HAL
open science

Leveraging the Honor Code: Public Goods Contributions under Oath

Jérôme Hergueux, Nicolas Jacquemet, Stéphane Luchini, Jason F Shogren

To cite this version:

Jérôme Hergueux, Nicolas Jacquemet, Stéphane Luchini, Jason F Shogren. Leveraging the Honor Code: Public Goods Contributions under Oath. 2016. halshs-01379060

HAL Id: halshs-01379060

https://shs.hal.science/halshs-01379060

Preprint submitted on 11 Oct 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2016 – 22

Leveraging the Honor Code: Public Goods Contributions under Oath

Jérôme Hergueux
Nicolas Jacquemet
Stéphane Luchini
Jason F. Shogren

JEL Codes: C72; D83
Keywords: WorkingPublic good game; Social Preference; Truth-telling Oath

PARIS-JOURDAN SCIENCES ECONOMIQUES
48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ECOLE DES HAUTES ETUDES EN SCIENCES SOCIALES
ÉCOLE DES PONTS PARISTECH – ECOLE NORMALE SUPÉRIEURE – INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE

Leveraging the Honor Code: Public Goods Contributions under Oath *

Jérôme Hergueux[†] Nicolas Jacquemet[‡] Stéphane Luchini[§] Jason F. Shogren[¶]

July 2016

Abstract

Real economic commitment (or the lack of it) of others affects a person's preferences to cooperate. But what if the commitment of others cannot be observed ex ante? Herein we examine how a classic non-monetary institution– a solemn oath of honesty –creates economic commitment within the public goods game. Commitment-through-the-oath asks people to hold themselves to a higher standard of integrity. Our results suggest the oath can increase cooperation (by 33%)– but the oath does not change preferences for cooperation. Rather people react quicker and cooperate, taking less time to ponder on the strategic free riding behavior.

Keywords: Public good game; Social Preference; Truth-telling Oath.

JEL Classification: C72; D83.

1 Introduction

Laboratory evidence on the voluntary provision of public goods has identified two main mechanisms that efficiently promote cooperation in situations in which private and public interests do not coincide: *(i)* Pricing – the "willing punishers" pay to inflict monetary punishment on non-cooperative

*We are grateful to Anne l'Hôte for outstanding research assistance and Maxim Frolov and Ivan Ouss for their help in running the laboratory experimental sessions. Financialsupport from the Health Chair, a joint initiative by PSL, Université Paris-Dauphine, ENSAE and MGEN under the aegis of the Fondation du Risque (FDR) ; the chair "Economie Publique et Développement Durable" (Aix-Marseille University) and the French National Research Agency, through the program Investissements d'Avenir (ANR-10–LABX-93-0) are gratefully acknowledged. Jacquemet thanks the *Institut Universitaire de France* for its support. Shogren thanks the Rasmuson Chair at the University of Alaska-Anchorage for the support while working on this project.

[†]ETH Zurich – Department of Humanities, Social and Political Sciences. jerome.hergueux@gess.ethz.ch

[‡]Paris School of Economics and University Paris I Panthéon–Sorbonne. Centre d'Economie de la Sorbonne, 106 Bd. de l'Hôpital, 75013 Paris. Nicolas.Jacquemet@univ-paris1.fr

[§]GREQAM-CNRS, Centre de la Vieille Charité, 13236 Marseille Cedex 02. stephane.luchini@univmed.fr

[¶]Department of Economics and Finance, University of Wyoming, Laramie, WY 82071-3985, United States. JRamses@uwyo.edu

free-riders (see e.g., Fehr and Gächter, 2000, 2002; Gürerk, Irlenbusch, and Rockenbach, 2006), and *(ii)* Sorting – a rule to cluster people based on their "social type" given pre-determined rules that banish free riders and weak reciprocators, either by fiat (see e.g., Gächter and Thöni, 2005; Burlando and Guala, 2005; Gunnthorsdottir, Houser, and McCabe, 2007) or by group vote (see e.g., Page, Putterman, and Unel, 2005; Cinyabuguma, Page, and Putterman, 2005; Charness and Yang, 2014).

While such mechanisms have been shown to work well in the lab, their field implementation raises several concerns. First, both mechanisms can be costly, because they both rely on creating real economic commitment at the personal level to promote beneficial outcomes at the social level. Pricing generates direct private costs for the punishers. In the voting case, the private cost is that the remaining collective is smaller so the public good benefits are fewer – involving private pain for social gain. Second, economic and social historians have shown that modern societies typically solve social dilemmas through institutions which coordinate punishment (Guala, 2012). This stands in sharp contrast with the decentralized punishment mechanisms implemented in the lab, whose efficacy relies on the strength of individuals' emotional responses to non-cooperative behavior, and on their degree of coordination, which makes the severity of sanctions hard to predict and fine-tune *a priori* (Posner, 2000). Third, while the threat of ostracism is an efficient tool used by collectives to enforce social norms and promote cooperation, such mechanisms are difficult and costly to generalize within legal systems.

Herein we step back and ask whether an alternative non-monetary legal instrument –an archetypal solemn oath of honesty– can create real economic commitment to cooperate within the public goods game. Commitment-through-the-oath creates an environment in which people are asked to hold themselves to a higher standard of integrity. Such procedures have a significant "cheap talk" component, which should limit their effectiveness (Farrell, 1987). In practice, however, their use is ubiquitous. Courts resort to truth-telling oath procedures in judicial hearings, administrative offices use them to elicit honest behavior from public servants (oath of office) and citizens (sworn statements) alike, various teaching institutions and military service academies rely on "honor codes" to foster cooperation among community members, and physicians are typically required to take a Hippocratic oath before they start practicing medicine.

Despite the popularity of this commitment device, its effectiveness remains unclear and is still an open question (Cohn, Fehr, and Maréchal, 2014). In this paper, we assess whether an oath of honesty can affect individual's propensity to cooperate in a public goods setting. Building on the social psychology theory of commitment (see, e.g., Kiesler, 1971; Joule, Girandola, and Bernard, 2007; Jacquemet, Joule, Luchini, and Shogren, 2013, for a detailled discussion of the fundations of the oath as a commitment device), we test whether a "pure" truth-telling oath procedure can create the commitment necessary to increase contributions in a Public Goods game. By "pure" oath procedure, we mean a procedure which strictly leverages one's intrinsic motivation to meet

a personal commitment, in the absence of punishment, reputation effects and social pressure. We therefore reduce the oath procedure to its simplest form: a non-enforceable, intrinsic commitment to holding oneself to a higher standard of integrity while performing a certain task. Such a design allows us to put a lower bound on the effectiveness of the oath procedure, and to focus our inquiry on what it is able to achieve by relying solely on subjects' good will. Our design departs from the way oath procedures are sometimes implemented in the field on several dimensions. First, unlike many field settings, breaking the oath is not associated with a formal threat of centralized punishment. Second, taking the oath is a private and non-mandatory choice, and we examine its impact on behavior in a one-shot anonymous Public Goods game. Those features of the design ensure that subjects have no incentive to meet the oath in order to maintain credibility in future interactions (Klein and Leffler, 1981; Bull, 1987; Kim, 1996; Levin, 2003; Resnick, Zeckhauser, Swanson, and Lockwood, 2006). Third, our oath procedure does not indicate what is the "desirable" behavior in the public good game, i.e it does not have any "expressive function" (Posner, 1998; McAdams, 2000).[1] For instance, Dal Bo and Dal Bo (2014) show that exposing subjects to moral messages (such as "*actions are moral to the extent that they contribute to maximizing collective payoffs*" and "*moral actions are those that treat others as you would like to be treated*") significantly increases cooperation levels. The honesty oath, by contrast, commits the person during the whole experiment to "tell the truth and always provide honest answers", leaving it to the person what constitutes honest behavior.

