Pre-Analysis Plan for "Mechanisms of Moral Motive Selection"

This document contains the pre-analysis plan for the data we plan to collect in additional sessions for the paper entitled "Moral Motive Selection" by Kai Barron, Robert Stüber, and Roel van Veldhuizen.

A. Original Experiment

We conducted a laboratory experiment in December 2018 in which participants took part in a "Lying Dictator Game" (LDG). In this game, the decision maker is asked to reveal a number between 0 and 10 by clicking on one of 11 boxes. The decision maker is then asked to report the number, which affects how a total of 10 euros is split between them and another participant. In particular, the decision-maker receives the number reported (e.g., EUR 6 after reporting 6), whereas the other participant receives the residual (e.g., 10-6= EUR 4). Participants then completed a spectator version of the same game (SLDG) where their choice affected how a payment of 10 euros was split between two other participants. The original pre-analysis plan for this experiment can be found here: https://www.socialscienceregistry.org/trials/3617. A working paper describing the design and results in greater detail can be found here: https://www.cesifo.org/DocDL/cesifo1_wp9911.pdf**.**

B. New Experiment

We plan to conduct a new experiment that differs from the original experiment in three main ways. First, it will be conducted as an online experiment on Prolific rather than in the lab. Second, we will now use the strategy method for the second game (the SLDG in our original experiment), where decision-makers are asked to decide on their report before they see the number (either the truth or a specific number). Third, we will now collect data for four treatments that vary two design features using a 2x2 between-subjects design.

The first treatment dimension we vary is the action space. In our original experiment, participants could always choose between four ways to make their report:

(i) tell the truth,
(ii) equalize payoffs,
(iii) maximize their own (LDG) / participant A's (SLDG) payoff, and
(iv) maximize the other participant (LDG) / participant B's (SLDG) payoff.

In the new experiment, we will replicate the experiment with this *restricted action space,* but we will now also consider a version of the experiment in which the decision-maker can choose any number between 0 and 10 (*full action space*). This will allow us to assess whether our results are replicated in a game that is closer to the way the dictator and lying

games are usually implemented. On the second treatment dimension, we will vary whether the second game is a spectator lying dictator game (SLDG, as before) or a regular lying dictator game (LDG). This will allow us to provide richer evidence on when spillovers occur.

We therefore have the following four treatments:

1. Restricted-FirstPerson: LDG/LDG (restricted action space)
2. Restricted-Spectator: LDG/SLDG (restricted action space)
3. FullSpace-FirstPerson: LDG/LDG (full action space)
4. FullSpace-Spectator: LDG/SLDG (full action space)

C. Main Outcomes

1. The decision maker's report in the first game (LDG).
2. The decision maker's report in the second game (either LDG or SLDG).

D. Main Explanatory Variables

1. The random draw in the first game
    a. We will typically compare decision-makers who received a LOW draw (0-4) to those who received a HIGH draw (6-10).
2. The two treatment variables
    a. Action space: full vs restricted
    b. Second game: LDG or SLDG

E. Main Analysis

We will start by replicating our main analysis from the earlier version of the experiment, as reported in our existing working paper:

1. Motive Selection: does the random draw affect the game 1 report?
    a. We will report the fraction of participants in the first game who choose (i) truth-telling, (ii) equality, (iii) payoff maximization, or (iv) any other report as a function of the random draw in the first game (either HIGH or LOW as per D1a above).
    b. We will test whether the truth-telling (equality) rate in the first game is significantly higher (lower) for HIGH rather than LOW random draws. In both cases, we will use a two-sided test of proportions.
    c. We will conduct these analyses (E1a and E1b) separately for each of the two action space treatments (full vs restricted) while pooling across the other treatment dimension (i.e., pooling the treatments with SLDG and LDG in game 2).
2. Spillovers: does the game 1 random draw affect the game 2 report?

a. We will report the fraction of participants in the second game who choose (i) truth-telling, (ii) equality, (iii) payoff maximization (in the LDG only), or (iv) any other report as a function of the random draw in the first game (either HIGH or LOW as per D1a above).
b. We will test whether truth-telling is chosen more often than equality in the second game for HIGH random draws from the first game. We will test whether equality is chosen more often than truth-telling in the second game for LOW random draws from the first game. We will use a two-sided test of proportions in both cases.
c. We will do this analysis (E1a and E1b) for each of the two spillover treatments (LDG vs SLDG), pooling across the other treatment dimension (i.e., pooling the two action spaces).

