

Barron, Kai; Nurminen, Tuomas

**Article — Accepted Manuscript (Postprint)**

## Nudging cooperation in public goods provision

Journal of Behavioral and Experimental Economics

**Provided in Cooperation with:**

WZB Berlin Social Science Center

*Suggested Citation:* Barron, Kai; Nurminen, Tuomas (2020) : Nudging cooperation in public goods provision, Journal of Behavioral and Experimental Economics, ISSN 2214-8043, Elsevier, Amsterdam, Vol. 88, pp. --,  
<https://doi.org/10.1016/j.socec.2020.101542>

This Version is available at:

<https://hdl.handle.net/10419/216878>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

# Nudging cooperation in public goods provision<sup>\*</sup>

Kai Barron,<sup>†</sup> Tuomas Nurminen<sup>‡</sup>

May 5, 2020

## Abstract

This paper experimentally studies two simple interventions that an authority figure might employ to promote cooperation in a public goods game when accurate feedback about contributions is not available. The first intervention aims to nudge participants to higher contribution levels by labeling contributions above a particular threshold as being “good”. Such a “norm-nudge” is intended to provide subjects with a clear, valenced focal point upon which they can coordinate. The second intervention aims to exploit lying aversion to induce higher contributions by requiring subjects to announce how much they contributed. We find that the nudge leads to a striking increase in the cooperation rate. By contrast, the ex post announcement mechanism does not have a significant effect on the cooperation rate. We present suggestive evidence that the nudge we use provides subjects with a focal point, helping conditional cooperators to coordinate their contributions.

**JEL Codes:** C91, C72, H41, Z13

**Keywords:** Cooperation, Nudge, Focal point, Public good, Lying, Experiment.

---

<sup>\*</sup>The authors would like to thank Philipp Albert, Puja Bhattacharya, Hendrik Bruns, Steffen Huck, Heather Jacklin, Felix Kölle, Erin Krupka, Topi Miettinen, and Grischa Perino as well as conference participants at the 15th Workshop on Social Economy for Young Economists for helpful comments and suggestions. We also thank the two anonymous referees and the editor who provided us with numerous valuable suggestions. Furthermore, we are grateful to Markus Konrad for help with programming in o-Tree, and Nina Bonge for excellent support in running the experiments. Tuomas Nurminen gratefully acknowledges financial support from the Emil Aaltonen Foundation (grant 1901710), the OP Group Research Foundation (grant 201500115), the Society of Swedish Literature in Finland (grant 1581), and the Yrjö Jahnsson Foundation (grants 20146643, 20166894 and 20177015).

<sup>†</sup>WZB Berlin. E-mail: kai.barron@wzb.eu.

<sup>‡</sup>Hanken School of Economics, Helsinki. E-mail: tuomas.nurminen@hanken.fi.

# 1 Introduction

A key task of a leader is to promote cooperation amongst those they lead. Achieving cooperation is particularly challenging when there is a tension between an individual’s private interests and the interests of the group (as in a social dilemma), and when it is simultaneously costly or impossible to monitor the actions of individuals. One can think of numerous scenarios that have these characteristics, including recycling, vaccinations, social distancing during epidemics, fishing, littering, employees’ effort provision, and pollution by companies and countries. While there exists a substantial literature that has sought to study which interventions can be used to sustain cooperation in social dilemmas (see, e.g., Ledyard (1995) and Chaudhuri (2011)), many of these interventions – such as sanctions – are ruled out when individuals’ actions are unobservable.

In this paper, we study two potential soft interventions that an authority figure could employ to encourage cooperation in this type of context. We do this by conducting an experimental study using the canonical linear public goods game paradigm. Subjects play the game once and we implement low observability by providing no feedback about contributions.

The experiment contains two treatment dimensions. The first treatment dimension, NUDGE,<sup>1</sup> builds on the observation that many individuals behave as conditional cooperators in the public goods game, contributing if (they believe) others do so (see, e.g., Fischbacher, Gächter, and Fehr (2001); Fischbacher and Gächter (2010)). This implies that for conditional cooperators the simultaneous-move public goods game presents a coordination problem. The treatment aims to exploit this coordination game structure by sending a public message to participants stating the minimum contribution level that the authority figure (here, the experimenter) views as “good”. Importantly, this provides subjects with a focal contribution level that unambiguously divides the decision space in two—the socially desirable contribution levels above the focal point, and the socially undesirable below. The primary objective of this nudge is to provide subjects with a clear focal point upon which they can coordinate (Schelling, 1960; Mehta et al., 1994; Brandts and MacLeod, 1995). However, due to the valenced nature of the signal, it may induce cooperation through other channels, such as subjects responding directly to the perceived expectations of the authority figure. Our central aim is to evaluate the joint treatment effect, but we also discuss and provide some evidence pertaining to potential mechanisms that might underlie any treatment effect we observe. Since the objective of the nudge is to increase cooperation, we chose the threshold for a “good” contribution level to be at 80% of the individual’s endowment (expecting this to be higher than the average contribution level in the control treatment, BASELINE).

The second treatment dimension, ANNOUNCE, investigates whether lying aversion might be leveraged to induce higher cooperation levels by requiring individuals to announce ex post how

---

<sup>1</sup>“Nudging” refers to the act of designing the non-pecuniary elements of the choice environment so as to increase the likelihood of desired behavior. A seminal reference is Thaler and Sunstein (2008).

much they contributed. Anticipating this announcement, an individual who dislikes lying and also feels discomfort about truthfully announcing having taken a selfish action (e.g. due to shame or guilt in response to the disapproval of her group members), may contribute more than she would have done in the absence of the announcement mechanism.

The rationale for studying these two treatment dimensions together is the following. Firstly, they are linked by the fact that they are both candidate interventions that someone in an authority position could use to induce cooperation in certain public goods provision settings, namely those characterized by a low observability of actions, implying that material punishments and legally binding constraints of the choice set are therefore impractical (see, e.g., Ambrus and Greiner (2012) for the ineffectiveness of costly punishment for social welfare under imperfect monitoring). Our objective here, therefore, is to design the choice architecture to test whether voluntary cooperation can be induced in such a setting by drawing on lessons from the recent behavioral economics literature. Secondly, in addition to evaluating each intervention in isolation, our design allows us to study whether the combination of the two interventions yields an interaction effect. The ANNOUNCE treatment relies on subjects experiencing disutility from announcing a selfish, low contribution. When the interventions are combined, this disutility could be magnified by a nudge that emphasizes the social merits of a high contribution. We therefore also explicitly study this interaction effect in the treatment NUDGE+ANNOUNCE.

The results from our experiment provide strong support for the hypothesis that the NUDGE treatment increases contributions. Subjects contribute over 40% more than in the BASELINE treatment, on average, when provided with the clear, common knowledge nudge towards cooperating. Furthermore, in the NUDGE treatment, the median subject contributes exactly the value of the focal point. This is in stark contrast with a median contribution of below 40% of the endowment in the two treatments without the nudge. We show that this upward shift in contributions is accompanied by an upward shift in beliefs about others' contribution levels. This is consistent with the explanation that the nudge helps participants to coordinate on cooperating.

By contrast, we find that the ANNOUNCE treatment has no influence on contribution levels. However, 27% of those contributing zero to the public good still lied about their contribution level when making their anonymous announcement. This lying without any prospect of financial rewards, but rather purely for image reasons, is notable in light of the burgeoning literature that finds that a large proportion of individuals prefer to tell the truth even when it is costly to them (Fischbacher and Föllmi-Heusi, 2013; Gneezy et al., 2018; Abeler et al., 2019). One potential explanation for our finding is that for these subjects the psychological cost of announcing their selfish choice is higher than the cost of telling an anonymous lie.

The main contribution of our paper is to demonstrate that in social dilemma contexts in which observing individuals' actions is difficult and in which it is thereby challenging to induce cooperation, the provision of a nudge that is transparent and common knowledge to the individuals may

be effective in encouraging cooperation. Our results relate to several threads of the large body of literature studying interventions aimed at promoting cooperation in social dilemma games, including studies considering the efficacy of *community framing* (Lieberman et al., 2004; Rege and Telle, 2004; Dufwenberg et al., 2011; Ellingsen et al., 2012), *moral suasion* (Dal Bó and Dal Bó, 2014; Fellner et al., 2013; Ito et al., 2018), *contribution recommendations or obligations* (Marks et al., 1999; Croson and Marks, 2001; Galbiati and Vertova, 2008, 2014), and *communication* (Isaac and Walker, 1988; Bicchieri, 2002; Bochet et al., 2006; Balliet, 2010; Irlenbusch and Ter Meer, 2013, 2015). This literature, and the relationship to our paper, are discussed in more detail in the next section.

One important concern that may be raised regarding the efficacy of the nudge we employ in our experiment is that the large treatment effect we observe could be driven by experimenter demand (Zizzo, 2010; De Quidt et al., 2018). In designing the experiment, our aim was to construct a transparent nudge towards a high contribution level by an authority figure. In evaluating the severity of the concern regarding experimenter demand, it is necessary to ask whether there are fundamental differences between the transparent nudge by an authority figure that we are trying to capture, and the nudge we employ. Our assessment of this question is that, while our nudge does bear the hallmark characteristics that are often associated with experimenter demand in experimental work, many of these characteristics are also intrinsic components of a transparent nudge by an authority figure. In other words, we specifically seek to harness an intervention similar in nature to the experimenter demand effect to investigate whether nudges by an authority figure can be used to foster cooperation. In this, our paper can be viewed as contributing to the recent literature studying *transparent nudges* (Loewenstein et al., 2015; Kroese et al., 2016; Steffel et al., 2016; Bruns et al., 2018). This literature demonstrates that making a nudge transparent need not reduce the effectiveness of the nudge. In addition, there are ethical arguments in favor of transparency of nudges. In our view, the tension between private and social interest makes social dilemmas a prime candidate for applying transparent nudges, and our paper aims to provide some initial evidence of this as a proof of concept. Nevertheless, this experimenter demand concern is certainly an important one and, as with the results of all experimental work conducted in a stylized setting in the lab, caution should be exercised when extrapolating our results to other settings. Future work, particularly in field settings, would be beneficial for evaluating whether the effectiveness of this simple nudge extends to more naturalistic settings.

The rest of the paper is organized as follows. The next section explains how the paper is related to existing literature. Section 3 outlines the experimental design, while the theoretical predictions are derived in Section 4. Section 5 presents the results. Section 6 concludes.

