Referenda under Oath*

Nicolas Jacquemet† Alexander G. James‡ Stéphane Luchini§ Jason F. Shogren¶

September 2013

Abstract

Herein we explore whether a solemn oath can eliminate hypothetical bias in a voting referenda, a popular elicitation mechanism promoted in non-market valuation exercises for its incentive compatibility properties. First, we reject the null hypothesis that a hypothetical bias does not exist. Second, we observe that people who sign an oath are significantly less likely to vote for the public good in a hypothetical referenda. We complement this evidence with self-reported measures of honesty which confirm that the oath increases truthfulness in answers. This result opens interesting avenues for improving the elicitation of preferences in the lab and beyond.

Keywords: Dichotomous Choice Mechanism; Hypothetical bias; Oath; Preference revelation.

JEL Classification: C9; H4; Q5.

1 Introduction

Stated preference methods remain a popular tool to value non-marketed goods such as environmental quality (e.g., Loureiro, Loomis, and Vazques, 2009), reduced risks to life and limb (Svensson, 2009), and recreation (e.g., Deisenroth and Bond, 2009). But stated preference methods remain susceptible

*We thank Maxim Frolov for his assistance in running the experiments and Ivan Ouss for his efficient research assistance. Financial support from the Health Chair, a joint initiative by PSL, Université Paris-Dauphine, ENSAE and MGEN under the aegis of the Fondation du Risque (FDR) and the Bugas Fund is gratefully acknowledged. Shogren thanks the Norwegian University of Life Sciences for the support. Nicolas Jacquemet acknowledges the *Institut Universitaire de France*.

†Université de Lorraine (BETA) and Paris School of Economics. 3 Place Carnot, 54035 Nancy. Nicolas.Jacquemet@univ-lorraine.fr

‡Center for the Analysis of Resource Rich Economies in the Department of Economics at the University of Oxford, Oxford, UK. alexander.james@economics.ox.ac.uk

§Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS, 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

to complaints of hypothetical bias – the gap between stated intentions and real economic commitments.[1] In general, the extant literature has collected a long line of evidence that hypothetical bias exists across numerous types of mechanisms designed to reveal preferences truthfully, including the popular valuation institution of binary referendum voting (see e.g., Cummings, Elliott, Harrison, and Murphy, 1997).[2]

Social psychology offers one explanation of hypothetical bias based on the lack of commitment to truth telling (Jacquemet, James, Luchini, and Shogren, 2011). Commitment theory posits a person is more likely to tell the truth after first making a strong promise (see Joule and Beauvois, 1998). Economic experiments support this idea. After pre-play communication, people who make verbal promises about future actions are more likely to keep them when playing in both hold-up and trust games (Ellingsen and Johannesson, 2004; Charness and Dufwenberg, 2006). The solemn oath is a time-tested mechanism used to promote commitment – the bond between a person and telling the truth (see e.g., Sylving, 1959; Kiesler and Sakumura, 1966; Schlesinger, 2008). According to this view, the oath acts as a foot-in-the-door that makes subjects more likely to comply with the content of their promise. In addition, the commitment is stronger when the promise is selected freely and voluntary (Jacquemet, James, Luchini, and Shogren, 2011). The oath is a pragmatic real-world commitment device that when publicly expressed, taken freely and signed, appears an extreme and more accentuated commitment device than a verbal promise or a written undertaking. Jacquemet, Joule, Luchini, and Shogren (2013) provide testbed experiments using Vickrey auctions and confirm the ability of a truth-telling oath to improve the reliability of revealed preferences. Extending the results to a referendum voting format is a key step from the perspective of applying the oath to CV studies in practice. According to the NOAA panel, for instance, a voting referendum has two advantages. First, relative to alternative formats, subjects may find it more familiar and realistic. For example, casting a vote on a proposition to raise taxes to pay for public roads is a CV survey in which the elicitation format is that of the voting referendum. Second, the NOAA report concluded the referendum format is less susceptible (though not immune) to hypothetical bias relative to other formats such as an open ended questionnaire.

Herein we use experimental methods to test whether the oath will act as a commitment device in referendum voting. In the "oath" treatments, subjects can freely sign a form by which they swear to tell the truth during the experiment. Our experimental results provide additional support for the

[1] While exceptions exist, hypothetical bias persists; see Diamond and Hausman (1994); Murphy, Stevens, and Weatherhead (2005); Jacquemet, Joule, Luchini, and Shogren (2011).

[2] In 1993 the National Oceanic and Atmospheric Administration (NOAA) appointed the Contingent Valuation Panel to determine under what conditions CV yields reliable value estimates. Among a host of recommendations, the panel reported that proper CV surveys should be administered via the voting referendum format (Arrow, Solow, Portney, Leamer, Radner, and Schuman, 1993). A voting referendum has two advantages. First, relative to alternative formats, subjects may find it more familiar and realistic. For example, casting a vote on a proposition to raise taxes to pay for public roads is a CV survey in which the elicitation format is that of the voting referendum. Second, the NOAA report concluded the referendum format is less susceptible (though not immune) to hypothetical bias relative to other formats such as an open ended questionnaire.

notion that the oath can reduce hypothetical bias. The oath reduces the proportion of "yes" votes, however not to the extent of observing no significant difference between real and hypothetical voting behavior. We complement this evidence with self-reported measures of honesty, which increase under oath as compared to the baseline hypothetical treatment. The oath seems to commit people to a behavioral act: to vote in a hypothetical referendum as if it were real. The oath appears to create the commitment needed to better link intentions and actions in non-market valuation, and perhaps beyond (also see Jacquemet, Luchini, Shogren, and Zylbersztejn, 2011).

2 Experimental Design

We use a 2x2 experimental design in which the treatments are: (i) hypothetical and real referenda and (ii) voting with and without a solemn oath to tell the truth. The design of the experiment closely follows Jacquemet, Joule, Luchini, and Shogren (2013), who studied the oath in both an induced value and a homegrown value context using a second price auction. We adapt the original design to the case of referendum voting.

