in the attached discussion paper), participants decided whether to permit or prohibit the experimenter to offer a third person (the “potential donor”) monetary compensation for registering as a stem cell and bone marrow donor. Participants decided for €10 (in treatment “*Low*”) and for €500 (in treatment “*High*”). The participants were then randomly assigned to one of two treatments in which they again decided whether to prohibit or permit a €500 offer to a third person. One half of the participants decided in treatment “*Reservation Price*”,<sup>1</sup> where they knew that the offer would only be made to a potential donor who was willing to sign up for €10 and therefore could not be persuaded by the high monetary payment of €500. The other half decided in treatment “*Probability*”, in which they knew that, even if they allowed the offer, the offer would only be made with a lowered probability.<sup>2</sup> Out of all participants, one decision and one participant was randomly selected and the decision of this participant was implemented. The experiment ended with an elicitation of participants’ behindness aversion, their beliefs about reservation prices, their risk-preferences, their self-reported envy, an elicitation of whether the participants consider the €500-offer to be coercive, and of basic demographics.

I also conducted two hypothetical survey experiments with members of ethics committees (members of the German Ethics Council and members of ethics committees affiliated to the ten biggest universities). Participants made hypothetical choices in treatments *Low*, *High*, and *Reservation Price*, as described above.<sup>3</sup>

---

<sup>1</sup> In the first part of the pre-analysis plan I used a different labelling of the treatments. Here, I stick to the labelling used in the discussion paper: the treatment labelled “TREATMENT” in the pre-analysis plan for part 1 is treatment “Reservation Price” and the treatment labelled “LOW\_PROB” in the pre-analysis plan for part 1 is treatment “Probability.”

<sup>2</sup> The probability was 30% and corresponds to the average belief about the share of individuals willing to sign up for €10. The control treatment therefore allowed to investigate the effect of reducing the probability of the offer has relative to the effect of ensuring that individuals cannot be persuaded by the monetary incentive.

<sup>3</sup> I did not preregister these survey experiments, mainly because I contacted the participants via email and did not know ex-ante whether I will be able to recruit a sufficiently large sample of members of ethics committees.

Based on these designs, I made two major findings. First, the majority of individuals who become less approving of the offer if the offer involves €500 rather than €10, labelled ethicists, are not prohibiting the €500-offer if it is made to a potential donor who reveals to have a low reservation prices for signing up (in *Reservation Price*).<sup>4</sup> Second, the share of individuals who are ethicists (individuals who become less approving of the offer if the offer involves a higher payment), seems to be higher in the ethics committees than in the general public.<sup>5</sup>

This pre-analysis plan specifies six additional treatments with which I explore these findings: four treatments to explore the finding that ethicists are not prohibiting offering a €500-payment to a potential donor who reveals to have a low reservation prices for becoming a donor, and two treatments for the finding that the share of ethicists seems to be higher in the ethics committees than in the general public.

The subsequent pre-analysis plan therefore consists of two parts. The first part provides the pre-analysis plan for the additional treatments conducted to explore treatment *Reservation Price*. The second part describes the additional control treatments to investigate the distribution of types between samples.

The table below provides an overview of the versions of the study and the treatments. In normal font, it shows the versions that I already have conducted. In bold, it shows the versions I preregister here.

Wave	Experiment 1 (incentivized choices)					Experiment 2 (unincentivized choices)			
Sample	Representative Sample							IRBs	
Version	Reservati on Price	Probabili ty	Social	Lottery	Medium	Hyp Rep	Hyp Rep IRB	Hyp IRB1	Hyp IRB2
(1) Feb '20	Low, High, Reservati on Price	Low, High, Probabili ty							
(2) Sep '20								Low, High	Low, High
(3) June '21	<b>Low, High, Reservati on Price Replicati on</b>		<b>Low, High, Social</b>	<b>Low, High, Lottery</b>	<b>Low, High, Medium</b>	<b>Low, High</b>	<b>Low, High</b>		

<sup>4</sup> Their willingness to pay shifts from €2.03 to prohibit the €500 offer made to everyone to more than €1 to facilitate the offer made to only individuals with a low reservation price. Equally, 52% of ethicists facilitate the conditional offer when looking at the second, binary, outcome.

