HAL
open science

Referenda Under Oath

Nicolas Jacquemet, Alexander James, Stéphane Luchini, Jason Shogren

To cite this version:

Nicolas Jacquemet, Alexander James, Stéphane Luchini, Jason Shogren. Referenda Under Oath. Environmental and Resource Economics, 2017, 67 (3), pp.479-504. 10.1007/s10640-016-0023-5 . halshs-01301784

HAL Id: halshs-01301784

https://shs.hal.science/halshs-01301784

Submitted on 12 Apr 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.

Referenda under Oath *

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

April 2016

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 a self-reported measure of honesty which confirms 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.

*This paper is a revised and extended version of GREQAM Working Paper n° 2010-15. 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), the Provence-Alpes-Côte d'Azur Regional Council and the Bugas Fund is gratefully acknowledged. Shogren thanks the Norwegian University of Life Sciences and the Rasmuson Chair of Economics at the University of Alaska-Anchorage for support. Jacquemet acknowledges the *Institut Universitaire de France* and the French National Research Agency, through the program Investissements d'Avenir, ANR-10–LABX-93-01.

†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

‡Department of Economics and Public Policy, University of Alaska Anchorage, Anchorage, AK 99508-4614 alex.james@uaa.alaska.edu

§Aix-Marseille University (Aix-Marseille School of Economics), CNRS and EHESS, Centre de la Vieille Charité, 13236 Marseille Cedex 02. stephane.luchini@univ-amu.fr

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

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 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., the views in McConnell, 1990; Cummings, Elliott, Harrison, and Murphy, 1997; Haab, Huang, and Whitehead, 1999; Champ and Bishop, 2001).

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).[2] One time-tested mechanism used to promote commitment is the solemn oath – a formal way to create 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 voluntary (Jacquemet, James, Luchini, and Shogren, 2011).

Jacquemet, Joule, Luchini, and Shogren (2013) make a full-blown argument for the oath-as-commitment device to reduce hypothetical bias in non-market preference elicitation. They designed a series of testbed experiments using both induced and homegrown values Vickrey second-price auctions and confirm the ability of a truth-telling oath to improve the reliability of elicited preferences. They find that the oath induced sincere bidding behavior in an induced value auction, and the oath induced bidders to take seriously both their budget and participation constraints in a homegrown value auction. Their results suggested the oath worked in non-market valuation because people were more committed to be truthful (unlike alternative cheap talk or guilt aversion explanations, see the detailed discussion in Jacquemet, Joule, Luchini, and Shogren, 2013).

The key question that the Jacquemet et al. study left open was how to apply the oath to stated preference studies in the field that use discrete choice mechanisms, like the classic referendum style vote on a public good. Carlsson, Kataria, Krupnick, Lampi, Lofgren, Qin, and Sterner (2013) and

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

[2] Also see Tavoni, Dannenberg, Kallis, and Löschel (2011) who design a public goods experiment to explore whether a nonbinding "pledge and review" commitment system can increase the coordination of voluntary carbon emission reductions. They find that subjects in the pledge treatments increased coordination significantly over the no-pledge baseline.

de Magistris and Pascucci (2014) address this gap by applying the oath to stated preference field studies in China/Sweden and the Netherlands. While the results of these studies are suggestive, they still do not allow an interested researcher to assess the ability of the oath to reduce hypothetical bias. Since the oath scripts were implemented only in hypothetical homegrown surveys, there is no control condition that elicits real values so one can assess the extent that hypothetical bias was removed or remains. We need to step back and ask the basic question: Can the oath help reduce hypothetical bias, if it exists, in a referendum vote?

Herein we extend our experimental work of the oath-as-commitment device to the classic referendum voting preference elicitation format (see e.g., Cameron, 1991; Arrow, Solow, Portney, Leamer, Radner, and Schuman, 1993; Cummings, Elliott, Harrison, and Murphy, 1997; Cummings and Taylor, 1999; List, Berrens, Bohara, and Kerkvliet, 2004; Carson, Groves, and List, 2014; James and Shogren, 2015). Our goal is to understand whether the individual-level oath (as administered by Jacquemet, Joule, Luchini, and Shogren, 2013) can induce more sincere voting behavior within a collective-style mechanism. We focus on the referendum voting format for three primary reasons. First, the referendum remains popular in non-market valuation exercises because it is familiar, realistic, strategy-proof, and straightforward to implement in the field (Carson and Groves, 2007). Second, an individual-level oath would likely still be used in any field application of this collective mechanism given each person would be voting in private, i.e., on-line survey, mail survey or mall-intercept techniques (see Carlsson, Kataria, Krupnick, Lampi, Lofgren, Qin, and Sterner, 2013).

Third, the accumulated evidence from induced valuation experiments suggests that it is not an empirically settled question as to whether the referendum format actually eliminates hypothetical bias in voting behavior (see for example Cummings, Harrison, and Rutström, 1995; Champ and Bishop, 2001; Polomé, 2003; Vossler and Kerkvliet, 2003; Schläpfer, Roschewitz, and Hanley, 2004). People voting in hypothetical referendum settings can be subject to social pressure bias and anchoring bias which can lead to "yea-saying", which can still create a gap between votes and actions (e.g., Green, Jacowitz, Kahneman, and McFadden, 1998; Cummings and Taylor, 1999; James and Shogren, 2015). Using an experimenter-controlled, pre-assigned induced value experimental design, Taylor, McKee, Laury, and Cummings (2001) and Vossler and McKee (2006) both observed no hypothetical bias in the referendum at the aggregate level, but they do find mistakes and mis-voting at the individual level (see Burton, Carson, Chilton, and Hutchinson, 2007; Mozumder and Berrens, 2007; Collins and Vossler, 2009; Murphy, Stevens, and Yadav, 2010). While several explanations have been offered and explored as to why people might mis-vote (i.e., preference uncertainty, other-regarding preferences), no one up to now has explored whether the oath-as-commitment-device might create the environment needed for more sincere voting behavior in the referendum.

We made four purposeful design choices in our experiment that differ from the implementation of a typical referendum style survey run in the field. First, we step back from the field and first test whether the oath works for the referendum in the lab, as originally suggested by Coursey and Schulze (1986). We pretest the elicitation mechanism before it is implemented in the field, both with and without monetary incentives. In our lab experiment, we can link the effect of the oath on hypothetical

referendum bidding to hypothetical bias, and then explore individual motivations behind it (for a review see, e.g., Shogren, 2005).

Second, we did not attempt to induce language to create "consequentiality" into our hypothetical baseline design. Recall "consequentiality" is the idea that participants should treat a stated preference style hypothetical valuation question as "real" if they believe some probability exists that the policy will actually be implemented (see e.g., Vossler, Doyon, and Rondeau, 2012). In contrast, our focus is on creating real economic commitment with a non-market mechanism-the oath. We chose not to conflate the issue with trying to induce consequentiality. We wanted to avoid adding another layer of language that might or might not affect our test of commitment. We do so because experiments that test for the impact of consequentiality essentially introduce a new lottery stage into the valuation exercise, with mixed results (see, e.g., Carson, Groves, and List, 2014, who argue that the votes for a referendum were consequential). In contrast, Shogren and Tadevosyan (2011) find insincere bidding behavior with "consequential" induced values in a second price auction–any uncertainty regarding implementation of the contract lead to non-optimal bids.