In addition, we will also look at the following treatment differences:

3. Action Space: does restricting the action space affect motive selection (E1)?
    a. We will test whether the frequency of each type of report (truth-telling, equality, payoff maximization, or other) differs between the full action space and restricted action space treatments, separately for LOW and HIGH random draws (per D1a). We will use a two-sided test of proportions in each case, and pool across the two treatments within each treatment arm (i.e., pool the treatments with SLDG and LDG in game 2).
    b. We will regress truth-telling (1-truth-telling, 0-other reports) on a dummy for HIGH draws (HIGH-1, LOW-0), a dummy for the full action space treatments (1-full, 0-restricted), and the interaction of the two. We will then test whether the interaction term is significant using a two-sided test. We will pool within the action space treatment dimension (i.e., for each action space treatment, we will pool the treatments with SLDG and LDG in game 2).
    c. We will repeat the previous test with equality (1-equality, 0-other reports) as the dependent variable instead.
    d. In case the frequency of "other" reports (i) differs by more than 15 percentage points and (ii) differs significantly by treatment (1-full, 0-restricted), we will also redo analysis E3b and E3c while removing participants that chose the "other" report.
4. Spillovers and Game Type: does the spillover effect (E2) depend on the game type (SLDG vs LDG)?
    a. We will regress truth-telling (1-truth-telling, 0-other reports) in the second game on a dummy for HIGH draws (HIGH-1, LOW-0) from the first game, a dummy for the game type (1-SLDG, 0-LDG) and the interaction of the two. We then test whether the interaction term is significant using a two-sided test.

b. We will repeat the previous test with equality (1-equality, 0-other reports) as the dependent variable instead.
c. If there is no significant effect of action space in E3, we will pool the data from both action space arms. If there is a significant effect of action space (at the 5% level in any of tests E3a-3c) we will instead conduct tests E4a and E4b separately in each treatment arm.

F. Power Calculations

We will aim to collect 400 participants per treatment, which would give us a power of 0.80 to detect motive selection or spillover effects of around 10pp.

1. Motive Selection (E1):

In our original experiment, we observed a 61pp increase in truth-telling (from 14% to 75%) and a 38pp decrease in equality choices (from 47% to 9%) for HIGH draws (relative to LOW draws). We expect smaller effect sizes in our online sample. Within each treatment arm (restricted vs full), our intended sample size (800 participants) would give us approximately 363 participants with LOW and HIGH draws, respectively (taking into account that we do not use participants with a draw of 5). This would allow us to detect effect sizes of approximately 12pp with a power of 0.9 and 10pp with a power of 0.8, respectively (e.g., “power twoproportions 0.4, n(726) power(0.9)” in Stata). Smaller effect sizes are detectable for proportions that lie closer to 0 or 1.

2. Spillovers (E2):

For the test that pools the data within each action space treatment arm, our power would be identical to the previous test.

3. Action Space (E3):

Here, we rely on a Difference-in-Difference test examining whether the difference in e.g., truth-telling rates between HIGH and LOW draws varies depending on the action space. Assuming an effect size of 15pp (HIGH 0.575 vs LOW 0.425) in the restricted action space treatments, we would have a power of approximately 0.80 to detect whether this effect size vanishes in the full action space treatments (with a sample size of 800 in each treatment arm). Hence, with the proposed sample size, we could reliably detect whether a motive selection effect of around 15pp vanishes completely after introducing a full action space.

4. Game Type (E4)

For the test pooling the data within each game type arm (both types of action spaces), the power would be the same as in part E3 above. If we look separately at each type of

action space, we would need an effect size of around 20pp to achieve a similar kind of power.