## 2 Related literature

This paper contributes to multiple strands of literature. First, it offers an evaluation of a particular type of frame, contributing to the literature investigating the influence of framing on cooperation in social dilemma games (e.g. Andreoni (1995); Ross and Ward (1996); Cookson (2000); Liberman et al. (2004); Dufwenberg et al. (2011)). These studies typically seek to enhance the social proximity among the subjects by using a “community” framing of the game in which terms such as “cooperate” and “free-ride” are emphasised. One contribution to this *community framing effects* literature that is particularly close in spirit to our paper is the study by Ellingsen et al. (2012). Using a one-shot prisoner’s dilemma, the authors build on earlier work by Deutsch (1958) and Liberman et al. (2004) to study the influence of framing on the decision to cooperate. After demonstrating that a “community” frame does shift cooperation rates, the authors use additional treatments to provide evidence that the framing effect operates via a coordination device channel (i.e. via shifting beliefs), rather than by shifting preferences. Similar to the underlying motivation for our nudge treatment serving as a focal point in the public goods game context, Ellingsen et al. (2012) draw on the observation that a prisoner’s dilemma game in material payoffs may be represented as a common interest game in utilities when individuals hold various types of social preferences.

We view our paper as complementing this study in several ways. Firstly, we present evidence that providing a nudge that can serve as a coordination device is also effective in increasing cooperation in the more complex context of a public goods game. Secondly, the beliefs we elicit about others’ contributions are consistent with the explanation that the treatment effect operates via shifting beliefs in our study, as in theirs. Finally, a key difference here between the prisoner’s dilemma (PD) and the public goods game (PGG) is the Boolean nature of choices in the PD in comparison to the larger choice space in the PGG. This difference may be of central importance when interpreting the evidence regarding the efficacy of the “community” frame in the two games. The reason is the following. Under the frame-as-a-coordination-device story posited by Ellingsen et al. (2012), providing a “community” frame can be sufficient as a coordination device when there are only two possible actions (as in the PD). However, even if the PGG is a coordination game in utilities, providing a “community” frame may still leave substantial residual uncertainty about others’ beliefs and behaviour. Interestingly, this provides a candidate explanation for the larger effectiveness of the “community” frame observed in PDs (Liberman et al., 2004; Ellingsen et al., 2012) in comparison to PGGs, where the evidence has been mixed (Rege and Telle, 2004; Brandts and Schwiieren, 2009; Dufwenberg et al., 2011). In the current paper, the nudge we adopt reduces this residual uncertainty by providing a clear candidate focal point.

This paper also relates closely to recent field work that has tried to leverage beliefs about social norms by nudging individuals towards pro-social behavior, such as electricity usage (Allcott, 2011; Allcott and Rogers, 2014) or water conservation (Ferraro and Price, 2013; Bhanot, 2018). The cen-

tral idea is to shift expectations about what other people do and/or approve of, thereby inducing behavioral change. This process of *norm-nudging* or *social norms messaging* relies on changing the perception of what the prevailing descriptive and/or social norms are in the context (Bhanot, 2018; Bicchieri and Dimant, 2019). Bicchieri and Dimant (2019) classify behavioral norms according to the motivations underlying the behavior and arrive at four distinct types of norms: customs, moral rules, descriptive norms, and social norms. The NUDGE treatment we adopt fits neatly into their taxonomy as a nudge that targets both moral rules and social norms. In particular, it aims both at changing the subjects' descriptive expectations about what others contribute or deem appropriate, but may also shift normative beliefs about what constitutes a socially desirable contribution level.

Our NUDGE treatment also complements the small but distinct literature that examines the impact of *recommended contributions* on cooperation (Marks et al., 1999; Croson and Marks, 2001). For example, Croson and Marks (2001) find weak or no effect of recommendations on contributions in a repeated threshold public goods game. In comparison, our NUDGE treatment implicitly recommends a contribution of at least 16, and can thus be seen as a special type of recommendation. However, the treatments differ in both spirit and implementation. The game in Croson and Marks (2001) has multiple Nash equilibria, and thus recommendations may serve as a useful coordination device even if all subjects were selfish. By contrast, in the linear public goods game we conduct, the unique Nash equilibrium is the one in which all players contribute zero. The effectiveness of recommendations in our experiment in increasing contributions thus rests on the assumption that at least some of the subjects have conditional preferences and can thereby benefit from the coordination device we provide.

Building on this idea, a separate strand of this literature tests whether recommendations can be strengthened through moral appeals. Dal Bó and Dal Bó (2014) evaluate the impact of *exogenous moral messages* on cooperation in the laboratory, while Fellner et al. (2013) and Ito et al. (2018) take this idea to the field. In a repeated linear public goods game, Dal Bó and Dal Bó (2014) find that recommendations backed up by moral appeals have a positive impact on contributions over and above the impact of neutral recommendations or experimenter demand effects. The nudge we employ may also be interpreted as containing an element of moral suasion, which may serve to strengthen subjects' motivation to coordinate on the focal point provided.

In a similar vein, our paper relates to the literature considering the effect of *obligations* on law-abidance in the context of public goods games (e.g., Galbiati and Vertova (2008, 2014)). For instance, Galbiati and Vertova (2008) note that legal rules often constitute obligations that are enforced by incentives. They use a repeated linear public goods game to study whether the obligation component of legal rules alone has an independent impact on contributions. They consider a setting in which the obligation consists of a required minimum contribution to the public good. They find that when the obligation is high, the average contribution is significantly higher than when there is no obligation or when the obligation is low. In complementary work, Lane and Nosenzo



(2019) study the causal relationship between laws and norms in the general population by eliciting the social appropriateness of actions just above and just below legal thresholds. Similarly, Kölle et al. (2019) provide field evidence that laws may shift norms towards cooperation, while monetary incentives may be counterproductive due to weakening cooperation norms. Our NUDGE treatment differs from these studies in that it does not directly *require* subjects to contribute anything, thereby differing in nature from a “law”. Indeed, in many real-world situations, appealing to legal rules is impossible simply because behavior cannot in practice be easily regulated (household energy and water consumption, and social distancing during epidemics are examples of such situations). In these settings socially beneficial behavior requires voluntary cooperation that can potentially be fostered using softer means, as in our experiment.

The ANNOUNCE treatment dimension in our experiment relates closely to the large literature on communication in social dilemma games (Isaac and Walker, 1988; Miettinen and Suetens, 2008; Bochet and Putterman, 2009; Balliet, 2010; Koessler et al., 2018). The main finding of this literature is that *pre-play communication* facilitates cooperation (Bochet et al., 2006). However, the objective of our ANNOUNCE treatment is somewhat different—we shut down many of the mechanisms that may be present in pre-play communication and focus only on the potential role played by intrinsic costs, such as shame and lying aversion, in helping to sustain cooperation in environments where it is difficult to observe others’ contribution levels. In particular, by focusing on anonymous, non-verifiable *ex post communication*, we remove strategic concerns, detection concerns, social image concerns, and the possibility that communication helps subjects to understand the game.

In the domain of communication, our paper relates most closely to two recent papers by Irlenbusch and Ter Meer (2013, 2015) who experimentally study the role played by non-verifiable ex post communication in a *repeated* public goods setting. Interestingly, they find high rates of lying in combination with low contribution choices, however subjects correctly anticipate that others are lying and accordingly form low beliefs about others’ contribution choices. The main difference between their setting and ours is that, since their game is repeated, subjects have a strong strategic incentive to try to convince others that they have been making high contributions. Therefore, the mechanisms explored in the two papers are very different. Our interest lies in the impact of the *existence* and *anticipation* of the ex post announcement mechanism rather than in the impact of the information communicated by the announcements and its influence on later rounds. Indeed, when strategic motives to lie are absent, we observe markedly lower lying rates than Irlenbusch and Ter Meer (2013, 2015) do. However, given the absence of material and social motives to lie, it is striking that we observe lying at all.



## 3 Experimental design

### 3.1 The public goods game

Our experiment is embedded in a standard linear public goods game. Public goods games reflect the central characteristics of many economically important situations in which individuals face a tension between advancing their private interests and cooperating for the benefit of the group. We study possible channels for inducing cooperation in the subset of social dilemma situations in which it is difficult to observe individual contribution levels since these are precisely the situations in which it is most challenging to induce cooperation.

In the experiment, subjects are randomly assigned to groups of four. Every subject is endowed with 20 experimental points and has to decide how many of these to contribute to a group project and how many to keep for herself. Every point kept increases individual earnings by one point. The sum of all points contributed to the group project is multiplied by 1.4 and divided equally among the four group members. Thus, subject  $i$ 's monetary payoff from the game is given by:

$$\pi_i^m = 20 - g_i + 0.35 \sum_{j=1}^4 g_j, \quad (1)$$

where  $g_j$  denotes the contribution to the group project by group member  $j$ .

The marginal per capita return (MPCR) of a contribution to the group is 0.35. Since it is less than unity, the dominant strategy for a selfish subject is to contribute nothing. Thus, free-riding by all group members is the unique Nash equilibrium of the game in terms of material gain. The equilibrium payoff is 20 points for all group members. By contrast, since  $4 \cdot 0.35 > 1$ , the socially efficient outcome is that every group member contributes their full endowment, which results in individual earnings of 28 points. An individual subject's payoff is maximized if she contributes zero points to the group project but the rest of her group contributes everything, in which case the free-rider earns 41 points. In the treatment conditions below, we study whether tailoring the choice environment can induce higher cooperation levels by magnifying various intrinsic motives.

### 3.2 Treatments

We implemented four treatments: BASELINE, ANNOUNCE, NUDGE, and NUDGE + ANNOUNCE.<sup>2</sup> In each treatment, subjects played the linear public goods game described above once. After the public goods game, subjects completed a short survey that gathered demographic information, as well as information about their social preferences. Feedback about earnings was only provided at

---

<sup>2</sup>See Appendix C for instructions for treatment NUDGE + ANNOUNCE. The full set of instructions is available from the authors.

the very end of the experiment. In all treatments, subjects were required to answer a series of control questions to ensure comprehension before making their contribution decisions.

Table 1: Treatment Conditions

	No announcement	Ex post announcement
Neutral instructions	<b>BASELINE</b>	<b>ANNOUNCE</b>
Nudging instructions	<b>NUDGE</b>	<b>NUDGE + ANNOUNCE</b>

In the **BASELINE** treatment, instructions for the game were written using neutral terms. The **NUDGE** treatment was exactly the same as **BASELINE** with the exception of the following adjustments aimed at inducing a valenced focal point. At the end of the instructions for **BASELINE**, subjects were additionally told that if everyone contributes to the group project it would be beneficial for the group. A contribution of “16 or greater” would be labelled as a “good” contribution, and a contribution of “15 or below” would be labelled as a “bad” contribution. Finally, “good” contributions were written in green, while “bad” contributions were written in red. This framing serves two purposes. First, it highlights the moral dimension of the contribution decision. Second, it introduces a salient focal contribution level of 16 that unambiguously divides the decision space into socially desirable and socially undesirable contributions. The reasons for choosing 16 as the focal point were twofold. Firstly, it was important to choose a focal point that was above mean and median contribution levels in **BASELINE**, such that coordination at the focal point would increase contribution levels. Secondly, given the recent literature on nudges backfiring,<sup>3</sup> we wanted to avoid using a nudge that required too extreme a shift in behavior. By this, we mean that we wanted to limit the size of the gap between contributions in **BASELINE** and the focal point. Requiring a larger increase in contributions from **BASELINE** levels (e.g. by using 20, the social optimum, as the focal point) would certainly be interesting, but would bear the risk of backfiring due to subjects being unwilling to make such a large increase in their contribution level—for example, due to creating doubts about whether others would comply with such an extreme contribution level. As a proof of concept, we therefore opted for an intermediate focal point that satisfied both these two criteria.