Third, we used a classic monetary vote-then-donate format rather than the "advisory referendum" format that has been used in some stated preference work (see Loomis, 2014, for a good review). Again since the focus of our paper is to create real economic commitment with a non-market incentive device, we want the participants to view the choice to vote "yay" as having real consequences under the oath, and not just advice to policymakers.

Fourth, given our laboratory design, we chose one monetary value to vote "yay" or "nay", and we use five rounds of voting rather than one, in which only one round was binding (for multi-price voting approach, see Messer, Poe, Rondeau, Schulze, and Vossler, 2010). Our aim is two-fold. This choice makes our referendum design as close as possible to our earlier work in Jacquemet, Joule, Luchini, and Shogren (2013). We can now calibrate and compare results across elicitation mechanisms given the same good, same number of rounds, one randomly chosen binding auction, and nearly identical instructions. In addition, by using several rounds with only one monetary value we now have an objective measure of the strength of preferences as defined by how consistently subjects respond (Rustichini, 2008). This allows us to assess the degree of hypothetical bias conditional on our participants' confidence levels of their own preferences. Some subjects were confident–5 "yay" or "nay" votes–while others were less confident. We now can correlate the effect of the oath with their confidence.[3]

We design our lab experiment to contrast preferences elicited in the referendum under both hypothetical and monetary incentives, with and without an oath. In the oath treatment, subjects choose to sign a form freely in which they swear to tell the truth during the experiment. The

[3]This is reminiscent of certainty questions in the stated preferences literature, in which a researcher examines the extent to which respondents are confident with their answers to willingness to pay questions by way of certainty scales (see, e.g., Luchini and Watson, 2013). The difference with our voting-preference-certainty approach is that our measure of confidence is an axiomatically grounded cardinal measure of the strength of preferences. In contrast, the answers to certainty questions are subjective measures of confidence that make interpersonal comparisons difficult.

working hypothesis is that the oath will induce a person to be consistent with his or her initial commitment to tell the truth in subsequent voting decisions. While our experimental design focuses on eliciting homegrown values, we calculate the referendum price that will be voted on from the real willingness to pay bids elicited in a second-price auction to buy a World Wide Fund (WWF) Adopt-a-Dolphin program, as found in Jacquemet, Joule, Luchini, and Shogren (2013). Our experimental results provide additional support for the idea that the truth-telling oath can reduce hypothetical bias. The oath created more real economic commitment – it reduces the proportion of "yes" votes, though it does not completely eliminate observed hypothetical bias. We also explored behavior at the individual level in more detail than previous referendum work by eliciting each voter's attitudes as reflected by: (a) the level of agreement with the WWF, (b) self-declared honesty, and (c) happiness. We observe that respondents in the hypothetical treatment more strongly agreed with WWF than in the real treatment (more "yea-saying"), but this result no longer held under oath. We then found the self-reported measure of honesty increases under oath as compared to the baseline hypothetical treatment. This suggests the oath incentivizes more subjects to be honest in the referendum than otherwise. When considering happiness (given response time), we find that the oath decreases the tendency to engage in self-serving assessments. Overall, our results further support the idea that signing a solemn oath helps create the commitment needed to better link intentions and actions in non-market valuation, and perhaps beyond (also see Jacquemet, Luchini, Shogren, and Zylbersztejn, 2015).

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".[4] 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

[4]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/`.

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 are 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.[5] 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 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 organization 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 organization 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-subject 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 five independent referenda. On the adoption page subjects are asked to vote by clicking on either of two buttons: YES or NO. We repeated five times the referendum in fixed groups: subjects take five successive votes in identical referenda before learning the result of each. At the end of the sequence, one 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.[6] We use this

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

[6] 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".

preliminary evidence to calibrate the amount of the donation to 11 Euros: this is a price at which *(i)* preferences exhibit a significant hypothetical bias and *(ii)* corner solutions in voting behavior are unlikely to be observed (see Fig.2, p.122, in Jacquemet, Joule, Luchini, and Shogren, 2013). 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.

2.2 Subject 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.[7] 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.[8] 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 labeled in ECU (*Experimental Currency Unit*) are proportional to 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.[9]

[7] In this context earned wealth is based more on knowledge and less so on effort which can influence a person's view of a "fair allocation" of wealth (see for example Konow, 1996). However, it is worth noting that, because the questionnaire is implemented across treatments, resulting bias may be differenced away.

[8] Our source is `http://pagesperso-orange.fr/bac-es/qcm/annales_c02_r01.html`, the full list of questions is available from the authors upon request.

[9] From our previous work using this procedure, the risk of a subject scoring in such a way that earned wealth is lower than 1 Euro is very unlikely. All subjects thus enter the referendum with earnings higher than 11 Euros.

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 makes 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:[10]

> *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 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 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 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 1, reads, "*I, the undersigned swear upon my honour that during the whole 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. Each subject enters alone and is directed to a monitor at the front of the laboratory. 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. Importantly, the monitor does not inform the subjects about the topic of the experiment when asked to take the oath. The

[10] 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 1: Oath form used in the experiments

PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

SOLEMN OATH

Topic: "JaJLucSh"; Research number 3627A

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.

subject is invited to read the oath, and decides whether to sign. Regardless of whether a subject signs the oath, he or she is thanked and invited to enter the lab. The exact wording used by the monitors to offer the oath to respondents was scripted to standardize the phrasing of the oath. One monitor stayed in the lab until all subjects had been presented with the oath. We did this to avoid communication prior to the experiment. Subjects waiting their turn could neither see nor hear what was happening at the oath-desk. Note, the 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.[11]

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 three different questions, always asked in the same order.[12] 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 societal 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. Mijoviċ-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). 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, 60 for each treatment), were conducted in the LEEP laboratory in Paris in May-July 2013.[13] 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 used in all experiments run in this laboratory to confirm the informed consent of each participant. The signing of the consent form happens 15-30 minutes before the oath is administered (in the oath treatment). Each subject receives the consent form upon arrival at the laboratory and reviews it while the monitor checks their identity. Subjects then wait until all participants arrive. The signed forms are then collected before subjects are directed to the laboratory. These procedures aim to make a clear distinction between the consent form and the oath. 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

[11] 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.

[12] The content of the post experimental questionnaire, including these three questions, is presented in the Appendix, Section A.

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

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 (gender, age, ...). 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 general knowledge 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.[14]

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 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. Subjects play each of the five successive referenda one after the other and do not receive information about the results until all referenda have been completed. If the vote passes in a group in the real treatment, each subject makes an 11 euros donation taken from experimental earnings, which we subsidize so as to reach the market clearing price of the donation. At the end of the experiment, subjects are asked to answer a computerized debriefing questionnaire. The questionnaire collects information such as whether subjects have participated in other experiments or not, their level of knowledge of the WWF and its actions, and the extent of their agreement with it. The questionnaire ends with the three happiness and honesty questions. Finally, the monitor pays each subject privately in cash.