2.1 Preference elicitation

We elicit preferences towards a private good with non-market attributes: adopting a dolphin through a monetary donation to the *World Wide Fund* (hereafter WWF), a well-known non-governmental organization devoted to "protecting the future of nature".[3] The WWF offers people the opportunity to "adopt" an endangered animal species. This adoption takes the form of an individual donation to a program aimed at fighting threats like habitat loss and poaching faced by endangered animals. Depending on the amount of the donation (among three possible values), donators are sent private gifts such as an adoption certificate, a photograph of the animal, a cuddly stuffed toy dolphin, a gift box, and so on. For the purpose of our experiment, this procedure has the attractive feature of ensuring the credibility of the donation, thanks both to the WWF label and to the documentation associated with donation. We chose the entry-level offer, *i.e.*, an adoption certificate and photograph are sent for each 25 USD (around 19 Euros when the experiments took place) donation to the WWF. Since the photograph and the adoption certificate is symbolic in nature, this reduces the risk of valuations being influenced by "by-product" goods, such as a cuddly stuffed toy or a gift box.

The adoption procedure is described to the subjects using a French-language, slightly modified version of the official web page set up by the WWF.[4] The page provides a short description of a dolphin's life and of the WWF and, more importantly, a detailed presentation of the donation program

[3]The WWF was formerly named the *World Wildlife Fund*, which remains its official name in the United States and Canada. Since 2001, the WWF has been named the *World Wide Fund* in all other countries. More information about the WWF can be found at `http://www.worldwildlife.org/about/`.

[4]The original page in English is available at `gifts.worldwildlife.org/gift-center/gifts/Species-Adoptions/Dolphin.aspx`..

and the documentation (gifts) sent should a subject adopt a dolphin. In the written experimental instructions, the good is described as follows:

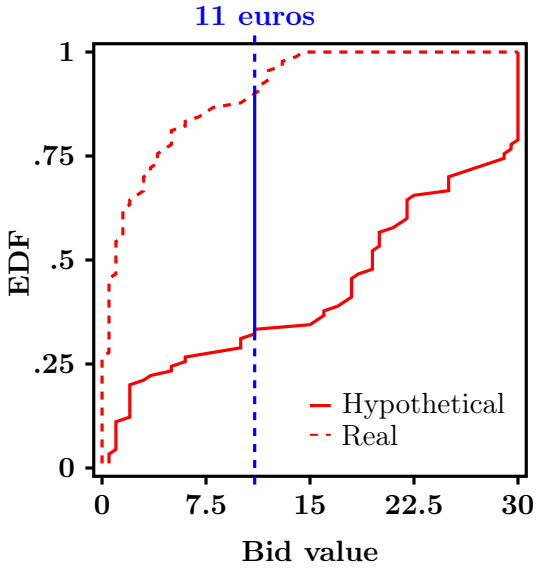
> *The World Wide Fund for Nature, better known as the WWF, is an international non-governmental organisation for the protection of nature and of the environment, fully committed to sustainable development. The head office is in Gland, Switzerland, and the association has more than 4.7 million members worldwide, with an operational network in 96 countries. It is a private organisation aimed at protecting wild animals and their habitats as well as nature in general, which it does by collecting funds for specific programs. Principally, it keeps a watchful eye on whether international regulations are being respected,restores damaged natural areas and provides training. As a way of financing its environmental protection activities, the WWF offers private individuals the opportunity to adopt an animal from an endangered species. The funds thereby collected enable the WWF to continue protecting the environment and preserving species diversity.*

The donation decision is taken within groups of five subjects, through majority voting. Since we divide each 20-subjects session into smaller groups of five subjects (once for all, *i.e.* groups remain the same for the whole experiment), four groups in each session are involved in four independent referenda. Subjects are asked a vote on the adoption page, by clicking on either of two buttons: YES or NO. We reduced the noise in elicited preferences by repeating five times the referendum in fixed groups. At the end of the sequence, one auction round out of the five is randomly drawn. The votes of the randomly drawn referendum decides whether the adoption passes: if more than 50% vote "yes", everyone in the group adopts a dolphin.

A well-known concern with dichotomous choice mechanisms is that it provides a point identification of the underlying preferences: if the price submitted to the vote is either too low or too high, then it is non-informative about the extent of a hypothetical bias, and how to reduce this bias – because elicited preferences become observationally insensitive to the elicitation environment. In the same context as the one we study here, Jacquemet, Joule, Luchini, and Shogren (2013) elicit the whole demand curve of their subjects regarding the adoption of a dolphin but preferences are elicited in a second price Vickrey auction rather than by a dichotomous choice mechanism.[5] Figure 1 provides the empirical distribution function (EDF) of individual bid values for the good in both the real (*i.e.* with monetary incentives) and hypothetical (baseline) conditions. We use this preliminary evidence to calibrate the amount of the donation to 11 Euros: this is a price at which preferences *(i)* exhibit a significant hypothetical bias and *(ii)* is likely to avoid corner solutions in observed votes. With a price set at this level, the good sold in the experiment is cheaper in the lab than in the market, so we subsidize the winning donation to reach the market price when monetary incentives are binding. Subjects are not told anything about this subsidy. This is enough to protect our data from the censoring issue raised by, e.g, Cherry, Frykblom, Shogren, List, and Sullivan (2004). To confirm that the observed values are independent of field opportunities, some items assessing subjects' knowledge about the procedure are included in a debriefing questionnaire.

[5]The only difference with the referendum setting is that subjects bid for adopting one dolphin and the final donation to the WWF equals the second highest bid value whereas in the referendum the donation to the WWF equals the sum of the contributions of the group, if a majority votes "yes".