In the **ANNOUNCE** treatment, subjects were told that they would be asked to announce how much they contributed after making their contribution decisions. This announcement was anonymous and non-verifiable by other subjects. Individual announcements were then revealed to the other group members on the computer screen, with anonymous player labels. Subjects were there-

<sup>3</sup>See, e.g., Damgaard and Gravert (2018); Bolton et al. (2020).

fore free to announce any integer in the interval  $[0, 20]$ . Importantly, the announcement procedure was described to the subjects in the instructions *before* the contribution decisions were made, so subjects could anticipate having to make the announcement.

The NUDGE + ANNOUNCE treatment was identical to the NUDGE treatment, with the addition of an announcement stage that was the same as the one in the ANNOUNCE treatment. The objective of this treatment was to test for the presence of an interaction effect between the two mechanisms targeted in the NUDGE and ANNOUNCE treatments.

### 3.3 Belief elicitation

In the NUDGE treatment, the usefulness of providing a focal point to shift contribution levels relies on subjects' beliefs regarding others' contribution levels being shifted towards the focal point thereby facilitating coordination by *conditional cooperators*.

In order to better understand the underlying reason for any change in contribution levels, we therefore elicited subjects' first order beliefs regarding the contribution choices of their group members. More specifically, in all treatments, after the contribution decision, we first elicited subjects' beliefs about the *average* contribution of the other three group members. In addition, in the ANNOUNCE and the NUDGE + ANNOUNCE treatments we also elicited subjects' beliefs about the *individual* contributions of each of their group members. This was done after the contribution announcements were made and anonymously revealed to the group. This second belief elicitation allows us to analyse how beliefs react to the announcements.

We incentivized both belief elicitation tasks using the commonly used quadratic scoring rule (QSR) reward function. Specifically, in the first belief elicitation the reward for subject  $i$  with a stated belief  $\bar{b}_i \in [0, 20]$  was

$$\pi_i^{B1} = 8 - 8 \left( \frac{\bar{b}_i - \bar{g}_{-i}}{20} \right)^2,$$

where  $\bar{g}_{-i}$  denotes the average contribution by the other three group members.

In the second belief elicitation, the reward for subject  $i$  with a stated belief  $b_i^j$  about group member  $j$ 's contribution was

$$\pi_{i,j}^{B2} = 8 - 8 \left( \frac{b_i^j - g_j}{20} \right)^2.$$

However, in the second belief elicitation subjects were informed that only one of the three stated beliefs would be payoff relevant. The payoff relevant belief was chosen randomly.

After the belief elicitation stage, the subjects were asked to complete a short survey on other-regarding preferences and demographics. Finally, each subject's earnings from the public goods game and the belief elicitation(s) were summed and converted to euros using an exchange rate of 4 points = 1 EUR.

### 3.4 Procedures

We conducted fourteen sessions of 24 subjects in the WZB-TU experimental laboratory in Berlin. In each session, we implemented one treatment condition. A total of 336 subjects, predominantly students in universities in Berlin, participated in the experiment. There were 96 participants in each of the two treatments with an announcement stage, and 72 participants in the other two treatments with no announcement stage. The rationale for having an additional session for each of the announcement treatments was to facilitate studying lying behavior, which could only be observed in these two ANNOUNCE treatments, and has a relatively low baseline level. Participants were solicited through an online database using ORSEE (Greiner, 2015), and the experiment was run using the experiment software oTree (Chen et al., 2016). In each session, subjects participated in two experiments. First, they participated in an investment game that is completely unrelated to the current paper, and second they played the one-shot public goods game studied here.<sup>4</sup> Sessions lasted up to 90 minutes in total, and participants earned, on average, 6 EUR for the public goods game, 1.8 EUR for the first belief elicitation, and 1.8 EUR for the second belief elicitation. In addition, they received a 5 EUR showup fee and their earnings from the investment game.

## 4 Hypotheses

In this section we outline our main hypotheses for each of the treatment conditions. To demonstrate how we think the treatment manipulations might translate into contribution decisions, we sketch a simple model of individual decision making. Throughout, we assume that an individual is risk neutral and cares about her monetary earnings,  $\pi_i^m$ , which is given by equation 1. The discussion below considers how additional motives, such as social preferences and lying aversion, could influence behavior.

### 4.1 Nudging

Building on the model of norm-dependent utility outlined by Kessler and Leider (2012), we assume that in addition to her monetary earnings, an individual may experience a disutility  $\lambda_i(g_i - N_i)^2$ , if her contribution,  $g_i$ , differs from a reference contribution,  $N_i$ . The parameter  $\lambda_i \geq 0$  represents the weight the individual puts on this intrinsic component of her utility relative to her monetary earnings. Since the literature on player types in public goods games documents the existence of

---

<sup>4</sup>The treatments of the two experiments were chosen independently of one another, and it was explicitly and clearly stated to subjects that the two games were completely independent of one another. The first game did not study social preferences nor social norms, and feedback on earnings from it was only disclosed at the end of the entire session. To check whether subjects' expectations of their earnings from the first game influenced their contribution choices, we regressed contribution choices on the investment game earnings. The relationship between the two variables was not statistically significant.

both *unconditional* and *conditional* cooperators, we want our model to accommodate both types (Fischbacher et al., 2001; Kocher et al., 2008; Fischbacher and Gächter, 2010; Thöni and Volk, 2018). In order to do that, we allow the personal reference contribution level,  $N_i$ , to take one of two forms, depending on the individual's preferences.

For the first type of individual, the *unconditional cooperator*,  $N_i$  reflects the individual's personal view of what is the morally appropriate or right contribution level (i.e. what Bicchieri and Dimant (2019) call a *moral norm*). For tractability, we denote  $N_i \equiv g_i^N$  in this case, where  $g_i^N$  reflects the individual's internal reference point for appropriateness. We assume that an individual maximizes the expectation (with respect to the contributions by her group members) of the utility function:

$$U_i(g_i, g_{-i}; g_i^N) = \pi_i^m(g_i, g_{-i}) - \lambda_i(g_i - g_i^N)^2, \quad (2)$$

in which  $(g_i, g_{-i})$  is the vector of contributions by  $i$ 's group. In this case the reference contribution level is fixed and known to the individual at the time of making the contribution choice. Such preferences are unconditional on the behavior of others, and would manifest as *unconditional cooperation* in the linear public goods game, provided that both  $g_i^N$  and  $\lambda_i$  are sufficiently high.

For the second type of individual, the *conditional cooperator*,  $N_i$  reflects the average contribution by the individual's group members:  $N_i \equiv \bar{g}_{-i} \equiv (\sum_{j \neq i} g_j)/3$ . In this case the reference contribution level is not known to the individual at the time of making the contribution choice which means she has to act on the basis of her expectation of  $\bar{g}_{-i}$ . She would thus maximize the expectation of the utility function:

$$U_i(g_i, g_{-i}) = \pi_i^m(g_i, g_{-i}) - \lambda_i(g_i - \bar{g}_{-i})^2. \quad (3)$$

An individual of this type has conditional preferences, and would behave as a *conditional cooperator* in the linear public goods game. The desire to comply with the perceived social norm is an example of a motive that could generate preferences of this form, in which case  $\bar{g}_{-i}$  represents a social norm (either a *descriptive norm* or a *social norm* in the terminology of Bicchieri and Dimant (2019)). Alternatively, one can think of the utility function as a reduced form representation of inequality averse preferences (Fehr and Schmidt, 1999), whereby the desire for an equal earnings distribution induces a desire for matching contributions.

In the instructions for the NUDGE treatment, contributions of "16 or greater" were labelled as "good" and contributions of "15 or below" as "bad". This nudge is designed to operate in two possible ways. Firstly, appealing to the moral nature of the contribution decision may affect subjects' preferences, independent of what they think others will do. Specifically, we posit that moral appeals could increase the contribution level that subjects consider appropriate, increasing  $g_i^N$  in the utility function given by equation 2. Secondly, the nudge could affect subjects' beliefs about others' contributions. Given the low average beliefs observed in the previous literature, we expect

beliefs to shift upwards in our NUDGE treatment, and at the same time largely cluster at the focal point of 16.<sup>5</sup> In terms of our model, we thus expect the expectation of  $\bar{g}_{-i}$  to increase and to be 16 for many individuals in the treatments with the nudge compared to the non-nudge treatments. This leads to an increase in the average contribution as individuals with conditional preferences as in equation 3 increase their contribution. The following summarizes our hypotheses regarding the treatment effect on contribution levels in the NUDGE treatment.

**Hypothesis 1.** (i) *The average contribution will be higher in the NUDGE treatment relative to the BASELINE treatment.* (ii) *The fraction of contributions above sixteen will be higher and the fraction of intermediate level contributions (i.e. between 1 and 15) will be lower in the NUDGE treatment relative to the BASELINE treatment.*

## 4.2 Ex post announcements

In the ANNOUNCE treatment, we add anonymous non-verifiable ex post announcements to test whether an aversion to lying can help sustain cooperation. The basic intuition for this treatment is that when we introduce the ex post announcements, individuals may face two additional psychological motives relative to the BASELINE treatment, namely (i) a cost of lying, and (ii) a psychological cost (e.g. shame or guilt) to announcing a selfish low contribution level.<sup>6</sup> The anticipation of these two costs may influence behavior. If the two psychological costs are sufficiently high to dominate the monetary benefits of making a low contribution, the individual may increase her contribution level. We summarize our hypothesis as follows:

**Hypothesis 2.** *The average contribution in the ANNOUNCE treatment will be higher than the average contribution in the BASELINE treatment.*

---

<sup>5</sup>Several studies have elicited subjects' beliefs concerning the average or the total contribution by their group members using either an incentivized (e.g., Croson (2007); Neugebauer et al. (2009); Fischbacher and Gächter (2010); Dufwenberg et al. (2011)) or an unincentivized (e.g., Cubitt et al. (2011)) elicitation mechanism. These studies typically find that beliefs about the contributions (in the case of a one-shot game) or about the first round contributions (in the case of a repeated game) are 40–50% of the maximum possible contribution. For example, in a design comparable to ours, Dufwenberg et al. (2011) found that subjects' average first-order belief in the Give-Neutral treatment was 40% of the maximum. Furthermore, most of the probability mass in the belief distribution is located below the level of 80%. Since the contribution level of 16 corresponds to 80% of the endowment in our experiment, we expect the average belief in the BASELINE treatment to be below 16. This implies that if the NUDGE treatment is effective in shifting beliefs to 16, this shift is an upward shift in beliefs compared to the BASELINE treatment.