<sup>5</sup> While 17.39% (35.65%) of participants are ethicists (ethicists or deontologists) in the representative sample, 35% (63.75%) are ethicists (ethicists or deontologists) in the ethics committees. By deontologists, I mean individuals who prohibit the offer independent of whether the payment is €10 or €500.

Across the different versions of the experiment, participants' choices vary in several dimensions. First, choices are incentivized or not. Second, the experiment is conducted with a representative sample or a sample of members of ethics committees. For the incentivized choices "Experiment 1 (incentivized choices)," I differentiate the following versions. In each of them, participants make exactly the same decision in *Low* (for €10) and *High* (for €500). They differ in the third choices participants make, that is, participants are randomly assigned to either treatment *Reservation Price Replication, Social, Lottery, or Medium*.<sup>6</sup> For the unincentivized choices "Experiment 2 (unincentivized choices)," I already obtained observations from members of the German Ethics Council (*Hyp IRB1*) and the ethics committees affiliated to the ten biggest German universities (*Hyp IRB2*). Here, I preregister to obtain hypothetical choices in *Low* and *High* from the representative sample (*Hyp Rep*) and from the representative sample when asked to decide as if they were members of an ethics council (*Hyp Rep IRB*). Below, I provide the pre-analysis plans for these new treatments.

## Part 2.1

### Experimental design

I conduct four additional treatments to further explore the result that individuals are much more favorable of the high payment offer in *Reservation Price*. The treatments fully replicate the already conducted treatments *Reservation Price* and *Probability* but differ in the third choices participants make (unless for *Reservation Price Replication*, with which I exactly replicate the original treatment *Reservation Price*).<sup>7</sup> As for the initial treatments, one decision and one participant will randomly be determined and her choice will be implemented.

First, I replicate the main treatment, treatment *Reservation Price*, with treatment *Reservation Price Replication*. Replicating the main treatment is helpful in several ways. It allows to derive all treatment comparisons using between-subjects designs and to compare them while using observations that were obtained at the same point at time: the comparison of *Reservation Price* and *Probability* based on data from Feb 2020, and the comparison of *Reservation Price* and *Social, Lottery, and Medium* based on data from June 2021.<sup>8</sup> Note that I preregister not to pool the data, but to report them separately,

---

<sup>6</sup> In addition, the versions may vary with respect to additional measures that I elicit at the end of the experiment (see below).

<sup>7</sup> For all treatments, I add a question concerning the efficiency of the lottery and the medium incentive (both see below) to the elicitation of participants' beliefs about reservation prices. In addition, I replace the question whether offering money for signing up as a donor is coercive and, if yes, starting from which amount, with the module to assess six moral principles proposed in Elias et al. (2019).

<sup>8</sup> (Self-)replicating the main treatment is in line with current efforts to focus on methodological rigor in social sciences (e.g., Camerer et al., 2018; Shah et al., 2019, Kessler and Exley, 2020). The initial treatments were run in February 2020. The corona pandemic made ethical trade-offs salient and might also have changed people's perception of incentives that are part of controversial transactions.

that is, I compare the results in *Reservation Price Replication* with the other three new treatments.

Second, I test the degree to which participants are permitting the €500 offer to potential donors with a reservation price of €10 or lower, because willing to sign up for €10 signals being a prosocial or altruistic person in treatment *Social*. The instructions equal the instructions of treatment *Reservation Price* except that, in the third part of the experiment, participants make a decision for a potential payment of €500, but they know that even if they permit the offer, the offer will only be made if the potential donor independently agrees to donate from €100 she receives, at least €50 to the German Red Cross. Thereby, the treatment ensures that the offer is made only to individuals who are sufficiently prosocial.<sup>9</sup> In addition, I obtain two comparisons of how social the participants perceive signing up as a donor for €10 relative to donating €50. First, participants are asked which of the actions they perceive as being more prosocial. Second, they are also asked to hypothetically allocate €10 to two fictitious individuals, one who is willing to sign up for €10 as a stem cell donor and the other willing to donate €50 to the Red Cross.

