

# Does context matter more for hypothetical than for actual contributions? Evidence from a natural field experiment

Francisco Alpizar · Fredrik Carlsson ·  
Olof Johansson-Stenman

Published online: 15 February 2008  
© Economic Science Association 2008

**Abstract** We investigated the importance of the social context for people's voluntary contributions to a national park in Costa Rica, using a natural field experiment. Some subjects make actual contributions while others state their hypothetical contribution. Both the degree of anonymity and information provided about the contributions of others influence subject contributions in the hypothesized direction. We found a substantial hypothetical bias with regard to the amount contributed. However, the influence of the social contexts is about the same when the subjects make actual monetary contributions as when they state their hypothetical contributions. Our results have important implications for validity testing of stated preference methods: a comparison between hypothetical and actual behavior should be done for a given social context.

**Keywords** Environmental valuation · Stated preference methods · Voluntary contributions · Anonymity · Conformity · Natural field experiment

**JEL Classification** C93 · Q50

## 1 Introduction

Context often matters even when conventional economic theory predicts that it should not (Tversky and Kahneman 1981). In this paper we aim to quantify the effect of two types of contexts on people's voluntary contributions to a national park in Costa Rica:

---

F. Alpizar  
Environment for Development Center, Tropical Agricultural and Higher Education Center (CATIE),  
7170, Turrialba, Costa Rica

F. Carlsson (✉) · O. Johansson-Stenman  
Department of Economics, Göteborg University, Box 640, 40530, Göteborg, Sweden  
e-mail: [Fredrik.Carlsson@economics.gu.se](mailto:Fredrik.Carlsson@economics.gu.se)

the degree of anonymity and information about the contributions of others. We use a natural field experiment to investigate whether the influence of social context is different for hypothetical contributions than for actual contributions.

In the literature, there is ample evidence of context effects on environmental valuation, for example that framing in terms of scenario description, payment vehicle, or the degree of anonymity influences survey responses (Blamey et al. 1999; Russel et al. 2003; List et al. 2004). Schkade and Payne (1994) used a verbal protocol methodology where they let people think aloud when answering a contingent valuation question, and concluded that people seem to base their responses on issues other than what the environmental valuation literature typically assumes. For example, the authors found that before the respondent provided an answer, more than 40% of the respondents considered how much others would be willing to contribute.

However, much of the experimental evidence suggests that context matters also in situations involving actual payments or contributions (Hoffman et al. 1994; Cookson 2000; McCabe et al. 2000). More specifically, there is ample support that so-called *conditional cooperation*, meaning that many people would indeed like to contribute to an overall good cause, such as a public good, but only if other people contribute their fair share (Fischbacher et al. 2001; Frey and Meier 2004; Gächter 2006; Shang and Croson 2006). In the light of this, the finding by Schkade and Payne (1994) may not be that surprising. One interesting question is whether respondent behavior is more sensitive to context (such as the perception of the behaviors of others) when making a hypothetical—but realistic—choice, compared to when making a choice that involves an actual payment. Some have suggested that this difference may be large (e.g. Bertrand and Mullainathan 2001), whereas others such as Haneemann (1994) believe that the difference is small (if it exist at all) and that context affects behavior generally and not just in survey-based valuation studies.<sup>1</sup> The empirical evidence for comparing the effects of context is rather scarce. Moreover, one may question the result of comparing lab experiments with hypothetical and actual money, if the purpose is to measure how closely they resemble real life behavior; see Levitt and List (2008) for a discussion.

This paper presents results of a natural field experiment—to use the terminology of Harrison and List (2004)—in Costa Rica, where we investigate the importance of (1) anonymity with respect to the solicitor and (2) information about the contributions of others.<sup>2</sup> In particular, we quantified and compared these effects for two samples: one based on hypothetical contributions and one on actual contributions.

The effect of anonymity has been investigated previously for both hypothetical and actual treatments (Legget et al. 2003; List et al. 2004; Soetevent 2005). For example, Legget et al. (2003) found that stated willingness to pay was approximately 23 percent higher when the contingent valuation survey was administered through face-to-face interviews rather than being self-administered by the respondents. List

---

<sup>1</sup>Note that we do not refer to the issue of hypothetical bias, i.e. that there is a difference between stated and real contributions for a given context. A large number of studies do find a hypothetical bias, although the occurrence and extent of it depends on a number of factors such as the type of good and the elicitation method. For an overview see List and Gallet (2001).

<sup>2</sup>For other recent field experimental studies on determinants of charitable giving, see e.g. List and Lucking-Reiley (2002), Landry et al. (2006) and Karlan and List (2007).

et al. (2004) looked at charitable contributions—both hypothetical and actual—to the Center for Environmental Policy Analysis at the University of Central Florida, using three different information treatments: (i) the responses were completely anonymous, (ii) the experimenter knew the response, and (iii) the whole group knew the response. While they found the largest share of “yes” responses when the whole group was informed of the response (followed by when only the experimenter knew the response), they also found that the differences among the information treatments were similar in the hypothetical and the actual voting treatments. A contribution of this present paper is to test whether this finding can be generalized to a field experiment setting.