<sup>6</sup>There are several micromotives for why this psychological cost might arise. First, one can view it as reflecting an aversion to revisiting a decision that one is ashamed of. This is akin to the cost one would experience if one refused to give money to a beggar and was afterwards reminded of the refusal. Second, one can think of the psychological cost as a reduced form of 'guilt-from-blame' (Battigalli and Dufwenberg, 2007), whereby one dislikes being blamed for having let someone else's payoff expectations down. In our setting, an individual announcing a low contribution level might think she is being blamed by her group members who are disappointed by her (perceived) low contribution.

### 4.3 Interaction effects

Lastly, we test for an interaction effect between the two mechanisms under consideration. We posit that there are two mutually exclusive possibilities. On the one hand, if the NUDGE treatment is effective at increasing the intrinsic incentives to choose a contribution level above 16, then by the same psychological rationale, it may also increase the psychological cost of (truthfully) announcing a contribution level below 16. This implies that the effect of introducing the announcement mechanism will be more pronounced when combined with the moral nudge than without it. This logic leads to Hypothesis 3a.

**Hypothesis 3a.** *The increase in average contributions in the NUDGE + ANNOUNCE treatment relative to the NUDGE treatment will be larger than the increase in average contributions in the ANNOUNCE treatment relative to the BASELINE treatment.*

On the other hand, if the NUDGE treatment by itself is highly effective, there may not be much scope for further improvement in the NUDGE+ANNOUNCE treatment. This implies that the effect of introducing the announcement mechanism will be more pronounced without the moral nudge than with it, generating Hypothesis 3b. We summarize this idea in the following hypothesis that is mutually exclusive with respect to 3a.

**Hypothesis 3b.** *The increase in average contributions in the ANNOUNCE treatment relative to the BASELINE treatment will be larger than the increase in average contributions in the NUDGE + ANNOUNCE treatment relative to the NUDGE treatment.*

Therefore, in the scenario where the NUDGE treatment alone leads to a small improvement in cooperation relative to the BASELINE, Hypothesis 3a should be viewed as the relevant one. However, in the scenario where the NUDGE treatment alone has a large effect relative to the BASELINE, Hypothesis 3b is the relevant one. Essentially, if the nudge alone succeeds in achieving high contribution levels, there remains no scope for a strong interaction effect.

Table 2: General descriptive statistics

	<i>N</i>	Average contribution	Average ex ante belief	Percentage lying	Average lie among liars	Average ex post belief
BASELINE	72	41.0 (36.0)	48.6 (25.2)			
ANNOUNCE	96	39.6 (38.8)	47.8 (27.4)	12.5	57.5 (31.9)	41.6 (24.1)
NUDGE	72	58.0 (40.7)	63.8 (26.6)			
NUDGE + ANNOUNCE	96	60.0 (39.0)	67.7 (24.6)	13.5	50 (41.6)	61.4 (22.4)

Notes: (i) Standard deviations are shown in brackets.

(ii) Contributions, beliefs and lie magnitudes are expressed as a percentage of the maximum.

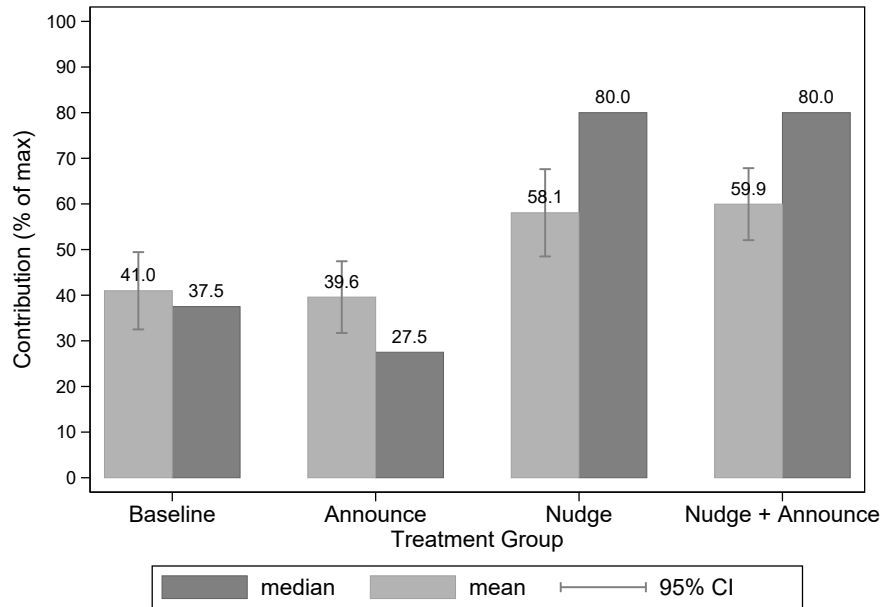


## 5 Results

### 5.1 Contributions

Table 2 provides a descriptive overview of the data from the four treatments. Figure 1 displays the mean and median contribution levels in each treatment, denoted as a percentage of the maximum possible contribution. Consistent with the prior literature, in the BASELINE treatment we observe the majority of subjects making positive contributions, with an average contribution level of 41%. The results show that providing the nudge has a large and significant influence on contribution levels, increasing the average contribution by over 40% [17 percentage points] to a contribution level of 58% in the NUDGE treatment (a one-sided Wilcoxon rank-sum test,  $p < 0.01$ ). It is also worth noting that the median contribution doubles from 40% to 80% to lie exactly at the focal contribution level of 16. In Table 3, we report the results from a set of OLS regression models with the contribution choice as the dependent variable. In Model 1 the independent variables are dummy variables for each of the treatment dimensions, as well as their interaction. The coefficient for the NUDGE treatment dimension is statistically significant, providing evidence in support of Hypothesis 1(i) (a one-sided Wald test,  $p < 0.01$ ).

Figure 1: Contribution levels across treatments (mean and median)



In contrast, the ex post announcements appear not to have shifted average contribution levels. There is no increase in mean contributions in the ANNOUNCE treatment relative to the BASELINE treatment (Wilcoxon rank-sum test, one-sided,  $p = 0.63$ ). In Model 1 in Table 3, the coefficient on the ANNOUNCE treatment dimension is not statistically significant (a one-sided Wald test,  $p =$

0.59). We therefore observe no evidence in support of Hypothesis 2. One explanation for this could be that since the announcements were anonymous, the intrinsic guilt or shame of announcing a low contribution level was not large enough to outweigh the material costs of choosing a higher contribution level. We return to this question in Section 5.3 where we study the announcements more closely.

Table 3: Treatment Effect on Contributions

	Model 1	Model 2	Model 3	Model 4	Model 5
NUDGE	17.08*** (6.45)	17.60*** (6.45)		1.60 (5.11)	2.62 (5.25)
ANNOUNCE	-1.39 (5.84)	1.69 (5.82)		-0.52 (4.35)	1.38 (4.46)
NUDGE * ANNOUNCE	3.28 (8.57)	2.10 (8.60)		-1.64 (6.28)	-2.26 (6.41)
EX ANTE BELIEF			1.03*** (0.053)	1.02*** (0.064)	0.99*** (0.066)
CONSTANT	40.97*** (4.28)	47.56*** (13.33)	-8.9** (3.16)	-8.74** (3.97)	0.91 (10.99)
CONTROLS	No	Yes	No	No	Yes
Observations	336	329	336	336	329
Adjusted $R^2$	0.049	0.097	0.500	0.496	0.514

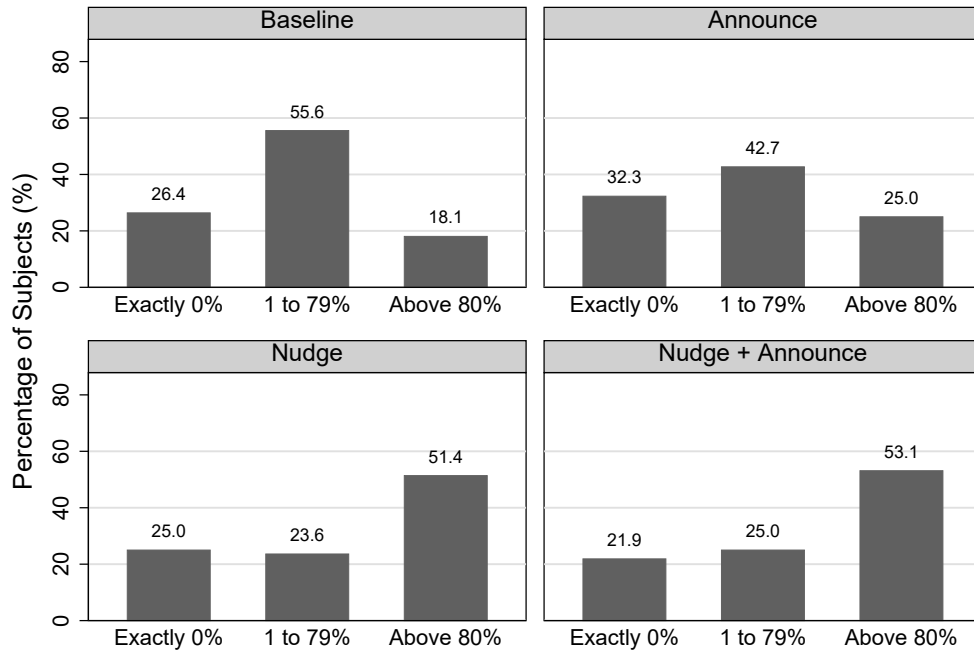
Notes: The table shows OLS regressions. Dependent variable is individual contribution. Independent variables are a dummy variable for the nudge treatment, a dummy for the announce treatment, a dummy for the interaction effect, and the elicited ex ante belief. In Models 2 and 5 we also include as control variables the individual's gender, age, squared age, a measure of English proficiency, and the answer to question "Imagine the following situation: Today you unexpectedly received 1,600 Euro. How much of this amount would you donate to a good cause?" Heteroskedasticity-robust standard errors are in parentheses (\*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

Turning to the interaction effect between the two treatment dimensions, we observe that when the announcement mechanism is in place, nudging cooperation still has a large positive impact on contributions: the average contribution in the NUDGE + ANNOUNCE treatment is 60% of the maximum, a statistically significant increase from the ANNOUNCE treatment (a Wilcoxon rank-sum test, one-sided,  $p < 0.01$ , a Wald test, one-sided  $p < 0.01$ ). Moreover, the median contribution is again exactly at the focal contribution level of 16. Comparing the NUDGE + ANNOUNCE treatment to the NUDGE treatment, however, reveals no significant difference (a Wilcoxon rank-sum test,

two-sided,  $p = 0.97$ , a Wald test, two-sided,  $p = 0.76$ ). As the announcement mechanism has no effect on contributions whatsoever independent of whether the instructions are neutral or nudging cooperation, we observe no support for either Hypothesis 3a. or 3b.