Third, I test a potential policy tool to use high economic incentives to incentivize stem cell donors. I derive this policy tool from the main treatment and the finding that ethicists do not object to high incentives if the potential has a low reservation price for becoming a donor. In the third part of treatment *Lottery*, participants decide whether to approve an offer that pays €500 with a probability of 50% and with 50% probability pays only €10. The decisive difference to treatment *Reservation Price* lies in the fact that an individual deciding whether to sign up as a stem cell donor for €10 does not know that signing up implies that they will receive a €500-offer in *Reservation Price*, while in treatment *Lottery* the individual knows that willing to sign up implies a coin-flip between a €10 and a €500-payment.

Fourth, I conduct a treatment in which the participants are offered the expected value of the lottery, that is, €255 (treatment *Medium*). While this treatment (in expectation) keeps the incentive for signing up as a donor the same as in *Lottery*, participants might evaluate the reservation price of a potential donor signing up to be higher (up to €255). Therefore, the participants might be much less likely to approve of the offer compared to *Lottery*. The treatment also allows me to investigate how strongly the participants object to a €255 relative to the €500 offer.

---

<sup>9</sup> In fact, according to the meta-analysis of Engel (2011), the average share of individuals who is willing to give half of the endowment in dictator games or more is exactly 30%. Therefore, (the average belief about) the probability of the offer being made in SOC should be the same as in *Reservation Price*. This implies that the signal of prosociality arising from donating half the endowment should be the same as for a participant willing to sign up as a donor for €10 in the sense that the share of individuals who are willing to engage in either action is equal.

## Main Research Questions

Q1. Is it possible to replicate the findings in *Reservation Price*, that is, what is the difference in the willingness to pay (WTP) of ethicists between *High* and *Reservation Price Replication*?

Q2. To what extent can the effect of *Reservation Price* be explained by the potential donor signaling to be prosocial, that is, what is the difference in the WTP of ethicists between *Reservation Price Replication* and *Social*?

Q3. Is a variant of *Reservation Price* successful in rescinding ethicists' objection to high monetary payments, that is, what is the difference in the WTP of ethicists between *Reservation Price Replication* and *Lottery*?

Q4. Is this variant of *Reservation Price* more accepted than paying the expected value of the lottery, that is, what is the difference in the WTP of ethicists between *Lottery* and *Medium*?

## Analysis

I plan to conduct exactly the same analyses as the once that I have preregistered in Part 1 of the pre-analysis plan (preregistered on February 05, 2020) now with the new control treatments *Social*, *Lottery*, or *Medium*. For the sake of clarity, I do not repeat these analyses here, but provide them in a separate document.<sup>10</sup> In this document treatment *Probability* (formerly labelled LOW\_PROB) is replaced by NEW TREATMENTS where NEW TREATMENTS is either treatment *Social*, *Lottery*, or *Medium*.<sup>11</sup> I plan to conduct some additional analysis for treatments *Lottery* and *Medium* that I specify in the document as well. In the new treatments (*Reservation Price Replication*, *Social*, *Lottery*, and *Medium*), I focus on participants' WTP to permit or prevent certain offers and do not obtain separate choices using the discrete decision whether to permit or prevent offers (as I had done for the original treatments). I can, however, recover the discrete decision from the WTP measure and preregister several analyses in this respect.

## Sample

I will employ the full sample of participants who complete the online experiment. As a robustness check, I plan to consider the sample consisting only of participants who pass

---

<sup>10</sup> I exchange the coercion question by the moral concerns module in analyses 5, 6, and 18. Because I additionally measure participants perception of how prosocial signing up as a donor, I update analyses 4 and 9. Finally, I correct an error in analyses 8.

<sup>11</sup> Note also again that in Part 1 of the pre-analysis plan, treatment *Reservation Price* was labelled "Treatment."

the screener, a question administered to check whether a participant was paying attention to the instructions.

## **Power Analysis**

I derive the sample size for a power of 0.8 and a significance level of 0.05 using two-sided tests. I engage in determining the sample sizes required to test both Hypothesis 2 and Hypothesis 3 (see Main Analysis 2 and 3 in the attachment). I will then follow the more restrictive hypothesis, that is, obtain the largest required sample, and apply it to all four new treatments. To derive the sample size, I ask how large the sample needs to be in order to be able to identify the effect of, for example, being prosocial compared to when deciding for a €500-offer without this information (*Social vs. High*) if the effect of being prosocial makes up for a quarter of the size of the effect of having a low reservation price (in treatment *Reservation Price*) on the one hand, and if one also wants to identify the difference between being prosocial and the reservation price (*Reservation Price Replication vs. Social*). Taking this approach implies that the first criterion is binding and that the sample size needs to allow to identify a difference in the WTP to prohibit the offer between the WTP in *High* (€2.030303) and the hypothesized WTP in *Social* (€(-2.030303-1.106061)\*0.5+(2.030303)=1.246212). This requires 19 ethicists per group and, with the share of ethicists being 16.71%, 114 observations per treatment. I plan to obtain a larger sample of 200 observations per treatment (implying 35 ethicists per treatment).