Our work is related to a growing literature on promises which relies on experimental methods to understand the design and assess the efficiency of existing relational contracting tools. Ellingsen and Johannesson (2004) conduct a one-shot, two-players experiment in which a buyer can generate surplus from an investment decision, but then faces a hold-up problem as the seller decides of its division. The authors find that sellers who are offered an *ex ante* opportunity to make a (non-enforceable) commitment regarding the division of the surplus typically keep their promises *ex post*. Charness and Dufwenberg (2006) show in the context of a one-shot modified Trust game that subjects who are offered the opportunity to make a promise *ex ante* tend to keep them *ex post*, which results in an increase in prosocial behavior and efficiency.[2]Similarly, taking an oath produces

[1]By making community standards explicit (*i.e.*, what is "desirable" behavior in a particular context), the law can incentivize compliance among individuals who exhibit social image and conformity concerns (Goldstein, Cialdini, and Griskevicius, 2008). For recent experimental evidence on the expressive function of the law, see McAdams and Nadler (2005); Galbiati and Vertova (2008); Galbiati, Schlag, and Van Der Weele (2013); Bremzen, Khokhlova, Suvorov, and Van de Ven (2015).

[2]A recent literature investigates the behavioral motives through which promises can become self-enforceable. Vanberg (2008) designs an experiment aimed at disentangling between two mechanisms : *(i)* lying aversion –individuals have an intrinsic preference for keeping their word– and *(ii)* guilt aversion –individuals experience disutility from a failure to live-up to the expectations that they created in others. Together with Ellingsen, Johannesson, Tjøtta, and Torsvik (2010), Vanberg (2008) reports evidence which he interprets in favor of the lying aversion mechanism. Ismayilov and Potters (2015) provide evidence of a selection effect –promises are more likely

individual commitment which results in a change in subsequent behavior. The oath works as a "*preparatory act*" put forward by the social psychology of commitment (Burger, 1999): compliance with an initial request induces behavioral changes in subsequent decision making situations.

Our results are three-fold. First, we find that the oath does help increase cooperation (by 33%). We moreover elicit both unconditional and conditional contributions so as to get subjects' full reaction function to every possible contribution of others. The oath shifts conditional behavior from free-riding and weak reciprocation to strong reciprocation and altruism. The increase in contribution thus goes through a change in reply functions, rather than a change in expectations about the behavior of others. Second, consistent with our hypothesis, we present evidence that the oath works not by impacting underlying social preferences, but rather by inducing subjects to put more weight on their internalized behavioral norms rather than strategic thinking. Under oath, many people go with their instincts to "do the right thing" (*i.e.*, cooperate at some "suitable" level). Conversely, subjects under oath are much less prone to reflect on the strategic interactions implied by rational choice theory and its subsequent prediction to free ride. This interpretation is further supported by the observation that, beyond social preferences, the oath has no impact on subjects' beliefs about the contribution level of others, nor does it modify their normative opinion as to how much people should typically contribute. Rather the oath procedure primarily works by increasing the level of agreement between subjects' actual contributions and their self-reported prescriptive norm (by 43% with respect to baseline). We provide in the last section of the paper a summary of possible psychological interpretations of such changes in behavior.

2 Experimental design I: implementing the oath

Our experiment has a control group (no oath) and a treatment group (under oath). For the treatment, we asked each subject to sign an explicit oath before entering the lab. Empirical evidence shows that signing creates a strong commitment to old promises (see, e.g., Pallack, Cook, and Sullivan, 1980; Katzev and Wang, 1994; Joule, Girandola, and Bernard, 2007; Shu, Mazar, Gino, Ariely, and Bazerman, 2012). Following Jacquemet, Joule, Luchini, and Shogren (2013), we implement the oath as follows: each subject enters alone and is directed to a monitor at the front of the laboratory. For the oath-desk, a monitor offers each subject a form to sign entitled "solemn oath" (an English translation of the original form in French is provided in Appendix A). The "oath" is written on the form and read by the subject, but never said aloud. To indicate that it is an official paper, we use the Paris School of Economics logo on the top of the form and the

to be sent by cooperators. By contrast, Ederer and Stremitzer (2016) develop a design in which subjects' second order beliefs are manipulated exogenously. They show that subjects who make a promise to perform do so at a significantly higher rate when they anticipate high expectations from their counterpart. Overall, the existing evidence supports the existence of an preference to meet other people's expectations, and most importantly for our work—to honor their commitment to keep one's word.

address at the bottom. The topic designation and the research number were added to ensure the credibility. The monitor explicitly points out to the subject before he or she reads the form that he or she is free to sign the oath or not, and that participation and earnings are not conditional on signing the oath. When asked to take the oath, however, subjects are not told what the actual experiment will be about. The subject reads the form, which asks whether he or she agrees "*to swear upon my honor that, during the whole experiment, I will tell the truth and always provide honest answers*" (emphasized in the original form). We chose the wording "solemn oath" and "upon honor" given the secularism of French modern society, in which law and political parties cannot be based on religion. Regardless of whether he or she signed the oath, the monitor thanks the subject and invites them to enter the lab. We scripted the exact wording for the monitor to read when offering the oath to standardize the phrasing of the procedure. To avoid subject communication prior to the experiment, one monitor stayed in the lab until all subjects had been presented with the oath. Subjects who were waiting for their turn could neither see nor hear what was happening at the oath-desk.

3 Experimental design II: decision tasks

Experimental design. Our experimental design follows a 4-step design for both the control and treatment.

Step 1. Roles: each subject is assigned to be either a participant A or B. The assigned role remains the same during the experiment.

Step 2. Public Goods Game (PG): participants participate a classic public goods game in groups of 4 players, each with a 10 Euros endowment. Each Euro invested into the common project yields a return of 0.4 Euro to each group member. Following Fischbacher, Gächter, and Fehr (2001), we elicit unconditional and conditional contributions to the common project. For the unconditional contribution, each subject makes his or her contribution (either 0, 1, 2, ... 10 Euros) to the common project. For the conditional contribution, each subject states his or her intended contribution for each possible value (0,1,2, ... 10) of the average contribution of the three other members. The monitor will randomly draw either the unconditional or conditional contribution to be binding, and will then determine individual earnings according to the payoff function:

$$\pi_i = 10 - contrib_i + 0.4 \sum_{j=1}^{4} contrib_j \tag{1}$$

Immediately after the decision screen, subjects reported *(i)* their normative opinions about how much people should contribute to the Public Good, *(ii.a)* whether they had an idea about how much the other group members would contribute, and if so *(ii.b)* their beliefs about how much the other members actually contributed on average.

Step 3. Social Preferences Games: to better understand the nature of each subject's social preferences, we then elicit decisions in three classic games found in the literature: dictator, ultimatum, and trust games (see, *e.g.*, Fehr and Camerer, 2004). In each game, *(i)* each A player is randomly paired with one B player, and *(ii)* at the end of each game, subjects answer non-incentivized questions about their beliefs and intentions in the game they just played.