To check the robustness of these results, we do three things. First, we extend Model 1 by adding a set of control variables. The estimates for the resulting model are displayed in Model 2 of Table 3. These estimates are very similar to those in Model 1. Second, we drop from our sample all subjects who provided more than one incorrect answer to the control questions measuring the understanding of the instructions (29 subjects). This serves to address the concern that our NUDGE treatment operates by enhancing the subjects' understanding of the game (e.g. by highlighting the tension between socially optimal and individually rational behavior). The results indicate that this is not the case (see Table 5 in Appendix A)—rather, the coefficients are very similar to those in Model 1. As the third robustness check, we estimate a set of Tobit models corresponding to the models in Table 3 (see Table 6 in Appendix A). The estimates represent marginal effects on the latent (desired) contribution and are thus quantitatively different from the corresponding OLS estimates, but qualitatively the results are highly consistent with our OLS estimates.

Figure 2: Distribution of contribution levels



Having examined the contributions at the average level, we next look at the distribution of contributions by treatment.<sup>7</sup> In Hypothesis 1(ii) we posited that the nudge would impact predom-

<sup>7</sup>In Figure 5 in Appendix B we provide the full distribution of contributions in each treatment. A Kruskal-Wallis test confirms that the (population) contribution distributions are significantly different across treatments ( $p < 0.01$ ).

inantly individuals who receive some intrinsic utility from contributing, and therefore would have been likely to contribute more than zero in the BASELINE treatment. This implies that the treatment effect would operate primarily by shifting individuals who would have been contributing between 1 and 15 (i.e. 5% to 75%) in the absence of the nudge to a contribution level of at least 16 (i.e. 80%). Since we employ a between-subjects design, we do not observe subjects choosing both with and without the nudge and thus must rely on a between-subject comparison. Figure 2 shows that the data is consistent with the idea. To examine this hypothesis more formally, we pool the data from: (i) the two treatments without the nudge, as well as (ii) the two treatments with the nudge. The proportion of individuals contributing zero is relatively stable across the two pooled samples, averaging 23.1% in the two treatments with the nudge, and 29.8% in the two treatments without the nudge (Fisher’s exact test,  $p = 0.22$ ).<sup>8</sup> This finding is consistent with our explanation that the NUDGE treatment has little impact on individuals who are not motivated by other-regarding concerns. By contrast, the proportion contributing above the focal point (i.e. at least 80%) more than doubles as a consequence of the nudge, averaging 52.4% in the two treatments with the nudge and 22.0% in the two treatments without it (Fisher’s exact test,  $p < 0.01$ ). This implies that the proportion contributing between 5% and 75% is significantly lower in the two treatments with the nudge than in the two treatments without it (Fisher’s exact test,  $p < 0.01$ ). The findings thus confirm Hypothesis 1(ii).

Rege and Telle (2004) consider a similar analysis of the heterogeneity of their treatment effect. Like us, they observe a fairly stable proportion of non-contributors, and that their social approval treatment effect operates predominantly by shifting intermediate contributors to high contribution levels. While their social approval treatment is very different from our NUDGE treatment, the parallels between the results in terms of how treatment influences individuals at different parts of the distribution are interesting. One explanation is that the intermediate contributors are *conditional cooperators* who receive intrinsic utility from cooperating and who simply hold low beliefs about the contribution levels of others. In this case, both the social approval treatment of Rege and Telle (2004) and our NUDGE treatment shift beliefs about others’ contribution levels, facilitating coordination.

## 5.2 Beliefs

By now a robust empirical finding in the literature on public goods games is that many subjects are *conditional cooperators*, making their contribution choices as an increasing function of (their beliefs about) others’ contributions (Keser and Van Winden, 2000; Fischbacher et al., 2001; Frey and Meier, 2004; Kocher et al., 2008; Fischbacher and Gächter, 2010; Thöni and Volk, 2018). Such

---

<sup>8</sup>The proportion of selfish money-maximizers in our experiment is in line with previous findings. For example, Fischbacher et al. (2001) observe that around 30% of their subjects were (unconditional) free riders. Similarly, Andreoni and Miller (2002) find that about 25% of their subjects were selfish money-maximizers.

behavior can be rationalized (with suitable beliefs) through several types of social preferences that differ in their underlying motivations. Examples include preferences that reflect *inequity aversion* (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000), *intentions-based reciprocity* (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004; Falk and Fischbacher, 2006), or *norm-dependence* (e.g., López-Pérez (2008); see also the discussion in Bicchieri and Dimant (2019)). For an individual with such preferences, beliefs about others' actions are an important determinant of contributions in the simultaneous-move public goods game. Moreover, depending on the distribution of preferences among the players, there may be multiple equilibria even in the standard linear public goods game. If this is the case, providing the subjects with a focal point would allow them to coordinate on it by shifting both their beliefs about others' contributions, and their own contributions to the focal point. In this section, we therefore study the role played by beliefs about others' contribution choices, in order to assess whether behavior is consistent with models of conditional preferences and to gain insight into what is generating the higher contributions in the NUDGE and the NUDGE+ANNOUNCE treatments.

Model 3 in Table 3 regresses contribution choices on ex ante beliefs. The results show a nearly one-to-one association between contributions and ex ante beliefs. Moreover, ex ante beliefs explain a large share of the variation in contribution choices (adjusted  $R^2 = 0.50$ ). After including dummies for the treatment dimensions and their interaction in Model 4, the estimate for ex ante beliefs remains practically unchanged and continues to be highly statistically significant, whereas estimates for the treatment dummy variables are no longer significant. This implies that nudging is not associated with contributions when ex ante beliefs are controlled for. These findings are robust to adding the usual set of covariates (Model 5).

In Table 7 in the Appendices, we assess the treatment effect on beliefs by reporting the results from regressing ex ante beliefs on the treatment dummy variables and their interaction. Model 16 shows that the nudge increased average ex ante beliefs by 15.14 percentage points compared to the BASELINE (a two-sided Wald test,  $p < 0.01$ ) and 19.95 percentage point compared to the ANNOUNCE treatment (a two-sided Wald test,  $p < 0.01$ ). Again, these estimates are robust to adding the set of covariates (Model 17). Although our design does not allow us to identify the causal mechanism underlying the effect of the nudge, these findings provide suggestive evidence that is consistent with the hypothesis that the treatment effect is mediated by shifting beliefs, which then induces higher contributions by the *conditional cooperators*.

### 5.3 Announcements

In our hypotheses section above, we discuss how individuals who anticipate a cost of announcing a selfish, low contribution have two avenues available to them to avoid this cost. They can either choose a high contribution level, and thereby avoid making a low announcement. Or, they can

make a low contribution and simply lie about it. Above, we showed that the ANNOUNCE treatment had no influence on contribution levels, suggesting that one (or both) of the two costs (i.e. lying aversion and shame or guilt) was not sufficiently high. One explanation for this is simply that the minimal intervention that we opted for, in which the announcements were anonymous and non-verifiable, did not induce sufficiently strong image costs. A stronger announcement intervention in which subjects' identities are linked to their announcements, thereby inducing social image concerns, could yield different results. However, it is interesting to note that despite the subtlety of our ANNOUNCE intervention, we find that a non-negligible proportion of individuals were willing to lie to avoid making a low announcement. Thirteen percent of all subjects lie in their announcement. However, Figure 3 in Appendix A.5 shows that the lying rate is strongly associated with contribution levels, with 22.6% in ANNOUNCE and 33.3% in NUDGE + ANNOUNCE of the subjects who contributed zero choosing to lie. Since these individuals have no monetary, strategic nor social image incentives for lying, this finding supports the assertion that some subjects lie in order to avoid a cost of announcing a low contribution. Appendix A.5 provides a more detailed discussion of this result.

## 6 Conclusion

Our experiment investigates the efficacy of two soft interventions an authority figure might use to employ lessons from behavioral economics to try to induce cooperation in a low-information public goods game. The results demonstrate that a valenced nudge can be effective in inducing a large increase in cooperation in such a context, while the announcement intervention we consider proved ineffective. The evidence presented above is consistent with the hypothesis that the nudge provides a focal point for coordination. However, our experiment is not designed to (and is unable to) rule out alternative mechanisms driving the treatment effect. We view our results as providing a proof of concept that such a nudge can be effective. Therefore, when interpreting the results, there are several caveats and considerations that are important to bear in mind.

First, our nudge can be viewed as a bundled treatment. It provides subjects with both a focal point, and a moral frame (making salient the tension between private and public interest). Furthermore, this nudge is transparently delivered by an authority figure (here, the experimenter). Our experiment evaluates this bundled intervention as a whole, and cannot isolate the effect of the individual components. This points towards an important role for future research. It seems desirable and of substantial value to learn both: (i) whether the type of simple nudge we adopt can be effective in a field setting, and (ii) to pin down the relative importance of each of the different components of the nudge.

Second, when interpreting the absence of a treatment effect of the announcement intervention, it is important to keep in the mind that the announcements were anonymous. Previous work has

demonstrated that *social-image* concerns can be a strong driver of cooperative behavior when actions are observable and can be linked to subjects' identities (see, e.g., Rege and Telle (2004)). It may be the case that a stronger announcement intervention that exploits *social-image* concerns by linking an individual's identity to her announcement would be effective in inducing cooperation even in settings with unobservable actions and non-verifiable announcements. This, too, appears a promising avenue for further work.

While these caveats indicate that caution should be exercised when drawing inference from the results presented here, situations requiring voluntary cooperation under low observability abound in the real world. It is therefore clearly of huge importance to find solutions that policy makers can employ in addressing these challenges. The two interventions we consider are both soft, low-cost, and transparent interventions. Providing even small increases to cooperation rates in some of these settings could yield considerable improvements to social welfare. We would argue that more work in this area is therefore warranted—evaluating bundled interventions to find the most effective ones, and isolating the role played by each of the underlying mechanisms.