Figure 1: Preferences for the good elicited in a Vickrey Auction (*Source:* Jacquemet, Joule, Luchini, and Shogren, 2013)



2.2 Subjects endowment

Our focus on donation behavior requires that subjects spend some money in the experiment. To improve the external validity of our design – in particular, to avoid an inflation in the number of false zeros – we want the subjects to enter the referendum with some positive experimental earnings to be spent on the donation. This would mean giving subjects a large show-up fee for participating in the experiment. Evidence suggests behavior can differ depending on whether wealth is "windfall" or is earned (also called *endowment effect*, see, among others, Cherry, Frykblom, and Shogren, 2002). In the context of demand revelation using Vickrey auctions, Jacquemet, Joule, Luchini, and Shogren (2009) show that earned money makes a significant difference to bidding behavior as compared to windfall wealth. In line with these results, and to be as close as possible to actual stated preference surveys in the field, we use an earned-wealth design. This also replicates a common feature of homegrown valuation experiments focusing on hypothetical bias (e.g., Cummings and Taylor, 1999; Cummings, Elliott, Harrison, and Murphy, 1997).

Earned wealth is implemented through a preliminary stage during which the subjects are asked to answer 20 general knowledge questions. Accompanying each question is a list of four possible answers. The set of questions was taken from the annals of the "Concours de Catégorie B de la fonction publique" which is a civil service entry test for those who hold at least the French baccalaureate.[6] This is appropriate to discriminate between undergraduate students. Accompanying each question is a list of four possible answers. Subjects are explicitly told that one and only one out of the four is true, and that monetary earnings labelled in ECU (*Experimental Currency Unit*) are proportional to

[6]Our source is `http://pagesperso-orange.fr/bac-es/qcm/annales_c02_r01.html`.

correct answers. The position of the correct answer is randomized between questions and the ordering of questions is kept the same for all subjects in all treatments.[7]

2.3 Experimental treatments

We use two treatment variables implemented through a factorial design – real /hypothetical combined with oath/no oath. All four treatments are performed using a between subjects design – each subject participates in only one out of the four treatments. Our benchmark situation is the hypothetical bias that arises in the standard laboratory situation, *i.e.* with no oath.

The REAL and HYPOTHETICAL treatments only differ regarding the monetary consequences of the adoption. In the real conditions, each subject belonging to groups in which the vote passed does make a donation. The donation is subtracted from subjects' earnings. In the hypothetical condition, by contrast, the donation votes are declarative – no funds are transferred to the WWF and no adoption certificate is sent to the adopters. The description of the donation to the subjects in the written experimental instructions as well as on the adoption page, is adapted accordingly:[8]

> *During this part, we ask you to imagine that you were taking part* **[REAL: you are going to take part]** *in this operation by making a donation, which would be* **[REAL: will be]** *deducted from your experimental earnings, to adopt a dolphin. The sums collected during this part would be [***REAL: will be]** *passed on by us to the WWF, to support their environmental protection activities. Your donation to the WWF would be* **[REAL: will be]** *recorded on an official certificate, which would be* **[REAL: will be]** *sent to your home address. We ask you to make your decisions as if, in this part, we were genuinely offering you the opportunity to adopt a dolphin, according to the procedure described below. The decisions made during this part are not, however, taken into account when calculating your Euro experimental earnings. In actual fact, regardless of your decisions, you will not be adopting a dolphin and your experimental earnings will not be affected.* **[REAL: We will genuinely make it possible for you to adopt a dolphin if you so decide, according to the procedure described below. The decisions made during this part are taken into account when calculating your Euro experimental earnings. This means that if you adopt a dolphin, your donation will be deducted from your earnings.]**

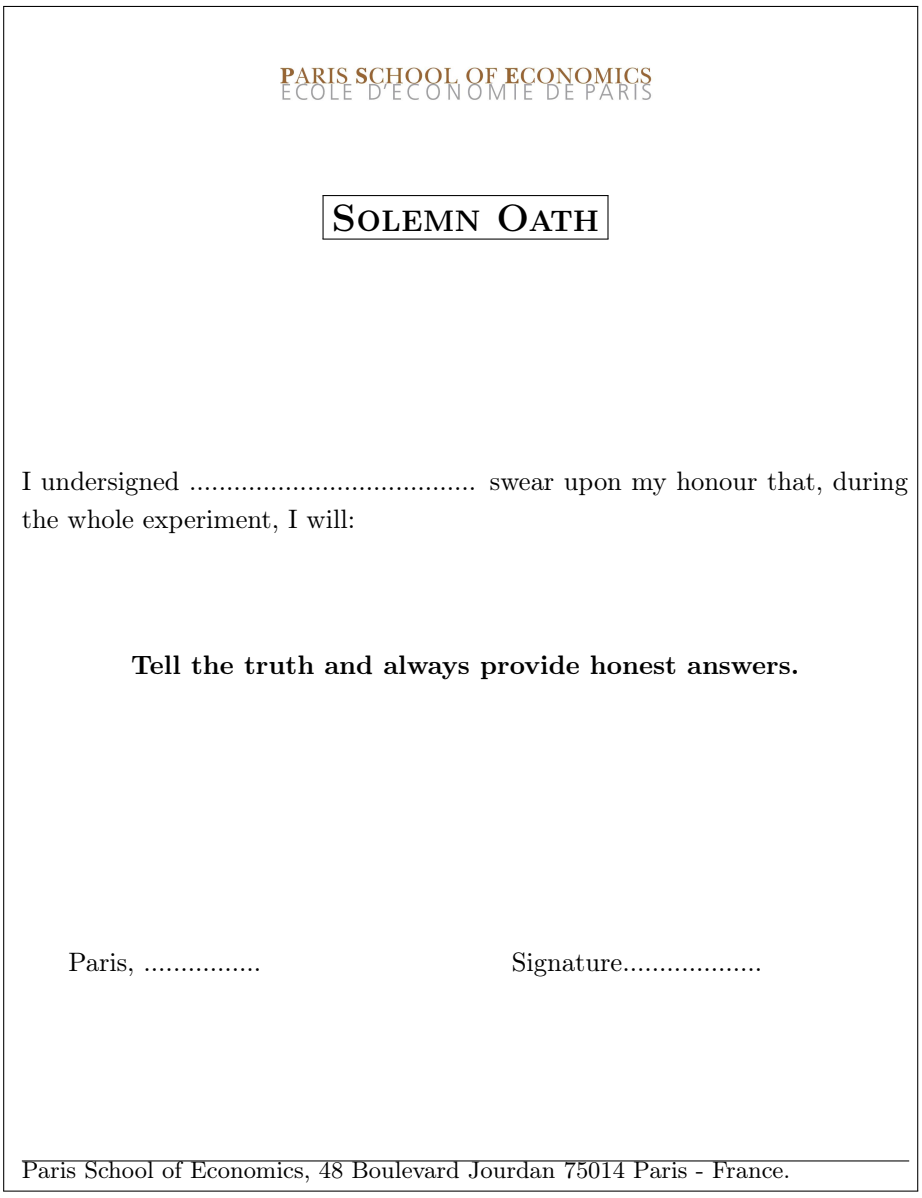
All other experimental features are kept the same in these two treatments. In particular, earnings stemming from the quiz are real in all treatments to avoid unwarranted wealth differences between our treatments.

