Journal of Economic Behavior & Organization 89 (2013) 105–121

Contents lists available at SciVerse ScienceDirect

Journal of Economic Behavior & Organization

journal homepage: www.elsevier.com/locate/jebo

The truth, the whole truth, and nothing but the truth—A multiple country test of an oath script

Fredrik Carlsson[a], Mitesh Kataria[b,*], Alan Krupnick[c], Elina Lampi[a], Åsa Löfgren[a], Ping Qin[d], Thomas Sterner[a]

[a] Department of Economics, University of Gothenburg, Box 640, SE-40530 Gothenburg, Sweden
[b] Max Planck Institute of Economics, Kahlaische Straße 10, D-07745 Jena, Germany
[c] Resources for the Future, 1616 P St., NW, Washington, DC 20036, USA
[d] Renmin University of China, School of Economics, 100872 Beijing, China

ARTICLE INFO

Article history:
Received 16 December 2011
Received in revised form 11 February 2013
Accepted 16 February 2013
Available online 27 February 2013

JEL classification:
D61
Q5

Keywords:
Oath script
Hypothetical bias
Willingness to pay

ABSTRACT

Hypothetical bias is one of the main issues bedeviling the field of nonmarket valuation. The general criticism is that survey responses reflect how people would like to behave, rather than how they actually behave. In our study of climate change and carbon emissions reductions, based on the increasing bulk of evidence from psychology and economics regarding the effects of making promises, we investigate the effect of an oath script in a contingent valuation survey. The survey was conducted in Sweden and China, and its results indicate that an oath script has significant effects on respondent behavior in answering willingness-to-pay (WTP) questions. In both countries, the shares of zero WTP responses and extremely high WTP responses decrease when an oath script is used, resulting in lower variance. Furthermore, the conditional WTP decreases in the Chinese but not in the Swedish sample. We also find that the effect of the oath script is not generally dependent on respondent characteristics, and the few differences we observe vary with the countries.

1. Introduction

Stated preference methods are frequently used for valuing improvements in public and quasi-public goods. In contrast to alternative approaches, only stated preference methods can reliably estimate non-use values, which represent a significant component of the total benefit for many public goods. One of the major concerns with stated preference approaches is to what extent survey responses are consistent with actual decision-making behavior. While the empirical evidence is mixed, most studies suggest that willingness-to-pay (WTP) estimates are higher in a hypothetical setting than the corresponding real setting, an inconsistency often called hypothetical bias[1] (see, e.g., Cummings et al., 1995, 1997; Frykblom, 1997; List and Gallet, 2001; and Murphy et al., 2005, for meta-analyses). Many studies in the contingent valuation literature explore different methods of mitigating hypothetical bias. One of the most successful and frequently imitated efforts is the use of a cheap talk script, initially suggested by Cummings and Taylor (1999). The cheap talk script aims to reduce hypothetical bias

* Corresponding author. Tel.: +49 3641 686 632; fax: +49 3641 686 667.
E-mail addresses: fredrik.carlsson@economics.gu.se (F. Carlsson), kataria@econ.mpg.de (M. Kataria), krupnick@rff.org (A. Krupnick), elina.lampi@economics.gu.se (E. Lampi), asa.lofgren@economics.gu.se (Å. Löfgren), pingqin@ruc.edu.cn (P. Qin), thomas.sterner@economics.gu.se (T. Sterner).

[1] The term hypothetical bias is somewhat misleading, as it suggests that one single real value exists with which we can compare the value obtained from a hypothetical situation. However, studies have shown that experimental situations involving real payoffs can also lead to biased estimates (e.g., Alpizar et al., 2008; List et al., 2004).

0167-2681/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jebo.2013.02.003

by describing the general propensity of survey respondents to exaggerate their stated willingness to pay. The underlying idea of the script is that by raising the issue in the survey, respondents should become aware of how the hypothetical nature of the inquiry might potentially influence their valuation, and therefore attempt to allocate sufficient effort to the choice problem and in doing so become less prone to hypothetical bias. The effect of a cheap talk script varies and, among other factors, its success seems to depend on the characteristics of the good, the length of the script, and the valuation method (see Aadland and Caplan, 2003, 2006; Carlsson et al., 2005; Cummings and Taylor, 1999; List, 2001; and Murphy et al., 2005).[2,3] Furthermore, respondents may still be tempted to answer insincerely due to motives such as projecting a "better" self-image, applying strategic behavior, or protesting against the survey.[4] Regardless of their motives, it is important to diminish the problem.

A more recent approach to addressing hypothetical bias is the use of an oath script. Under this approach, respondents are asked to swear (or promise) to answer truthfully, mimicking the act of taking an oath in a courtroom. An oath can be viewed as an active commitment, in contrast to a cheap talk script, which is merely informative. Whereas a cheap talk script informs about the general tendency of overstating the willingness to pay in a hypothetical study, an oath may actually induce more thoughtful and truthful answers to hypothetical valuation questions, thus mitigating hypothetical bias, by appealing to respondents' moral incentives.

The seminal paper by Jacquemet et al. (2013) on the effects of oaths explores the impact of an oath script in a laboratory setting. Subjects, drawn from a sample of university students in France, participated in an incentive-compatible induced-value second-price auction or a home-grown value auction. The study employed a between-subject design with four induced value treatments that varied by the nature of the decision (real or hypothetical) and the presence or absence of the oath script. The two oath treatments presented subjects with the opportunity to sign an oath document prior to participating in the auction. By signing the oath document, subjects agreed to "swear on their honor" to tell the truth and provide honest answers. Participation in the experiment was not conditional on taking the oath. Two important results emerge from the experiment. First, subjects who took the oath were on average less likely to overstate or understate their bids; that is, the variance of bids was reduced. Second, the hypothetical payment treatment with the oath outperformed both non-oath treatments and the real payment treatment with the oath.[5] In a subsequent experimental study that compares voting behavior in real and hypothetical referenda, Jacquemet et al. (2010) find that the oath script eliminated hypothetical bias.

The aim of our study is to investigate the effect of an oath script outside the laboratory, in a relatively large-scale contingent valuation survey conducted in two countries, Sweden and China.[6] The survey elicits hypothetical values for three different levels of reductions in greenhouse gas emissions. Our study makes two primary contributions to the literature. First, it represents the first field test of an oath script. In contrast to the previous laboratory tests that involved comparisons between behavior in hypothetical and real payment settings with and without an oath in an induced value setting, a field test of the oath script simply relies on comparisons between stated choices with and without the oath script. That is, since we cannot observe true WTP in a field survey (indeed this limitation necessitates the field survey), our test of the oath mechanism involves comparing estimated WTP across treatments with and without an oath.

Second, by testing the oath script in Sweden and China, we explore the potential for the oath script to mitigate hypothetical bias in two culturally different settings. Ehmke et al. (2008) compare hypothetical and real behavior in five different countries/states, including China,[7] and find that both the direction and degree of hypothetical bias differ across countries and states. This suggests that the effect of an oath script might differ depending on the cultural context. Unfortunately, the evidence on cross-country differences is rare, and we are not aware of any other studies that have investigated differences in hypothetical bias between different cultural contexts. However, there are related studies that show that the effect on respondent cheating behavior can be affected by respondents' attitudes to cheating and lying, which we believe could alter the impact of an oath on respondent behavior. Studies analyzing cross-country differences in cheating behavior and attitudes among students (see, e.g., Rawwas et al., 2007; Chapman and Lupton, 2004; Lupton et al., 2000) suggest that there are significant cross-national differences in cheating behavior, and there are also studies indicating that the concept of lying

[2] There are also other suggestions for how to reduce hypothetical bias, e.g., ex-post calibration of the WTP responses with follow-up questions on certainty (e.g., Champ et al., 1997; Champ and Bishop, 2001; Johannesson et al. 1999) and time-to-think protocols (Cook et al., 2007; Whittington et al., 1992).

[3] Aadland and Caplan (2006) provide evidence concerning the effects of a neutral cheap-talk script. They find that a neutral cheap-talk script, which informs the respondents that people have a tendency to misstate their true WTP instead of telling them that people have a tendency to overstate their WTP, actually exacerbates the hypothetical bias. Ideally, one would therefore like to inform respondents about the tendency of people to bias their WTP. This could help respondents in their cognitive process, but also ensure that each individual gives a well-considered response and not heuristically reduce their stated WTP in an attempt to please the survey administrators.

[4] Lusk and Norwood (2009) found that people get utility not only from being generous (as in the case of warm glow) but also from "looking good" to others. Champ and Bishop (2001) trace the hypothetical bias to respondents who are not sure about their hypothetical yes-answers and emphasize the role of uncertainty.

[5] Jacquemet et al. (2013) find that monetary incentives weaken the positive impact of oath-taking and discuss that as an extrinsic motivation. In other words, money may have a crowding-out effect on an intrinsic motivation, such as keeping one's promises after taking an oath.

[6] What works in lab experiments does not necessary work outside the lab. There are many reasons for this. For example, differences in population profiles (experiments in economics and psychology often use students as subjects) and demand effects (i.e., subjects want to comply with what the experimenter expects from them) are two well-known sources of possible divergence between lab and field behavior. Furthermore, an auction-based value revelation mechanism may trigger different behavior and feelings (for example a pleasure of winning) than a non-market valuation survey.