> *Dictator Game (D).* Player A is the Dictator with a 10 Euro endowment. Each A decided how much of his or her endowment to transfer (if any) to B. The dictator's split determines the subjects' earnings for the game.
>
> *Ultimatum Bargaining Game (U).* Player A is the Proposer with a 10 Euro endowment. Each A makes an offer to B (the responder) on how to split the endowment. Each responder simultaneously states the minimum threshold below which he or she would reject the offer. If A offers more than B's threshold, the players earn the proposed split; otherwise, they both earn zero.
>
> *Trust Game (T).* Now both player A and B receive a 10 Euros endowment. A is the trustor and chooses how much of his endowment to transfer to B, the trustee. The trustee receives three times the amount sent by the trustor, and then chooses how much to send back to A. We elicit this decision through the strategy method: for each possible transfer from the trustor (1 to 10) the trustee chooses how much will be returned without knowing the trustor's actual choice. Right after the decision screen, we ask the trustors *(i)* whether they had an idea about how much the trustee would return to them, and if so *(ii)* their beliefs about the amount that the trustee would return.

Social values survey. After all games were played, subjects answered a questionnaire asking *(i)* standard demographic questions, and *(ii)* social preferences questions taken from the *World Value Survey* (WVS), the *General Social Survey* (GSS) and the *German Socio-Economic Panel* (GSEP) – the three commonly used sources in the empirical literature. All questions are mandatory and none is remunerated:

(i) To what extent they consider it justifiable to free-ride on public social allowances (cooperation variable; 10 points scale, WVS question);

(ii) Do you think people are mostly looking out for themselves as opposed to trying to help each other (altruism variable; 10 points scale, WVS question);

(iii) Do you think people would try to take advantage of them if they got a chance as opposed to trying to be fair (fairness variable; 10 points scale, WVS question);

(iv) Do you think most people can be trusted or that one needs to be very careful when dealing with people (trust variable; binary answer, WVS and GSS question);

(v) How much do you trust people in general (general trust variable; 4 points scale, GSEP question);

(vi) How much do you trust people you just met (trust in strangers variable; 4 points scale, GSEP question);

(vii) Do you generally see yourself as fully prepared to take risks as opposed to generally trying to avoid taking risks (risk aversion variable; 10 points scale question taken from Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner, 2011).

Response times. An established body of research in psychology indicates that shorter decision times are likely to be associated with *instinctive and emotional reasoning processes* rather than *cognitive and rational ones* (Kahneman, 2003). Such differences in decision times can have a significant impact on the elicited social preferences (Rubinstein, 2007; Piovesan and Wengstrom, 2009; Rand, Greene, and Nowak, 2012; Lotito, Migheli, and Ortona, 2013). We control for any potential interaction between the oath treatment and decision times by having the platform record detailed data on the time in seconds that subjects spend on each screen of the interface (this timer was not visible to the subjects).

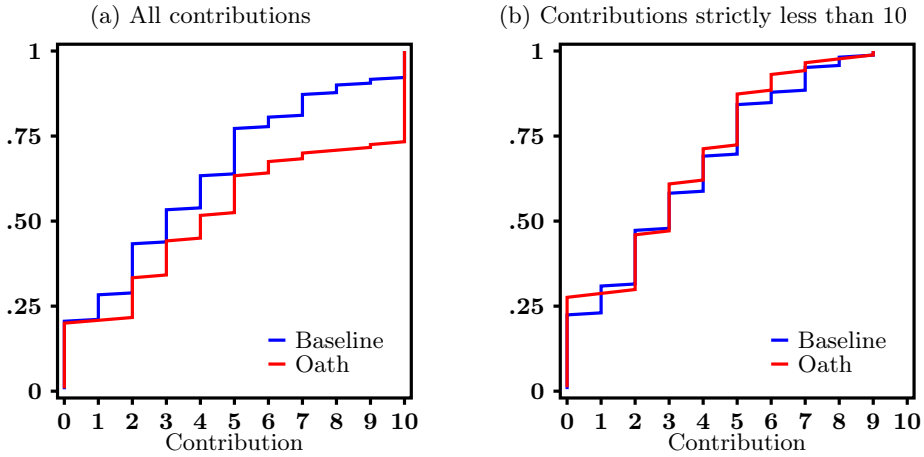
Procedures. All participants were contacted through the subject database of the experimental economics laboratory of University Paris 1 Pantheon-Sorbonne. Given our goal to elicit social preferences in isolation from learning effects and strategic concerns, each game is only played once.[3] To neutralize reputation effects, we match subjects in each game according to a perfect stranger procedure. Since all games are played one after the other, we implement three orderings: (1): *PG-D-U-T*; (2): *PG-T-U-D*; (3): *PG-U-D-T*. We conducted 3 and 6 no-oath sessions with 180 subjects in November 2010 and November 2011, and 6 oath sessions with 120 subjects in June 2013. We conducted 6 sessions with Order (1) (80 participants in the baseline treatment, 40 in the oath treatment), 5 sessions with Order (2) (60 and 40) and 4 sessions with Order (3) (40 and 40).

To break any possible correlation between games, we randomly draw only one game out of the entire session as binding to compute each subject's earnings. Final payoffs equal the earnings from the corresponding decision plus a 5 Euros show-up fee. Subjects are only informed of their earnings in each game at the end of the experiment. Subjects' earnings are paid privately in cash. Subjects earned 21.10 Euros on average in the control and treatment sessions.

Subjects were randomly assigned to a computer on arrival. The instructions for the experiment were read aloud, and subjects were then left to use all devices at their disposal to check

[3]The experiment is computerized based on an own-developed decision interface described in Hergueux and Jacquemet (2015). All sessions took place at the *Laboratoire d'Economie Experimentale de Paris* (LEEP) at Paris School of Economics. Subjects were recruited via an on-line registration system based on Orsee (Greiner, 2015).

Figure 1: Contributions to the public good by treatment (Empirical distribution functions)



their own understanding (access to the text, earnings calculators, etc.). On the decision interface, the first screen provides subjects with general information about the experiment. Next the screen describes the game, followed by examples, and the payoffs for each player. The next screen introduces the earnings calculators, which is an interactive page that lets subjects explore hypothetical scenarios of interest before making decisions in the Public Good and Trust games. On all screens, including decision-making ones, a "review description button" gives subjects a direct access to the instructions displayed at the beginning of the game.

4 Main results: Observed contributions in the public good game

First, consider the unconditional contributions. At the aggregate level, taking an oath induces a 33.1% increase in unconditional contributions to the public good: mean unconditional contribution is 4.85 in the oath treatment, up from 3.65 in the baseline (Mean difference test $p = .003$). Figure 1.a presents the empirical distribution function (EDF) of contributions by treatment. Contributions are on the x-axis and the cumulated probability of observing a given contribution is on the y-axis. In both treatments, although subjects use the entire range of possible contributions, the EDF of contributions under oath first order dominates the EDF of contributions in the baseline ($p = .001$).[4]

Why does the oath work to increase average unconditional contributions? The main reason is that the oath induces significantly more subjects to increase their contributions at the top 3 levels (8-9-10). Otherwise behavior is about the same as in the baseline. We see this by examining the

[4]This result comes from a bootstrap version of the univariate Kolmogorov-Smirnov test. This modified test provides correct coverage even when the distributions being compared are not entirely continuous and, unlike the traditional Kolmogorov-Smirnov test, allows for ties (see Abadie, 2002; Sekhon, 2011).