The effect of information about the contributions or behaviors of others has been investigated in several field experiments (Alpizar et al. 2008; Frey and Meier 2004; Shang and Croson 2006; Heldt 2005; Martin and Randal 2005). For example, Shang and Croson (2006) investigated how information about a typical contribution to a radio station affects subject contributions. They found that their highest reference amount (\$300) implied a significantly higher contribution than giving no information at all. The direction for smaller amounts (\$75 and \$180) was the same, although not statistically significant. As far as we know, no previous study has looked directly at how information about the contributions of others affect stated contributions.<sup>3</sup> Consequently, the present paper is also the first to analyze the difference between a hypothetical and actual treatment with respect to the influence of provided information about the contributions of others.

We find that, both the degree of anonymity and information provided about the contributions of others influence contributions in the hypothesized direction. We also find a substantial hypothetical bias with regard to the amount contributed, which is consistent with earlier results. The most important finding is that the influence of the social contexts is similar when the subjects make actual monetary contributions as when they state their hypothetical contributions. Thus, we do not find that people are significantly more vulnerable to framing effects in the hypothetical treatment. Our results have important implications for valuation methods, including validity testing of stated preference methods. For example, our results suggest that a comparison between hypothetical and actual behavior should be done for a given social context. The body of this paper is organized as follows. Section 2 presents our field-experimental design, Sect. 3 the corresponding results, while Sect. 4 concludes the paper.

## 2 Design of the experiment

The experiment/survey look at contributions by visiting international tourists to the Poas National Park (PNP) in Costa Rica in 2006. We put great effort into ensuring that the situation was realistic and credible; there was nothing indicating that this was

---

<sup>3</sup>However, one explanation of so-called yea-saying—the tendency of some respondents to agree with an interviewer’s request regardless of their true views (Mitchell and Carson 1989)—is that respondents believe that the suggested bid in a contingent valuation survey contains information about the behaviors of others. If so, one may interpret observed yea-saying bias as an indication of the influence of the contributions of others. Several papers have investigated the presence of yea-saying; see for example Blamey et al. (1999), Holmes and Kramer (1995).

a university study to analyze people's behavior. This is potentially very important since, as noted by Levitt and List (2008), a perceived experimental situation may highlight people's sense of identity or self-image to a larger extent than outside the experimental situation; cf. Akerlof and Kranton (2000).

Our five solicitors were officially registered interviewers of the Costa Rican Tourism Board. We began by inviting all potential interviewers by email to a first screening meeting where we evaluated their personalities and abilities to speak fluently in both Spanish and English. Of ten possible solicitors interviewed, we chose five who fulfilled all our requirements. The five solicitors participated randomly in all parts of the experiment. Nevertheless, we control for solicitor effects in the regression analysis. The solicitors underwent extensive, paid training sessions both in the classroom and in the field. Once they were ready to start, we dedicated a whole week to testing their performance and to making small adjustments in the survey instrument. In addition, there were daily debriefing questions and regular meetings with the whole team to make sure that all solicitors were using the same exact wording of the scenarios.

The solicitors approached international tourists after they had visited the volcano crater, which is the main attraction of the park. The tourists were approached at a "station" outside the restaurant and souvenir shop, which was decorated with the logos of the PNP, the National System of Protected Areas (SINAC), and Tropical Agricultural Research and Higher Education Center (CATIE). The solicitors wore uniforms with the logos of the PNP and CATIE, and carried formal identification cards that included a photo and signatures of park authorities. The uniforms were very similar to those used by the PNP park rangers. A formal letter authorizing the collection of contributions/the survey was also clearly visible.

Only international tourists who could speak either Spanish or English participated in the experiment. The subjects were approached randomly, and only one person in the same group of visitors was approached. The selection was a key element of the training sessions, and we checked daily for subject selection biases. No corrections were required after the pilot sessions.

Subjects were first asked if they were willing to participate in an interview about their visit to the PNP. No mention of voluntary contributions took place at this stage, so we expect that participation was not affected by monetary considerations. Overall participation rates were high (above 85% each day). Once it was established that the subjects were international tourists and that they had already visited the crater, the solicitors proceeded with the interview. Before the experiment, subjects were asked a few questions regarding their visit to Costa Rica and to the national park. The solicitors were provided with standardized replies to the most common questions regarding the survey, the experiment, the institutions involved, etc. For further information the participants were advised to talk to the main supervisor of the contribution campaign.

In total 991 subjects participated in the experiments. We conducted experiments both with hypothetical and with actual contributions. For each type of experiment, we used anonymous and non-anonymous treatments as well as three different reference levels for the stated contributions of others. Table 1 summarizes the experimental design for all treatments. To avoid cross-contamination we decided to conduct the hypothetical and actual treatments during the same period, but never simultaneously.