The only change to the procedure in the OATH treatments is a preliminary stage based on an oath form. The oath, provided in Figure 2, reads, "*I, the undersigned swear upon my honor that during the entire experiment, I will tell the truth and always provide honest answers.*" This solemn oath is distributed for signing before any information is provided about the experiment. Note, the

[7] The data on observed answers are not commented on here; the full list of questions and data are available from the authors upon request.

[8] We follow Cummings and Taylor (1999) in replacing the affirmative language used in real conditions ("you *will* participate in the adoption procedure", "you *will* adopt a dolphin", "we *commit* ourselves to sending your donation to the WWF") with a subjunctive language in the hypothetical ones: "we want you to *suppose you were to* participate in the adoption procedure", "you *would* adopt a dolphin", "we *would commit* ourselves to sending your donation to the WWF" (italics added). The changes specific to the REAL treatment appear in brackets.

Figure 2: Oath form used in the experiments



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

SOLEMN OATH

I undersigned ..... swear upon my honour 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.

informational content of the oath focuses on truth-telling in itself, and does not describe either the hypothetical bias issue or the potential shortcomings of CV studies under hypothetical incentives. Signing the oath was not required to participate in oath treatments: the refusal rate is 1.7% (1 subject over 60) in the HYPOTHETICAL-OATH treatment and 3.3% (2 subjects over 60) in the REAL-OATH treatment. The acceptance rate prevents the results from being influenced by endogenous selection of subjects into the truth-telling promise.[9]

[9]We examine the data according to an intention to treat procedure. Note that none of the results are sensitive to this choice given so few subjects refused to sign the voluntary oath.

2.4 Self-reported honesty and happiness

At the end of the experiment, we complement our data with a set of self-reported attitudinal and happiness questions – note that none of these questions are incentivized. In all treatments and all sessions, we use two different sets of questions asked in the same order.[10] In the oral instructions, we insist that this questionnaire is only declarative although we expect subjects to take it seriously.

First, it has been argued that respondents in CV surveys express positive attitudes towards public goods or concerns for society problems rather than preferences (Kahneman and Sugden, 2005). To control for this dimension, we ask subjects to answer a set of questions to elicit their attitudes towards the WWF and control their knowledge about the WWF's adoption procedure. Second, subjects are asked how honest they think they were in their votes in the experiment on a numerical scale ranging from 1 (totally dishonest) to 7 (totally honest) (see, e.g. Mijovič-Prelec and Prelec, 2010). Finally, to assess whether subjects felt more pressure under oath, we also elicit the level of happiness measured on a typical 7-point scale (from very happy to very unhappy). We know that such self-reported attitudinal information should be interpreted with caution. For each of these variables, we focus on treatment variations rather than absolute values.

2.5 Experimental procedures

Three 20-subject sessions per treatment (12 sessions and 240 subjects overall), were conducted in the LEEP laboratory in Paris in May-July 2013.[11] Since each subject posts five votes for adopting a dolphin, this provides 300 observations for each treatment. On arrival, each subject signs an individual consent form and enters the lab. This form is mandatory for participation in the experiment. In the oath treatments only, subjects are then asked to take a truth-telling oath. A computer is then randomly assigned to each subject and a monitor distributes and reads aloud the instructions.

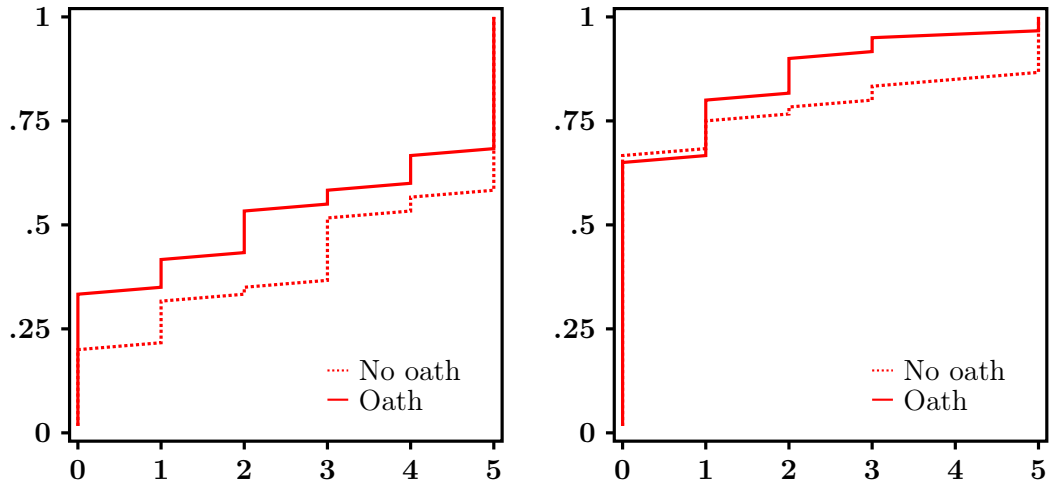
The experiment begins by asking the subjects to fill out a computerized questionnaire about socio-economic characteristics. The instructions of each part of the experiment are distributed and read aloud just before it starts and participants are encouraged to ask clarifying questions, privately answered by the monitor. The first part of the experiment is the quiz (questions along with the four possible answers are displayed one after the other). Subjects are provided information on their score only at the end of the quiz along with their corresponding earnings in ECU. The payment rate is 2 ECU *per* correct answer and the exchange rate is 3 ECU for 1 €. With an expectation of ten correct answers out of 20, the average monetary earnings for the quiz would be 7 €, (payment is rounded up to the next 50 cents), which makes 17 € in total once added to the 10 € show-up fee.