[7] The five different countries/states in the Ehmke et al. (2008) study are Kansas and Indiana in the U.S. and China, Niger, and France.

differs between cultures (Fu et al., 2001). Fu et al. (2007) find that there seem to be few cross-cultural differences regarding whether or not a lie is categorized as a lie; however, whether it is morally acceptable to lie in particular circumstances is affected by the cultural context. The authors compare Chinese and Canadian children and show that when evaluating a lie, the Chinese children tended to emphasize the effects on the collective rather than on the individual to a larger extent than did the Canadian children, which affected the moral judgment about the use of deception. This is argued to be well in line with the Chinese cultural values, i.e., to be group-oriented.

Sociologists identify several factors that affect ethical attitudes (Hofstede, 2001; House et al., 2004) and that could, by extension, affect cheating behavior and the perception of when it is morally acceptable to lie (such as the argument above that Chinese are group-oriented). Cultural differences that could influence the effect of an oath in different countries include respect for authorities, trust, and religious beliefs, all factors that may vary considerably both between and within countries. The consequences of breaking laws, e.g., perjury, may also vary quite substantially between different legal systems and hence influence everyday behavior differently. Finally, particularly relevant for this study is that we know from Krupnick et al. (2010) that a tendency for yea-saying in stated preference surveys is particularly strong in China relative to Western countries. It is plausible that an oath could reduce this tendency, with a greater effect in China than in Sweden. By testing the oath script in China and Sweden, we explore the potential for the oath script to mitigate hypothetical bias in two culturally different settings.

In addition to a direct comparison between the two samples, we can identify which socio-demographic groups of respondents are more likely to be affected by the oath script and whether the results vary significantly within the Chinese and Swedish samples.

We find several notable results. In both countries, the oath results in lower variance in WTP, a result that is in line with the findings by Jacquemet et al. (2013). In both countries, respondents are less likely to state a zero WTP in the oath treatment. The effect of oath on conditional WTP (i.e., the WTP of those with a positive WTP) differs by country: it reduces the average conditional WTP substantially in the Chinese sample, while the effect is insignificant in the Swedish sample. Turning to unconditional WTP, for Sweden we find that the oath script significantly increases average unconditional WTP for two of the three CO_2 reduction levels, while for China the increase is significant for one reduction level. Finally, with a few exceptions, the effect of the oath script did not systematically vary with the respondent characteristics that we were able to measure.

The paper is organized as follows. Section 2 describes the survey and the oath script, as well as hypotheses on the effect of the oath script. Section 3 gives the characteristics of the respondents. Section 4 presents the WTP results and econometric analysis, including cross-country differences and an analysis of which groups of respondents are more likely to change their behavior as a result of the oath script. Section 5 concludes the paper.

2. Descriptions of the survey and the oath script

The survey consisted of four sections. The first section elicited general attitudes about climate change, e.g., whether climate change is in fact occurring and if so whether humans are causing it. In the second section, the survey provided information on the effects of climate change, based on IPCC (Intergovernmental Panel on Climate Change) reports. Respondents were told that a future temperature increase will depend on the amount of future global CO_2 emissions; specifically, if CO_2 emissions are reduced from current emissions levels by 30 percent, 60 percent, or 85 percent, then the temperature increase can be limited to 4 °F, 3 °F, or 2 °F, respectively. If the world instead does not reduce emissions, but continues with "business as usual," the temperature is expected to increase by more than 4 °F by 2050. We explained that this would most likely imply large changes in the global ecosystems and that most countries would be negatively affected. An information screen (Table A1 in Appendix A) summarized the effect of the temperature increases on harvests, flooding, storms, and ecosystems by the year 2050. Subsequently, questions about the respondents' attitudes on reducing global CO_2 emissions were included, e.g., how countries should use their resources to reduce CO_2 emissions and where the reductions should be made. For a full description of the survey and the implementation, see Carlsson et al. (2012).

The end of the second section included a brief cheap talk script, followed by the oath script (only in the oath treatment) and the WTP questions. The cheap talk script read: "Before making your choices, please consider how an increased cost would affect your possibilities of buying other things. Previous studies of this kind have shown that people generally claim to be willing to pay more money than they actually would in a real situation. Given this, it is important for you to answer these questions as truthfully as possible." The oath script is discussed in more detail in Section 2.1.

The survey elicited three distinct WTP values using the payment card method. For each of the described reductions in global greenhouse gas emissions, each respondent chose the number among those presented in a matrix that represented their maximum WTP. We opted for the payment card approach over the incentive-compatible dichotomous choice method based on the previous work by Krupnick et al. (2010), which found that the dichotomous choice method failed some important validity tests in China, mostly caused by frequent yea-saying. The values in the payment card matrix, which were the same for each of the three WTP questions, ranged from US$0 to US$220.[8] Respondents who selected the highest value had the

[8] More specifically, the values in the matrix ranged from SEK 0 to SEK 2000 in the Swedish survey, and from CNY 0 to CNY 740 in the Chinese survey. SEK = Swedish kronor; CNY = Chinese yuan. For both countries, this corresponds to PPP-adjusted US$0–220; SEK 9.3 = PPP-adjusted US$1, and CNY 3.4 = PPP-adjusted US$1.

option to state their maximum WTP in a subsequent open-ended payment question. The first valuation question elicited WTP for a 30 percent reduction in global emissions, compared with doing nothing (no reduction). The second WTP question asked the respondents how much more than indicated for question 1 they would pay for an *additional* 30 percent reduction (i.e., a total reduction of 60 percent). Finally, the third WTP question asked the respondents how much more than indicated for question 2 they would pay for an *additional* 25 percent reduction (i.e., a total reduction of 85 percent). To ensure that respondents would understand this correctly, both the actual value chosen for a specific reduction level (e.g., X for a 60% reduction) ***and*** the total WTP value for the same reduction level (i.e., $Y+X$, where Y was the WTP for the 30% reduction level) were shown on the screen. In Table A2 in Appendix A, we show the WTP question for the 30 percent reduction.

Our design is somewhat unusual in the sense that we have repeated contingent valuation questions. Since the three questions are linked in a logical way, by asking for additional willingness to pay, the cognitive burden on the respondents should not be too large. The main advantage of letting the respondents answer all three questions is that it increases statistical efficiency compared with a between-subject design. The main drawback is that there is a risk of order effects in the sense that the responses to the second and third question depend on the fact that the respondents received the first question.[9] The reason for why we use the payment card instead of a choice experiment is the increased amount of information we receive from respondents using the payment card method rather than dichotomous choices in a choice experiment.

The payment was expressed as a monthly cost for the household until 2050, and examples of the typical ways households would pay were listed, such as increased energy and gasoline prices.[10] Moreover, we asked the respondents to assume that the costs of reducing CO_2 emissions are shared among all countries according to their respective shares of the total emissions in the world.[11] When the respondents chose their WTP values, both the monthly payment and the corresponding annual cost for the household were shown on the screen. The purpose of this was to ensure that the respondents clearly understood how much they said they were willing to pay.

The third section of the survey was a choice experiment on rules for allocating the responsibilities for CO_2 reductions across countries. This topic is not addressed in the present paper. Finally, the fourth section asked questions about the respondent's socioeconomic characteristics.

2.1. The oath script

The surveys with and without the oath were identical except for the inclusion of the oath script. Before inviting respondents to take the oath, we informed them that they would be asked questions about their household's willingness to pay for CO_2 emissions reductions.

In our oath design, respondents were asked, "Do you feel you can promise us to answer the questions that will follow as truthfully as possible?" The alternatives were (1) "Yes, I promise to answer the questions in the survey as truthfully as possible," or (2) "No, I cannot promise this." From a practical viewpoint, such a commitment device is reminiscent of a common practice in Anglo-Saxon courts, where a witness is instructed to take an oath "to tell the truth, the whole truth, and nothing but the truth." However, asking a witness to swear to tell the truth is not customary in the Swedish and Chinese court systems. Respondents might also feel uncomfortable and regard it as strange to be forced to swear an oath in a survey. Therefore, we opted for a more neutral wording of the question, using "promise" instead of "swear."

The essence of an oath script is to increase the respondent's commitment and attention to the survey, diminishing the propensity of light-hearted or "insincere" responses. This could be especially important in combination with the cheap-talk script since the cheap-talk script can tempt some respondents to arbitrarily reduce their willingness to pay.[12]

Theoretical support for the oath can be found in the theory of commitment in social psychology (Kiesler, 1971; Jacquemet et al., 2013; Joule and Beauvois, 1998). There is also evidence in economic research suggesting that a promise can induce

[9] Although the empirical evidence is mixed, giving more than one valuation question to the same respondent implies a risk of order and learning effects (see, e.g., Bateman et al., 2008; Day et al., 2010; Day and Pinto, 2010). In order to reduce this risk, respondents were presented with all three scenarios of emissions reduction before they answered any WTP questions. In addition, the second and third questions were related to the preceding questions in the sense that respondents were asked to state how much more they would pay for an additional reduction. Therefore, we would expect value learning during the course of the three valuation questions to be fairly limited. This design was also more natural than beginning with the largest reduction and then asking how much less a respondent would pay for a smaller reduction. An alternative would have been to use a split sample design, but given the small number of observations we thought the benefits outweighed the potential costs.

[10] Wiser (2007) finds that the WTP to support climate change policy depends on the payment vehicle used in the study. For example, the elicited WTP is higher with a collective payment mechanism than with a voluntary payment mechanism. In our study, we did not test for different payment vehicles, but rather made clear how the payments were made and kept this constant across countries.

[11] The text read: "We will now ask you about your household's willingness to pay for CO_2 emissions reductions. Reducing emissions will be costly for households, mainly because of increased energy costs, such as higher electricity and gasoline prices. Your household and your descendants will have to pay a monthly cost until the year 2050. Moreover, the cost will be adjusted for inflation over time. Also, assume that the costs of reducing CO_2 emissions are shared among the countries according to their emissions today."