**Table 1** Experimental design for all treatment combinations

	Hypothetical contributions		Actual contributions		Total
	Anonymous	Non-anonymous	Anonymous	Non-anonymous	
No reference contribution	62 observations	62 observations	62 observations	63 observations	250
Reference contribution: \$2	63 observations	62 observations	61 observations	63 observations	249
Reference contribution: \$5	60 observations	61 observations	62 observations	62 observations	249
Reference contribution: \$10	62 observations	62 observations	62 observations	62 observations	249
Total	247 observations	247 observations	247 observations	250 observations	991

This means that all solicitors worked on hypothetical contributions during one part of the day and actual contributions during the other part of the day. This ordering was randomly decided. All the other different treatments were conducted simultaneously, and they were randomly distributed both in terms of time of day and among solicitors.

The different treatments required slight modifications of the interviewing script, as outlined below, but we were very careful to limit the differences between the treatments. Subjects also received a card where they could read the scenario and the instructions for the voluntary contribution. The experiment began with the following sentence for all treatments:

I will now read to you some information about the funding of national parks in Costa Rica. Here is a paper with the information I will read.

After this, the participants were told about the main purpose of the request for a contribution. The wording that is unique for the hypothetical treatment is in parentheses, whereas the corresponding wording for the actual treatment is in brackets.

The system of national parks in Costa Rica is now suffering from the lack of funds to achieve a good management of the parks, both for biodiversity conservation and tourism. Available funds are simply not enough and national parks are trying to obtain new funds. We are now (researching) [testing] a system at Poas National Park where visitors can make donations to the park. The entrance fee (would remain) [remains] the same seven dollars, but people (would have) [have] the possibility to make voluntary donations to the park in addition to the fee. Contributions (would) [will] be used to improve the standard of living of park rangers, to provide for better trails and to make sure that this beautiful and unique ecosystem is well taken care of.

The effect of a social reference point was investigated by providing the subjects with information about a typical previous contribution by other visitors. If a reference point was provided, the following sentence was read:

We have interviewed tourists from many different countries and one of the most common donations has been 2/5/10 US dollars.

We obtained the monetary reference values from a pilot study conducted at the same park before our main experiment; thus, the reference information is not based on

deception. In the treatments with no mentioned reference amount, we simply omitted the above sentence.

Finally, the actual request for a contribution differed depending on whether the contribution was to be anonymous or not. In the anonymous treatments, subjects were asked to go into a private area that was part of our interviewing station and write down their contribution on a piece of paper or put their contribution (if any) in a sealed envelope and then into a small ballot box. This way their contribution was completely anonymous to the solicitor.<sup>4</sup> The following text was then read:

(If there was a possibility, how much would you donate?) [How much are you willing to donate to this fund?] Please go to the booth and (write down the amount of money you would like to donate if you had the possibility) [put the amount of money you would like to donate in the envelope]. Remember that donations will be used exclusively to maintain and improve the Poas National Park, as described before. When you are done, (please fold it up twice) [please seal the envelope] and put it in this box. Do not show it to me, because your (stated donation) [donation] should be completely anonymous. Please put the (paper) [envelope] in the box even if you do not wish to donate anything.

We provided a locked ballot box into which the contributions were put. This box was actually part of the interviewing station used for the experimental session. In the non-anonymous setting, the following text was read:

(If there was a possibility, how much would you donate?) [How much are you willing to donate to this fund?] Remember that donations will be used exclusively to maintain and improve the Poas National Park, as described before. When you are done reading, please (tell me the amount of money you would like to donate if you had the possibility) [give the envelope and your contribution to me so that I can count and register your donation before sealing the envelope. Please return the envelope even if you do not wish to donate anything].

Thus, in this treatment the subjects were well aware that the solicitor was observing each contribution. In addition to the differences described above, everything else was identical in all interviews and we expected the typical variations of a field experiment (weather, type of tourist, etc.) to affect our results randomly.

### 3 Experimental results

Table 2 presents the full basic results of the experiments in terms of the share with a positive contribution, the average conditional contribution given a positive contribution, and the resulting sample average contribution, for each of the cells in Table 1.

---

<sup>4</sup>In order for us to identify the contributions and link them to the other questions in the questionnaire, an ID number was written on the envelope. The subjects were informed about the ID number and the reason for using it. The important feature is that the solicitor was not able to observe the contribution, not even afterwards.

**Table 2** Summary results for all treatments

	Anonymous			Non-anonymous		
	Share pos. contribution	Conditional average contribution (std)	Sample average contribution (std)	Share pos. contribution	Conditional average contribution (std)	Sample average contribution (std)
Hypothetical contributions						
No Reference	0.86	11.27	9.64	0.81	12.27	9.86
Reference: \$2	0.89	6.54	5.82	0.87	5.44	4.7
Reference: \$5	0.85	6.68	5.67	0.91	7.44	6.71
Reference: \$10	0.94	11.22	10.5	0.84	9.11	7.64
Actual contributions						
No Reference	0.43	7.72	3.37	0.46	5.31	2.45
Reference: \$2	0.57	2.57	1.48	0.54	4.36	2.36
Reference: \$5	0.47	5.57	2.6	0.4	3.96	1.59
Reference: \$10	0.42	4.78	2	0.52	6.85	3.53

Naturally, since there is as many as 16 treatments in total, the number of observations becomes quite limited (between 61 and 63).