The second part of the experiment is the adoption referendum. The subjunctive language illustrated in the previous section is used throughout the instructions to differentiate the REAL and

[10] The full list of all sets of questions is presented in the Appendix, Section A.

[11] Please visit `http://leep.univ-paris1.fr/accueil.htm` for details. The experiment were computerized using a software program developed under REGATE (Zeiliger, 2000) and participants were recruited based on ORSEE (Greiner, 2004).

Figure 3: Distribution of "yes" responses by treatment



(a) Hypothetical (b) Real

HYPOTHETICAL treatments. Once the instructions have been read aloud, subjects are offered to answer a questionnaire to check their understanding. Once all questions have been answered, the second part starts. At the end of the experiment, subjects are asked to answer a computerized debriefing questionnaire. The questionnaire collects various individual information (e.g., gender and age) as well as, among other things, the level of knowledge and the level of agreement of the subjects regarding the WWF and its actions, their level of honesty (and that of other subjects) during the experiment, their level of happiness and whether they have participated in other experiments or not.[12] Finally, the monitor pays each subject privately in cash.

3 Results

To summarize individual behavior, we compute the total number of "yes" responses for each subject, which varies from zero (if the subject votes "no" in all five rounds) to five (if the subject votes "yes" in all five rounds). Figure 3 presents the empirical distribution functions (EDF) of the total number of "yes" responses by treatment. In the hypothetical treatments (Figure 3.a), we observe that the EDF in hypothetical significantly first order dominates the EDF in hypothetical under oath with $p = .065$.[13]

[12]The list of questions is provided as supplementary material, Section **??**.

[13]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). The Bootstrap is implemented by drawing observations under the null that votes are identical in both treatments. The procedure accounts for potential correlation between the five votes of the same subject and for asymmetry in the empirical distribution of votes. The procedure is based on bootstrapping subjects and their five votes in the sample, instead of considering independent votes, i.e.,

Table 1: Treatments and Summary Statistics

Treatment		Round 1	2	3	4	5	All
Hypothetical no oath	Yes	71.7%	56.7%	53.3%	65.0%	58.3%	61.0%
	Adoptions (#)	9	9	6	9	8	41 (68.3%)
Hypothetical with oath	Yes	55.0%	50.0%	46.7%	48.3%	46.7%	49.3%
	Adoptions (#)	7	5	6	6	5	29 (48.3%)
Real no oath	Yes	27.1%	18.6%	25.4%	18.6%	23.7%	22.7%
	Adoptions (#)	0	0	2	0	2	5 (8.3%)
Real with oath	Yes	11.7%	11.7%	16.7%	20.0%	15.0%	15.0%
	Adoptions (#)	1	1	0	1	1	3 (5.0%)

Note. For each treatment, the Table provides the percentage of "yes" votes observed by period and overall as well as the number of adoptions realized. There were three sessions (60 subjects and 12 groups) per treatment.

"Yes" responses are significantly shifted down by signing the oath at the individual level. This shift is explained by an increasing number of subjects always voting "No" under oath (20% of subjects in hypothetical doing so and 33.3% under oath) and a decreasing number of subjects always voting "Yes" (43% in hypothetical and 33.3% under oath). EDF in real treatments (Figure 3.b) exhibit the same shift of "yes" responses under oath but to a lesser degree ($p = .162$). The shift is now explained by a decreasing number of subjects always voting "Yes" (15% in real and .5% under oath). These figures support our main result: having subjects sign a truth telling oath before participation to a dichotomous choice mechanism significantly shifts hypothetical voting behavior downwards. In the next sections, we turn to two additional outcomes of the experiment: the resulting aggregate behavior and the effect of the treatments on self-reported attitudes.

3.1 Aggregate outcomes

Table 1 summarizes the votes elicited in each treatment and the resulting number of adoptions.[14] First, we reject the null hypothesis that a hypothetical bias does not exist when voting over contributions to the WWF – confirming our *ex ante* presumptions. Overall, 61.0% of the subjects voted "yes" in the Hypothetical no oath treatment; whereas about 22.7% voted "yes" in the Real no oath treatment. The difference is significant according to a bootstrap proportion test with $p < .001$. At the group level, observed votes lead to 41 adoption decisions (68.3%) in hypothetical whereas only 5 adoption decisions (8.3%) are made in real no oath. Second, signing the oath leads to a 20% decrease in the "yes" responses in the hypothetical condition, from 61% to 49.3%. The p-value of this decrease according to a unilateral bootstrap proportion test is $p = .125$. This results in a 30% drop in the

bootstrapping on votes. The number of replications is 9999.

[14]Recall that in each treatment, the actual decision to adopt depends on a random draw of one of the 5 voting rounds. In the real conditions, 3 groups of 5 subjects finally made a donation at the end of the experiment.