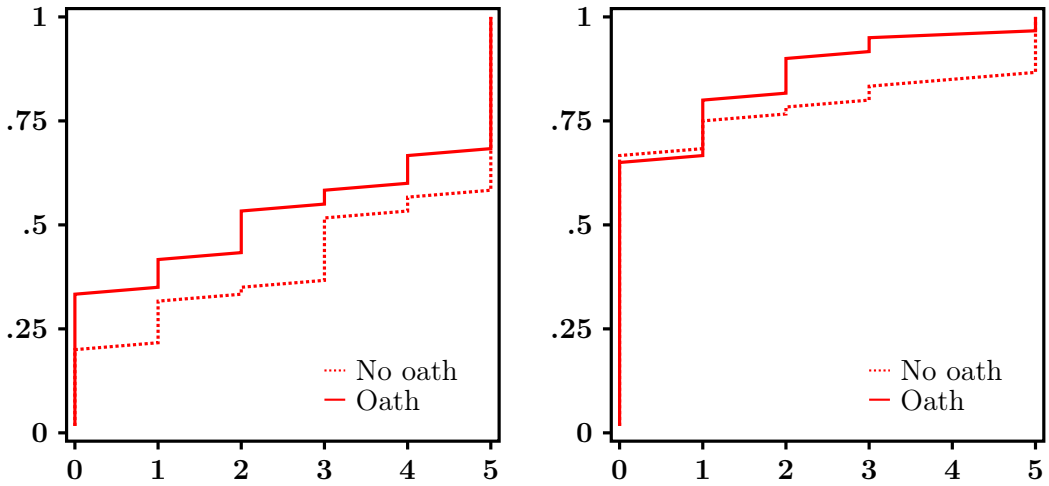
3 Results 1: How does the oath affect referendum voting?

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 2 presents the empirical distribution functions (EDF) of the total number of "yes" responses by treatment. In the hypothetical treatments (Figure 2.a), we observe that the EDF in hypothetical significantly first-order dominates the EDF in hypothetical under oath with $p = .065$.[15] "Yes" responses are significantly shifted down by signing the oath at the individual level.

[14] In the experiment, the lower bounds on the earnings from the quiz are 15.5 in the hypothetical no oath treatment, 14 in hypothetical with oath, 12.5 in real no oath and 13.5 in real with oath treatment. Total earnings are always higher than the laboratory price of the donation.

[15] 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

Figure 2: Distribution of "yes" responses by treatment



(a) Hypothetical (b) Real

This shift is explained by an increase in the number of subjects always voting "No" under oath (20% in hypothetical and 33.3% in hypothetical under oath) and a decrease in the number of subjects always voting "Yes" (43% in hypothetical and 33.3% in hypothetical under oath). The EDF in real treatments (Figure 2.b) exhibits the same shift in "yes" responses under oath but to a lesser degree ($p = .162$). The shift is now explained by a decrease in the number of subjects always voting "Yes" (15% in real and 0.5% in real under oath). These figures support our main result: having subjects sign a truth-telling oath before participation in a dichotomous choice mechanism significantly shifts hypothetical voting behavior downwards. Prior to signing an oath, hypothetical affirmative responses are 38% greater than real ones. After signing an oath, this figure drops to 26%, implying hypothetical bias is reduced by about 30%. 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.

Table 1 summarizes the votes elicited in each treatment and the resulting number of adoptions. 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 22.3% 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 no oath treatments whereas only 5 adoption decisions (8.3%) were made in real no oath treatments. Second, signing the oath leads to an 11.7% decrease in the "yes" responses in the hypothetical condition, from 61% to

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., bootstrapping on votes. The number of replications is 9999.

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%	22.3%
	Adoptions (#)	1	0	2	0	2	5 (8.3%)
Real with oath	Yes	11.7%	11.7%	16.7%	20.0%	15.0%	15.0%
	Adoptions (#)	0	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.

49.3%. The p-value of this decrease according to a unilateral bootstrap proportion test is $p = .125$.[16] This results in a 20% drop in the 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 22.3% to 15%, a difference that is not statistically significant ($p = .198$).[17]

By construction, round effects are driven solely by uncertain subjects—subjects who are not confident and change their vote from one round to the other. In the following, we examine how uncertain subjects vote across rounds. In all treatments, except Real with Oath, we observe that uncertain subjects are more likely to cast a "yes" vote in the first round than in forthcoming rounds 2 to 5. Proportions of "yes" votes in round 1 are 77.2%, 56.0% and 63.7% in Hypothetical no Oath, Hypothetical with Oath and Real no Oath, whereas they are, on average, 40.9%, 43.8% and 34.1% in rounds 2 to 5. In Real Oath, we observe a small increase in uncertain subjects voting in favor of the adoption (22% in round 1, 36.1% in rounds 2 to 5). Bootstrap proportion tests indicate that there are no significant differences in "yes" votes between Hypothetical no Oath and Real no Oath when looking separately at round 1 (p=0.482) and at rounds 2 to 5 (p=0.428). In addition, we find

[16] In our setting, we cannot implement standard Mann-Whitney tests since observations are independent at the individual level but not at the vote level. Each subject votes 5 times and this can induce within subject correlation. We therefore cannot carry out standard Mann-Whitney tests, which rely on the assumption of independence of observations. Our bootstrap procedure consists of bootstrapping on subjects and their five votes rather than bootstrapping on votes. The bootstrap proportion test accounts for within subject correlation without specifying it parametrically (see Jacquemet, Joule, Luchini, and Shogren 2013 for more details). In addition, bootstrap tests have more statistical power for small sample sizes.

[17] 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 observable ranges. We note, however, that the preferences we elicit are statistically very close to those observed in Jacquemet, Joule, Luchini, and Shogren (2013)—see the Appendix, Section B.

no effect of the oath on hypothetical votes in round 1 (p=0.379) and in rounds 2 to 5 taken together (p=0.723). The small difference between the proportion of "yes" votes in real no oath and real with oath is also not significant with p=0.271.

We now examine the degree of hypothetical bias and the effect of the oath conditional on the strength of preferences of subjects. In line with stochastic choice theory (see Köbberling, 2006; Rustichini, 2008), we split the sample into two groups: confident subjects who vote identically in all 5 rounds (either all "yes" or all "no") and non-confident subjects who change their vote in at least one round.[18] A large majority of subjects are confident with their choice in all 4 treatments: 63.3% in hypothetical, 66.7% in hypothetical under oath, 70.0% in real and 70.0% in real under oath. Further, this result indicates the oath has no effect on the strength of the preferences. Our results are informative–we observe that hypothetical bias only exists for confident subjects and the oath only decreases bias for these confident subjects. The proportion of "yes" votes in hypothetical for confident subjects is 68.4% and 18.3% in real, a significant 50.1% decrease ($p < 0.001$). Under oath, the proportion of "yes" for confident subjects decreases to 50% ($p = 0.092$). In contrast, now consider the sample of non-confident subjects. Here the proportion of "yes" votes in hypothetical is 48.2% and 40% in real–we find no hypothetical bias in this group (the difference is not significant with $p = 0.335$). In addition, the oath has no effect on the votes of non-confident respondents with 48.0% voting "yes" in hypothetical under oath ($p = 0.981$). This finding is re-assuring. Based on Commitment Theory, there is no reason why a truth-telling commitment device such as the oath should work when a person has difficulties knowing his or her own preferences, i.e. asking a subject to be truthful when she does not even know the truth, or she is experimenting with the truth, is futile in our setting.