Table 1: Conditional contributions by treatment

Others' average contribution		0	1	2	3	4	5	6	7	8	9	10
Baseline	Average contribution	0.94	1.46	1.91	2.54	3.00	3.61	3.99	4.29	4.66	5.09	5.36
	Contribute 10 (%)	5.0	1.7	1.1	0.6	1.1	0.6	2.2	1.7	3.9	10.0	35.0
Oath	Average contribution	0.86	1.56	2.13	2.62	3.14	3.78	4.39	4.66	5.34	5.69	6.57
	Contribute 10 (%)	4.2	3.3	3.3	3.3	3.3	4.2	3.3	5.8	8.3	9.2	55.0

Note. For each treatment in row, the first row provides the average contribution observed against the average level of contribution of other group members (in column). The second row provides the share of subjects who contribute their full endowment.

distribution at the top and lower ends. First-order dominance arises under oath because more subjects contribute the maximal amount (10) to the public good–27.5% in the oath treatment relative to 8.3% in the baseline: a 230.0% increase, which is significant at the 1% threshold (proportion test: $p < .001$). At the lower-end of the distribution, we observe no change in zero contributions: 20.0% under oath, 20.5% in the baseline. Figure 1.b shows the EDF of contributions by treatment for the sub-sample of contributions strictly less than 10. We see the two treatment are similar - no first-order dominance of contributions in the oath treatment (first order bootstrap dominance test $p = .734$). Mean contributions for contributions less than 10 are also similar: 3.1 and 2.9 in the oath and baseline ($p = .621$). The oath seems to work by transforming less than 10 unconditional contributions into a maximal amount contribution of 10.

Now consider the levels of conditional contributions. Table 1 presents mean contributions by average contribution of other group members and treatment. The table also shows the proportion of subjects who contribute their whole endowment (*i.e.*, 10). The results indicate that subjects under oath contribute more to the public good for all average contribution considered. The increase in mean behavior, however, is significant only for the top 3 average contributions of other group members with $p = 0.043$, $p = 0.085$ and $p = 0.008$. The proportions of subjects who contribute 10 to the public good do not differ by average contribution of other group members, except when the average contribution is 10. In this case, 35.0% of subjects contribute 10, whereas 55.0% do so in the oath treatment ($p < 0.001$). The remaining subjects (those contributing less than 10) are close to self-interested rational behavior in both treatments with a mean contribution of 2.37 and 2.85 in the oath and baseline (the decrease in mean contribution is not significant, $p = 0.158$). As a result, the oath seems to work by inducing higher levels of reciprocity when subjects are exposed to highly contributing group members.

To understand better the mechanism by which the oath modifies subjects' contributing patterns, we build upon Fischbacher, Gächter, and Fehr (2001)'s approach and classify subjects into four exclusive social groups, depending on the contributing preferences that they reveal in the conditional Public Goods game. We compute *(i)* the slope of subjects' reaction functions to the possible average contributions of the other group members ("reciprocity" r) and *(ii)* the average

proportion of the endowment that is conditionally contributed across all 11 conditional contributions decisions ("mean contribution" m). We then classify subjects according to the following rule:

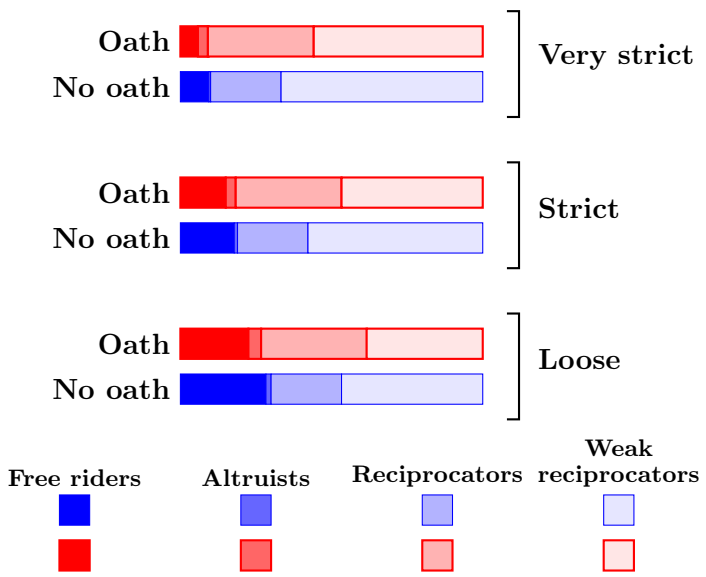
Free riders:	$\{r < 1\}$ & $\{m = 0\}$
Weak reciprocators:	$\{r < 1\}$ & $\{0 < m < 1\}$
Reciprocators:	$\{r \geq 1\}$
Altruists:	$\{r < 1\}$ & $\{m = 1\}$

We label this classification rule as "very strict", in the sense that it requires that free riders satisfy $m = 0$ (*i.e.*, the subject never makes a positive contribution, irrespective of the average contribution of the other members of the group) and that altruists satisfy $m = 1$ (*i.e.*, the subject always contributes all of his endowment, irrespective of the average contribution of the other members of the group). To check for the consistency of our results, we also apply less stringent classifications rules, which notably allow subjects for some range of decision error. Specifically, the "strict" classification rule requires that free riders satisfy $m \leq 0.1$ as opposed to $m = 0$ (and, conversely, that altruists satisfy $m \geq 0.9$ as opposed to $m = 1$), while the "loose" classification rule requires that free riders satisfy $m \leq 0.2$ (and, conversely, that altruists satisfy $m \geq 0.8$).

In all cases, our classification distinguishes between "weak" and "non weak" reciprocators. Distinguishing between those two types of reciprocators is important, as groups constituted of reciprocators typically succeed in sustaining their contribution levels in a repeated Public Goods experiment. In contrast, the presence of free-riders and weak reciprocators in a group usually triggers the progressive decline of cooperation that is typical of lab experiments (Chaudhuri, 2011). In Fischbacher, Gächter, and Fehr (2001)'s setting, our "weak reciprocators" would be labeled "hump shaped" contributors. We depart slightly from their approach, however, in the sense that they identify hump shaped contributors through a visual examination of their conditional contribution patterns, whereby contributions increase up to a point with that of group members, and then tend to decrease as the financial stakes become higher.

Figure 2 presents the distribution of social types in the conditional Public Goods game by treatment (*i.e.*, "no oath" versus "oath") and classification rule (*i.e.*, very strict, strict and loose). Irrespective of the classification rule, the oath appears to induce a reduction in the proportion of weak reciprocators (and, to a lesser extent, free riders) in the sample, while increasing the proportion of reciprocators and altruists. Table 2 performs a formal test of the significance of those oath-induced changes in the proportions of each social type in the sample, based on non-parametric Wilcoxon-Mann-Whitney two-sided tests. If we are to adopt a very strict classification rule, we can see that the oath induces a significant increase in the proportion of altruists in the sample from 0.6 to 3.3% (*i.e.*, a 550% increase, $p = 0.066$). Similarly, the proportion of reciprocators increases from 23 to 35% (*i.e.*, a 52% increase, $p = 0.028$). This increase is almost fully compensated by the decrease in the proportion of weak reciprocators in the sample from 67 to

Figure 2: Distribution of types in the conditional Public Goods game



56% (*i.e.*, a 20% decrease, $p = 0.058$). Adopting a strict or loose classification rule does not change the nature of those results. It seems that the oath works primarily by turning weak reciprocators into perfect reciprocators in the game. To a lesser extent, the oath also increases the proportion of altruists (almost non existent in the "no oath" treatment) and decreases the proportion of free riders.