Table 2: Probit regression on treatment variables

	Coefficient	Marginal effects	p-value
Treatment effects			
Constant	-1.716	-	.114
Hypothetical	2.942	0.693	.000
Hypothetical × Oath	-0.817	-0.172	.078
Real × Oath	-0.278	-0.067	.633
Controls			
Age	0.014	0.026	.543
Male	0.102	0.003	.809
Occupational status (ref. is employed)			
Unemployed	-1.196	-0.184	.167
Student no grant	-0.454	-0.116	.566
Student with a job	0.431	0.130	.733
Student with a grant	-1.606	-0.226	.114

Note. Individual random effect panel Probit model of individual yes vote on treatment dummies and individual characteristics ($n = 239 \times t = 5$). The *endogenous* variable is the "yes" vote. *Round* (fixed) effects are controlled for in the estimation but omitted here. Joint nullity test: Wald = 51.91 with $p<$ 0.001.

adoption rate when subjects are under oath (from 68% to 48%), a figure however still greater than in the real condition (8%). In the real treatment, the oath induces a slight decrease in the "yes" response rate, from 22.7% to 15%, a difference that is not statistically significant ($p = .198$).[15]

We assess the robustness of the results by conditioning the effect of the treatments on participants' characteristics.[16] Table 2 provides the results from a random effect panel probit regression of the decision to vote "yes" on individual characteristics, round dummies and treatment effects measured by three dummy variables (*Hypothetical*, *Hypothetical* × *Oath* and *Real* × *Oath*). The reference

[15] Overall, our estimates of the treatment effects on subjects' preferences remain quite imprecise despite important quantitative differences. This illustrates the difficulty to testbed elicitation mechanisms based on discrete choice elicitation formats – in which continuous underlying preferences are reduced to 2 observables ranges. We note, however, that the preferences we elicit are statistically very close to the one observed in Jacquemet, Joule, Luchini, and Shogren (2013). To that end, we simulate how subjects observed in Jacquemet, Joule, Luchini, and Shogren (2013) would have voted in a 11 euros contribution referendum mechanism: a bid equal or higher to 11 euros is interpreted as a "yes" vote and "no" otherwise. We then generate confidence intervals from both experiments based on a bootstrap procedure. With a 5% level of confidence the results are:

	Hypothetical		Real	
	No oath	With Oath	No oath	With Oath
Referendum	[52.3%; 69.6%]	[40.3%; 58.3%]	[14.6%; 30.7%]	[9.6%; 21.0%]
Vickrey Auction	[52.2%; 84.0%]	[23.0%; 49.0%]	[2.0%; 20.0%]	—

[16] Participants characteristics include gender, age, occupational status and whether or not the subject attended to lab experiments in the past.

observation is a subject in the real treatment. The coefficient associated with the dummy variable *Hypothetical* is positive and significant at the 1% level, indicating a clear hypothetical bias. Being in the hypothetical treatment induces a 69.5% increase in the probability of voting "yes" as compared to real. The conditioning further weakens the effect of the oath in the real context: according to the interaction term *Real* × *Oath*, the oath induces a slight decrease of "yes" answers (6.7%), which is far from being significantly different from 0 ($p = .675$).

The interaction term *Hypothetical* × *Oath* measures the treatment effect of implementing an oath in the hypothetical treatment. The effect is negative (-0.809) and significant at a 10% threshold with $p = .078$ –restricting the test to the working hypothesis of a decrease of hypothetical bias, the effect of the oath is significant at a 5% threshold with $p = .039$. Conditional on observable heterogeneity of the subject pool, the oath induces a 17.2% decrease in the probability of voting "yes" in the hypothetical context. A Wald test, however, rejects the null hypothesis that *Hypothetical* × *Oath* + *Hypothetical* =0 (Wald=18.72 with $p = .000$). This means the oath does decrease "yes" votes in hypothetical but not to the extent that it eliminates completely the gap between the proportion of "yes" votes in hypothetical under oath and the proportion of "yes" votes with monetary incentives. Note there is a drawback of enhanced external validity through eliciting preferences for an homegrown good, such as the WWF donation–there is a loss of control over the true underlying preferences. Subjects enter the laboratory with their own, privately known valuation for the good. As a long-standing consequence, there is no obvious way to choose the benchmark situation to which one should compare the variation in elicited preferences. Under monetary incentives, in particular, subjects may undermine their true preferences by voting no as a way to opt-out from the elicitation mechanism (Smith, 1994; Jacquemet, Joule, Luchini, and Shogren, 2011). To get further evidence on the content of the effect of the oath on preference elicitation in this context, we now turn to the correlation of the variation in self-reported honesty with the observed changes in voting behavior.

3.2 The effect of the oath on self-reported attitudes

To obtain some insight into why the oath induced the variations in voting behavior observed above, we now explore how it affects responses to attitudinal questions.[17] Table 3 reports a set of separate ordered probit regression on three variables: the level of agreement with the WWF, self-declared honesty and happiness.[18]

The left-hand side of Table 3 reports the results of an ordered probit regression using subjects' level of agreement with WWF actions as the dependent variable. This question aims to assess whether subjects use the elicitation exercise to express positive attitudes towards public goods, or

[17] We restrict this presentation to those questions for which we do observe some significant differences between treatments.

[18] While treatments are truly exogenous, allowing for regressions of attitudes on treatment variables, there are obvious endogeneity issues if one explains votes by self-reported attitudes. It is not our aim to disentangle the respective effect of the oath on attitudes and reported preferences ; but rather to gather some information on the channel through which the oath changes behavior.

Table 3: Treatment effects on self reported honesty and care questions

	Agreement with WWF		**Honesty self**		**Happiness**	
	Parameter estimates	p-value	Parameter estimates	p-value	Parameter estimates	p-value
Treatment effects						
Hypothetical	.389	.054	-.769	.003	-.033	.115
Hypothetical × Oath	-.019	.923	.031	.791	-.265	.012
Real × Oath	.035	.854	.317	.196	-.273	.202
Controls						
Age	-.017	.042	.016	.120	.019	.004
Male	.087	.538	-.205	.348	-.367	.003
Occupational status (ref. is employed)						
Unemployed	.271	.318	-.267	.351	.170	.253
Student no grant	-.481	.050	.037	.868	.219	.372
Student with a job	-.371	.306	-.231	.490	.511	.205
Student with a grant	-.530	.074	.406	.112	.578	.073
Cutoff points (st.error)						
cut 1	-3.044 (.459)		-2.342 (.388)		-1.733 (.358)	
cut 2	-2.685 (.422)		-1.589 (.413)		-1.441 (.303)	
cut 3	-2.484 (.410)		-1.142 (.418)		-0.726 (.259)	
cut 4	-1.460 (.384)		-0.310 (.393)		-0.087 (.296)	
cut 5	-0.678 (.378)		—		0.803 (.311)	
cut 6	0.056 (.378)		—		1.749 (.352)	