[12] In experimental economics and psychology this is a well-known phenomenon known as a demand effect, which refers to an experimental artifact where participants form an interpretation of the experimenters' purpose and change their behavior accordingly (Zizzo, 2010).

an emotional commitment to fulfill the promise (Braver, 1995; Ostrom et al., 1992; Ellingsen and Johannesson, 2004) or that simply reading an honor code can reduce unethical behavior (Mazar et al., 2008; Shu et al., 2011a). Ellingsen and Johannesson (2004) proposed a model that includes preferences for keeping one's word. In another study, Vanberg (2008) examined whether people keep their promises because they have preferences for keeping their word or because they dislike letting others down. His results suggest that people have preferences for promise keeping per se. Jacquemet et al. (2013) tested whether their results changed when the respondents were first asked to read a warning that lying might cause negative consequences for other people. They find that the results of the oath are independent of this warning and conclude that the oath works through the intrinsic motives of a person to tell the truth. In our case, the respondents were anonymous survey participants, meaning that any effects of the oath would be through internal, not external, processes. Shu et al. (2011b) find that signing a form at the beginning of a self-report leads to significant reductions in dishonesty than does signing at the end, which is the current common practice. In our case it is of course natural to include the oath script before the actual valuation questions.

An oath script might also be seen as a signal that the topic of the study is important and that one's answers matter more than in a survey without an oath script. We do not assume, however, that all people not telling the truth will admit it. What is crucial is that we emphasize that it is important to tell the truth, and thus the answer to the oath question is of secondary importance. In our study, most subjects chose to promise to tell the truth. To estimate treatment effects of including the oath script, we included all subjects in the analysis irrespective of their responses.[13]

2.2. Hypotheses on the effect of an oath script

What is the expected effect of an oath script in a payment card setting? Previous literature motivates our hypotheses. If hypothetical bias leads some respondents to overstate their WTP and the oath script reduces this tendency, then one might expect that an oath would shift the whole WTP distribution to the left, including an increase in the number of zero values.[14] This would lead to a lower valuation in general and a lower average WTP in particular. Alternatively, some respondents may understate their WTP or state a zero WTP in a hypothetical setting (see Carson and Groves, 2007, for a detailed discussion). For example, respondents might state a zero WTP as a way of protesting against the whole, or parts, of the survey instrument. Alternatively, a respondent with a relatively low WTP might state zero WTP because it is a simple way to answer a complex question as opposed to devoting the required attention to consider the question carefully. If the oath script alters these incentives, then we would expect to observe a decrease in the number of zero WTP responses with the oath script. Jacquemet et al.'s (2013) experimental results, i.e., that taking an oath decreased both the frequency of zero bids (interpreted as a decrease in opting-out behavior) and extreme bids that violated the earnings budget constraint, are consistent with this intuition. Consequently, if our sample includes both respondents who tend to overstate WTP in a hypothetical setting and respondents who tend to understate WTP (in particular to state a zero WTP), then we expect the oath script to reduce the number of extreme WTP values and the number of zero WTP values. Thus, we hypothesize that including an oath script will lower conditional WTP (i.e., $WTP|WTP > 0$) and decrease the share of respondents stating zero WTP, leaving the effect on average WTP indeterminate but the variance of WTP lower.[15]

3. Administration of surveys and description of samples

The two surveys were conducted in November and December 2009. The questionnaire was designed and tested on several focus groups and pilot studies. The survey was designed to be self-administered on the computer to eliminate interviewer bias and strategic answering to please the interviewer. In China, the survey was conducted on laptops in special rooms with invited respondents. In Sweden, the survey was taken online. The survey yielded 2406 responses in Sweden (1176 for the version with and 1230 for the version without an oath script) and 550 responses in China (273 for the version with and 277 for the version without an oath script).[16] The Chinese survey was administered in the city of Nanning, the capital of the autonomous Guangxi region in southwest China. Respondents were randomly selected to participate in the survey using a neighborhood-based sampling strategy that has been used in previous surveys (Krupnick et al., 2010).[17] The respondents in the Swedish survey were reached using panel members of "Panel.se," Sweden's largest survey panel with around 100,000

[13] We also conducted an analysis that excluded the few respondents who answered *no* to the oath question, and our results remained robust.

[14] Warm-glow effects (Andreoni, 1990), moral satisfaction (Kahnemann and Knetsch, 1992), or strategic behavior could lead respondents to exaggerate their stated WTP. In a payment card setting, respondents may be tempted to overstate their WTP if they assume that a favorable decision depends on whether or not the survey produces a sufficiently large average value.

[15] As expressed in Carson (2012): "Which position does the empirical evidence support: the 'hypothetical bias' prediction that surveys will overestimate true willingness to pay, or the Samuelson's prediction that strategic behavior will lead to an underestimate? The answer is 'both'."

[16] In both countries, this corresponds to a response rate of around 25%.

[17] To develop a sample that was reasonably representative of the Nanning population, we randomly selected 273 respondents to take the survey with and 277 respondents to take the survey without the oath script. Although samples within any community should ideally be stratified for income and education, there is no public information database in Nanning that includes education, income, gender, and other important demographic characteristics. Therefore, we randomly chose families through a population information network set up by the Nanning government. Specifically, the survey teams adopted a five-stage random sampling method to select respondents. The primary sampling unit was the city district, the second sampling unit the urban sub-district, the third the neighborhood committee, the fourth households, and the fifth the individual.

Table 1
Descriptive statistics.

Variable (description)	Sweden				China			
	Without oath	*Std. dev.*	*With oath*	*Std. dev.*	*Without oath*	*Std. dev.*	*With oath*	*Std. dev.*
Female (1 if female)	0.483		0.490		0.484		0.440	0.497
Age (in years)	49.679	15.431	49.821	15.366	53.950	13.965	53.934	14.106
No. of adults in household (>18 years)	1.868	0.677	1.845	0.632	3.256	1.350	3.223	1.200
No. of children in household (<18 years)	0.522	0.911	0.526	0.946	0.596	0.809	0.531	0.697
Household members with university education (1 if at least one)	0.374		0.412		0.274		0.253	
Income (thousand SEK/CNY)	3.386	1.526	3.413	1.492	1.034	0.952	0.915	0.784
Household active in religious organization (1 if active)	0.064		0.065		0.025		0.011	
Left-wing party	0.315		0.305					
Green party	0.119		0.134					
Other party	0.154		0.150					
Center/right-wing party (ref. category)	0.413		0.411					
Communist					0.170		0.304	
No temperature increase (1 if agree)	0.059		0.037		0.036		0.103	
Humans affect temperature increase (1 if agree)	0.939		0.946		0.964		0.908	
No. in sample	1230		1176		277		273	

members. The panel members are recruited by telephone via random digit dialing and through online recruitment. Panel members 18–74 years old were randomly selected to participate in the survey.

Descriptive statistics of the four samples are presented in Table 1. Using a chi-square test (binary variables) and a *t*-test (remaining variables), we cannot reject the hypothesis of equal means/distributions for any of the socio-economic variables in Sweden, except for university education. In China, the same is true for all socio-economic variables other than members of the Communist Party and attitudes to climate change. Thus, the two samples in the two different treatments are generally comparable to each other, but there is a need to control for the differences that do exist since they could have important effects on the WTP for reduction of greenhouse gases. Also note that a large fraction of the respondents in the oath treatment group promised to answer the survey truthfully: 98 percent in the Swedish sample and 95 percent in the Chinese sample. This is in line with Jacquemet et al. (2013), where only one respondent elected not to sign the oath in one of the experiments, corresponding to a refusal rate of just over 5 percent.

4. Willingness to pay results

We analyze responses from the three payment card questions using several different methods. We begin by exploring the raw data. We calculate the mean monthly household WTP for three levels of reduction of carbon dioxide emissions, 30%, 60%, and 85%, all relative to a zero percent reduction baseline. For respondents whose selected WTP value was lower than $220, we assign a WTP equal to the midpoint of the payment card interval. For example, if a respondent selected a WTP value of $40 for the first payment card question, then we assigned a household WTP of $45 for the 30% reduction (since $50 is the value closest to but greater than $40 on the payment card). For respondents who provided open-ended WTP values, we assigned actual point values. Table 2 reports these unconditional mean WTP results in PPP-adjusted U.S. dollars for the Swedish and Chinese samples with and without the oath script for the three reduction levels.[18] The table also reports conditional mean monthly household WTP as well as the share of respondents reporting a zero WTP. In Appendix B we present the full distribution of responses for the various reduction levels and treatments (Tables B1 and B2) followed by histograms for the 0–85% reduction scenario.

A number of regularities are apparent in Table 2. First, Swedes have a substantially higher WTP for reductions in CO_2 emissions at all reduction levels, whether on a conditional or unconditional basis. The share of Swedes who selected a zero WTP value is lower than the corresponding share of Chinese for all reduction levels. Turning to differences in the treatments, we had two hypotheses about the effects of an oath: on the conditional WTP and on the share of respondents with zero WTP. We find ample evidence supporting the hypothesis that the share of zero bids is lower for oath-taking subjects. Indeed, for every CO_2 reduction level in both countries, the share of subjects stating a zero WTP is lower in the oath script treatment. Proportion tests indicate that there are significant differences between the shares of zero WTP for all reduction levels in Sweden, while the difference is not significant in China for any reduction level. The raw data indicate a significant and negative treatment effect of the oath script on conditional WTP in the Chinese sample but not in the Swedish sample. In fact,

[18] There are some differences between the respondents who agreed to answer truthfully and those who did not. In both the Swedish and Chinese samples, there is a larger fraction of zero WTP responses and lower WTP responses for those who did not agree to answer truthfully, but no significant differences. Given the small sample sizes for those who did not agree to answer, the comparisons and tests are not very reliable.