5 Results II: Diagnostic experiments–why does the oath work?

We see that the oath increases contributions to the public good by 33 percent. The open question is why: does the oath change other-regarding behavior, or does the oath get people to go with their instincts to "do the right thing" (*i.e.*, cooperate) rather than ponder and reflect on the strategy of rational choice and its subsequent consequences.

5.1 Does the oath affect other-regarding behavior?

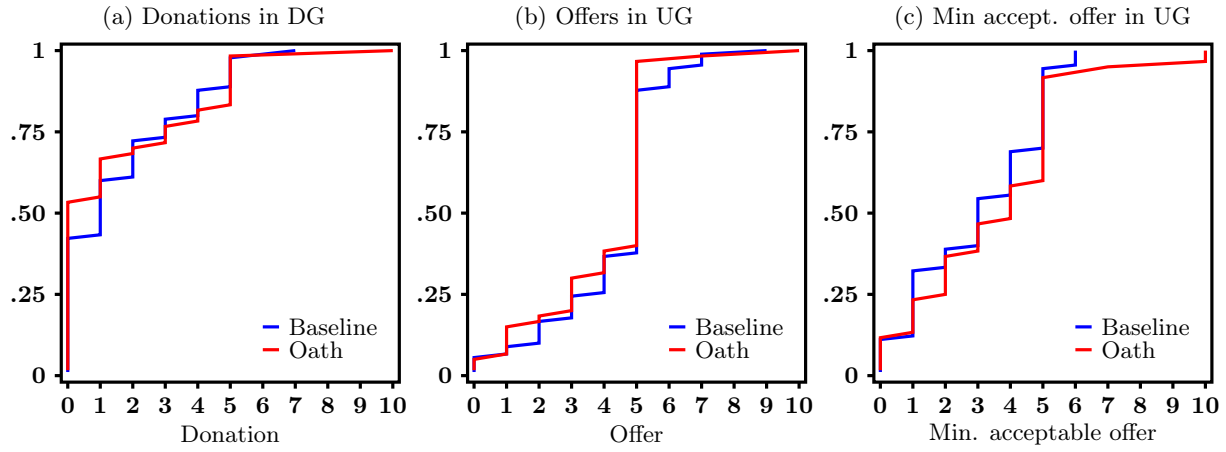
We first assess whether the oath works by directly impacting other-regarding preferences. We address this question in three steps. We first investigate altruistic and fairness preferences based on two classical zero-sum games – the Dictator game and the Ultimatum bargaining game. We then turn to the Trust game and consider reciprocity preferences (based on responders' behavior) and efficiency concerns (based on senders' behavior). In both cases, our null hypothesis is that the

Table 2: Distribution of types in the conditional Public Goods game by classification rule

Classification rule	Social type	No oath	Oath	Δ (Oath – No oath)
Very strict	Free rider	0.094	0.058	-0.036
		(0.0219)	(0.0215)	($p = .260$)
	Altruist	0.006	0.033	0.027*
		(0.0056)	(0.0165)	($p = .066$)
	Reciprocator	0.233	0.35	0.117**
		(0.0316)	(0.0437)	($p = .028$)
	Weak reciprocator	0.667	0.558	-0.109*
		(0.0352)	(0.0455)	($p = .058$)
Strict	Free rider	0.178	0.15	-0.028
		(0.0286)	(0.0327)	($p = .528$)
	Altruist	0.011	0.033	0.022
		(0.0078)	(0.0165)	($p = .179$)
	Reciprocator	0.233	0.35	0.117**
		(0.0316)	(0.0437)	($p = .028$)
	Weak reciprocator	0.578	0.467	-0.111*
		(0.0369)	(0.0457)	($p = .059$)
Loose	Free rider	0.283	0.225	-0.058
		(0.0337)	(0.0383)	($p = .260$)
	Altruist	0.017	0.042	0.025
		(0.0096)	(0.0183)	($p = .189$)
	Reciprocator	0.233	0.35	0.117
		(0.0316)	(0.0437)	($p = .154$)
	Weak reciprocator	0.467	0.383	-0.084**
		(0.0373)	(0.0446)	($p = .028$)

Note. Each column provides the distribution of subjects in each treatment over the four cooperation types defined by the classification in row (standard errors in parenthesis). Classification rules differ according to how liberal is the identification of free-riding behavior based on the average level of individual contributions in the conditional public good game, m: very strict ($m = 0$); strict ($m < 0.1$) or loose ($m < 0.2$). The last column provides mean comparisons between treatments, with p-values from Mann-Withney tests in parenthesis.

Figure 3: Behavior in the dictator and ultimatum games, EDF by treatment



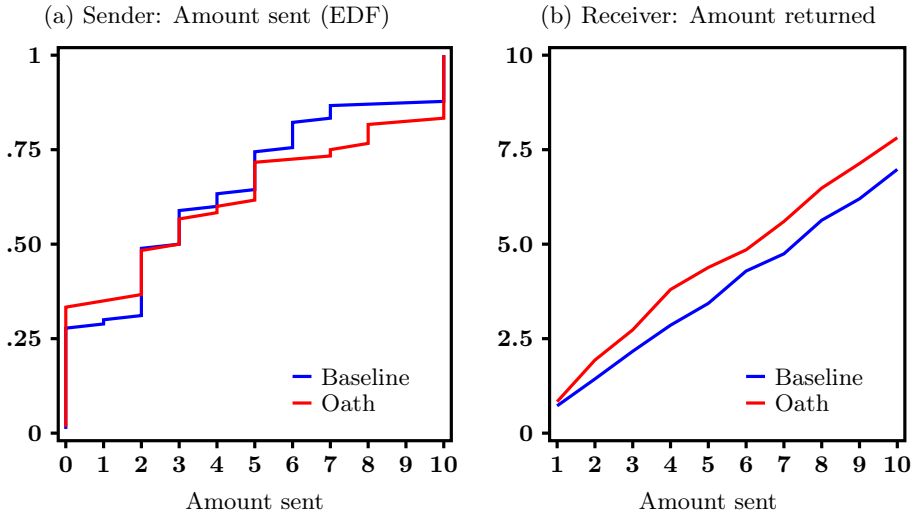
oath does not change other-regarding behavior. Finally, we investigate whether the oath procedure has an impact on subjects' self-reported measures of social preferences and risk aversion. To do so, we use the final questionnaire described in section 3, in which subjects are asked to answer a set of traditional survey questions aimed at eliciting those preferences.

We cannot reject the null hypothesis for the Dictator game. We see the EDFs are similar in the oath and Dictator baseline. ($p = 0.379$). Figure 3.a presents the EDF of donations of dictators for both treatments. The other share of rational self-interested dictators – those who gave zero – is 11.1% higher in the oath treatment, 53.3% in oath treatment and 42.2% in the baseline. This increase, however, is not significant ($p = 0.121$), which suggest the oath did not foster pro-social behavior in dictators game. If anything, the results suggest that the oath might have reinforced rational self-interested behavior.