We assess the robustness of the results by conditioning the effect of the treatments on participants' characteristics. Participant characteristics include gender, age, occupational status and whether or not the subject attended lab experiments in the past. 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*).[19] The reference observation is a subject in the real no oath treatment. The coefficient

[18] In theory, the observed frequency of choices provides an axiomatic cardinal and continuous measure of the strength of preferences. We split the sample into two groups to ensure large enough sample sizes.

[19] Our intent in estimating the panel Probit model is to check if the result is robust to controlling for observed heterogeneity that would have been missed in our unconditional statistics (although one may argue the purpose of a lab experiment is to testbed mechanisms/theories on a properly randomized homogeneous population so that conditional statistics are not needed). When one assumes that there exists a latent variable that generates voting behavior as it is assumed in a Probit model, one may wonder whether the total effect of the oath or monetary incentives come from a change in the mean preference or the sample variance (see Haab, Huang, and Whitehead, 1999) —note that this distinction may not be relevant if one is only interested in the marginal effect of the experimental treatment such as we are – see, e.g., Harrison (2006). Likelihood ratio Heteroskedasticity tests (available from the authors on request) indicate the treatment variables *hypothetical* and *oath* have no or barely significant effect on the variance of the latent variable in basic Probit regressions carried out for each round independently (LR joint tests p-value for each round are $p = 0.360$, $p = 0.123$, $p = 0.100$, $p = .0.732$ and $p = .0.121$). Further research choosing to explore this issue in more

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

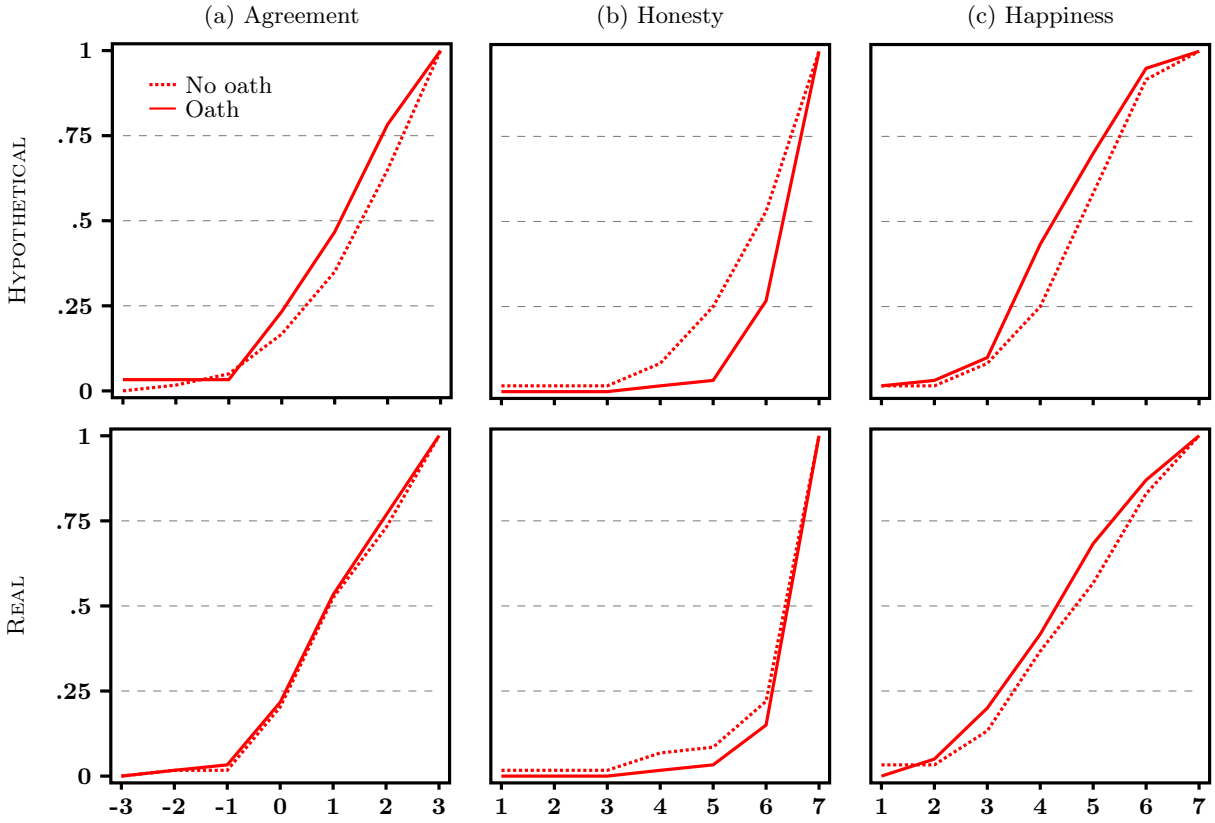
Note. Individual random effect panel Probit model of individual yes vote on treatment dummies and individual characteristics ($n = 239 \times t = 5$). The *endogenous* variable is the "yes" vote. *Round* (fixed) effects are controlled for in the estimation but omitted here. Joint nullity test: Wald = 51.91 with $p < 0.001$. Marginal effects are computed at the means of continuous independent variables, for binary covariates a discrete change from 0 to 1 is considered (see Williams, 2012).

associated with the dummy variable *Hypothetical* is positive and significant at the 1% level, indicating a clear hypothetical bias in the baseline condition. Being in the hypothetical treatment induces a 69.3% 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 ($p = .633$).

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 observed 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 $p = .000$), implying that while the oath decreases "yes" votes in hypothetical treatments, it fails to completely eliminate the observed hypothetical bias. Note there is a drawback of enhanced external validity through eliciting preferences for a 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 private valuation of the good. As a long-standing consequence, there is no obvious way

detail would need to define an appropriate structural form that integrated more explicitly a theory of incentives and oath on preferences, which is beyond the scope of this current project.

Figure 3: Empirical Distribution Function (EDF) of self-reported attitudes



Note. Answers are on a 7-point scale: *What is your opinion of the WWF's activities?* (from totally opposed to totally in favor); *Please rate how honest you think you were in your votes* (from *Not at all honest* to *Totally honest*) and *Please indicate how happy you are at the moment.* The first line of Figures presents EDF for the hypothetical treatments whereas the second line presents the EDF for real treatments.

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 of the elicitation mechanism (Smith, 1994; Jacquemet, Joule, Luchini, and Shogren, 2011). For further contextual evidence of the effect of the oath on preference elicitation, we now turn to the correlation of the variation in self-reported honesty with the observed changes in voting behavior.