Table 2
Monthly WTP in PPP U.S. Dollars and Share of Subjects with Zero WTP.

CO_2 reduction	Sweden					
	Without oath (1230 obs.)			With oath (1176 obs.)		
	All *Mean (std. dev.)*	*WTP > 0* *Mean (std. dev.)*	*Share zero WTP*	*All* *Mean (std. dev.)*	*WTP > 0* *Mean (std. dev.)*	*Share zero WTP*
0–30%	24.08 (40.68)	26.49 (41.91)	0.091	28.08 (40.51)	30.16 (41.23)	0.069
0–60%	43.87 (83.15)	47.84 (85.73)	0.083	49.00 (72.19)	52.25 (73.40)	0.062
0–85%	61.01 (133.23)	66.29 (137.61)	0.080	64.98 (100.38)	69.28 (102.20)	0.062
CO_2 reduction	China					
	Without oath (277 obs.)			With oath (273 obs.)		
	All *Mean (std. dev.)*	*WTP > 0* *Mean (std. dev.)*	*Share zero WTP*	*All* *Mean (std. dev.)*	*WTP > 0* *Mean (std. dev.)*	*Share zero WTP*
0–30%	4.48 (8.43)	6.46 (9.48)	0.307	3.51 (3.91)	4.67 (3.87)	0.249
0–60%	7.87 (17.53)	10.58 (19.62)	0.256	5.45 (5.80)	6.95 (5.70)	0.216
0–85%	11.10 (28.00)	14.71 (31.41)	0.246	7.02 (7.82)	8.87 (7.81)	0.209

the conditional WTP is *higher* in the oath treatment for the Swedish sample (although this difference is significant only for the 30% reduction).

While informative, our analysis of the raw data is limited since we set WTP to the midpoint of the interval, and the description is incomplete since it does not address the effect of the oath script on the variance of the WTP distribution. In addition, our analysis up to this point has not controlled for individual characteristics that may drive the observed differences between treatments, nor has it allowed for the possibility that the oath script has a differential impact on different respondent groups. To overcome these weaknesses, we estimate the effects of the oath treatment using an interval regression approach (Cameron, 1989).

We estimate three censored (at zero) interval regression models, one for each of the cumulative reduction levels (i.e., 30%, 60%, and 85%), for each country using maximum likelihood estimation. The bounds used to form the contribution to the log likelihood function for a given respondent are based on the sequence of responses provided to the three WTP questions. For example, consider a respondent who selected 20 SEK for the 0–30% reduction question and 10 SEK for the 30–60% reduction question. Based on this sequence of responses, we would infer that his total WTP to reduce emissions by 60% lies between 15 (=15 + 0) and 40 SEK (=25 + 15).[19] In U.S. dollars, this corresponds to a WTP between $1.6 and $4.3. The model of interest is specified using an unobserved *latent* variable, WTP_j, for subject j,

$$WTP_j = x_j'\beta + \varepsilon_j, \quad j = 1, \ldots, N,$$

which under specific conditions is partly or fully observable. The model assumes that ε_j is normally distributed with mean zero and variance σ^2. Our likelihood function reflects the three types of data we obtain based on responses to the valuation questions; (i) point data, represented by C; (ii) left-censored data, represented by L; and (iii) interval-censored data, represented by I. For observations $j \in C$, we observe WTP_j individual j's response to an open-ended question. As discussed before, respondents belonging to this group had a WTP above the maximum amount given in the payment cards and were therefore asked an open-ended WTP question. For (at zero) left-censored observations $j \in L$, we know only that the unobserved WTP_j is less than or equal to $WTP_{Lj} = 0$. For interval-censored observations $j \in I$, we know that the unobserved WTP_j lies in the interval $[WTP_{1j}, WTP_{2j}]$. The log likelihood function is

$$\ln L = -\frac{1}{2}\sum_{j\in C}\left\{\left(\frac{WTP_{Cj} - x_j'\beta}{\sigma}\right)^2 + \log 2\pi\sigma^2\right\} + \sum_{j\in L}\log\left\{\Phi\left(\frac{WTP_{Lj} - x_j'\beta}{\sigma}\right)\right\}$$
$$+ \sum_{j\in I}\log\left\{\Phi\left(\frac{WTP_{2j} - x_j'\beta}{\sigma}\right) - \Phi\left(\frac{WTP_{1j} - x_j'\beta}{\sigma}\right)\right\}$$

where $\Phi(\cdot)$ is the standard cumulative normal distribution function.

[19] Since the interval (data) bids used in the survey were 10, 20,. . ., 2000 SEK, and 10 SEK was the lowest bid that was chosen in this example.

Table 3
Results of interval regression: marginal effects on the probability of a positive WTP.

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
No temperature increase	−0.159***	−0.124***	−0.125***	−0.007	−0.004	0.008
	(0.045)	(0.044)	(0.044)	(0.051)	(0.044)	(0.042)
Humans affect temperature increase	0.242***	0.249***	0.249***	0.023	0.051	0.052
	(0.042)	(0.042)	(0.042)	(0.056)	(0.051)	(0.049)
Female	−0.048***	−0.054***	−0.062***	−0.009	0.009	0.020
	(0.014)	(0.014)	(0.014)	(0.025)	(0.022)	(0.022)
Age	−0.001***	−0.002***	−0.002***	0.002*	0.002**	0.002**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Adults	−0.027**	−0.017	−0.011	−0.004	−0.004	−0.000
	(0.012)	(0.012)	(0.012)	(0.010)	(0.009)	(0.009)
Children	−0.004	−0.006	−0.005	0.034**	0.021*	0.017
	(0.008)	(0.008)	(0.008)	(0.017)	(0.015)	(0.015)
University education	0.085***	0.071***	0.067***	0.051**	0.038	0.039
	(0.014)	(0.014)	(0.015)	(0.030)	(0.027)	(0.026)
Income	0.046***	0.043***	0.040***	0.303***	0.241***	0.205***
	(0.006)	(0.006)	(0.006)	(0.020)	(0.017)	(0.016)
Religious	0.049**	0.032	0.022	−0.099	−0.140	−0.132
	(0.026)	(0.027)	(0.028)	(0.129)	(0.120)	(0.115)
Left party	0.000	0.005	0.013			
	(0.017)	(0.017)	(0.017)			
Green party	0.092***	0.101***	0.108***			
	(0.196)	(0.019)	(0.019)			
Other party	−0.069***	−0.063***	−0.056**			
	(0.023)	(0.023)	(0.023)			
Communist party				0.043*	0.045*	0.042*
				(0.028)	(0.024)	(0.024)
Oath	0.039***	0.038***	0.039**	0.261***	0.292***	0.288***
	(0.015)	(0.015)	(0.015)	(0.034)	(0.032)	(0.033)
No. of observations	2406	2406	2406	550	550	550

Notes: Standard error is in parentheses.
* Coefficient is statistically significant at 10%.
** Coefficient is statistically significant at 5%.
*** Coefficient is statistically significant at 1%.

Our WTP regression allows for multiplicative heteroskedasticity. The variance of the disturbance term, ε_j, is given by $Var[\varepsilon_j] = \sigma^2 \exp(\gamma' Oath_j)$, where the disturbance variance for the treatment without an oath is given by σ^2. That is, the models allow for different variance in the two treatments, where the free parameter γ measures this difference (see Harvey, 1976 and Greene, 2003).

Tables 3 and 4 report two sets of marginal effects from each interval regression model: (i) the marginal effects on the probability of a positive WTP (i.e., the probability that the variable is uncensored) and (ii) the marginal effects on the conditional WTP, respectively.[20]

The estimated marginal effect of the oath script on the probability of a positive WTP is consistently positive and significant for all reduction levels in both countries. In Sweden the oath increases the probability of a positive WTP by about 4% for all reduction levels. In the Chinese sample, the effect is stronger and the oath increases the probability of positive WTP by at least 26% for all reduction levels. As reported in Table 4, the oath reduces conditional WTP for all reduction levels in the Chinese sample, consistent with our hypothesis. For the 30%, 60%, and 85% reductions, the conditional willingness to pay falls by about $1.1, $3.6, and $6.7, respectively.[21] The estimated marginal effect of the oath script on conditional WTP in Sweden is statistically insignificant for the three reduction levels.[22]

The estimated effect of the oath script on unconditional WTP reflects the combined effects on the probability of a positive WTP and on conditional WTP. Table B3 in Appendix B reports these results. For both countries and all reduction levels, the oath script increases unconditional WTP. However, the effect is significant only for the 30% and 60% reductions in the Swedish sample and for the 30% reduction in the Chinese sample.

Our final set of results, reported in Table 5 (also included in the lower panel of Table B3 in Appendix B), explores the impact of the oath script on the variance of WTP.

[20] The regression coefficients are reported in Table B3 in Appendix B together with the heteroskedastic term.
[21] Notably, while the conditional WTP is roughly doubled comparing a 30% to an 85% emission reduction (see Table 2), the reduction in conditional WTP using the oath script is about six times greater for an 85% than a 30% emission reduction.
[22] If anything, there was a slight tendency in the opposite direction, i.e., to report a *higher* conditional WTP. One explanation could be that subjects understated their WTP for reasons similar to why they report a zero WTP in the treatment without an oath script.