A similar result emerges in the Ultimatum game. Figure 3.b, shows the EDFs of offers for both treatments. While the share of senders who propose equal splits increases relative to the Dictator game, we see that the oath has no impact on ultimatum-style offers. EDF are alike with $p = 0.939$. Figure 3.c presents the EDF of minimum acceptable offers of receivers for both treatments. Except for 3 subjects in the oath treatment who decide on a 10 minimum acceptable offer (maximum is 6 in the baseline), EDF are again similar ($p = 0.498$). The oath did not affect the behavior of receivers.[5]

[5]Note the use of the strategy method (elicitation of the minimum acceptable offer) rather than a sequential game ("yes" or "no" to the offer) may be of importance here. The oath might affect the behavior of receivers in a sequential situation as receivers *communicate* with senders through acceptance or refusal of the offer. Whereas, with the strategy method, receivers do not send negative signals by refusing unfair offer. Experimental evidence has shown that allowing receivers to send messages as well as accepting or refusing the offer decreases significantly refusal rates in a sequential setting (Xiao and Houser, 2005).

Figure 4: Behavior in the trust game, by treatment



Turning our attention to the Trust game, we again conclude that the oath does not directly impact social preferences. Figure 3.d shows the EDFs of the amount sent by senders, and Figure 3.d the amount returned by amount sent (recall that we also use the strategy method for the trust game). Results suggest the oath has no significant effect on the behavior of receivers. The amount sent by senders under oath are similar to those observed in the baseline (see Figure 3.d, $p = .221$). As for the behavior of senders, we observe a small increase in the amount returned (conditional on the amount sent considered). Figure 3.e exhibits a constant upward shift for amounts sent greater than 4. This upward shift is only significant for amounts sent ranging from 2 to 5 ($p = 0.028$, $p = 0.066$, $p = 0.031$ and $p = 0.057$). The oath did not make senders more trustful, but reciprocity levels among receivers seem to be marginally improved.

Finally, Table 3 reports the result of the comparison in self-reported social preferences between treatments (based on two-sided t-tests) –together with an experimentally validated question on risk aversion (Dohmen, Falk, Huffman, Sunde, Schupp, and Wagner, 2011). We can see that no statistically or economically significant difference arises between treatments in those self-reported measures, which we interpret as further evidence that the oath procedure does not directly affect subjects' social preferences or risk aversion levels.

5.2 Strategic thinking versus Norms

As a second diagnostic test, we now investigate whether the oath gets people to put more weight on their internalized social norms rather than strategic thinking. We find this to be the case, based on four pieces of evidence.

Our first piece evidence comes from self-reported expectations elicited at the end of the un-

Table 3: Self-reported social preferences between treatments

		No oath	Oath	Δ(Oath - No oath)
(i)	Cooperation	3.71	4.01	p=.342
(ii)	Altruism	3.90	3.71	p=.426
(iii)	Fairness	5.69	5.39	p=.228
(iv)	Trust (WVS)	0.66	0.73	p=.230
(v)	General trust	2.43	2.34	p=.301
(vi)	Trust in strangers	2.04	1.93	p=.184
(vii)	Risk aversion	6.12	5.69	p=.126

Note. Each column reports the mean answer (on a 0-10 scale for questions *(i)*-*(iii)* and *(vii)*; a binary answer for question *(iv)*; a 0-4 scale for questions *(v)* and *(vi)*) among each treatments subjects to the value survey questions. The last column reports t-tests of mean difference between the two treatments.

conditional public good contribution decision. We asked each subject to report *(ii.a)* whether he thought about the behavior of the other group members while making his own contribution decisions and, if so, *(ii.b)* how much he believed the others members of his group actually contributed on average. We observe no change in the responses to these two questions. The proportion of subjects who declared thinking about the contributions of others while making their own decision is 74.4% in the baseline and 74.2% in the oath treatment. For those who declared that the behavior of other subjects mattered, Table 4 presents mean estimated contributions by actual contribution and treatment. We can see that the more subjects expect free-riding the less they contribute, but beliefs are not significantly different between the two treatments.[6] Estimated average contribution is 4.34 in baseline and 4.80 in the oath treatment. The increase is not significant with $p = 0.135$.

The second piece of evidence comes from the answer to the normative question *(i)* how much do you think people should contribute to the common project. We observe that, under oath, significantly more subjects make a contribution to the public good that matches what they think people should contribute –50.0% under oath as opposed to 35.0% in the baseline– a 42.9% significant increase in proportion ($p = .014$).

We can interpret this result in two ways. First, subjects could act more consistently with their normative opinion about how much one should contribute to the common project. Second, normative opinions could be more consistent with actual contributions because of some form of cognitive dissonance (Festinger, 1957), *i.e*, subjects would *rationalize* their behavior in the public good game by stating a contribution in the normative question which equals their actual contribution. By doing so, subjects under oath would seek to avoid the cognitive dissonance

[6]No estimated contributions are available for actual contribution 8 and 9 in the oath treatment because of the few subjects who contributed 8 and 9 in the public good game, none of them thought about the contribution of other subjects. Although, the results suggest that thinking about others' behavior does not depend on one's actual contribution (results are available upon request).

Table 4: Beliefs and normative opinions about contribution – by actual contribution and treatment

Actual contribution	0	1	2	3	4	5	6	7	8	9	10	Mean	St.d.
Estimated average contribution of others													
Baseline	2.7	3.2	2.9	3.1	5.4	4.3	5.4	6.0	7.7	6.3	8.8	4.3	2.56
Oath	2.2	6.0	2.4	3.0	3.7	4.6	6.0	6.0	-	-	8.6	4.8	3.26
Normative opinion on contribution level													
Baseline	4.4	4.4	4.4	5.0	6.2	6.0	7.7	7.0	8.2	7.7	9.5	5.7	3.22
Oath	3.4	5.0	3.9	3.6	5.1	6.6	6.4	7.7	8.0	9.0	9.6	6.1	3.41

Note. For each treatment in row, and each level of actual contribution observed from the respondent, the table reports the average amount subjects report to the question about: their estimated average contribution of others (top panel); their normative opinion on what should be the level of contribution.

induced by any discrepancy between their normative and actual contributions.[7] Our data tends to favor the first interpretation, since we observe that the oath did not affect subjects' normative opinions on how much people should contribute *per se* (the mean value is 6.08 under oath as opposed to 5.73 in the baseline, with $p = 0.383$), but only the consistency between those normative opinions and actual behavior.[8]