4 Results 2: Additional insight into why the oath affects voting

To gain insight into why the oath induced the observed variations in voting behavior, we now explore how the oath affected responses to the attitudinal questions.[20] Figure 3 presents the EDF of answers

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

to three questions: the level of agreement with the WWF, self-declared honesty, and happiness.[21] Table 3 reports a set of separate ordered Probit regressions on the same three variables.

4.1 Agreement with WWF actions and surrogate voting

Figures 3.a present the EDF's of subject's level of agreement with WWF actions in the real (bottom part) and hypothetical (upper part) treatments. This question addresses the long-standing question of whether subjects use the elicitation exercise to express generic positive attitudes towards public goods, or concerns for societal problems, rather than their true underlying preferences (see Kahneman and Knetsch, 1992).

This general concern about surrogate voting is supported by comparing behavior in the two benchmark treatments: the EDF in hypothetical without oath first-order dominates the EDF in real without oath ($p = .024$). This result implies that respondents in hypothetical exhibit a stronger agreement with WWF than in the real treatment, which is in line with a higher willingness to vote "yes" in the hypothetical condition. Interestingly, this result no longer holds when subjects are under oath. The EDF in hypothetical under oath and real are not significantly different ($p = .697$). We find similar results for (i) the EDF in hypothetical under oath and real under oath ($p = .834$) and (ii) the EDF in real and real under oath ($p = .998$). This means that the level of agreement with WWF actions is only greater in the hypothetical no oath treatment. Our econometric regressions confirm these results. The left-hand side of Table 3 reports the results of an ordered Probit regression using subjects' level of agreement with the WWF actions as the dependent variable. Subjects in hypothetical treatments exhibit stronger agreement with WWF than in real ones ($p = .054$). Subjects under oath with and without monetary incentives express the same level of agreement ($p = .923$ and $p = .854$). The oath seems to correct for a positive shift in 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 result also suggests that this shift is not the main explanation of hypothetical bias.

4.2 The effect of the oath on self-declared honesty

We now turn to self-declared honesty. The main goal of this question is to elicit the degree of strategic voting or conscious manipulation of the elicitation exercise. Figures 3.b present the EDF of how honest subjects thought they were in their votes. We find evidence that subjects know they are reporting insincere preferences more frequently without monetary incentives. The EDF in real

[21] Since these measures are self reported and hypothetical, they may be noisy signals of the underlying attitudes of interest and we acknowledge that such self-reported attitudinal information should be interpreted with caution. The changes according to the treatment however, remain informative about underlying changes in attitudes as there is no reason to expect a correlation between the noise and the treatments. 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
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-point scale to the question: *What is your opinion of the WWF's activities?* (from totally opposed to totally in favor); the second model relies on answers from a 7-point 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 from a 7-point scale to the question *Please indicate how happy you are at the moment.* The top rows report the results of treatment dummies and individual characteristics ($n = 239$). The bottom part of the Table reports the cutoff parameters, *i.e.* the thresholds associated by Probit model to each possible answer in order to map the discrete data into the latent variable—for the honesty question, only 4 cut points (instead of 6) are estimated as no subjects answered neither 2 nor 3 on the scale. Results are conditioned on *Round* (fixed) effects, subject age, gender, grant, and employment status. A full set of results are available from the authors upon request. Joint nullity test: Wald = 51.91 with $p < 0.001$. Marginal effects are computed at the means of continuous independent variables, for binary covariates a discrete change from 0 to 1 is considered (see Williams, 2012).

first order dominates the EDF in hypothetical ($p < .001$): subjects rate themselves as significantly less honest in hypothetical than in real treatments. For instance, 46.7% of subjects in hypothetical declare themselves as totally honest, whereas 77.9% do so in real.

Regarding the effect of the oath, two interesting results emerge. First, the EDF's are statistically the same for hypothetical and real treatments when subjects are under oath ($p = .133$). In terms of self-perceived honesty, the two oath treatments thus achieve the same outcome. Additionally, the EDF for the real no oath treatment is statistically indifferent from that for the hypothetical oath treatment ($p = .631$). Again, conditional analysis confirms these results. The model in the center column of Table 3 reports the results of an ordered Probit regression for the honesty question. The coefficient associated with the hypothetical dummy is negative and significant ($p = .003$) and the coefficients for the hypothetical oath and real oath dummies are not significant ($p = .791$ and $p = .196$). In hypothetical under oath treatments, 77.3% of subjects for instance declare to be totally honest, compared to 85% in real under oath treatments.

Despite the discrepancy in actual voting behavior, this outcome variable thus shows that the oath achieves to elicit the same level of self-perceived honesty in hypothetical as in the baseline, real-no oath and real-oath treatments. Since subjects perceive themselves to be as honest in hypothetical-oath treatments as they are in a real referendum but at the same time vote yes more often, this suggests

one possible reason why a hypothetical bias remains after signing the oath. In our hypothetical referendum, subjects have to form an assessment about what they would do if votes were real. But their vote has no immediate consequences on protecting the environment; even if a majority votes "yes", no contribution is made to the WWF. In that context, voting "yes" in the referendum can serve as a self-serving assessment, making them too optimistic about their chances of voting "yes" in a real referendum.[22] Self-serving may occur here because people *learn* that they are nature preserving persons when they vote "yes" in the referendum. This induces self-deception, *i.e.* people think they would vote "yes" more often than what they actually do. Mijovič-Prelec and Prelec (2010), building on self-signaling theory, shows that self-deception arises if people not only derive utility from actions (*outcome utility*) but also from learning about their inaccessible characteristics by doing those actions (*diagnostic utility*)—see also, *e.g.*, Benabou and Tirole (2002); Caplin and Leahy (2001). It is this diagnostic utility that makes people engage in self-deception, whether they are aware of it or not. What the answers to our honesty question suggest is that even subjects under oath engage in self-deception (and are not aware of it as in Quattrone and Tversky, 1984).

4.3 Does the oath simply elicit higher cognitive effort?

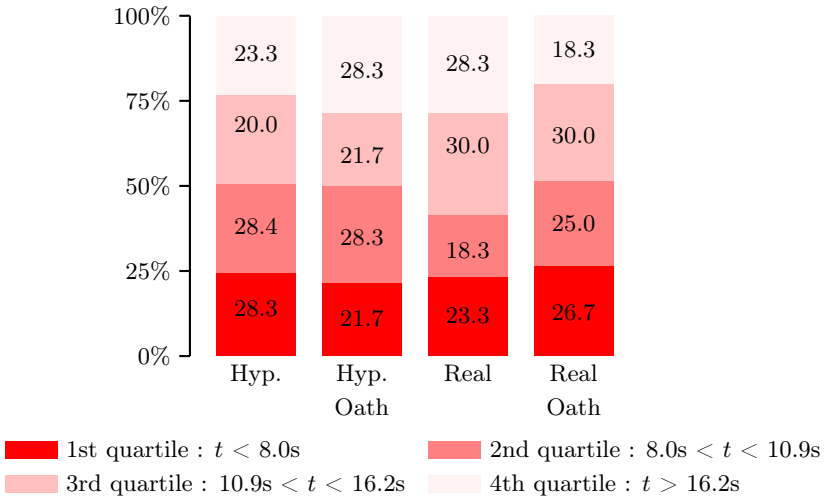
Third, we assess the degree of pressure imposed on subjects signing an oath by measuring their happiness on a 7-point scale. Assuming we can measure happiness as a cardinal scalar, we observe that subjects under oath are less happy than subjects in the hypothetical condition: mean happiness is 5.13 in hypothetical and 4.76 and 4.78 in hypothetical with oath and real with oath. A Mann-Whitney test indicates that happiness tends to have larger values in hypothetical than in hypothetical with oath ($p = .018$) and real with oath ($p = .062$). Mean happiness in real no oath treatments, by contrast, is comparable to that in hypothetical no oath treatments ($p = .122$). When no oath treatments are pooled against oath treatments, a Mann-Whitney test confirms the positive shift in happiness induced by the oath ($p = 0.018$). The EDF in Figures 3.c better illustrates this phenomenon. The EDF in non oath treatments first-order dominates the EDF in oath treatments, significantly in the hypothetical treatments ($p = .051$) and not significantly in the real treatments ($p = .110$).[23]