Table 4
Results of interval regression: marginal effects on conditional WTP.

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
Regression function						
No temperature increase	−11.229***	−16.055***	−21.997***	−0.075	−0.074	0.193
	(2.663)	(4.889)	(6.739)	(0.520)	(0.777)	(1.084)
Humans affect temperature increase	15.819***	28.583***	38.982***	0.234	0.850	1.232
	(2.169)	(3.736)	(5.126)	(0.549)	(0.797)	(1.093)
Female	−4.066***	−8.201***	−12.782***	−0.093	0.167	0.490
	(1.209)	(2.123)	(2.924)	(0.258)	(0.391)	(0.542)
Age	−0.124***	−0.243***	−0.398***	0.016*	0.030**	0.044**
	(0.043)	(0.075)	(0.104)	(0.010)	(0.015)	(0.021)
Adults	−2.293**	−2.583	−2.365	−0.039	−0.079	−0.002
	(1.038)	(1.826)	(2.522)	(0.104)	(0.158)	(0.219)
Children	−0.373	−0.929	−1.108	0.351**	0.381	0.417
	(0.714)	(1.253)	(1.727)	(0.176)	(0.268)	(0.371)
University education	7.490***	11.020***	14.214***	0.549	0.700	1.104
	(1.295)	(2.227)	(3.129)	(0.335)	(0.503)	(0.697)
Income	3.943***	6.525***	8.266***	3.128***	4.311***	5.140***
	(0.479)	(0.844)	(1.163)	(0.189)	(0.291)	(0.401)
Religious	4.560*	5.034	4.846	−0.893	−2.090	−2.811
	(2.607)	(4.482)	(6.113)	(1.102)	(1.515)	(2.109)
Left party	0.007	0.685	2.773			
	(1.438)	(2.531)	(3.504)			
Green party	8.899***	17.516***	25.526***			
	(2.086)	(3.724)	(5.163)			
Other party	−5.510***	−8.909***	−10.914***			
	(1.760)	(3.075)	(4.259)			
Communist party				0.461	0.828*	1.085*
				(0.306)	(0.460)	(0.636)
Oath	2.006	2.087	0.048	−1.058***	−3.566***	−6.680***
	(1.244)	(2.184)	(3.023)	(0.374)	(0.000)	(1.127)
No. of observations	2406	2406	2406	550	550	550

Notes: Standard error is in parentheses.
* Coefficient is statistically significant at 10%.
** Coefficient is statistically significant at 5%.
*** Coefficient is statistically significant at 1%.

The estimated coefficient γ is consistently negative and significant in five of the six interval regressions, which means that the variance of WTP is lower in the oath treatment. This is presumably our most important finding and fully in line with the findings of Jacquemet et al. (2013). A natural interpretation is that the oath leads to more careful reflection, which reduces both high and low (particularly zero) bids. The model estimates also suggest that the estimated impact of the oath on the variance of WTP is larger in the Chinese than in the Swedish sample.

4.1. The oath script: differences between Sweden and China

That the effect of oath is stronger in China than in Sweden is intriguing, but our study is not designed to give a definitive or exhaustive explanation to this difference. If we think about what factors could influence the effect of an oath, we realize that it is tied into a whole web of factors influencing stated preferences in general. We believe that the socio-psychological

Table 5
Impact of the oath script on the variance of WTP.

Variance function	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
Ln(sigma)	3.698***	4.271***	4.626***	2.084***	2.817***	3.299***
	(0.022)	(0.022)	(0.215)	(0.054)	(0.052)	(0.051)
γ	−0.044	−0.068**	−0.108***	−1.038***	−1.375***	−1.499***
	(0.031)	(0.031)	(0.031)	(0.078)	(0.076)	(0.074)
No. of observations	2406	2406	2406	550	550	550

Notes: Standard error is in parentheses.
* Coefficient is statistically significant at 10%.
** Coefficient is statistically significant at 5%.
*** Coefficient is statistically significant at 1%.

mechanisms behind truth-telling and the possible design of mechanisms to enhance the disclosure of true individual values may vary significantly between cultures (as discussed Section 1), and this may indeed be a nascent and worthwhile field for further research. There could be factors that relate to the size of the hypothetical bias as well as factors relating specifically to the effect of an oath on this bias. International comparisons of values and characteristics are notoriously complex. In this paper, we can only offer some plausible explanations to our results that the effect is stronger in China.

Previous studies show that there exist arguments for lower hypothetical bias in both Sweden and China. For example, Hofstede (2001) finds much more "power distance" (which measures the degree of human equality in a society) in China than in Sweden, and Ehmke et al. (2008) show that people with high power distance are less likely to have hypothetical bias than people with lower power distance. However, after controlling for power distance, Ehmke et al. (2008) still find a hypothetical bias in their China sample.

Chinese are said to be collectivists rather than individualists (Hofstede, 2001; Rawwas et al., 2007). Ehmke et al. (2008) find that collectivism decreases the hypothetical bias, while individualism might increase it. On the other hand, the oath script might have had a greater impact on the Chinese sample just because of the country's collectivistic culture: Differences in collectivism/individualism can affect cheating behavior (Fu et al., 2007; Rawwas et al., 2007) and perception of when it is morally acceptable to lie (Fu et al., 2007). Moreover, Ehmke et al. (2008) argue that people in individualistic countries (such as Sweden) have a larger gap between the private and the social self than people living in collective cultures such as China. Following their arguments, Swedes are more free than Chinese to express their private wants and opinions. Thus, if the default implies seeking to say something to fit into a group or to promote group harmony, then the effect of an oath to tell the truth might be stronger in China than in Sweden.

Compared with their Chinese counterparts, Swedes are found to have higher trust in others.[23] It is possible the Swedes trusted us (e.g., trusted that the stated aim of the survey was honest) more and therefore their baseline was to provide more honest answers. Swedes are also more used to participating in surveys and being asked for their opinions. The "good" described may also be more familiar in Sweden. Few countries have such a long history of public debate of environmental, energy, and climate issues as Sweden. Thus it is plausible that the issues addressed by the survey were more familiar to the average Swede than the average Chinese. There is ample evidence that experience makes people less prone to hypothetical bias (List, 2001). If there is less hypothetical bias in the absence of an oath script, then the effect of such a script will by necessity be smaller.

Finally, another factor that might have enhanced the effect of the oath in the Chinese context was that, since computers are unusual in randomly selected homes in China (making it impossible to conduct the surveys in the same way in China and Sweden), the survey was not, as in Sweden, completed in the privacy of the home. Instead it was undertaken in a laboratory with computers. It is possible that this made the Chinese respondents more alert throughout the whole survey, including when taking the oath.

4.2. Who is affected by the oath script?

Thus far, we have shown that the oath script affects survey behavior in both countries albeit to different extents. It would also be interesting to determine whether different respondent groups are more or less affected by the oath script. Hence we look for differences that are consistent across respondent groups and between the two countries. In order to examine the effect of the oath script, we estimate the same models as before, but include interaction terms between the oath treatment dummy variable and respondent characteristics. In other words, for each group of respondents, we estimate two sets of parameters: One set represents the main effect, while the other represents the change due to the oath script. Because we still include the treatment dummy variable for the oath treatment, these interaction terms show whether a certain group of respondents is more or less affected by taking the oath, holding other respondent characteristics constant. Table 6 reports only on the significant coefficients from the interval regression, i.e., the effects on unconditional WTP, and only for the 30% reduction. The full set of results with all included variables and for the three reduction levels can be found in Table B4 in Appendix B.

For both the Swedish and Chinese samples, only a few of the interaction terms are significant. This indicates that the effect of the oath script is generally not dependent on respondent characteristics. For the Swedish sample, there are three significant interaction terms: older respondents, religious respondents, and supporters of "other" party, where in all three cases the interaction terms are positive. These three groups are also generally groups of respondents with a low WTP without the oath script. The interpretation of the positive interaction terms is that older respondents, voters for "other" parties, and religious respondents are the groups that increase their WTP with an oath script, all else equal. For example, the WTP decreases with age by about \$0.31 per year in the treatment without oath, while in the oath treatment age has a much smaller effect since the interaction term between age and oath is positive and \$0.26 per year. The effect of oath on religious respondents is that it increases the WTP by \$12 compared with religious respondents in the treatment without an oath. One might argue that an oath gets much of its emotional and psychological charge from religion, and this explains why religious respondents are to such a large extent affected by the oath. Another interesting group here is the supporters of "other" party. The political parties in this group are all small political parties in Sweden, usually formed in protest to the more established political

[23] The World Values Survey found that 68% of Swedes think that most people can be trusted, compared to 52% for Chinese (World Values Survey, 2010).

Table 6
Coefficient estimates interval regression with interaction terms.

Variable	Sweden *0–30%*	China *0–30%*
Age	−0.307***	
	(0.084)	
Age × oath	0.261***	
	(0.119)	
Income		5.927***
		(0.585)
Income × oath		−1.932***
		(0.634)
Religious	−0.285	
	(4.858)	
Religious × oath	11.853*	
	(6.790)	
Other party	−14.677***	
	(3.736)	
Other party × oath	12.746***	
	(5.150)	
Oath	−22.605**	7.025*
	(12.093)	(4.060)
Constant	6.398	−9.217**
	(8.580)	(3.882)
No. of observations	2406	550

Standard error is in parentheses.
* Coefficient is statistically significant at 10%.
** Coefficient is statistically significant at 5%.
*** Coefficient is statistically significant at 1%.

parties.[24] It is therefore possible that supporters of these "other" parties are more likely to protest against the survey. With an oath script, the WTP of this group increases by almost $13. Considering that the average WTP for a 30% reduction is about $24, the effect of oath on this specific group of respondents is considerable.