Due to our randomized design, it makes sense to compare the results of different treatments in one dimension aggregated over different treatments in the other dimension. For example, we can compare the effect of anonymity versus non-anonymity in the actual money treatment by aggregating over the different reference information treatments. In Table 3 we therefore provide more aggregate results, which facilitate straightforward interpretations and comparisons.

The most striking finding is the large amount of hypothetical bias. In the actual contribution treatment, 48 percent of the subjects chose to contribute and the average contribution was \$2.43, while in the hypothetical contribution treatment, 87 percent of the subjects stated that they would contribute an average of \$7.58.<sup>5</sup> Thus, the average contribution in the hypothetical treatment was more than three times as large as in the actual treatment, and the difference is highly significant using a simple *t*-test. The large hypothetical bias came as no surprise. First, there is much evidence that suggests the existence of a hypothetical bias (List and Gallet 2001) unless certain measures are taken, e.g. the use of so-called cheap-talk scripts (e.g. Cummings and Taylor 1999). We did not take any such measures. Second, there is also evidence that the hypothetical bias is particularly large for public goods, compared to private goods (List and Gallet 2001; Johansson-Stenman and Svedsäter 2007).

The signs of the effects of different social contexts are largely as expected. For example, if people choose to donate, they will donate substantially more if they are

<sup>5</sup>As always in stated preference surveys with an open-ended question, a number of respondents state very high numbers. These responses have a strong influence on the average contribution. We have therefore dropped observations stating contributions larger than \$100. The lowest contribution we deleted was \$450. In the actual contribution experiment, the highest contribution was \$50.

**Table 3** Summary results of contributions for different treatments

Treatment	Nobs.	Share pos. contribution	Conditional average contribution (std)	Sample average contribution (std)
Hypothetical contributions				
Total	494	0.87	8.73 (10.56)	7.58 (10.27)
Anonymous	247	0.88	8.97 (11.69)	7.92 (11.35)
Non-anonymous	247	0.85	8.49 (9.26)	7.25 (9.07)
No Reference	124	0.83	11.76 (15.81)	9.77 (15.07)
Reference: \$2	125	0.88	6.00 (6.94)	5.28 (6.80)
Reference: \$5	121	0.88	7.08 (5.82)	6.20 (5.92)
Reference: \$10	124	0.89	10.22 (10.08)	9.07 (10.03)
Actual contributions				
Total	497	0.48	5.09 (5.74)	2.43 (4.70)
Anonymous	247	0.47	5.00 (5.65)	2.37 (4.62)
Non-anonymous	250	0.48	5.17 (5.84)	2.48 (4.80)
No Reference	125	0.45	6.48 (7.45)	2.90 (3.58)
Reference: \$2	124	0.56	3.46 (3.81)	1.92 (3.32)
Reference: \$5	124	0.44	4.82 (3.24)	2.10 (3.21)
Reference: \$10	124	0.47	5.92 (7.05)	2.78 (5.20)

given a \$10 reference point instead of a \$2 reference point. This holds for both the hypothetical and the actual treatments.<sup>6</sup> The effect of anonymity is less clear. In the case of actual contributions, the conditional contribution is larger in the non-anonymous

<sup>6</sup>As noted by a referee, people's willingness to contribute may depend on the baseline quality of the park without the contribution. This quality, in turn, will then be affected by whether others contribute a lot or not. This is therefore another potential reason why the reference information may matter. However, the direction of such an effect is not clear, and one could equally well argue that people would be willing to pay more if the park has a large financial need, which would be amplified if others were willing to contribute very little. Overall we therefore doubt that this motivation can explain why people want to contribute more when they are informed about the high reference contribution by others. The good here also have similarities to the one considered by Champ et al. (1997), the removal of roads near Grand Canyon, in that it is scalable with the contribution.

case, as one might expect, whereas the opposite pattern holds in the hypothetical case. This is perhaps a bit surprising, since one would expect that individuals would feel more social pressure to contribute in the non-anonymous setting. There are two possible explanations. First, it is easier to exaggerate when making anonymous statements. You can state a high number for your own sake and pleasure without having to face the potential incredulity of the interviewer should this number be made public. In our data this shows as a somewhat larger fraction of extreme (very high) contributions in the anonymous-hypothetical treatment, and correspondingly a higher standard deviation, compared to the non-anonymous-hypothetical scenario. Moreover, this effect is not significant in the robust regression, once outliers are accounted for. Second, some subjects may have suspected that the hypothetical question was going to be followed by an actual request from the solicitor for contribution if they stated a positive amount. In the anonymous setting, on the other hand, they simply walked away and put the answer in the box. This effect could perhaps in particular explain the smaller fraction of positive contribution in the non-anonymous hypothetical setting, although this difference is small.

However, the main purpose here is neither to investigate the extent of hypothetical bias nor to quantify the importance of various kinds of social contexts, but instead to investigate the response differences between the hypothetical and actual treatments with respect to these social contexts. Table 4 summarizes these differences.