## **Part 2.2**

### **Experimental design**

My results indicate that the share of true ethicists (true ethicists plus deontologists) is higher in a representative sample than in IRBs. However, while I incentivized participants' choices in the representative sample, the results for the members of IRBs were based on hypothetical choices. To investigate further whether the share of ethicists and the share of individuals prohibiting the €500 offer (true ethicists plus deontologists) varies between the public and members of ethics committees, I conduct the following two versions of the experiment. In the first one, I plan to conduct the hypothetical version of the experiment with the general public. Thereby, I can rule out that differences in the share of ethicists (and ethicists plus deontologists) are driven by a difference in the methodology. In the other new treatment, I investigate why the preferences of the general public and members of IRBs may differ. While there are several reasons for a potential difference – members of IRBs might be better informed, they might be more experienced in making ethical decisions, or they might be a selection of individuals with specific ethical attitudes – I test

whether members of IRBs decide differently because they are deciding as members of an IRB.<sup>12</sup>

### *Hyp Rep:*

The survey experiment mirrors the incentivized experiment: Participants decide to approve an offer concerning monetary compensation for becoming a stem cell and bone marrow donor. They decide for a €10-compensation in *Low* and a €500-compensation in *High*.<sup>13</sup> Participants make simple discrete decisions as to whether to approve the offers or not. Choices are not incentivized. A post-experimental questionnaire asks participants to explain their choices, whether and how they would have behaved differently as part of an IRB, elicits a moral concerns module (Elias et al., 2019), and basic demographics.

### *Hyp Rep IRB:*

Treatment *Hyp Rep IRB* resembles treatment *Hyp Rep* with the only difference that the participants are asked to make their decision as if they were a member of an IRB. The instructions also make clear that the participants should make the decision in the way they would make the decision if they were a member of an IRB and not how they think members of an IRB would decide. The post-experimental questionnaire is analogous to the one for treatment *Hyp Rep*.

## **Main Research Questions**

Q1. Is there a difference in the share of true ethicists (true ethicists plus deontologists) between the public and ethic committees, when the same methodology is used to elicit these shares, that is, between *Hyp Rep* and *Hyp IRB*?

Q2. Is the potential difference between the general public and members of IRBs driven by individuals deciding differently as members of IRBs? That is, is there a difference in the share of true ethicists (true ethicists plus deontologists) between *Hyp Rep* and *Hyp Rep IRB*?

## **Analysis**

### ***Preliminaries***

1. The outcome of interest is the decision whether to approve the offer or not. I define four types of participants: A “true ethicist” approves of making the offer when there are no monetary consequences in *Low* but not in *High*; a “true economist” approves of the offer in *High* but not in *Low*; a “libertarian” approves of the offer in

---

<sup>12</sup> The IRB might serve the purpose of protecting potential donors, deciding as a member of an IRB might involve a specific assessment of the risks of being a donor etc.

<sup>13</sup> As for the main experiment, I varied the order of the incentives.

both *Low* and *High*; a “deontologist” disapproves of making the offer in both *Low* and *High*.

2. *Hyp Rep*, *Hyp IRB*, and *Hyp Rep IRB*: Analysis of the distribution of the ethical types using histograms.

### **Main Analysis**

1. *Hyp IRB* vs. *Hyp Rep*: Two-sided test of proportions testing whether the share of true ethicists is higher in *Hyp IRB* than in *Hyp Rep*. Analogously for share of true ethicists plus deontologists.<sup>14</sup> In addition, test of equality of distributions.
2. *Hyp Rep IRB* vs. *Hyp Rep*: Two-sided test of proportions testing whether the share of true ethicists (true ethicists plus deontologists) is higher in *Hyp Rep IRB* than in *Hyp Rep*. In addition, test of equality of distributions.