Such a decrease in average happiness induced by the oath suggests the oath may not be an

[22] In their well-known experiment on self-deception, Quattrone and Tversky (1984) explained to subjects that a certain medical condition was associated with cold tolerance (high or low depending on the experimental condition considered). Subjects were then asked to put their hands in very cold water as long as they could. Experimental results show that subjects tend to let their hands in the cold water longer (or less depending on the experimental condition) although this does not change the fact that they have the medical condition or not. In debriefing interviews, subjects are often not aware that they were self-deceiving themselves about this.

[23] The third column of Table 3 provides further statistical evidence on that result. We take happiness as an ordinal measure and estimate an ordered Probit similar to that applied to previous attitude variables. The regression 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 ($p = .202$). A test of equality of parameters associated with the hypothetical under oath and real under oath treatments however cannot reject the null of equality with $p = .972$. We also estimate the unique effect of being under oath, whether with or without monetary incentives. Results indicate that the parameter associated with the oath is negative and significant with $p = .015$.

Figure 4: Response time in the first round by treatments



innocuous instrument. Given the challenges of measuring happiness, a decrease in happiness 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 people are asked to form and declare their preferences. Our low refusal rate seems to contradict the first explanation.

A study of response time in the experiment provides further insights on the cognitive effort explanation.[24] If the oath induces more cognitive efforts because subject take the overall valuation task more seriously, we should observe an increase in response time under oath whatever the vote made by subjects. To avoid mixing up cognitive effort and learning, we focus only on the first voting round.[25] Figure 4 compares the observed response across treatments. The quartiles of the distribution of response times for all treatments together are first computed in round 1. We then compute the proportion of response times in period 1 in each treatment that fall in each quartile of the distribution. We observe an upward shift in response time in the hypothetical under oath treatment as compared to the hypothetical treatment. The proportion of response times in the first quartile decreases from 28.3% in hypothetical to 21.7% in hypothetical under oath whereas the proportion in the fourth quartile increases from 23.3% to 28.3%. The shift in response times induced by the oath seems to go the other way around in real treatments, with subjects in the real under oath treatment responding quicker than in the real treatment. There is an increase in the share of response times

[24] In using response times as a measure of cognitive effort, we follow Rubinstein (2007) who shows that short response time is usually associated with instinctive response and longer response time with choices based on more active cognitive reasoning.

[25] The results are fairly similar when considering all 5 votes. It is worth noting that in doing so, we loose a great deal of statistical power. Given that study of response time usually requires large sample sizes, the results can only be interpreted with caution. Statistics on response time for all 5 votes are available upon request.

Table 4: Happiness by vote in the first round and treatment

	Hyp.	Hyp Oath	Real	Real Oath
Vote "Yes"	5.26	4.82	5.44	4.29
Vote "No"	4.82	4.70	4.91	4.85

Note. For each treatment in column and each individual vote in the first round in row, the table provides the average level of happiness as measured by the post-experiment questionnaire (on a scale ranging from 1 to 7).

falling in the first quartile (21.7% in real and 26.7% in real under oath) and a decrease of the share in the fourth quartile (from 28.3% to 18.3%)—in which people take more than 16 seconds to vote. The small upward shift in response times in hypothetical and the small downward shift in real appear to go against the idea that the oath induces increased cognitive effort for all subjects under oath.

This asymmetric effect rather suggests another explanation. The oath may make self-serving assessments more costly, and affects only those subjects who are willing to vote yes. Thus, the change in happiness only comes from the reduction in the tendency to vote "yes". If this is the case and if we take self-declared happiness as a proxy for utility, we should observe that the oath only decreases happiness of subjects who vote "Yes" in oath treatments. Table 4 presents mean happiness levels by vote and treatment. We observe an asymmetry between "Yes" and "No" votes in non-oath treatments, subjects declare themselves happier when they voted "Yes" in the first round compared to when they voted "No". The difference is significant in hypothetical ($p = .0719$) and in real ($p = .055$) according to Mann-Whitney tests. When non-oath treatments are merged together, the difference is even more significant with $p = 0.018$. This is not the case for oath treatments. All subjects, whether they voted "Yes" or "No" express similar levels of happiness. The happiness level of subjects in the oath treatments corresponds to that of subjects who voted "No" in the first round in non-oath treatments.

To summarize our results to these three attitude questions, the correlations between self-reported measures with the treatment effects suggest truthfulness improves under oath: subjects are less prone to use the vote to express positive general attitudes towards public goods and they see themselves as more sincere in their answers. The study of happiness combined with that of response time supports the idea that the oath decreases the tendency to engage in self-serving assessments. That would also explain the small, albeit insignificant, decrease in "Yes" votes in real under oath as compared to the real treatment. This all suggests future avenues to explore the power and limits of using a rare but powerful social mechanism like the oath in an everyday setting like a survey.

5 Conclusion

Preference elicitation methods—even a straightforward approach like a binary voting referenda—can suffer from hypothetical bias. Evidence from laboratory experiments exploring hypothetical bias in

induced-value referendum voting behavior finds mixed results depending on whether one considers the aggregate vote or individual behavior: weak hypothetical bias at the aggregate level, but many mistakes and mis-voting at the individual level (Taylor, McKee, Laury, and Cummings, 2001). Without real economic commitment, people can allow their minds to wander towards social preferences, preference uncertainty, or inconsistent choices. When our choices do not have real consequences, we can experiment with our hypothetical votes or hypothetical dollars. We float trial balloons, we take straw polls, we test the wind. Without market pressure (i.e, money pump or arbitrage) to help concentrate the mind on the real consequences of the vote at hand, people are free to wander from the consistent and rational choice. The challenge is that rationality in economics is a social construct based on active market exchange, not an individual concept based on isolated introspection (see for example Arrow, 1987). For a more detailed discussion in the context of non-market valuation, see Shogren (2006, 2012) who discusses how to identify and create the missing institutional context, or "money pumps", to induce more rational choice in non-market valuation and environmental protection. These market or non-market institutions can help people help themselves by learning what it means to be the rational agents economists presume to live in our models (also see e.g., Cherry, Crocker, and Shogren, 2003). Herein we ask whether we can create this real economic commitment in a referendum vote with a non-market commitment device—the solemn oath.