The first part in Table 4 reports the comparison between non-anonymous and anonymous treatments. For example, for hypothetical contributions, the share of people contributing is 3 percentage points lower in the non-anonymous treatment, and the sample average contribution is \$0.67, or 8 percent, lower. By comparing the second and third columns, we can compare the response difference between hypothetical and actual contributions for a given social context treatment. Although there are indeed differences between the hypothetical and actual treatments, they are rather small (particularly compared to the hypothetical bias). More importantly, although we exclude some extreme outliers, the mean values are still rather sensitive to a few observations.

In order to deal with the outlier problem, we also present the results from a regression analysis. The dependent variable, contribution, is censored since it equals zero for a substantial fraction of the subjects. In addition, there are two issues of interest here: whether to contribute anything at all and how much to contribute, given a positive contribution. Since there are good reasons to consider these as two different decisions, a basic Tobit model would be inappropriate. We therefore used a simple two-stage model. The decision whether to contribute anything or not is modeled with a standard Probit model. The decision concerning how much to contribute, given a positive contribution, is modeled with a regression model that used only subjects with a positive contribution. We tested for correlation between the two stages based on a standard sample selection formulation, but the parameter that reflects correlation was never significant at conventional levels; therefore, we only report the independent model where there is no correlation between the two stages. For completeness, we present both a standard OLS regression and a robust regression, where the latter puts a lower weight on outliers.<sup>7</sup> The base case in the regression models is given

---

<sup>7</sup>We use the `rreg` command in STATA. First a standard regression is estimated, and observations with a Cook's distance larger than one are excluded. Then the model is estimated iteratively: it performs a

**Table 4** Contribution differences between different treatments divided along hypothetical and actual contribution treatments

	Contribution differences between samples	
	Hypothetical contributions	Actual contributions
	Non-anonymous–anonymous	
Share positive contribution	–3 percentage points	1 percentage point
Conditional contribution	–\$0.48 (–5%)	\$0.17 (3%)
Sample contribution	–\$0.67 (–8%)	\$0.11 (5%)
	Reference \$2–No reference	
Share positive contribution	5 percentage points	8 percentage points
Conditional contribution	–\$5.76 (–49%)	–\$3.02 (–47%)
Sample contribution	–\$4.49 (–46%)	–\$0.98 (–34%)
	Reference \$5–No reference	
Share positive contribution	5 percentage points	–1 percentage point
Conditional contribution	–\$4.66 (–40%)	–\$1.66 (–26%)
Sample contribution	–\$3.57 (–36%)	–\$0.80 (–28%)
	Reference \$10–No reference	
Share positive contribution	6 percentage points	2 percentage points
Conditional contribution	–\$1.54 (–13%)	–\$0.56 (–9%)
Sample contribution	–\$0.7 (–7%)	–\$0.12 (–4%)

by actual contributions in the anonymous treatment with no mention of a reference contribution. In Table 5, marginal effects for the two estimated models are presented together with the total marginal effect, i.e. including the effects of the Probit stage. All marginal effects are calculated at sample means.<sup>8</sup> The total marginal effect is calculated as:

$$\frac{\partial E[C_i]}{\partial x_i} = \frac{\partial P[C_i > 0]}{\partial x_i} E[C_i | C_i > 0] + \frac{\partial E[C_i | C_i > 0]}{\partial x_i} P[C_i > 0], \quad (1)$$

regression, calculates weights based on absolute residuals, and regresses again using those weights (STATA 2005). See Rousseeuw and Leroy (1987) for a description of the robust regression model.

<sup>8</sup>For the probit model, the marginal effect for dummy variables is for a discrete change of the variable from zero to one.

**Table 5** Regression analysis of hypothetical and actual contributions to the national park. The coefficients reflect marginal effects evaluated at sample means. All models include an intercept, solicitor dummy variables and subject characteristics variables. *P*-values in parentheses