### **Additional Analysis**

3. *Hyp Rep*: Distribution of the six moral principles (see Elias et al., 2019, Figure 4). Pearson correlation coefficients  $r$  between the six moral principles.
4. *Hyp Rep*: Regression of the ethical type on controls (gender, age, employment status, educational attainment, income, role of religion in life, political party, own donor status, the six moral principles) in order to analyze whether there are partial correlations between the type and individual characteristics. I also plan to conduct multiple separate regressions for the controls to investigate the correlations with the types.<sup>15</sup> I will also derive the distribution of types that would result if the general public would resemble the ethics committee with respect to observables (like age, gender, occupation).
5. *Hyp Rep* and *Hyp Rep IRB*: Regression of the type on a dummy for the treatment condition controlling for all controls.
6. To analyze the open comments that participants can leave at the end of the experiment participants can leave to explain their choices and to discuss whether they would have behaved differently had they been in the other treatment (*Hyp Rep IRB* or *Hyp Rep*, respectively), I plan to conduct an explanatory analysis.

### **Sample**

I will employ the full sample of participants who complete the online experiment. As a robustness check, I plan to consider the sample consisting only of participants who pass

---

<sup>14</sup> In case of a difference, it will be interesting to see whether the share of true ethicists (true ethicists plus deontologists) is the same when using incentivized and when using unincentivized choices in the representative sample.

<sup>15</sup> For the moral concerns module, I will test them separately and derive a linear index. As in Elias et al., 2019, I also plan to once use the principal component(s) of the moral principles.

the screener, a question administered to check whether a participant was paying attention to the instructions.

## **Power Analysis**

I again derive the values for a power of 0.8 and a significance level of 0.05 using two-sided tests. When deriving the sample size, I pool the results from the two survey experiments (the survey experiment with members of the German Ethics Council and the survey experiment with members of IRBs). I engage in determining the sample sizes required to test both Hypothesis 1 and 2 (see Main Analysis 1 and 2) and obtain the larger sample.

For Hypothesis 1, the hypothesized share of true ethicists (true ethicists plus deontologists) is given from the findings using incentivized choices and equals 17.39% (35.65%). Deriving a required sample size requires an alternative hypothesis that in this case can be based on the share I observe in the IRB sample. I therefore derive one sample power calculations using 17.39% (35.65%) as the shares of true ethicist (true ethicists plus deontologists) and alternative values of 35.00% (63.75%) leading to required sample sizes of 43 participants (23 participants).

For Hypothesis 2, asking the participants to decide as a member of an IRB (in *Hyp Rep IRB*) might bring their choices closer to the choices of members of IRBs (in *Hyp Rep*). However, as several factors might be driving the observed difference between the representative sample and the members of IRBs (see above), asking the participants to make their decisions as an IRB member might not make up for the full difference. I therefore derive the sample size necessary if the effect makes up for two-thirds of the difference between the representative sample and the IRB sample. The identified difference is  $35.00 - 17.39 = 17.61$  ( $63.75 - 35.65 = 28.10$ pp) for true ethicists (true ethicists plus deontologists). The hypothesized differences between *Hyp Rep* and *Hyp Rep IRB* are therefore 11.74pp (18.73pp) leading to required sample sizes of 234 participants (110 participants) per treatment. I plan to obtain a slightly larger sample of 250 for both versions.

## References

Camerer, Colin F., et al. "Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015." *Nature Human Behaviour* 2.9 (2018): 637-644.

Exley, Christine L. and Judd B. Kessler. 2020. "The Gender Gap in Self-Promotion." Unpublished Manuscript.

Elias, J. J., Lacetera, N., and Macis, M. (2019). "Paying for kidneys? A randomized survey and choice experiment." *American Economic Review* 109(8), 2855-88.

Shah, Anuj K., Sendhil Mullainathan, and Eldar Shafir. 2019. "An exercise in self-replication: Replicating Shah, Mullainathan, and Shafir (2012)." *Journal of Economic Psychology*, 75, 102127.

Stüber, Robert. 2020. "Why High Incentives Cause Repugnance: A Framed Field Experiment." Available at SSRN: <https://ssrn.com/abstract=3850618>. May 15, 2021.