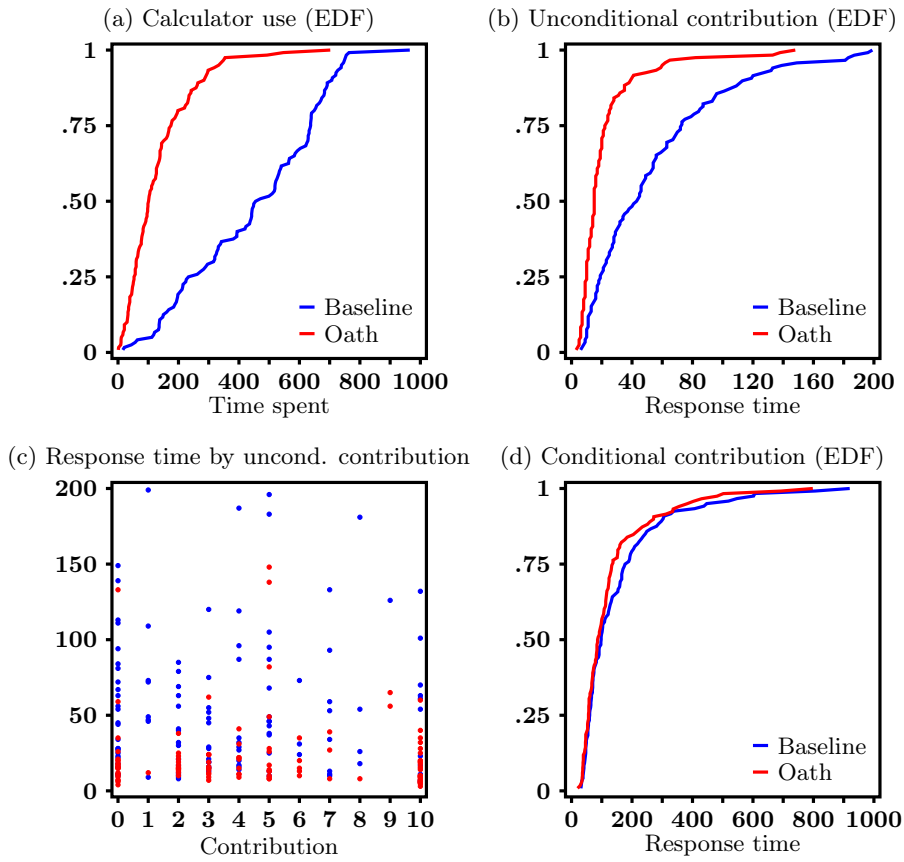
Our third piece of evidence comes from section 4, in which we observed that the oath had the main effect of turning weak reciprocators into perfect reciprocators – as opposed to dramatically impacting subjects' social preferences by turning, e.g., free riders into reciprocators or altruists. This observation is consistent with the finding that the oath has a limited impact on subjects' preferences and normative opinions, but increases the level of agreement between individual norms and observed behavior. To the extent that weak reciprocators can be usefully described as "*conditional cooperat[ors] with a self-serving bias*" (Fischbacher, Gächter, and Fehr, 2001, p. 401), subjects who go through the oath procedure and commit to holding themselves to a higher standard of integrity may put more weight on their own internal norm (*i.e.*, reciprocity) while making their decision, at the expense of strategic thinking and profit maximization.

Our final piece of evidence comes from the effect of the oath on decision times, *i.e.*, the time elapsed between the appearance of the decision screen and subjects' actual choice. We also recorded the total time spent experimenting with various scenarios on the earnings calculator screen. If the oath works mainly by inducing subjects to put more weight on internal norms rather than strategic thinking, then we should expect them to make more "intuitive" decisions, resulting in a significant decrease in decision times and in the time spent on the earnings calculator (Kahneman, 2003). Figure 5 shows the EDF of response time for unconditional and conditional decisions, as

[7] See Beauvois and Joule (1996) for a detailed study of *rationalization*.

[8] For those subjects who give an unconditional contribution strictly less than 10 we however observe a slight change in opinions. For these subjects, mean response to the normative question is 5.34 in the baseline and 4.72 in the oath treatment. The decrease, albeit small, is significant with $p = .068$. This suggests that the oath induces subjects who still engage in free-riding under oath into (more) rationalization of their behavior.

Figure 5: Time spent using the earnings calculator and response time in the public good game



well as the time spent on the earnings calculator screen. We indeed observe substantial differences.

The time spent on the earnings calculator screen under oath was considerably lower than without oath. In Figure 5.a, the EDF in the baseline first order dominates the oath treatment with $p < 0.001$. The median time spent on the calculator in baseline is 463 seconds but only 101 seconds in the oath treatment, a 80% decrease in median time.

Figure 5.b shows the EDF of response time for unconditional contributions. Again we see that the response time under oath decreases significantly relative to the baseline. Median response time is 15 seconds in the baseline and 44.5 seconds under oath, while the EDF of response time in baseline first order dominates the oath treatment ($p < 0.001$). The quicker response time does not depend on the contribution level. Figure 5.c shows the response time by unconditional contribution and treatment (unconditional contribution on the x-axis and response time on the y-axis). There is no clear pattern between response time and unconditional contribution in either treatment [9] and the downward shift in response time for subjects under oath emerges for every level

[9] The correlation coefficient between response time and unconditional contribution is 0.12 in the baseline and -0.05 under oath.

of unconditional contribution. The decrease in time spent using the calculator and in response time suggests that subjects under oath adopt a simplifying decision rule. This result is in line with the existing literature. Longer decision times are associated with choices that require more cognitive efforts (e.g. "profit maximization") while shorter decision times are associated with choices that are more intuitive and where subjects apply decision heuristics – which usually promotes cooperation (Rubinstein, 2007; Rand, Greene, and Nowak, 2012; Lotito, Migheli, and Ortona, 2013).

We do not find a difference in response time, however, for the conditional contribution decisions. Figure 5.d presents the EDF of response time in this task for both treatments. Elapsed time is longer than for the unconditional contribution task. This is because subjects give 10 contribution levels as compared to one only in the unconditional contribution task. There is now no difference between the EDF in the baseline and in the oath treatment ($p = 0.268$) suggesting that the conditional decisions are more straightforward than unconditional ones.

6 Conclusion

The literature on voluntary contribution in public good games has identified several mechanisms achieving higher levels of cooperation in social dilemma (Chaudhuri, 2011). Such improvement is obtained by implementing real economic commitment at the cost of private losses due to punishment, or welfare losses due by sorting. The aim of this paper is to design a non-monetary institution aimed at fostering cooperation in a social dilemma without affecting the social benefits of the situation. To that end, we asses whether a solemn oath of honesty is enough to create commitment towards cooperation.

We provide evidence on behavior in a one-shot, linear, voluntary contribution mechanism. In the main experimental treatment, subjects are offered to sign an oath to tell the truth and provide honest answers before the experiment takes place. We observe a significant rise in cooperation – one third more under oath as compared to no oath. This mainly comes from weak reciprocators more often behaving as strong reciprocators and altruists . Our experiment also provides evidence on the channels through which the effect occurs. First, we observe no change in subjects' social preferences based on either behavioral measures (dictator game, ultimatum bargaining game, and sender behavior in the trust game) or self-stated social preferences measures through social values survey like questions. Rather, we provide evidence that subjects more often behave consistently with their internalized social norms when under oath.

References

Abadie, A. (2002): "Bootstrap Tests for Distributional Treatment Effects in Instrumental Variable Model," *Journal of the American Statistical Association*, 97(457), 284–292.

Beauvois, J., and R.-V. Joule (1996): *A Radical Dissonance Theory*. Taylor and Francis, London.

Bremzen, A., E. Khokhlova, A. Suvorov, and J. Van de Ven (2015): "Bad News: An Experimental Study on the Informational Effects of Rewards," *Review of Economics and Statistics*, 97(1), 55–70.

Bull, C. (1987): "The Existence of Self-Enforcing Implicit Contracts," *Quarterly Journal of Economics*, 102(1), 147–159.

Burger, J. M. (1999): "The Foot-in-the-Door Compliance Procedure: A Multiple-Process Analysis and Review," *Personality and Social Psychology Review*, 3(4), 303–325.

Burlando, R. M., and F. Guala (2005): "Heterogeneous agents in public goods experiments," *Experimental Economics*, 8(1), 35–54.

Charness, G., and M. Dufwenberg (2006): "Promises and partnership," *Econometrica*, 74(6), 1579–1601.

Charness, G., and C.-L. Yang (2014): "Starting small toward voluntary formation of efficient large groups in public goods provision," *Journal of Economic Behavior & Organization*, 102, 119–132.

Chaudhuri, A. (2011): "Sustaining cooperation in laboratory public goods experiments: a selective survey of the literature," *Experimental Economics*, 14(1), 47–83.

Cinyabuguma, M., T. Page, and L. Putterman (2005): "Cooperation under the threat of expulsion in a public goods experiment," *Journal of Public Economics*, 89(8), 1421–1435.