	Probit	Dependent variable: contribution			Dependent variable: log(contribution)				
		OLS-regression		Robust regression		OLS-regression		Robust regression	
		Conditional effect	Total effect	Conditional effect	Total effect	Conditional effect	Total effect	Conditional effect	Total effect
Hypothetical contribution (HC)	0.388 (0.000)	5.808 (0.001)	6.775 (0.000)	1.979 (0.002)	4.210 (0.000)	0.628 (0.000)	1.042 (0.000)	0.432 (0.003)	0.910 (0.000)
Non-anonymous treatment	0.012 (0.764)	0.021 (0.986)	0.075 (0.945)	-0.121 (0.790)	0.008 (0.987)	0.017 (0.877)	0.030 (0.789)	-0.065 (0.527)	-0.024 (0.825)
Treatment with a \$2 reference contribution	0.092 (0.068)	-3.012 (0.068)	-1.337 (0.367)	-2.091 (0.001)	-0.716 (0.281)	-0.502 (0.000)	-0.256 (0.096)	-0.726 (0.000)	-0.339 (0.021)
Treatment with a \$5 reference contribution	-0.015 (0.795)	-1.450 (0.408)	-1.082 (0.494)	-0.022 (0.983)	-0.126 (0.862)	-0.099 (0.535)	-0.090 (0.588)	-0.095 (0.528)	-0.087 (0.584)
Treatment with a \$10 reference contribution	0.021 (0.708)	-0.107 (0.951)	0.082 (0.958)	0.096 (0.883)	0.218 (0.756)	-0.138 (0.376)	-0.060 (0.714)	-0.060 (0.683)	-0.007 (0.963)
HC*Non-anonymous treatment	-0.056 (0.397)	-0.320 (0.829)	-0.633 (0.647)	0.305 (0.589)	-0.214 (0.759)	0.010 (0.943)	-0.158 (0.597)	0.112 (0.377)	-0.015 (0.923)
HC*Treatment with a \$2 reference contribution	-0.003 (0.973)	-2.772 (0.181)	-1.835 (0.338)	-0.701 (0.374)	-0.447 (0.639)	-0.030 (0.875)	-0.015 (0.945)	0.070 (0.694)	0.051 (0.806)
HC*Treatment with a \$5 reference contribution	0.077 (0.335)	-3.032 (0.160)	-1.459 (0.458)	-0.758 (0.356)	0.064 (0.945)	-0.180 (0.359)	0.003 (0.990)	-0.121 (0.512)	0.042 (0.837)
HC*Treatment with a \$10 reference contribution	0.077 (0.339)	-1.385 (0.514)	-0.356 (0.855)	1.511 (0.062)	1.585 (0.086)	0.209 (0.280)	0.263 (0.214)	0.194 (0.286)	0.253 (0.214)
Solicitor dummy variables	Included	Included	Included	Included	Included	Included	Included	Included	Included
Subject characteristics variables	Included	Included	Included	Included	Included	Included	Included	Included	Included
Number of obs	900	666	666	666	666	666	666	666	666
R2/pseudo R2	0.17	0.10	0.23	0.23	0.22	0.22	0.23	0.23	0.23

where  $E[C_i]$  is the expected contribution of individual  $i$ ,  $P[C_i > 0]$  is the probability that individual  $i$  contributes anything at all, and  $x_i$  is a covariate. Both the probit model and the regression models include a constant.

We present four different models for the contribution decision: two where the dependent variable is the contribution (one with a standard OLS regression and one with a robust regression), and two where the dependent variable is the natural logarithm of the contribution (one with a standard OLS regression and one with a robust regression). In all models we pool the hypothetical and actual contribution data.

In order to correct for an overall hypothetical bias we include a dummy variable for the hypothetical experiment. To be able to identify response differences between the hypothetical and actual contribution treatments with respect to the different social contexts (the main task of this paper), we create interaction variables between the dummy variable for hypothetical treatment and the dummy variables for each social context. The results are presented in Table 5, where the total marginal effects are computed from the probit and regression models using the expression in (1); since we assume independence between the probit and regression model, the standard error of the total marginal effect is simply a weighted sum of the standard errors of the marginal effects in the two models.  $P$ -values are reported for a two-sided  $t$ -test.

The coefficient associated with the hypothetical experiment is, as expected, large and highly significant in all models, reflecting a large hypothetical bias. The following four coefficients in Table 5 show the influence of the different social contexts for the actual contribution experiment. Interestingly, there is no statistically significant difference between the anonymous and non-anonymous treatments.<sup>9</sup> These results can be compared to List et al. (2004) who found that the proportion of subjects voting in favor of a proposal to finance a public good was significantly lower in a treatment where subjects were completely anonymous (20 percent) compared with a treatment where the solicitor observes the behavior (38 percent). The likelihood of a positive contribution is also higher in the treatment with a \$2 reference contribution compared with giving no reference information at all, whereas the corresponding effect on conditional contributions is negative. It thus appears that while providing a low reference point increases the probability of a positive contribution, the average size of the contribution is lower when compared to not providing a reference point.

Our main interest lies in the last four coefficients. They reflect the difference in social context effects between the hypothetical and actual experiments, where we have controlled for an overall difference between the two experiments. For non-anonymity we do not find any significant difference between the hypothetical and actual experiments for any of the presented models. For reference contributions, we do not find any significant difference between the hypothetical and actual experiments for the \$2 and \$5 reference contributions; this applies both for the probability of a positive contribution and for the size of the conditional contribution. For the \$10 reference contribution, we do not find any significant difference in most models. However, in the case of a robust regression where the dependent variable is the contribution, we

<sup>9</sup>We also estimated models where the difference between anonymous and non-anonymous treatment was allowed to vary among the different reference contribution treatments. The results were the same in these models, with the exception of one interaction term in the OLS regression model.

do find a significant difference (at the 10 percent level). For the \$10 reference level, the increase in contributions is \$1.40 higher in the hypothetical experiments, compared to the actual experiments. This finding is far from robust, and in the standard OLS regression the sign is reversed (although the effect is insignificant).<sup>10</sup> In the two models with the log of contribution as the dependent variable, both the OLS and the robust regression show that the influence of the \$10 reference level on the conditional contribution is about 20 percent higher in the hypothetical compared to the actual treatment, but the coefficient is insignificant in both cases.