Note. Ordered probit models on self reported attitudes: the left-hand side uses answers on a 7-points scale to the question: *What is your opinion of the WWF's activities?* (from totally opposed to totally in favour); the model relies on answers on a 7-points scale to the question *Please rate how honest you think you were in your votes* (from *Not at all honest* to *Totally honest*) and the third model relies on answers on a 7-points scale to the question *Please rate how honest you think others were in their votes.* The top rows report the results of treatment dummies and individual characteristics ($n = 239$). The bottom part of the Table reports the cutoff parameters – for the honesty question, only 4 cut points (instead of 6) are estimated as no subjects answered neither 2 nor 3 on the scale.

concerns for society problems, rather than their true underlying preferences. This view is supported by a comparison of the two benchmark treatments: respondents in hypothetical exhibit a stronger agreement with WWF than in the real ($p = .054$) – which is in line with a higher willingness to vote "yes" in the hypothetical condition. Interestingly, this is no longer the case when subjects are under oath with and without monetary incentives –one cannot reject the null of no effect with $p = .923$ and $p = .854$. The oath seems to correct a positive shift of agreement with the WWF induced by the absence of monetary incentives. Because a large discrepancy between hypothetical and real votes remains when subjects are under oath, this also suggests that this shift is not the main explanation of hypothetical bias.

Second, the model presented in the center column of Table 3 reports the results of an ordered probit regression on the question about how honest subjects think they were in their votes. Our concern with this question is to elicit the degree of conscious manipulation of the elicitation exercise.

We find evidence that subjects know they are reporting insincere preferences more often without monetary incentives: subjects rate themselves as significantly less honest in hypothetical than in real ($p = .001$). For instance, 46.7% of subjects in hypothetical declare themselves as totally honest whereas 77.9% do so in real. This is no longer the case for subjects under oath without monetary incentives when compared with subjects in real ($p = .899$). This is also true for subjects under oath with monetary incentives ($p = .230$). In hypothetical under oath, self-declared totally honest subjects are 77.3% in hypothetical with oath and 85% in real with oath. This suggests the remaining discrepancy between hypothetical under oath and real may result from subjects' self-deception, *i.e.* subjects do not realize they would behave differently under binding monetary incentives.

Last, to assess the extent of the pressure imposed on subjects by having them sign an oath before voting, we introduced a standard measure of happiness on a 7-point scale. Taking happiness as a scalar cardinal measure, we observe that subjects under oath are less happy than subjects in the hypothetical condition: mean happiness is 5.13 in hypothetical and 4.76 and 4.78 in hypothetical with oath and real with oath. According to a mean difference tests, these differences are significant with $p = .046$ and $p = .070$. Mean happiness under monetary incentives, by contrast, is only slightly (but not significantly) lower than without monetary incentives ($p = .150$). In the third column of Table 3, we take happiness as an ordinal measure and estimate an ordered probit similar to that applied to previous attitude variables. Results indicate that happiness decreases significantly in the hypothetical under oath treatment ($p = .012$) whereas the decrease is not significant in the real under oath treatment ($p = .176$). A test of equality of parameters associated with the hypothetical under oath and real under oath however cannot reject the null of equality with $p = .972$. It is therefore interesting to estimate n unique parameter when subjects are under oath, whether with or without monetary incentives. Results indicate that the parameter associated with the oath is negative and significant with $p = .015$ (results are given in appendix B). This suggests that the oath is not an innocuous instrument. Such a decrease in happiness however remains hard to interpret in terms of improved or lower internal validity of the results. For instance, it can either reflect that subjects feel uncomfortable with the experimental exercise after the oath – and maybe over-react to the environment – or it could reflect the opposite: the oath elicits higher cognitive efforts when subjects are asked to form and declare their preferences.

To sum up, the correlation of self-reported measures with the treatment effects suggests truthfulness improves under oath: subjects are less prone to use the vote to express positive general attitudes towards public goods and see themselves as more sincere in their answers. This comes at the price of decreasing well-being during the experiment.

4 Conclusion

Preference elicitation methods – even a straightforward approach like a binary voting referenda – can suffer from hypothetical bias. Herein we explore in a referendum experiment whether signing a solemn oath to tell the truth can reduce hypothetical bias (also see recent work by others who have tested

the robustness of the oath idea using alternative elicitation mechanisms, and sample populations, e.g., Carlsson, Kataria, Krupnick, Lampi, Löfgren, Qin, Sterner, and Chung, 2013). Our results suggest the oath can work to fill the gap between stated intentions and real economic commitments: the oath causes hypothetical "yes" response rates to significantly decrease, while real "yes" response rates remained statically identical. As we elicit preferences for a homegrown good, the results cannot be related to the true underlying preferences for the good. The correlation of the observed variation in stated preferences with self-reported measures of honesty however supports the idea that the oath enhances the truthfulness of the votes. Having subjects (freely) sign an oath to provide honest answers makes them more likely to do so even without any actual economic commitment.

Beyond the particular application of our results to contingent valuation studies, this evidence suggests one can improve the accuracy of preferences elicited in the lab through commitment devices such as an oath. This point remains a speculative interpretation of our results as long as the oath has not been applied to a wider range of experimental applications. Further research will explore this avenue.

References

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

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner, and H. Schuman (1993): "Report of the NOAA Panel on Contingent Valuation," *Federal Register*, 58(10), 4601–4614.

Carlsson, F., M. Kataria, A. Krupnick, E. Lampi, Å. Löfgren, P. Qin, T. Sterner, and S. Chung (2013): "The Truth, the Whole Truth, and Nothing but the Truth - A Multiple Country Test of an Oath Script," *Journal of Economic Behavior & Organization*, Forthcoming.