## References

- Abeler, J., D. Nosenzo, and C. Raymond (2019). Preferences for truth-telling. *Econometrica* 87(4), 1115–1153.
- Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics* 95(9-10), 1082–1095.
- Allcott, H. and T. Rogers (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review* 104(10), 3003–37.
- Ambrus, A. and B. Greiner (2012). Imperfect public monitoring with costly punishment: An experimental study. *American Economic Review* 102(7), 3317–3332.
- Andreoni, J. (1995). Warm-glow versus cold-prickle: The effects of positive and negative framing on cooperation in experiments. *Quarterly Journal of Economics* 110(1), 1–21.
- Andreoni, J. and J. Miller (2002). Giving according to GARP: An experimental test of the consistency of preferences for altruism. *Econometrica* 70(2), 737–753.
- Balliet, D. (2010). Communication and cooperation in social dilemmas: A meta-analytic review. *Journal of Conflict Resolution* 54(1), 39–57.
- Battigalli, P. and M. Dufwenberg (2007). Guilt in games. *American Economic Review* 97(2), 170–176.
- Bhanot, S. P. (2018). Isolating the effect of injunctive norms on conservation behavior: New evidence from a field experiment in California. *Organizational Behavior and Human Decision Processes* (forthcoming).
- Bicchieri, C. (2002). Covenants without swords: Group identity, norms, and communication in social dilemmas. *Rationality and Society* 14(2), 192–228.
- Bicchieri, C. and E. Dimant (2019). Nudging with care: The risks and benefits of social information. *Public Choice* (forthcoming).
- Bochet, O., T. Page, and L. Putterman (2006). Communication and punishment in voluntary contribution experiments. *Journal of Economic Behavior & Organization* 60(1), 11–26.
- Bochet, O. and L. Putterman (2009). Not just babble: Opening the black box of communication in a voluntary contribution experiment. *European Economic Review* 53(3), 309–326.

- Bolton, G., E. Dimant, and U. Schmidt (2020). When a nudge backfires: Combining (im)plausible deniability with social and economic incentives to promote pro-social behavioral change. *Working Paper*.
- Bolton, G. E. and A. Ockenfels (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review*, 166–193.
- Brandts, J. and W. B. MacLeod (1995). Equilibrium selection in experimental games with recommended play. *Games and Economic Behavior* 11(1), 36–63.
- Brandts, J. and C. Schwieler (2009). Frames and economic behavior. *Working Paper*.
- Bruns, H., E. Kantorowicz-Reznichenko, K. Klement, M. L. Jonsson, and B. Rahali (2018). Can nudges be transparent and yet effective? *Journal of Economic Psychology* 65, 41–59.
- Chaudhuri, A. (2011). Sustaining cooperation in laboratory public goods experiments: A selective survey of the literature. *Experimental Economics* 14(1), 47–83.
- Chen, D. L., M. Schonger, and C. Wickens (2016). oTree - An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance* 9, 88–97.
- Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics* 3(1), 55–79.
- Croson, R. and M. Marks (2001). The effect of recommended contributions in the voluntary provision of public goods. *Economic Inquiry* 39(2), 238–249.
- Croson, R. T. (2007). Theories of commitment, altruism and reciprocity: Evidence from linear public goods games. *Economic Inquiry* 45(2), 199–216.
- Cubitt, R. P., M. Drouvelis, and S. Gächter (2011). Framing and free riding: emotional responses and punishment in social dilemma games. *Experimental Economics* 14(2), 254–272.
- Dal Bó, E. and P. Dal Bó (2014). Do the right thing: The effects of moral suasion on cooperation. *Journal of Public Economics* 117, 28–38.
- Damgaard, M. T. and C. Gravert (2018). The hidden costs of nudging: Experimental evidence from reminders in fundraising. *Journal of Public Economics* 157, 15–26.
- De Quidt, J., J. Haushofer, and C. Roth (2018). Measuring and bounding experimenter demand. *American Economic Review* 108(11), 3266–3302.
- Deutsch, M. (1958). Trust and suspicion. *Journal of Conflict Resolution* 2(4), 265–279.

- Dufwenberg, M., S. Gächter, and H. Hennig-Schmidt (2011). The framing of games and the psychology of play. *Games and Economic Behavior* 73(2), 459–478.
- Dufwenberg, M. and G. Kirchsteiger (2004). A theory of sequential reciprocity. *Games and Economic Behavior* 47(2), 268–298.
- Ellingsen, T., M. Johannesson, J. Mollerstrom, and S. Munkhammar (2012). Social framing effects: Preferences or beliefs? *Games and Economic Behavior* 76(1), 117–130.
- Falk, A. and U. Fischbacher (2006). A theory of reciprocity. *Games and Economic Behavior* 54(2), 293–315.
- Fehr, E. and K. M. Schmidt (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, 817–868.
- Fellner, G., R. Sausgruber, and C. Traxler (2013). Testing enforcement strategies in the field: Threat, moral appeal and social information. *Journal of the European Economic Association* 11(3), 634–660.
- Ferraro, P. J. and M. K. Price (2013). Using nonpecuniary strategies to influence behavior: Evidence from a large-scale field experiment. *Review of Economics and Statistics* 95(1), 64–73.
- Fischbacher, U. and F. Föllmi-Heusi (2013). Lies in disguise - An experimental study on cheating. *Journal of the European Economic Association* 11(3), 525–547.
- Fischbacher, U. and S. Gächter (2010). Social preferences, beliefs, and the dynamics of free riding in public goods experiments. *American Economic Review* 100(1), 541–556.
- Fischbacher, U., S. Gächter, and E. Fehr (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economics Letters* 71(3), 397–404.
- Frey, B. S. and S. Meier (2004). Social comparisons and pro-social behavior: Testing “conditional cooperation” in a field experiment. *American Economic Review* 94(5), 1717–1722.
- Galbiati, R. and P. Vertova (2008). Obligations and cooperative behaviour in public good games. *Games and Economic Behavior* 64(1), 146–170.
- Galbiati, R. and P. Vertova (2014). How laws affect behavior: Obligations, incentives and cooperative behavior. *International Review of Law and Economics* 38, 48–57.
- Gneezy, U., A. Kajackaite, and J. Sobel (2018). Lying aversion and the size of the lie. *American Economic Review* 108(2), 419–53.

- Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association* 1(1), 114–125.
- Irlenbusch, B. and J. Ter Meer (2013). Fooling the nice guys: Explaining receiver credulity in a public good game with lying and punishment. *Journal of Economic Behavior and Organization* 93, 321–327.
- Irlenbusch, B. and J. Ter Meer (2015). Lying in public good games with and without punishment. Technical report, Cologne Graduate School in Management, Economics and Social Sciences.
- Isaac, R. M. and J. M. Walker (1988). Communication and free-riding behavior: The voluntary contribution mechanism. *Economic Inquiry* 26(4), 585–608.
- Ito, K., T. Ida, and M. Tanaka (2018). Moral suasion and economic incentives: Field experimental evidence from energy demand. *American Economic Journal: Economic Policy* 10(1), 240–67.
- Keser, C. and F. Van Winden (2000). Conditional cooperation and voluntary contributions to public goods. *Scandinavian Journal of Economics* 102(1), 23–39.
- Kessler, J. B. and S. Leider (2012). Norms and contracting. *Management Science* 58(1), 62–77.
- Kocher, M. G., T. Cherry, S. Kroll, R. J. Netzer, and M. Sutter (2008). Conditional cooperation on three continents. *Economics Letters* 101(3), 175–178.
- Koessler, A.-K., L. Page, and U. Dulleck (2018). Public statements of good conduct promote pro-social behavior. *Working Paper*.
- Kölle, F., T. Lane, D. Nosenzo, and C. Starmer (2019). Promoting voter registration: The effects of low-cost interventions on behaviour and norms. *Behavioural Public Policy*, 1–24.
- Kroese, F. M., D. R. Marchiori, and D. T. de Ridder (2016). Nudging healthy food choices: A field experiment at the train station. *Journal of Public Health* 38(2), e133–e137.
- Lane, T. and D. Nosenzo (2019). Law and norms: Empirical evidence. *Working Paper*.
- Ledyard, J. (1995). Public goods: A survey of experimental research. In J. H. Kagel and A. E. Roth (Eds.), *Handbook of Experimental Economics*. Princeton University Press.
- Liberman, V., S. M. Samuels, and L. Ross (2004). The name of the game: Predictive power of reputations versus situational labels in determining prisoner’s dilemma game moves. *Personality and Social Psychology Bulletin* 30(9), 1175–1185.
- Loewenstein, G., C. Bryce, D. Hagmann, and S. Rajpal (2015). Warning: You are about to be nudged. *Behavioral Science & Policy* 1(1), 35–42.

- López-Pérez, R. (2008). Aversion to norm-breaking: A model. *Games and Economic Behavior* 64(1), 237–267.
- Marks, M. B., D. E. Schansberg, and R. T. Croson (1999). Using suggested contributions in fundraising for public good. *Nonprofit Management and Leadership* 9(4), 369–384.
- Mehta, J., C. Starmer, and R. Sugden (1994). The nature of salience: An experimental investigation of pure coordination games. *American Economic Review* 84(3), 658–673.
- Miettinen, T. and S. Suetens (2008). Communication and guilt in a prisoner’s dilemma. *Journal of Conflict Resolution* 52(6), 945–960.
- Neugebauer, T., J. Perote, U. Schmidt, and M. Loos (2009). Selfish-biased conditional cooperation: On the decline of contributions in repeated public goods experiments. *Journal of Economic Psychology* 30(1), 52–60.
- Rabin, M. (1993). Incorporating fairness into game theory and economics. *American Economic Review* 83(5), 1281–1302.
- Rege, M. and K. Telle (2004). The impact of social approval and framing on cooperation in public good situations. *Journal of Public Economics* 88(7), 1625–1644.
- Ross, L. and A. Ward (1996). Naive realism in everyday life: Implications for social conflict and misunderstanding. In E. S. Reed, E. Turiel, and T. Brown (Eds.), *Values and Knowledge*, pp. 103–135. Lawrence Erlbaum Associates.
- Schelling, T. C. (1960). *The strategy of conflict*. Harvard University Press: Cambridge, MA.
- Steffel, M., E. F. Williams, and R. Pogacar (2016). Ethically deployed defaults: Transparency and consumer protection through disclosure and preference articulation. *Journal of Marketing Research* 53(5), 865–880.
- Thaler, R. H. T. and C. R. Sunstein (2008). *Nudge: Improving decisions about health, wealth, and happiness*. Yale University Press.
- Thöni, C. and S. Volk (2018). Conditional cooperation: Review and refinement. *Economics Letters* 171, 37–40.
- Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics* 13(1), 75–98.

# APPENDICES

## Appendix A.1: Summary Statistics

Table 4: Summary statistics of demographic variables

	Female	Age	English Skill	German	Charity	Wrong answers
BASELINE						
Mean	0.40	25.33	8.15	0.81	266.11	0.71
St. Dev.	-	3.94	1.45	-	358.90	1.96
N. of Subjects	72	72	72	72	72	72
NUDGE						
Mean	0.33	25.69	7.97	0.71	279.03	0.64
St. Dev.	-	10.09	1.44	-	328.35	1.08
N. of Subjects	72	72	72	72	72	72
ANNOUNCE						
Mean	0.42	24.55	7.90	0.71	211.88	1.18
St. Dev.	-	5.82	1.48	-	257.09	2.33
N. of Subjects	96	96	96	96	96	96
NUDGE + ANNOUNCE						
Mean	0.35	24.51	8.16	0.73	255.80	1.19
St. Dev.	-	4.83	1.40	-	285.86	2.65
N. of Subjects	96	96	96	96	96	96
P-value	0.65	0.20	0.46	0.48	0.42	0.052

Notes: Female is an indicator that is equal to 1 if the subject is female. English Skill is the subject's answer to the question: "On a scale of 1 to 10, how would you evaluate your English language skills? (where "1" is "very basic understanding of English" and 10 is "fluent in English")". German is an indicator that is equal to 1 if the subject answered "Yes" to the question: "Was German spoken frequently in your home while you were growing up?" Charity is the subject's answer to the question: "Imagine the following situation: Today you unexpectedly received 1,600 Euro. How much of this amount would you donate to a good cause?" For Age, English Skill, Charity and Wrong answers, the p-value is for the Kruskal-Wallis test. For Female and German the p-value is for the  $\chi^2$  test. For each test, the null hypothesis is that there are no differences between the four treatment groups.