In the Chinese sample, there is one significant interaction effect: high-income respondents. In the treatment without an oath, there is a positive and significant relationship between income and WTP. This holds for the treatment with an oath script as well, but since the interaction effect is negative, the correlation between income and WTP is weaker.

5. Conclusions

The aim of our study was to implement and investigate the effects of an oath script in a stated preference survey. The main advantage of our study over existing work is that we test the effect of an oath script in the field. An additional advantage is that our relatively large sample size allows us to identify groups of respondents that are more likely to change their behavior in the oath treatment. A third advantage is that we are able to test these effects in two different cultural contexts: China and Sweden. A further advance is our decision to separately analyze the probability of bidding zero and the conditional WTP, given that a positive bid was offered. This disaggregation of average WTP revealed quite interesting, though complex, patterns.

We find that the oath treatment produces lower variances in WTP, which is in line with the results found by Jacquemet et al. (2013) in a lab environment. A possible explanation for the reduced variance is a generally lower percentage of zero bidders and extremely high WTP responses with the oath than without oath. Moreover, taking an oath also has a clear effect on the willingness to state a positive WTP: The probability of stating a positive WTP increases in both countries. However, the results for the conditional WTP are more mixed. The magnitude of the WTP among those with positive WTP decreases significantly in the Chinese sample but increases in the Swedish sample. However, the Swedish results are not statistically significant. Considering average WTP, for Sweden we find that taking the oath significantly increases the average WTP while for China the increase is never significant. Hence, those hoping that administering an oath will unilaterally work to reduce WTP in all situations and partly answer critics of hypothetical bias in stated preference surveys will not find comfort in our results. While we cannot test why we observe differences between China and Sweden, we can offer some ideas. One reason may be that Swedes (but not Chinese) are more used to participating in surveys and being asked for their opinions, and are also relatively familiar with the issues related to climate change and hence with the "good" in question, causing them to react differently to the oath. Swedish culture (with openness, egalitarianism, trust, and lack of hierarchy) may favor the individual stating his opinion, while the Chinese culture may emphasize more the importance of "in-group" cohesion and

[24] Common supporters of these parties are young males (Holmberg and Weibull, 2009; Oscarsson and Holmberg, 2008).

hierarchy, which might favor trying to find answers to fit expectations. If so, then the effect of the oath might be stronger in the Chinese case.

Disaggregating the oath effects by respondent characteristics opens another window to understand these results. The largest effects are found for the high-income respondents in China and for those who vote for smaller "alternative" political parties in Sweden. For example, the voters for these alternative parties increase their WTP by almost $13 for a 30% reduction when taking the oath compared with not taking the oath.

We believe that we can discern the contours of an exciting new research field in which much more research could be dedicated to understanding both the subtle psychological and cultural determinants of stated preferences responses as well as the degree to which these can be mediated or controlled through cultural devices that favor the elicitation of truthful individual values and responses.

Thus, while we find the results of lower variance of WTP on the oath treatment for hypothetical valuation questions encouraging, we are still aware that challenges remain regarding how an oath script can reduce hypothetical bias. In line with the work by Aadland and Caplan (2006) on the effect of a cheap talk script on hypothetical bias, we note that there is a need to further understand how the oath script influences the cognitive process of the survey respondents.

Acknowledgments

We have received valuable comments from the editor, associate editor, and two anonymous reviewers. Financial support from the Adlerbertska Research Fund, Formas, Formas through the program "Human Cooperation to Manage Natural Resources" (COMMONS), Sida (through the Environment for Development Initiative), Jan Wallander and Tom Hedelius Foundation, and the Mistra Clipore Program is gratefully acknowledged.

Appendix A.

Tables A1 and A2.

Table A1
Global CO_2 emissions reduction, temperature increases, and their effects as described in the survey.

Global CO_2 emissions reduction	85% reduction	60% reduction	30% reduction
Temperature increase	2 °F increase	3 °F increase	4 °F increase
Harvest	Harvests in countries near the equator decrease by 4–6%. Harvests in countries in the northern hemisphere *increase* by 1–3%.	Harvests in countries near the equator decrease by 10–12%. Harvests in countries in the northern hemisphere are not affected.	Harvests in countries near the equator decrease by 14–16%. Harvests in the northern hemisphere decrease by 0–2%.
Increased flooding and storms	Small tropical islands and lowland countries, such as Bangladesh, experience increased flooding and storms.	Additional low-lying areas in the Americas, Asia, and Africa experience increased flooding and storms.	Populous cities face increased flood risks from rivers and ocean storms. Existence of small island countries is threatened.
Threatened ecosystems	Sensitive ecosystems, such as coral reefs and the Arctic, are threatened.	Most coral reefs die. Additional sensitive ecosystems and species around the world are threatened.	Sensitive and less-sensitive ecosystems and species around the world are threatened.

Table A2
Contingent valuation question: 30% reduction.

Question 1: How much would your household pay for the 30% reduction?

Global emission reduction	No reduction	30% reduction
Temperature increase	More than 4 °F increase	4 °F increase
Amount your household is willing to pay per month until 2050	$0	**$_?_**

Click here to see the comparison table again.

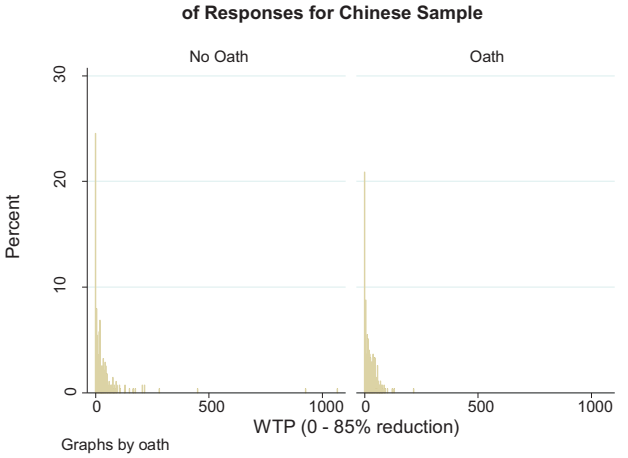
To fill in the blank above, select the amount that is the highest monthly amount your household would pay.

[5pt] $0	$1	$2	$3	$4	$5	$6	$7	$8
$9	$10	$15	$20	$25	$30	$40	$50	$60
$80	$100	$130	$170	$220	>$220			

You clicked on **[X1]** per month. This means that your household would be willing to pay **[12 times X1]** per year until the year 2050 to reduce emissions by 30 percent.

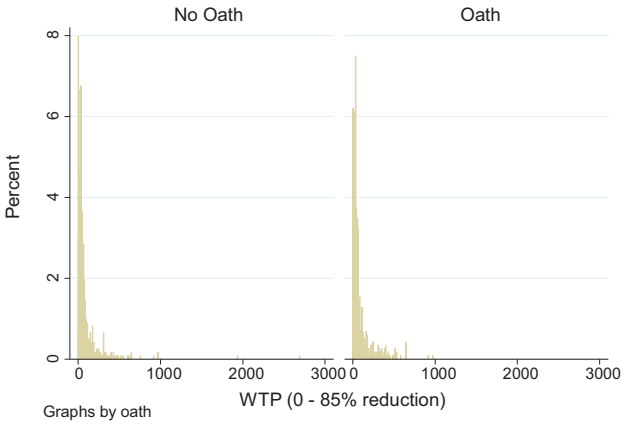
Appendix B.

Figs. B1 and B2.



Note that the share of respondents with WTP =0 is higher in the no oath treatment. Also note that a small share of respondents in the no oath treatment have a WTP>250 U.S. dollars. This is never the case in the oath treatment.

Fig. B1. Histogram of responses for Chinese sample.



Note that while the peak in the no oath treatment is at WTP =0, in the oath treatment the peak is at a value where WTP>0. Also note that a small share of respondents in the no oath treatment have a WTP>1000 U.S. dollars. This is never the case in the oath treatment.

Fig. B2. Histogram of responses for Swedish Sample.

Tables B1–B4.

Table B1
Distribution of responses for the Chinese sample.

Bid	0–30%		30–60%		60–85%	
	No oath	*Oath*	*No oath*	*Oath*	*No oath*	*Oath*
0	30.69	24.91	44.77	39.93	49.82	54.21
4	19.49	17.58	17.33	21.61	16.97	15.02
7	1.44	9.52	2.17	10.26	3.25	9.52
11	17.33	10.99	14.08	6.23	13	5.86
15	5.42	12.09	3.97	10.62	3.25	7.69
19	5.42	5.86	6.14	2.93	5.78	1.83
22	1.08	5.13	1.08	5.13	0.36	2.20
26	0.72	1.47	0.00	0.73	0	1.10
30	6.86	6.59	5.05	2.20	1.81	1.83
33	0.72	0.73	0.00	0.00	0.00	0.00
37	1.08	0.73	0.00	0.00	0.36	0.37
45	4.33	1.47	1.44	0.00	1.81	0.00

Table B1 (*Continued*)

Bid	0–30%		30–60%		60–85%	
	No oath	*Oath*	*No oath*	*Oath*	*No oath*	*Oath*
55	0.72	2.20	1.81	0.37	1.08	0.00
75	3.61	0.73	0.72	0.00	1.08	0.00
130	0.36	0.00	0.36	0.00	0.36	0.37
165	0.00	0.00	0.36	0.00	0.00	0.00
210	0.00	0.00	0.36	0.00	0.36	0.00
270	0.72	0.00	0.00	0.00	0.00	0.00
350	0.00	0.00	0.00	0.00	0.36	0.00
445	0.00	0.00	0.36	0.00	0.36	0.00

Table B2
Distribution of responses for the Swedish sample.