Cohn, A., E. Fehr, and M. Maréchal (2014): "Business culture and dishonesty in the banking industry," *Nature*, 516, 86–89.

Dal Bo, E., and P. Dal Bo (2014): ""Do the right thing": The effects of moral suasion on cooperation," *Journal of Public Economics*, 117, 28–38.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. G. Wagner (2011): "Individual risk attitudes: Measurement, determinants, and behavioral consequences," *Journal of the European Economic Association*, 9(3), 522–550.

Ederer, F., and A. Stremitzer (2016): "Promises and Expectations," *Cowles Foundation Discussion Paper*, (1931).

Ellingsen, T., and M. Johannesson (2004): "Promises, threats and fairness*," *Economic Journal*, 114(495), 397–420.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010): "Testing guilt aversion," *Games and Economic Behavior*, 68(1), 95–107.

Farrell, J. (1987): "Cheap talk, coordination, and entry," *RAND Journal of Economics*, 18(1), 34–39.

Fehr, E., and C. F. Camerer (2004): "Measuring Social Norms and Preferences Using Experimental Games: A Guide for Social Scientists," *Foundations of Human Sociality*, 1(9), 55–96.

Fehr, E., and S. Gächter (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90(4), 980–994.

——— (2002): "Altruistic punishment in humans," *Nature*, 415(6868), 137–140.

Festinger, L. (1957): *A theory of cognitive dissonance*. Stanford University Press, CA.

Fischbacher, U., S. Gächter, and E. Fehr (2001): "Are people conditionally cooperative? Evidence from a public goods experiment," *Economics Letters*, 71(3), 397–404.

Gächter, S., and C. Thöni (2005): "Social learning and voluntary cooperation among like-minded people," *Journal of the European Economic Association*, 3(2-3), 303–314.

Galbiati, R., K. H. Schlag, and J. J. Van Der Weele (2013): "Sanctions that signal: An experiment," *Journal of Economic Behavior & Organization*, 94, 34–51.

Galbiati, R., and P. Vertova (2008): "Obligations and cooperative behaviour in public good games," *Games and Economic Behavior*, 64(1), 146–170.

Goldstein, N. J., R. B. Cialdini, and V. Griskevicius (2008): "A room with a viewpoint: Using social norms to motivate environmental conservation in hotels," *Journal of consumer Research*, 35(3), 472–482.

Greiner, B. (2015): "Subject pool recruitment procedures: organizing experiments with ORSEE," *Journal of the Economic Science Association*, 1(1), 114–125.

Guala, F. (2012): "Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate," *Behavioral and Brain Sciences*, 35(01), 1–15.

Gunnthorsdottir, A., D. Houser, and K. McCabe (2007): "Disposition, history and contributions in public goods experiments," *Journal of Economic Behavior & Organization*, 62(2), 304–315.

Gürerk, Ö., B. Irlenbusch, and B. Rockenbach (2006): "The competitive advantage of sanctioning institutions," *Science*, 312(5770), 108–111.

Hergueux, J., and N. Jacquemet (2015): "Social Preferences in the online laboratory: A randomized Experiment," *Experimental Economics*, 18(2), 252–283.

Ismayilov, H., and J. Potters (2015): "Why do promises affect trustworthiness, or do they?," *Experimental Economics*, Forthcoming.

Jacquemet, N., R.-V. Joule, S. Luchini, and J. F. Shogren (2013): "Preference Elicitation under Oath," *Journal of Environmental Economics and Management*, 65(1), 110–132.

Joule, R.-V., F. Girandola, and F. Bernard (2007): "How Can People Be Induced to Willingly Change Their Behavior? The Path from Persuasive Communication to Binding Communication," *Social and Personality Psychology Compass*, 1(1), 493–505.

Kahneman, D. (2003): "Maps of Bounded Rationality: Psychology for Behavioral Economics," *American Economic Review*, 93(5), 1449–1475.

Katzev, R., and T. Wang (1994): "Can commitment change behavior ? A case study of environmental actions," *Journal of Social Behavior and Personality*, 9, 13–26.

Kiesler, C. (1971): *The psychology of commitment. Experiments liking behavior to belief.* Academic Press, New York.

Kim, J.-Y. (1996): "Cheap talk and reputation in repeated pretrial negotiation," *RAND Journal of Economics*, 27(4), 787–802.

Klein, B., and K. B. Leffler (1981): "The role of market forces in assuring contractual performance," *Journal of Political Economy*, 89(4), 615–641.

Levin, J. (2003): "Relational incentive contracts," *American Economic Review*, 93(3), 835–857.

Lotito, G., M. Migheli, and G. Ortona (2013): "Is cooperation instinctive? Evidence from the response times in a public goods game," *Journal of Bioeconomics*, 15(2), 123–133.

McAdams, R. H. (2000): "An attitudinal theory of expressive law," *Oregon Law Review*, 79, 339–390.

McAdams, R. H., and J. Nadler (2005): "Testing the Focal Point Theory of Legal Compliance: The Effect of Third-Party Expression in an Experimental Hawk/Dove Game," *Journal of Empirical Legal Studies*, 2(1), 87–123.

Page, T., L. Putterman, and B. Unel (2005): "Voluntary association in public goods experiments: Reciprocity, mimicry and efficiency," *Economic Journal*, 115(506), 1032–1053.

Pallack, M., D. Cook, and J. Sullivan (1980): "Commitment and energy conservation," in *Applied Social Psychology Annual*, ed. by L. Bickman, pp. 235–253. Beverly Hills, CA: Sage.

Piovesan, M., and E. Wengstrom (2009): "Fast or fair? A study of response times," *Economics Letters*, 105(2), 193–196.

Posner, E. A. (1998): "Symbols, signals, and social norms in politics and the law," *The Journal of Legal Studies*, 27(S2), 765–797.

——— (2000): "Law and social norms: The case of tax compliance," *Virginia Law Review*, pp. 1781–1819.

Rand, D. G., J. D. Greene, and M. A. Nowak (2012): "Spontaneous giving and calculated greed," *Nature*, 489(7416), 427–430.

Resnick, P., R. Zeckhauser, J. Swanson, and K. Lockwood (2006): "The value of reputation on eBay: A controlled experiment," *Experimental Economics*, 9(2), 79–101.

Rubinstein, A. (2007): "Instinctive and Cognitive Reasoning: A Study of Response Times," *Economic Journal*, 117(523), 1243–1259.

Sekhon, J. (2011): "Multivariate and Propensity Score Matching Software with Automated Balance Optimization," *Journal of Statistical Software*, 42(7), 1–52.

Shu, L., N. Mazar, F. Gino, D. Ariely, and M. Bazerman (2012): "Signing at the beginning makes ethics salient and decreases dishonest self-reports in comparison to signing at the end," *PNAS*, 109(38), 15197–15200.

Vanberg, C. (2008): "Why do people keep their promises? an experimental test of two explanations," *Econometrica*, 76(6), 1467–1480.

Xiao, E., and D. Houser (2005): "Emotion expression in human punishment behavior," *Proceedings of the National Academy of Sciences*, 102, 7398–7401.

Appendix

A Oath form used in the experiment

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

SOLEMN OATH

I undersigned ..... swear upon my honor that, during the whole experiment, I will:

Tell the truth and always provide honest answers.

Paris, ... Signature.....

Paris School of Economics, 48 Boulevard Jourdan 75014 Paris – France.