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

Cherry, T., P. Frykblom, J. Shogren, J. List, and M. Sullivan (2004): "Laboratory Testbeds and Non-Market Valuation: The Case of Bidding Behavior in a Second-Price Auction with an Outside Option," *Environmental & Resource Economics*, 29(3), 285–294.

Cherry, T. L., P. Frykblom, and J. F. Shogren (2002): "Hardnose the Dictator," *American Economic Review*, 92(4), 1218–1221.

Cummings, R. G., S. Elliott, G. W. Harrison, and J. Murphy (1997): "Are Hypothetical Referenda Incentive Compatible?," *Journal of Political Economy*, 105(3), 609–621.

Cummings, R. G., and L. O. Taylor (1999): "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method," *American Economic Review*, 89(3), 649–665.

Deisenroth, D., J. L., and C. Bond (2009): "Non market valuation of off-Highway vehicle recreation in Larimer County, Colorado: Implications of trail closures," *Journal of Environmental Economics and Management*, 90(11), 3490–3497.

Diamond, P. A., and J. A. Hausman (1994): "Contingent Valuation: Is Some Number better than No Number?," *Journal of Economic Perspectives*, 8(4), 45–64.

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

Greiner, B. (2004): "An Online Recruitment System for Economic Experiments.," *University of Cologne, Working Paper Series in Economics*, 10, 79–93.

Jacquemet, N., A. James, S. Luchini, and J. Shogren (2011): "Social psychology and environmental economics: a new look at ex ante corrections of biased preference evaluation," *Environmental & Resource Economics*, 48(3), 411–433.

Jacquemet, N., R.-V. Joule, S. Luchini, and J. Shogren (2009): "Earned Wealth, Engaged Bidders? Evidence from a second price auction," *Economics Letters*, 105(1), 36–38.

——— (2011): "Do people always pay less than they say? Testbed laboratory experiments with IV and HG values," *Journal of Public Economic Theory*, 13(5), 857–882.

——— (2013): "Preference Elicitation under Oath," *Journal of Environmental Economics and Management*, 65(1), 110–132.

Jacquemet, N., S. Luchini, J. Shogren, and A. Zylbersztejn (2011): "Coordination with Communication under Oath," *GREQAM WP*, 2011(49).

Joule, R., and J. Beauvois (1998): *La soumission librement consentie*. Presses Universitaires de France, Paris.

Kahneman, D., and R. Sugden (2005): "D. Kahneman and R. Sugden. Experienced utility as a standard of policy evaluation," *Environmental & Resource Economics*, 32, 161–181.

Kiesler, C., and J. Sakumura (1966): "A test of a model for commitment," *Journal of Personality and Social Psychology*, 3(3), 349–353.

Loureiro, M., J. Loomis, and M. Vazques (2009): "Economic valuation of environmental damages due to the Prestige oil spill in Spain," *Environmental & Resource Economics*, 4(4), 537–553.

Mijovič-Prelec, D., and D. Prelec (2010): "Self-deception as self-signalling: a model and experimental evidence," *Philosophical Transactions of the Royal Society*, 365(1538), 227–240.

Murphy, J. J., T. Stevens, and D. Weatherhead (2005): "Is Cheap Talk Effective at Eliminating Hypothetical Bias in a Provision Point Mechanism?," *Environmental & Resource Economics*, 30(3), 327–343.

Schlesinger, H. J. (2008): *Promises, Oaths, and Vows: On the Psychology of Promising*. Analytic Press, New-York (NJ).

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

Smith, V. L. (1994): "Economics in the Laboratory," *Journal of Economic Perspectives*, 8(1), 113–131.

Svensson, M. (2009): "The value of a statistical life in Sweden: Estimates from two studies using the "certainty approach" calibration," *Accident Analysis and Prevention*, 41(3), 430–437.

Sylving, H. (1959): "The Oath: I," *The Yale Law Journal*, 68(7), 1329–1390.

Zeiliger, R. (2000): "A presentation of Regate, Internet based Software for Experimental Economics," *http://regate-ng.gate.cnrs.fr/ sferriol/*.

Appendix

A Post-experimental questionnaire

The voting stage is now over. Before proceeding to the payment of your winnings, we would be grateful if you could answer a series of questions. There are no right or wrong answers; take your time in answering the questions because your answers are important for the research in which you are taking part. It is impossible to connect these answers to your identity.

(The questionnaire starts with the following debriefing questions that appear one at a time on the screen).

1. Do you belong to an environmental association? (YES / NO)
2. Did you know of the WWF before taking part in this experiment? (YES / NO)
3. Did you know of the WWF's dolphin adoption programme before taking part in this experiment? (YES / NO)
4. Have you previously adopted an animal to help an association for the protection of nature? (YES / NO)
5. What is your opinion of the WWF's activities? (Totally opposed, opposed, moderately opposed, no opinion, moderately in favour, in favour, totally in favour)

(The attitudinal questions below straightly follow, with each question and the scale on each screen).

1. Using the scale below, please indicate how happy you are at the moment:

+	+	+	+	+	+	+
1	2	3	4	5	6	7
	Very				Very	
	unhappy				Happy	

2. Using the scale below, please indicate how honest you were during the experiment

+	+	+	+	+	+	+
1	2	3	4	5	6	7
	Not at			Totally		
	all honest				honest	

3. Using the scale below, please indicate how honest the other participants were during the experiment

+	+	+	+	+	+	+
1	2	3	4	5	6	7
Not at all honest					Totally honest	

B Happiness ordered probit with oath as a single variable

	Parameter estimates	p-value
Treatment effects		
Hypothetical without oath	-0.032	.778
Oath	-0.269	.015
Controls		
Age	0.019	.004
Male	-0.367	.003
Occupational status (ref. is employed)		
Unemployed	0.171	.267
Student no grant	0.219	.370
Student with a job	0.510	.198
Student with a grant	0.577	.081
Cut points (s.e)		
cut 1	-1.733 (.359)	
cut 2	-1.440 (.303)	
cut 3	-0.727 (.258)	
cut 4	0.086 (.296)	
cut 5	0.802 (.311)	
cut 6	1.749 (.353)	