## Appendix A.2: Contribution Regressions

Table 5: Treatment Effect on Contributions (restricted sample)

	Model 6	Model 7	Model 8	Model 9	Model 10
NUDGE	16.40** (6.68)	15.49** (6.68)		1.39 (5.35)	1.66 (5.40)
ANNOUNCE	-0.86 (6.11)	1.84 (6.03)		-0.64 (4.61)	1.02 (4.67)
NUDGE * ANNOUNCE	4.02 (8.96)	3.83 (8.98)		-0.84 (6.67)	-0.69 (6.75)
EX ANTE BELIEF			1.01*** (0.057)	1.00*** (0.069)	0.96*** (0.070)
CONSTANT	40.14*** (4.37)	45.52*** (13.82)	-8.01** (3.38)	-7.87* (4.13)	5.23 (11.57)
CONTROLS	No	Yes	No	No	Yes
Observations	307	301	307	307	301
Adjusted R <sup>2</sup>	0.047	0.107	0.480	0.475	0.500

Notes: The table shows OLS regressions estimated on a sample from which 29 subjects were removed due to answering more than one time incorrectly to the control questions. Dependent variable is individual contribution. Independent variables are a dummy variable for the nudge treatment, a dummy for the announce treatment, a dummy for the interaction effect, and the elicited ex ante belief. The control variables in Models 7 and 10 include the individual's gender, age, squared age, a measure of English proficiency, and the answer to question "Imagine the following situation: Today you unexpectedly received 1,600 Euro. How much of this amount would you donate to a good cause?" For Models 7 and 10, 6 subjects were further excluded from the sample since they did not report their gender. Heteroskedasticity-robust standard errors are in parentheses ( \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).



Table 6: Tobit Estimates

	Model 11	Model 12	Model 13	Model 14	Model 15
NUDGE	27.05** (11.96)	28.88** (12.04)		-0.70 (8.56)	1.88 (8.66)
ANNOUNCE	-1.95 (11.14)	3.94 (11.21)		1.84 (7.86)	5.41 (7.95)
NUDGE * ANNOUNCE	2.91 (15.81)	-0.71 (15.85)		-6.39 (11.17)	-8.93 (11.29)
EX ANTE BELIEF			1.73*** (0.12)	1.76*** (0.13)	1.72*** (0.13)
CONSTANT	34.04*** (8.38)	32.65 (25.31)	-50.65*** (7.46)	-51.38*** (8.84)	-47.65** (19.29)
CONTROLS	No	Yes	No	No	Yes
Observations	336	329	336	336	329
Left-censored	89	89	89	89	89
Right-censored	71	70	71	71	70

Notes: The table shows coefficient estimates for Tobit regressions (marginal effects on the latent dependent variable). Dependent variable is individual contribution. Independent variables are a dummy variable for the nudge treatment, a dummy for the announce treatment, a dummy for the interaction effect, and the elicited ex ante belief. In Models 7 and 10 we also include as control variables the individual's gender, age, squared age, a measure of English proficiency, and the answer to question "Imagine the following situation: Today you unexpectedly received 1,600 Euro. How much of this amount would you donate to a good cause?" Standard errors are in parentheses ( \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

## Appendix A.3: Ex ante beliefs

Table 7: Treatment Effect on Ex Ante Beliefs

	Model 16	Model 17
NUDGE	15.14*** (4.35)	15.12*** (4.44)
ANNOUNCE	-0.85 (4.10)	0.31 (4.20)
NUDGE * ANNOUNCE	4.81 (5.76)	4.40 (5.90)
CONSTANT	48.61*** (2.99)	47.10*** (10.37)
CONTROLS	No	Yes
Observations	336	329
Adjusted $R^2$	0.102	0.108

Notes: The table shows OLS regressions. Dependent variable is ex ante belief. Independent variables are a dummy variable for the nudge treatment, a dummy for the announce treatment, a dummy for the interaction effect, and the elicited ex ante belief. In Model 17 we include as control variables the individual's gender, age, squared age, a measure of English proficiency, and the answer to question "Imagine the following situation: Today you unexpectedly received 1,600 Euro. How much of this amount would you donate to a good cause?" Heteroskedasticity-robust standard errors are in parentheses ( \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ ).

## Appendix A.4: Ex post beliefs

In treatments ANNOUNCE and NUDGE + ANNOUNCE, we elicited every subject's beliefs about the contribution of each of her group members *after* the announcements were made and revealed. This allows us to analyze how subjects update their beliefs after seeing the announcements. While belief updating is not the main focus of our study, it is interesting to see how subjects assess the credibility of others' announcements when they know that there do not exist monetary or strategic incentives to lie.

The average ex post belief is 41.6% in the ANNOUNCE treatment, and 61.4% in the NUDGE + ANNOUNCE treatment (a two-sided WRS,  $p < 0.01$ ).<sup>9</sup> The difference is hardly surprising, considering that both contributions and announcements were higher in NUDGE + ANNOUNCE than in ANNOUNCE.

More to the point, ex post beliefs are lower than ex ante beliefs in the ANNOUNCE treatment (a two-sided Wilcoxon signed-rank test (WSR),  $p = 0.08$ ), as well as in the NUDGE + ANNOUNCE treatment (a two-sided WSR,  $p = 0.03$ ). This suggests that ex post beliefs tend to be more accurate than ex ante beliefs. We test this claim formally by first constructing for each subject a measure of ex ante belief accuracy and three measures of ex post belief accuracy (one for each group member). We do this by subtracting the relevant belief from the actual contribution. Within each group, we then calculate the average ex ante belief accuracy and the average ex post belief accuracy. Average ex ante belief accuracy in the two treatments with announcements is -1.59 points, while average ex post belief accuracy is -0.34 points (a two-sided WSR,  $p < 0.01$ ). While both ex ante and ex post beliefs tend to be, on average, too optimistic, ex post beliefs are generally much more accurate.<sup>10</sup>

This finding is interesting in that it shows that announcements do carry informational content. In general, subjects see the announcements as credible and update their beliefs accordingly, which results in largely realistic ex post judgments.<sup>11</sup>

---

<sup>9</sup>We respect the independence assumption by basing all our tests in this section on group averages.

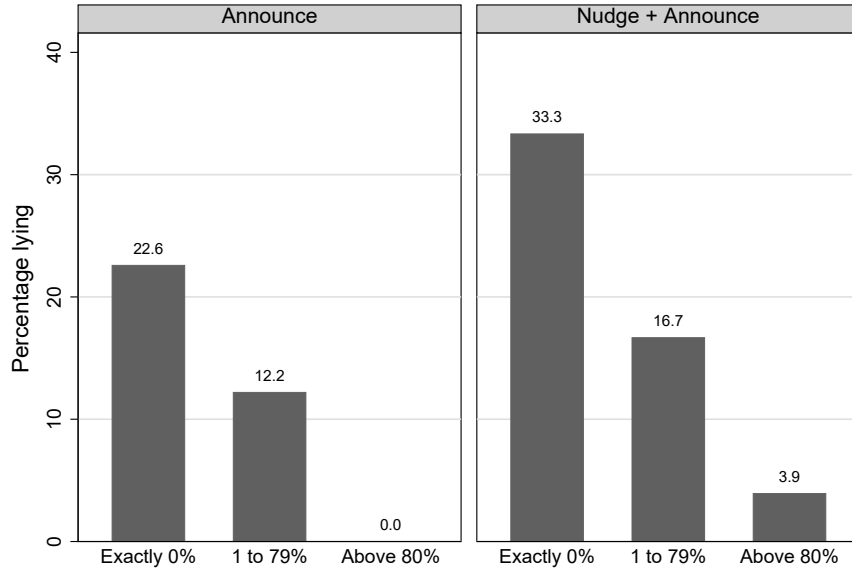
<sup>10</sup>These observations hold even if we compare ex ante beliefs to ex post beliefs by treatment. In addition, both ex ante and ex post beliefs are equally accurate *across* the two treatments.

<sup>11</sup>As an anonymous referee pointed out, the objects of elicitation were slightly different in the two elicitation stages. We cannot completely rule out the possibility that the observed difference in ex ante and ex post beliefs is an artefact of differences in the elicitation stages themselves.

## Appendix A.5: Lying in announcements

Despite the subtlety of the ANNOUNCE intervention we find that a non-negligible proportion of individuals were willing to lie to avoid making a low announcement. In particular, 13% of all subjects lie in their announcement. However, this includes all the individuals who chose high contribution levels and had no motive to lie. Therefore, it is more informative to examine Figure 3 which displays the proportion of individuals lying, split by their contribution choice.

Figure 3: Lying as a function of contribution choice



Focusing on the individuals who contributed zero, the rate of lying is 22.6% in ANNOUNCE and 33.3% in NUDGE + ANNOUNCE. The rate of lying among free riders differs significantly from zero in both these treatments (two Binomial tests,  $p < 0.01$ ). Since these individuals have no monetary, strategic nor social image incentives for lying, this finding supports our presumption that some of the subjects lie in order to avoid a cost of announcing a low contribution. Hence, for these individuals the costs of announcing their selfish choice of contributing nothing seems to have outweighed any aversion to lying they may have had.<sup>12</sup> On the other hand, both in the ANNOUNCE and in the NUDGE + ANNOUNCE treatments the majority of free-riders chose a truthful announcement. It seems therefore that the way the announcement mechanism was implemented in our design was

<sup>12</sup>As pointed out to us by an anonymous referee, in our setting subjects who only care about money are indifferent between lying and telling the truth, which implies that our observation is consistent with the possibility that these agents are choosing their announcements randomly. However, the fact that lying is much more prevalent among free riders than among those contributing a positive amount suggests that the announcement decision is not random. Furthermore, 96% of those who lied announced a higher contribution than their true contribution, which also speaks against random announcements.

not stark enough to induce sufficiently high costs of announcing a low contribution. It would be intriguing to see how this would change if anonymity and non-verifiability of the announcements were (sequentially) removed. We leave this for future research to explore.