Under normal conventions (money on the barrelhead), the oath acts as a substitute for real economic commitment (albeit an imperfect one). No real money is at risk. But if one broadens out the definition to include non-market goods lacking an exchange institute, the oath can be considered a real economic commitment within this non-market space. Commitment-via-oath helps a person match his or her words with deeds. While no real money is at risk, his or her real honor is—which to us is a real economic commitment through a non-market device in a non-market institutional setting. A Jeremy Bentham quote we have used before is worth repeating here: "What gives an oath the degree of efficacy it possesses, is, that in most points, and with most men, a declaration upon oath includes a declaration upon honor: the laws of honor enjoining as to those points the observance of an oath. The deference shown is paid in appearance to the religious ceremony: but in reality it is paid, even by the most pious religionists, much more to the moral engagement than to the religious" (Bentham, 1827).

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, Lofgren, Qin, and Sterner, 2013; de Magistris and Pascucci, 2014). 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 statistically identical. As we elicit preferences for a homegrown good, the results may not 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 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.

What are potential field applications of the referendum under oath? First, taking the oath from the lab to the field can be done within stated preference surveys designed to elicit truthful preferences for non-market goods, *e.g.*, Carlsson, Kataria, Krupnick, Lampi, Lofgren, Qin, and Sterner (2013). Using the popular referendum method with and without the oath we can test for potential impacts, although we have no rigorous baseline to judge complete success. Second and more broadly, the use of the voting-referendum under oath has been implemented for those registering to vote in Vermont. Either taken in-person or in front of a notary, every person registering to vote must take the Voter's Oath (previously called the Freeman's Oath):

> **The Voter's Oath**: You solemnly swear or affirm that whenever you give your vote or suffrage, touching any matter that concerns the State of Vermont, you will do it so as in your conscience you shall judge will most conduce to the best good of the same, as established by the Constitution, without fear or favor of any person. [*Voter's Oath, Vermont Constitution, Chapter II, Section 42*]

This seems to be the only explicit Oath asking the voter to vote based on his or her conscience. Other US states require each voter to swear under penalty of perjury that he or she meets the qualifications to vote (*e.g.*, not voted elsewhere, a state citizen, not convicted of a felony). An idea worthy of future research would be to explore whether a Vermont-style Voter's Oath could be implemented in other field applications.

On a related note regarding the use of a commitment device in a real world policy, consider the recent Paris Agreement on climate change. In Paris, countries used a "pledge and review" model as the primary mechanism to coordinate collective action on carbon emission reductions (see e.g., Aldy, Pizer, and Keigo, 2016; Fawcett et al., 2015). Paralleling the oath that we examine, the pledge is a voluntary commitment made by each country to develop and implement a domestic action plan subject to international review to evaluate the adequacy of the action. The pledge-as-commitment device is designed to act as a focal point mechanism to increase trust and to coordinate actions better. In a recent lab experiment on coordination under oath, we found that this was the case (Jacquemet, Luchini, Shogren, and Zylbersztejn, 2015). We observed that coordination with communication under oath improved by over fifty percent relative to the no oath baseline, and that senders were more likely to send truthful signals under oath, while receivers were slightly more likely to believe the signals sent. Further studies into how commitment and coordination under alternative allocation/sharing rules with the oath/pledge seems most worthwhile.

Finally, we acknowledge that we run risk of alienating people by using such a strong commitment device like the oath for a non-market valuation exercise. Our experience with the oath over the last decade however suggests the opposite: we observe a high acceptance rate (95 to 100%) and subjects say to the monitor while signing the oath in our experiments that "of course I will tell the truth" (we have had about 1000 subjects take the oath in the lab). Recall we designed our oath procedure based on the "compliance without pressure" literature: (1) subjects are free to sign, (2) participation and monetary gains are not conditional on signing the oath, and (3) subjects are unaware about what is going to happen in the lab afterwards. Social psychology tells us that under these three conditions people are most likely to comply without pressure or reactance, which appears to be what they do in our experiment. That said, one future experimental direction could be to see if one could use the oath to generate sincere behavior without the same level of moral pressure implied by "truth-telling". One idea that seems worthy to explore in future work is examine whether an oath that commits each subject to "take the valuation exercise seriously by thinking about my answers carefully and providing honest answers" would generate similar behavior as the truth-telling oath.

Our focus on the solemn oath was driven by the aim of investigating the strongest real-world commitment device we could replicate in the lab. Now that we have established the oath can create commitment, the next step is to back down from this strong position and explore how implementing weaker forms of commitment like a promise or pledge or even an honor code will affect behavior in a survey. This is an area worthy of further research.

References

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

Aldy, J., W. Pizer, and A. Keigo (2016): "Comparing emissions mitigation efforts across countries," *Climate Policy*, pp. 1–15.

Arrow, K. (1987): "Rationality of self and others in an economic system," in *Rational Choice*, ed. by R. Hogarth, and M. Reder. University of Chicago Press.

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.

Benabou, R., and J. Tirole (2002): "Self-confidence and personal motivation," *Quarterly Journal of Economics*, 117(3), 817–915.

Bentham, J. (1827): *Rationale of judicial evidence: Specially applied to english practice: in five volumes.* Hunt and Clarke, UK.

Burton, A. C., K. S. Carson, S. M. Chilton, and W. Hutchinson, George (2007): "Resolving questions about bias in real and hypothetical referenda," *Environmental and Resource Economics*, 38(4), 513–525.

Cameron, T. A. (1991): "Interval estimates of non-market resource values from referendum contingent valuation surveys," *Land Economics*, 67, 413–421.

Caplin, A., and J. Leahy (2001): "Psychological expected utility theory and anticipatory feelings," *Quarterly Journal of Economics*, 116(1), 55–79.

Carlsson, F., M. Kataria, A. Krupnick, E. Lampi, A. Lofgren, P. Qin, and T. Sterner (2013): "The Truth, the Whole Truth, and Nothing but the Truth - A Multiple Country Test of an Oath Script," *Journal of Economic Behavior & Organization*, 89, 105–121.

Carson, R., T. Groves, and J. List (2014): "Consequentiality: A Theoretical and Experimental Exploration of a Single Binary Choice," *Journal of the Association of Environmental and Resource Economists*, 1, 171–207.

Carson, R. T., and T. Groves (2007): "Incentive and informational properties of preference questions," *Environmental & Resource Economics*, 37(1), 181–210.

Champ, P. A., and C. Bishop, Richard (2001): "Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias," *Environmental & Resource Economics*, 19(4), 383–402.

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

Cherry, T., P. Frykblom, J. F. 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., T. D. Crocker, and J. F. Shogren (2003): "Rationality spillovers," *Journal of Environmental Economics and Management*, 45(1), 63–84.

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

Collins, J. P., and C. A. Vossler (2009): "Incentive compatibility tests of choice experiment value elicitation questions," *Journal of Environmental Economics & Management*, 59(2), 226–235.