Bid	0–30%		30–60%		60–85%	
	No oath	*Oath*	*No oath*	*Oath*	*No oath*	*Oath*
0	9.11	6.89	15.69	13.01	27.24	26.28
10	5.77	4.85	6.5	6.38	8.29	6.63
20	3.74	3.06	5.45	4.85	6.91	5.95
30	2.11	2.04	3.9	2.13	2.76	3.15
40	1.06	1.36	1.38	1.62	1.63	1.45
50	12.03	9.78	14.07	15.22	13.33	14.03
60	0.49	0.77	0.81	0.94	0.81	0.51
70	0.73	0.34	1.06	1.19	0.41	1.02
80	1.22	0.85	0.98	0.94	0.57	0.68
90	0.41	0.09	0.57	0.34	0.24	0.17
100	22.85	22.79	15.77	16.41	13.17	14.88
125	0.89	1.28	2.03	2.38	1.79	1.79
150	3.41	3.23	5.69	6.55	4.07	4.08
200	14.15	16.67	8.86	9.10	6.26	6.21
275	2.03	1.36	2.36	3.15	1.06	2.21
350	4.07	6.55	3.74	3.32	2.76	2.04
450	5.61	5.10	3.17	3.49	1.95	1.70
575	3.01	2.38	2.11	2.30	1.71	1.96
725	0.89	1.62	1.79	1.36	0.65	1.02
950	2.52	4.00	1.06	1.53	1.54	1.45
1200	2.52	2.89	1.3	1.87	1.06	1.11
1550	0.08	0.43	0.81	0.77	0.57	0.60
2000	1.06	1.45	0.49	0.94	0.73	0.94
>2000	0.24	0.26	0.4	0.26	0.49	0.18

Table B3
Coefficient estimates interval regressions.

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
Regression function						
No temperature increase	−17.160***	−23.884***	−33.191***	−0.105	− 0.103	0.279
	(4.587)	(7.957)	(11.141)	(0.727)	(1.109)	(1.561)
Humans affect temperature increase	25.512***	46.152***	63.876***	0.329	1.222	1.851
	(4.245)	(7.396)	(10.317)	(0.783)	(1.185)	(1.701)
Female	−5.617***	−11.294***	−17.830***	−0.130	0.232	0.711
	(1.672)	(2.926)	(4.082)	(0.358)	(0.545)	(0.786)
Age	−0.171***	−0.334***	−0.554***	0.023*	0.042**	0.064**
	(0.059)	(0.104)	(0.145)	(0.014)	(0.021)	(0.030)
Adults	−3.165**	−3.553	−3.293	−0.054	−0.110	−0.003
	(1.432)	(2.511)	(3.513)	(0.150)	(0.221)	(0.319)
Children	−0.514	−1.278	−1.543	0.486**	0.531*	0.606
	(0.986)	(1.723)	(2.405)	(0.243)	(0.374)	(0.539)
University education	10.215***	15.001***	19.607***	0.748	0.964	1.454
	(1.751)	(3.063)	(4.273)	(0.449)	(0.684)	(0.985)
Income	5.442***	8.975***	11.513***	4.335***	6.013***	7.471***
	(0.661)	(1.157)	(1.616)	(0.235)	(0.359)	(0.518)
Religious	6.109*	6.793	6.656	−1.324	−3.210	−4.493
	(3.400)	(5.939)	(8.284)	(1.633)	(2.601)	(3.754)

Table B3 (*Continued*)

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
Left party	0.010	0.941	3.848			
	(1.985)	(3.473)	(4.846)			
Green party	11.717***	22.882***	33.631***			
	(2.655)	(4.644)	(6.472)			
Other party	−7.864***	−12.632***	−15.625***			
	(2.581)	(4.503)	(6.277)			
Communist party				0.629	1.136**	1.553
				(0.422)	(0.622)	(0.895)
Oath	3.564**	5.022*	4.989	0.932*	0.930	1.106
	(1.640)	(2.871)	(4.017)	(0.564)	(1.113)	(1.783)
Variance function						
Ln(sigma)	3.698***	4.271***	4.626***	2.084***	2.817***	3.299***
	(0.022)	(0.022)	(0.215)	(0.054)	(0.052)	(0.051)
Oath	−0.044	−0.068**	−0.108***	−1.038***	−1.375***	−1.499***
	(0.031)	(0.031)	(0.031)	(0.078)	(0.076)	(0.074)
No. of observations	2406	2406	2406	550	550	550

Notes: Standard error is in parentheses.
* Coefficient is statistically significant at 10%.
** Coefficient is statistically significant at 5%.
*** Coefficient is statistically significant at 1%.

Table B4
Coefficient estimates interval regression with interaction terms.

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
No temperature increase	−17.579***	−26.232**	−36.054**	0.391	−0.660	−0.907
	(6.157)	(10.843)	(15.478)	(2.821)	(5.713)	(9.164)
No temperature increase × oath	0.491	5.022	7.528	−0.579	0.482	1.054
	(9.309)	(16.122)	(22.509)	(2.917)	(5.817)	(9.296)
Humans affect temp. increase	20.379***	39.762***	59.223***	0.616	2.299	3.079
	(6.252)	(11.059)	(15.800)	(3.048)	(6.200)	(9.966)
Humans affect temp. increase × oath	10.505	12.443	9.232	−0.211	−1.248	−1.092
	(8.512)	(14.889)	(20.885)	(3.150)	(6.313	(10.110)
Female	−4.087*	−7.239*	−11.400*	−0.202	−0.296	−0.301
	(2.397)	(4.253)	(6.063)	(1.048)	(2.091)	(3.343)
Female × oath	−2.954	−7.600	−11.634	−0.061	0.551	1.023
	(3.336)	(5.850)	(8.184)	(1.113)	(2.164)	(3.437)
Age	−0.307***	−0.508***	−0.859***	0.065	0.172**	0.266**
	(0.084)	(0.149)	(0.212)	(0.041)	(0.083)	(0.132)
Age × oath	0.261***	0.321	0.543*	−0.045	−0.135	−0.209
	(0.119)	(0.208)	(0.290)	(0.044)	(0.085)	(0.136)
Adults	−2.567	−2.604	−1.337	0.064	0.041	−0.327
	(1.920)	(3.416)	(4.872)	(0.390)	(0.772)	(1.238)
Adults × oath	−1.197	−1.937	−3.819	−0.152	−0.188	0.306
	(2.876)	(5.035)	(7.026)	(0.418)	(0.804)	(1.279)
Children	−1.137	−1.764	−2.910	0.500	0.247	1.007
	(1.426)	(2.531)	(3.609)	(0.625)	(1.257)	(2.010)
Children × oath	1.188	0.879	2.419	0.003	0.329	−0.392
	(1.968)	(3.454)	(4.835)	(0.676)	(1.314)	(2.084)
University education	11.451***	17.220***	25.149***	2.626**	3.969	6.675***
	(2.532)	(4.496)	(6.409)	(1.291)	(2.598)	(4.150)
University education × oath	−2.744	−4.912	−10.845	−2.143	−3.229	−5.542
	(3.502)	(6.143)	(8.596)	(1.373)	(2.699)	(4.267)
Income	4.726***	8.275***	11.454***	5.927***	10.791***	15.804***
	(0.912)	(1.620)	(2.309)	(0.585)	(1.170)	(1.869)
Income × oath	1.350	1.303	−0.114	−1.932***	−5.2283***	−9.028***
	(1.322)	(2.317)	(3.236)	(0.634)	(1.224)	(1.939)
Religious	−0.285	−1.457	−2.145	2.103	3.667	4.925
	(4.858)	(8.606)	(12.273)	(3.422)	(6.925)	(11.116)
Religious × oath	11.853*	15.124	15.417	−4.120	−7.407	−9.939
	(6.790)	(11.887)	(16.632)	(3.891)	(7.461)	(11.791)
Left party	1.257	2.530	5.070			
	(2.820)	(5.005)	(7.137)			
Left party × oath	−2.727	−2.867	−2.355			
	(3.960)	(6.956)	(9.712)			

Table B4 (*Continued*)

Variable	Sweden			China		
	0–30%	*0–60%*	*0–85%*	*0–30%*	*0–60%*	*0–85%*
Green party	7.714**	18.638***	29.373***			
	(3.863)	(6.864)	(9.782)			
Green party × oath	7.442	7.821	7.770			
	(5.298)	(9.302)	(13.021)			
Other party	−14.677***	−24.081***	−30.973***			
	(3.736)	(6.615)	(9.422)			
Other party × oath	12.746***	21.256***	28.111**			
	(5.150)	(9.013)	(12.611)			
Communist party				1.238	1.964	4.625
				(1.452)	(2.920)	(4.675)
Communist party × oath				−0.705	−0.920	−3.240
				(1.512)	(2.986)	(4.760)
Oath	−22.605**	−22.645	−21.032	7.025*	16.233**	24.197*
	(12.093)	(21.176)	(29.635)	(4.060)	(8.047)	(12.855)
Constant	6.398	3.391	2.346	−9.217**	−19.827***	−30.031***
	(8.580)	(15.221)	(21.721)	(3.882)	(7.852)	(12.604)
Variance function						
Ln(sigma)	3.694	4.269***	4.624***	2.072***	2.777***	3.249***
	(0.022)	(0.022)	(0.021)	(0.053)	(0.051)	(0.050)
Oath	−0.043	−0.073**	−0.110***	−1.045***	−1.352***	−1.444***
	(0.031)	(0.031)	(0.031)	(0.076)	(0.073)	(0.071)
No. of observations	2406	2406	2406	550	550	550

Standard error is in parentheses.