## Appendix B: Additional figures

Figure 4: CDF of contribution levels across treatments

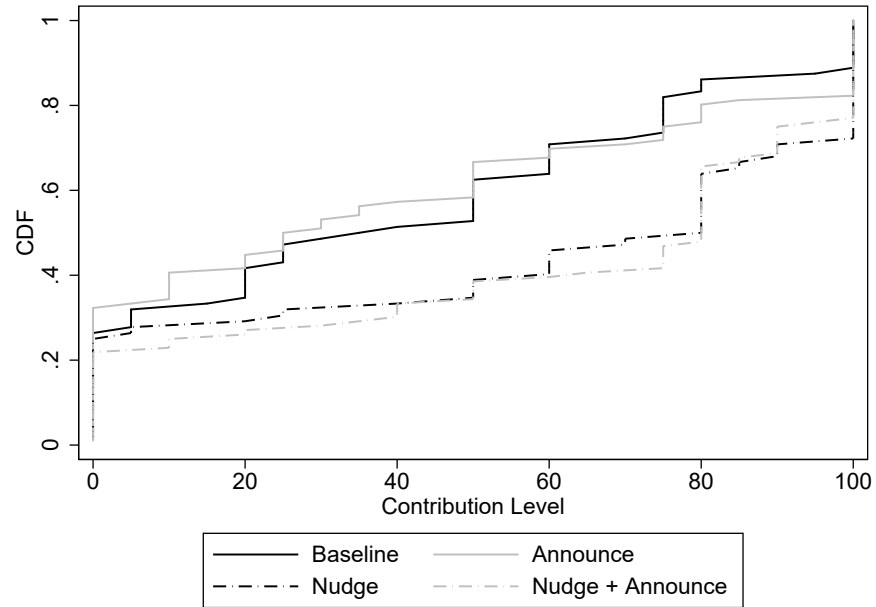


Figure 5: Distribution of contribution levels

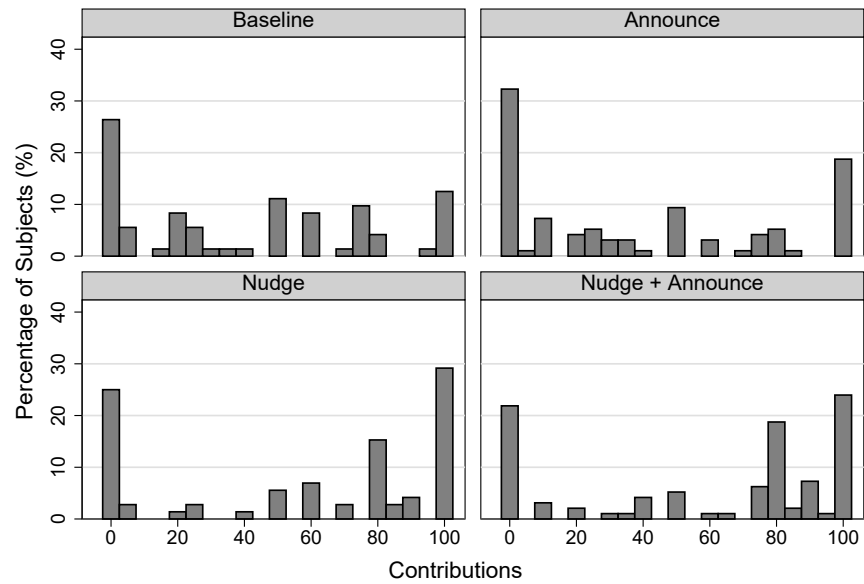


Figure 6: Ex ante beliefs across treatments (mean and median)

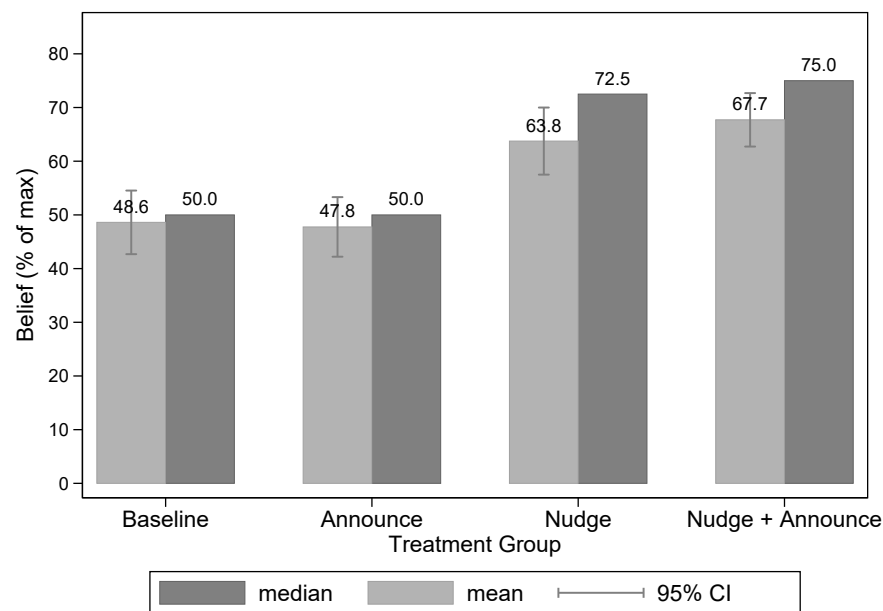
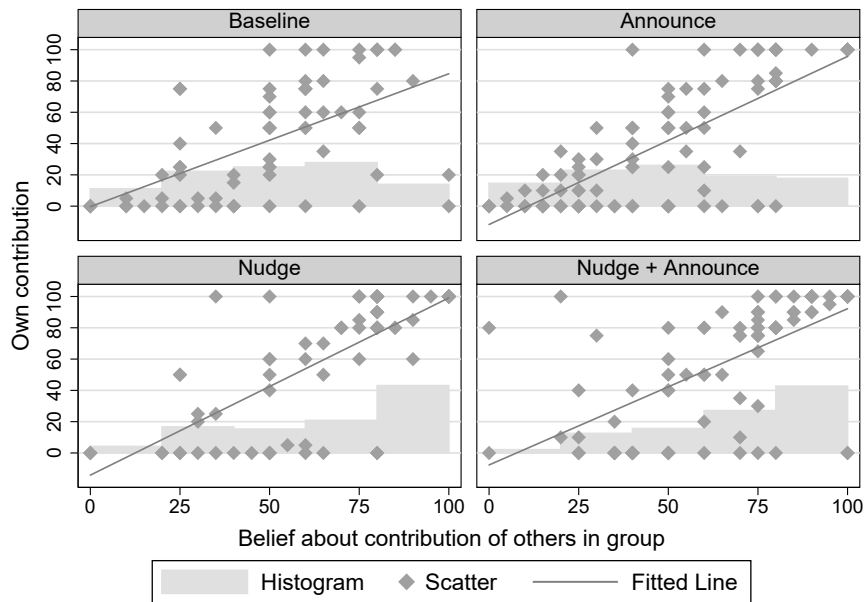


Figure 7: Beliefs and contributions





# Appendix C

*Instructions for the NUDGE + ANNOUNCE treatment.*

## Game II: Overview

In Game II, all participants are divided into groups of four. These are completely new groups and are not related to the groups from Game I in any way. (Recall that Game II is not related in any way to Game I). You will therefore be in a group with three other participants (called partners later on). However, you will not learn the identities of your partners, nor will they learn yours.

Game II of the experiment consists of two stages. At the beginning of the first stage you (and each of your partners) will receive an endowment of 5 EUR. This endowment will be converted into experimental points using the following exchange rate:

$$4 \text{ points} = 1 \text{ EUR}$$

Therefore, you will start Game II of the experiment an endowment of 20 points. At the end of the experiment, we will convert your points back to Euro using the same exchange rate as indicated above and it will be added to your earnings from the rest of the experiment and paid to you in cash.

## The Contribution Stage:

In the contribution stage, you will need to decide how many of the 20 points you want to contribute to a group project, and how many you would like to keep yourself. Every point that you keep yourself, increases your final earnings by 1 point. Every point that you contribute to the group project will increase your own earnings by 0.35 points, and will also increase the earnings of each member of your group by 0.35 points. Similarly, every point that another group member contributes to the group project will also increase your earnings by 0.35 points. At the end of the experiment, you will learn the value of your earnings, but the amount contributed by each participant will not be revealed.

Your payoff from the group project will be calculated as follows.

First, all the group contributions in your group are summed up. You (and each of your partners) then get  $0.35 \times$  the sum of your group contributions. This will be added to the amount that you kept for yourself.

Hence your total payoff from the Contribution Stage consists of two components:

- 1) the fraction of your endowment you have kept for yourself,

2) your share of the total group project contributions.

Your earnings can therefore be calculated as follows:

$$\text{Earnings} = (20 - \text{your group contribution}) + (0.35 * \text{total group contribution})$$

Examples:

1. Imagine that you and all your group members contribute the full endowment (i.e. 20 points); the sum of group contributions in your group is 80 points. You would get  $0.35 * 80 = 28$  points from the group project. Hence your total payoff is  $(20 - 20) + 28 = 28$  points. This is also what your partners get.
2. Imagine that you and all your group members contribute nothing to the group project (i.e. 0 points); the sum of group contributions in your group is 0. You get  $0.35 * 0 = 0$  points from the group project. Hence your total payoff is  $(20 - 0) + 0 = 20$  points. This is also what your partners get.

Before making your contribution decisions, we will ask you some questions on the computer to ensure that you have fully understood these instructions. In addition, during this initial clarification phase, there will be a calculator on the screen that allows you to input hypothetical contribution values for you and all your group members and the calculator will tell you how much each person would earn.

### **Contributing to the group project is beneficial for everyone**

As you can see from the examples above, if everyone contributes to the group project, you will all earn more points than if you all contribute nothing and keep all your points to yourself. It is therefore beneficial to everyone to contribute more to the group project.

Of course, if everyone else contributes to the group project and you keep all your points to yourself, then you will earn even more, but this will harm the other members of your group. They will lose more than you gain.

### **A contribution of 16 or above is “good”**

While you may choose any contribution level you would like, in this experiment, we will call a contribution of 16 or larger a “good” contribution, and a contribution of 15 or below a “bad” contribution, since contributing to the group is good for everyone. On your decision screen, you will see that the “good” contribution values are denoted in green, while the “bad” contribution values are denoted in red.

## **The Announcement Stage:**

After you have made your contribution decision, every member of the group will make an announcement to the other members of the group regarding the amount that he / she has contributed. Your final payment will depend on the contribution you made and not on your announcement.

The announcement will take the following form and you will have to choose what message to send in place of  $x$  and  $20-x$  :

**“I contributed  $x$  points to the group project and kept  $20-x$  points for myself”**

Once all the group members have made their announcements, you will observe the announced contributions of your group members.

**We will now proceed to Game II. Before we do, if you have any questions at this moment, please raise your hand. The experimenter will come to you.**