Coursey, D., and W. Schulze (1986): "The Application of Laboratory Experimental Economics to the Contingent Valuation of Public Goods," *Public Choice*, 49(1), 47–68.

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., G. W. Harrison, and E. E. Rutström (1995): "Homegrown Values and Hypothetical Surveys: Do Dichotomous Choice Questions Elicit Real Economic Commitments?," *American Economic Review*, 85(1), 260–266.

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.

de Magistris, T., and S. Pascucci (2014): "Does "solemn oath" mitigate the hypothetical bias in choice experiment? A pilot Study," *Economics Letters*, 123(2), 252–255.

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.

Fawcett, A. A., et al. (2015): "Can Paris pledges avert severe climate change?," *Science*, 350, 1168–1169.

Green, D., K. E. Jacowitz, D. Kahneman, and D. McFadden (1998): "Referendum contingent valuation, anchoring, and willingness to pay for public goods," *Resource and Energy Economics*, 20(2), 85–116.

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

Haab, T., J. Huang, and J. Whitehead (1999): "Are hypothetical referenda incentive compatible? A comment," *Journal of Political Economy*, 107(1), 186–196.

Harrison, G. (2006): "Experimental Evidence on Alternative Environmental Valuation Methods," *Environmental & Resource Economics*, 34(1), 125–162.

Jacquemet, N., A. James, S. Luchini, and J. F. 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. F. 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 (2015): "Coordination with communication under oath," *Working Paper, Paris School of Economics.*

James, A. G., and J. F. Shogren (2015): "Revisiting the effect of voter isolation," in *Replication in Experimental Economics*, ed. by C. Deck, E. Fatas, and T. Rosenblat. Emerald Group Publishing Limited, Bingley, UK.

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

Kahneman, D., and J. Knetsch (1992): "Valuing public goods: The purchase of moral satisfaction," *Journal of Environmental Economics and Management*, 22(1), 57–70.

Kahneman, D., and R. Sugden (2005): "Experienced utility as a standard of policy evaluation," *Environmental & Resource Economics*, 32(1), 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.

Köbberling, V. (2006): "Strength of preferences and cardinal utility," *Economic Theory*, 27(2), 375–391.

Konow, J. (1996): "A positive theory of economic fairness," *Journal of Economic Behavior & Organization*, 31(1), 13–35.

List, J. A., R. P. Berrens, A. K. Bohara, and J. Kerkvliet (2004): "Examining the role of social isolation on stated preferences," *American Economic Review*, 94, 741–752.

Loomis, J. (2014): "Strategies for Overcoming Hypothetical Bias in Stated Preference Surveys," *Journal of Agricultural and Resource Economics*, 39(1), 34–46.

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.

Luchini, S., and V. Watson (2013): "Uncertainty and framing in a valuation task," *Journal of Economic Psychology*, 39, 204–214.

McConnell, K. K. (1990): "Models for referendum data: the structure of discrete choice models for contingent valuation," *Journal of Environmental Economics and Management*, 18(1), 19–34.

Messer, K., G. Poe, W. Rondeau, W. Schulze, and C. Vossler (2010): "Social preferences and voting: An exploration using a novel preference revealing mechanism," *Journal of Public Economics*, 94(3-4), 308–317.

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.

Mozumder, P., and R. P. Berrens (2007): "Investigating hypothetical bias: induced-value tests of the referendum voting mechanism with uncertainty," *Applied Economics Letters*, 14(10), 705–709.

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.

Murphy, J. J., T. H. Stevens, and L. Yadav (2010): "A comparison of induced value and home-grown value experiments to test for hypothetical bias in contingent valuation," *Environmental and Resource Economics*, 47(1), 111–123.

Polomé, P. (2003): "Experimental evidence on deliberate misrepresentation in referendum contingent valuation," *Journal of Economic Behavior & Organization*, 52(3), 387–401.

Quattrone, G., and A. Tversky (1984): "Causal versus diagnostic contingencies: on self-deception and on the voter's illusion," *Journal of Personality and Social Psychology*, 46(2), 237–248.

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

Rustichini, A. (2008): "Neuroeconomics: formal models of decision making and cognitive neuroscience," in *Neuroeconomics: Decision making and the brain*, ed. by P. W. Glimcher, and E. Fehr. Academic Press.

Schläpfer, F., A. Roschewitz, and N. Hanley (2004): "Validation of stated preferences for public goods: a comparison of contingent valuation survey responses and voting behaviour," *Ecological Economics*, 51(1), 1–16.

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.

Shogren, J. F. (2005): "Experimental Methods and Valuation," in *Handbook of Environmental Economics*, ed. by K.-G. Mäler, and J. Vincent. Elsevier, Amsterdam.

Shogren, J. F. (2006): "A Rule of One," *American Journal of Agricultural Economics*, 88(5), 1147–1159.

——— (2012): "Behavioral Environmental Economics: Money Pumps & Nudges," *Journal of Agricultural and Resource Economics*, 37(3), 349–360.

Shogren, J. F., and L. Tadevosyan (2011): "Le Comportement d'Enchérisseur dans une Enchère Conséquentialiste au Second Prix," *Revue française d'économie*, 2011, 13–28.

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.

Tavoni, A., A. Dannenberg, G. Kallis, and A. Löschel (2011): "Inequality, communication, and the avoidance of disastrous climate change in a public goods game," *Proceedings of the National Academy of Sciences*, 108(29), 11825–11829.

Taylor, L. O., M. McKee, S. K. Laury, and R. G. Cummings (2001): "Induced-value of the referendum voting mechanism," *Economics Letters*, 71(1), 61–65.

Vossler, C. A., M. Doyon, and D. Rondeau (2012): "Truth in consequentiality: theory and field evidence on discrete choice experiments," *American Economic Journal: Microeconomics*, 4(4), 145–171.

Vossler, C. A., and J. Kerkvliet (2003): "A criterion validity test of the contingent valuation method: comparing hypothetical and actual voting behavior for a public referendum," *Journal of Environmental Economics and Management*, 45(3), 631–649.

Vossler, C. A., and M. McKee (2006): "Induced-value tests of contingent valuation elicitation mechanisms," *Environmental and Resource Economics*, 35(2), 137–168.

Williams, R. (2012): "Using the margins command to estimate and interpret adjusted predictions and marginal effects," *Stata Journal*, 12(2), 308–331(24).

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 unhappy					Very Happy	

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

+	+	+	+	+	+	+
1	2	3	4	5	6	7
Not at all honest					Totally 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 Simulations of voting behavior in the referendum and of bidding behavior in Vickrey auctions

Herein, 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 to or greater than 11 euros is interpreted as a "yes" vote and "no" vote otherwise.

We proceed by random sampling of subjects with replacement as for the boostrap proportion tests to account for within-subject correlation. In each simulation, we draw subjects and their five bids in the Vickrey auctions sample. We proceed similarly for the referendum experiment, we draw subjects and their five votes. In a typical auction simulation, we compute the proportion of bids equal to or greater than 11 euros whereas in a referendum simulation we obtain directly this proportion by counting "Yes" votes . We then generate confidence intervals from both experiments based these simulations. 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%]	—

C 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)	