* Coefficient is statistically significant at 10%.

** Coefficient is statistically significant at 5%.

*** Coefficient is statistically significant at 1%.

References

Aadland, D., Caplan, A., 2003. Willingness to pay for curbside recycling with detection and mitigation of hypothetical bias. American Journal of Agricultural Economics 85, 492–502.

Aadland, D., Caplan, A., 2006. Cheap talk revisited: new evidence from CVM. Journal of Economic Behavior and Organization 60, 562–578.

Alpizar, F., Carlsson, F., Johansson-Stenman, O., 2008. Does context matter more for hypothetical than for actual contributions? Evidence from a natural field experiment. Experimental Economics 11, 299–314.

Andreoni, J., 1990. Impure altruism and donations to public goods: a theory of warm glow giving. The Economic Journal 100, 464–477.

Bateman, I., Burgess, D., Hutchinson, G., Matthews, D., 2008. Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness. Journal of Environmental Economics and Management 55, 127–141.

Braver, S., 1995. In: Schroeder, D. (Ed.), Social Contracts and the Provision of Public Goods, in Social Dilemmas: Perspectives on Individuals and Groups. Praeger, New York, pp. 69–86.

Carlsson, F., Frykblom, P., Lagerkvist, C.J., 2005. Using cheap-talk as a test of validity in choice experiments. Economics Letters 89, 147–152.

Carlsson, F., Kataria, M., Krupnick, A., Lampi, E., Löfgren, Å., Qin, P., Chung, S., Sterner, T., 2012. Paying for mitigation: a multiple country study. Land Economics 88, 326–340.

Carson, R.T., 2012. Contingent valuation: a practical alternative when prices aren't available. Journal of Economic Perspectives 26, 27–42.

Carson, R., Groves, T., 2007. Incentive and information properties of preference questions. Environmental and Resource Economics 37, 181–210.

Cameron, T., 1989. OLS versus ML estimation of non-market resource values with payment card interval data. Journal of Environmental Economics and Management 17, 230–246.

Champ, P., Bishop, R., 2001. Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias. Environmental and Resource Economics 19, 383–402.

Champ, P., Bishop, R., Brown, T., McCollum, D., 1997. Using donation mechanisms to value nonuse benefits from public goods. Journal of Environmental Economics and Management 33, 151–162.

Chapman, K.J., Lupton, R.A., 2004. Academic dishonesty in a global educational market: a comparison of Hong Kong and American university business students. International Journal of Educational Management 18 (7), 425–435.

Cook, J., Whittington, D., Canh, D., Johnson, F.R., Nyamete, A., 2007. Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue Vietnam. Economic Inquiry 45, 100–114.

Cummings, R., Elliot, S., Harrison, G., Murphy, J., 1997. Are hypothetical referenda incentive compatible. Journal of Political Economy 105, 609–621.

Cummings, R., Harrison, G., Rutstrom, E., 1995. Homegrown values and hypothetical surveys: is the dichotomous choice approach incentive compatible? American Economic Review 85, 260–266.

Cummings, R., Taylor, O., 1999. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. American Economic Review 89, 649–665.

Day, B., Bateman, I., Carson, R., Scarpa, R., Louviere, J., Wang, P., Dupont, D., 2010. Task independence in stated preference studies: a test of order effect explanations. In: Paper Presented at the WCERE Conference 2010, Montreal, Canada.

Day, B., Pinto, J., 2010. Ordering anomalies in choice experiments. Journal of Environmental Economics and Management 59, 271–285.

Ehmke, M., Lusk, J., List, J., 2008. Is hypothetical bias a universal phenomenon? A multinational investigation. Land Economics 84, 489–500.

Ellingsen, T., Johannesson, M., 2004. Promises, threats and fairness. The Economic Journal 114, 397–420.

Frykblom, P., 1997. Hypothetical question modes and real willingness to pay. Journal of Environmental Economics and Management 34, 275–287.

Fu, G., Xu, F., Cameron, C.A., Heyman, G., Lee, K., 2007. Cross-cultural differences in children's choices, categorizations, and evaluations of truths and lies. Developmental Psychology 43 (2), 278–293.

Fu, G., Lee, K., Cameron, C.A., Xu, F., 2001. Chinese and Canadian adults' categorization and evaluation of lie- and truth-telling about prosocial and antisocial behaviors. Journal of Cross-Cultural Psychology 32 (6), 720–727.

Greene, W., 2003. Econometric Analysis. Prentice-Hall, New Jersey.

Harvey, A., 1976. Estimating regression models with multiplicative heteroskedasticity. Econometrica 44, 461–465.
Hofstede, G., 2001. Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations, 2nd ed. Sage Publications, Inc., Thousand Oaks, California.
Holmberg, S., Weibull, L., 2009. 46, Svensk höst [Swedish autumn]. SOM report no. 46. SOM Institute, University of Gothenburg, Gothenburg, Sweden (in Swedish).
House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W., Gupta, V., 2004. Culture, Leadership, and Organizations: The GLOBE Study of 62 Societies. Sage, Thousand Oaks, CA.
Jacquemet, N., Joule, R.V., Luchini, S., Shogren, J., 2013. Preference elicitation under oath. Journal of Environmental Economics and Management 65, 110–132.
Jacquemet, N., Jameas, A., Luchini, S., Shogren, J., 2010. Referenda under oath. Unpublished working paper. GREQAM-CNRS, Universités d'Aix-Marseille II et III.
Johannesson, M., Blomquist, G.C., Blumenschein, K., Johansson, P.-O., Liljas, B., O'Conor, R.M., 1999. Calibrating hypothetical willingness to pay responses. Journal of Risk and Uncertainty 8, 21–32.
Joule, R., Beauvois, J., 1998. La soumission librement consentie. Presses Universitaires de France, Paris.
Kahnemann, D., Knetsch, J., 1992. Valuing public goods: the purchase of moral satisfaction. Journal of Environmental Economics and Management 22, 57–70.
Kiesler, C., 1971. The Psychology of Commitment. Experiments Linking Behavior to Belief. Academic Press, New York.
Krupnick, A., Hoffmann, S., Qin, P., 2010. The willingness to pay for mortality risk reductions in China. RFF Working Paper. Resources for the Future, Washington, DC.
List, J., 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auction experiments. American Economic Review 91, 1498–1507.
List, J., Berrens, R., Bohara, A., Kerkvilet, J., 2004. Examining the role of social isolation on stated preferences. American Economic Review 94, 741–752.
List, J., Gallet, C., 2001. What experimental protocol influence disparities between actual and hypothetical stated values? Environmental and Resource Economics 20, 241–254.
Lupton, R.A., Chapman, K.J., Weiss, J.E., 2000. International perspective: a cross-national exploration of business students' attitudes, perceptions, and tendencies toward academic dishonesty. Journal of Education for Business 75 (4), 231–235.
Lusk, J.L., Norwood, F.B., 2009. An inferred valuation method. Land Economics 85, 500–514.
Mazar, N., Amir, O., Ariely, D., 2008. The dishonesty of honesty people: a theory of self-concept maintenance. Journal of Marketing Research 45, 633–644.
Murphy, J., Allen, B., Stevens, T., Weatherhead, D., 2005. A meta-analysis of hypothetical bias in stated preference valuation. Environmental and Resource Economics 30, 313–325.
Oscarsson, H., Holmberg, S., 2008. Regeringsskifte: Väljarna och valet 2006 (Change of government: The voters and the 2006 election). Norstedts Juridik, Stockholm, Sweden (in Swedish).
Ostrom, E., Walker, J., Gardner, R., 1992. Covenants with and without the sword: self-governance is possible. American Political Science Review 86, 404–417.
Rawwas, M., Swaidan, Z., Isakson, H., 2007. A comparative study of ethical beliefs of Master of Business Administration students in the United States with those in Hong Kong. Journal of Education for Business 82, 146–158.
Shu, L., Gino, F., Bazerman, M., 2011a. Dishonest deed, clear conscience: when cheating leads to moral disengagement and motivated forgetting. Personality and Social Psychology Bulletin 37, 330–349.
Shu, L., Mazar, N., Gino, F., Ariely, D., Bazerman M., 2011b. When to sign on the dotted line? Signing first makes ethics salient and decreases dishonest self-reports. Working Paper 11-117. Harvard Business School.
Vanberg, C., 2008. Why do people keep their promises? An experimental test of two explanations. Econometrica 76, 1467–1480.
Whittington, D., Smith, V.K., Okorafor, A., Okore, A., Liu, J.L., McPhail, A., 1992. Giving respondents time to think in contingent valuation studies: a developing country application. Journal of Environmental Economics and Management 22, 205–225.
Wiser, R., 2007. Using contingent valuation to explore willingness to pay for renewable energy: a comparison of collective and voluntary payment vehicles. Ecological Economics 62, 419–432.
World Values Survey, 2010. World Values Survey 2005–2008. ICPSR, Ann Arbor, MI.
Zizzo, D., 2010. Experimenter demand effects in economic experiments. Experimental Economics 13, 75–98.