In the regression models we corrected for individual characteristics in terms of gender, and age of the subjects, whether they are members of an environmental organization, whether they saw the volcano (the main attraction of the park), a dummy variable for US subjects, and a dummy variable for European subjects. The corresponding parameters (not reported) revealed small and statistically insignificant effects on behaviour. We also corrected for solicitor effects by including solicitor dummy variables. Moreover, previous studies have shown that the solicitor effects may differ systematically with respect to the respondent characteristics. For example, Landry et al. (2006) investigated the effect of physical attractiveness of female solicitors and found a much larger contribution effect for Caucasian males than for other groups. Therefore we also interacted the solicitor dummy variables with the gender and age of the respondent. Some of the solicitor coefficients were significant, and some of the interactions between solicitors and age of the respondent had significant effects.

#### 4 Conclusions

This paper discusses a test for whether people are more influenced by social contexts in a hypothetical experiment than in an experiment with actual monetary implications. We base the test on a natural field experiment with voluntary contributions to a national park in Costa Rica. Although we find a large hypothetical bias, we did not find any significant differences between hypothetical and actual contributions with respect to the effects of social context, except for one treatment and one regression model for which a significant effect at the 10 percent level was observed. The results suggest that social context is important in general, and is not a phenomenon that is primarily present in situations that do not involve tradeoffs with actual money. This can be compared to List et al. (2004), who observed similar effects of different information treatments for hypothetical and actual voting treatments. Our results consequently provide empirical support to the findings by List et al. in the context of field experiments.

Our results also have important implications for validity tests of stated preference methods, such as the contingent valuation method. A frequently used test, which is often considered reliable, is to compare the hypothetical responses from a stated preference method with a corresponding set-up that involves actual money (e.g. Cummings

---

<sup>10</sup>The underlying reason for this rather large difference between the robust regression and the OLS results is of course the influence of a few large contributions.

et al. 1997; Blumenschein et al. 2008). However, it follows from the results here that treatments that involve actual monetary payments are in general also vulnerable to framing effects. This questions then the general validity of such tests; this conclusion parallels List et al. (2004). If the ultimate purpose of the validity test is to find out to what extent the results from stated and the revealed preference methods differ—what Carson et al. (1996) denote a *convergence validity test*—it is important that the social context is as similar as possible in the two settings.

The ultimate purpose of valuation methods, including stated preference methods, is some kind of welfare analysis of a non-market (e.g. environmental) good, with the aim of providing useful information related to a public policy issue. The results here indicate that one should be careful when making such analysis, both for the reason of hypothetical bias and due to the framing effects. Moreover, since the results from the actual contribution experiment were equally vulnerable to the framing and to the context in which the preferences were elicited, one must be cautious when making welfare analysis also based on revealed behavior. Consequently, it is in general not straight forward based on either a stated preference or a revealed preference method to generalize the findings obtained in one context/domain to another. Finally, one has also to consider whether the framing effects reflect what Kahneman et al. (1997) denote *experienced utility*, i.e. the kind of well-being that we presumably would like the welfare analysis to reflect, or whether they just reflect *decision utility* so that choices are affected but that well-being is not; see also Kahneman and Thaler (2006).

However, as noted by a referee, the framing effects do not only cause problems; sometimes they may be seen as an asset for the researcher. Assume that the framing effects provide some element of information that respondents use to update beliefs over uncertain or ambiguous outcomes. The researcher could then use a variation of contexts in order to identify belief structures, and about how variations in the degree of uncertainty or ambiguity affect the stated value. If our finding that actual contribution experiments are equally sensitive to framing holds more generally, we could in principle generalize the insights from the framing effects in a stated preference experiment to an actual payment setting. Future research based on other samples and different situations is encouraged in order to test the extent to which the findings here are robust.

**Acknowledgements** Financial support from the Swedish Research Council and from Sida to the Environmental Economics Unit at Göteborg University and to the Environment for Development Center at CATIE is gratefully acknowledged. We are also grateful for the comments from John List and two anonymous referees. Our gratitude goes to the park authorities at Poas National Park and its conservation area (ACCVC).

## References

- Akerlof, G., & Kranton, R. (2000). Economics and identity. *Quarterly Journal of Economics*, *115*, 715–753.
- Alpizar, A., Carlsson, F., & Johannson-Stenman, O. (2008, in press). Anonymity, reciprocity, and conformity: evidence from voluntary contributions to a National Park in Costa Rica. *Journal of Public Economics*.
- Bertrand, M., & Mullainathan, S. (2001). Do people mean what they say? Implications for subjective survey data. *American Economic Review, Papers and Proceedings*, *91*, 67–72.

- Blamey, R. K., Bennett, J. W., & Morrison, M. D. (1999). Yea-saying in contingent valuation surveys. *Land Economics*, 75, 126–141.
- Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N., & Freeman, P. (2008). Eliciting willingness to pay without bias: evidence from a field experiment. *Economic Journal*, 118, 114–137.
- Carson, R., Flores, N., Martin, K. M., & Wright, J. L. (1996). Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. *Land Economics*, 72, 80–99.
- Champ, P. A., Bishop, R. C., Brown, T. C., & McCollum, D. W. (1997). Using donation mechanisms to value nonuse benefits from public goods. *Journal of Environmental Economics and Management*, 33, 151–162.
- Cookson, R. (2000). Framing effects in public goods experiments. *Experimental Economics*, 3, 55–79.
- Cummings, R. G., & Taylor, L. O. (1999). Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *American Economic Review*, 89, 649–665.
- Cummings, R., Elliot, S., Harrison, G., & Murphy, J. (1997). Are hypothetical referenda incentive compatible. *Journal of Political Economy*, 105, 609–621.
- Fischbacher, U., Gaechter, S., & Fehr, E. (2001). Are people conditionally cooperative? Evidence from a public goods experiment. *Economic Letters*, 71, 397–404.
- Frey, B., & Meier, S. (2004). Social Comparisons and pro-social behavior: testing “Conditional Cooperation” in a field experiment. *American Economic Review*, 94, 1717–1722.
- Gächter, S. (2006). *Conditional cooperation: behavioral regularities from the lab and the field and their policy implications*. CeDEX Discussion Paper No. 2006-03, University of Nottingham.
- Hanemann, W. M. (1994). Valuing the environment through contingent valuation. *Journal of Economic Perspectives*, 8, 19–43.
- Harrison, G., & List, J. (2004). Field experiments. *Journal of Economic Literature*, 42, 1009–1055.
- Heldt, T. (2005). *Conditional cooperation in the field: cross-country skiers’ behavior in Sweden*. Working Paper, Department of Economics and Society, Dalarna University.
- Hoffman, E., McCabe, K., Shachat, J., & Smith, V. (1994). Preferences, property rights, and anonymity in bargaining games. *Games and Economic Behavior*, 7, 346–380.
- Holmes, T., & Kramer, R. (1995). An independent sample test of yea-saying and starting point bias in dichotomous-choice contingent valuation. *Journal of Environmental Economics and Management*, 29, 121–132.
- Johansson-Stenman, O., & Svedsäter, H. (2007). *Self image and the valuation of public goods*. Working Paper, Department of Economics, Göteborg University.
- Kahneman, D., & Thaler, R. (2006). Anomalies: utility maximisation and experienced utility. *Journal of Economic Perspectives*, 20, 221–234.
- Kahneman, D., Wakker, P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics*, 112, 375–406.
- Karlan, D., & List, J. (2007). What do laboratory experiments tell us about the real world? *American Economic Review*, 97, 1774–1793.
- Landry, C., Lange, A., List, J., Price, M., & Rupp, N. (2006). Toward an understanding of the economics of charity: evidence from a field experiment. *Quarterly Journal of Economics*, 121, 747–782.
- Legget, C., Kleckner, N., Boyle, K., Duffield, J., & Mitchell, R. (2003). Social desirability bias in contingent valuation surveys administered through in-person interviews. *Land Economics*, 79, 561–575.
- Leviitt, S., & List, J. (2008, in press). What do laboratory experiments tell us about the real world? *Journal of Economic Perspectives*.
- List, J. A., & Gallet, C. A. (2001). What experimental protocol influence disparities between actual and hypothetical values? *Environmental and Resource Economics*, 20, 241–254.
- List, J. A., & Lucking-Reiley, D. (2002). The effects of seed money and refunds on charitable giving: experimental evidence from a university capital campaign. *Journal of Political Economy*, 110, 215–233.
- List, J. A., Berrens, A. P., Bohara, A. K., & Kerkvliet, J. (2004). Examining the role of social isolation on stated preferences. *American Economic Review*, 94, 741–752.
- Martin, R., & Randal, J. (2005). *Voluntary contributions to a public good: a natural field experiment*. Working Paper, Victoria University, New Zealand.
- McCabe, K., Smith, V., & LePore, M. (2000). Intentionality detection and “Mindreading”: why does game form matter? *Proceedings of the National Academy of Sciences*, 97, 4404–4409.
- Mitchell, R., & Carson, R. (1989). *Using surveys to value public goods: the contingent valuation method*. Washington: Resources for the Future.
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. New York: Wiley.

- Russel, C., Bjorner, T., & Clark, C. (2003). Searching for evidence of alternative preferences, public as opposed to private. *Journal of Economic Behavior and Organization*, *51*, 1–27.
- Schkade, D. A., & Payne, J. W. (1994). How people respond to contingent valuation questions—a verbal protocol analysis of willingness-to-pay for an environmental regulation. *Journal of Environmental Economics and Management*, *26*, 88–109.
- Shang, J., & Croson, R. (2006). *Field experiments in charitable contribution: the impact of social influence on the voluntary provision of public goods*. Working Paper.
- Soetevent, A. R. (2005). Anonymity in giving in a natural context: an economic field experiment in thirty churches. *Journal of Public Economics*, *8*, 2301–2323.
- STATA (2005). *STATA base reference manual*. College Station: Stata Press.
- Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, *211*, 453–458.