HOW THE SCOPE AND METHOD OF
PUBLIC FUNDING AFFECT
WILLINGNESS TO PAY
FOR PUBLIC GOODS

DONALD PHILIP GREEN
DANIEL KAHNEMAN
HOWARD KUNREUTHER

Abstract This article examines the sensitivity of survey measures of willingness to pay for public goods. Visitors to a science museum in San Francisco were asked to provide estimates of their willingness to pay for saving seabirds from oil spills and for teaching English to immigrants under various experimental conditions. Willingness to pay was substantially reduced by a seemingly innocuous reminder about how many individuals would be affected by a tax or would be asked to contribute to a given cause. This finding, which cannot be explained by standard economic interpretations of willingness to pay, is consistent with previous studies showing that subtle changes in question order and wording can affect the nature of the responses.

Despite a tradition of skepticism about verbal expressions of preference, a group of economists has recently invested increasing intellectual and financial resources in public opinion surveys. The enterprise that brings economists into the orbit of survey research is the valuation of nonmarket goods. Resource economists define the social value of a public good—such as improved water quality, cleaner air, or the preservation of an endangered species—by the aggregate amount that all individuals in society are willing to pay for that good. Over the past...

DONALD PHILIP GREEN is associate professor, Department of Political Science, Yale University; DANIEL KAHNEMAN is professor, Woodrow Wilson School of Public and International Affairs, Princeton University; HOWARD KUNREUTHER is professor, Wharton School of Business. The authors thank the San Francisco Exploratorium for allowing them to conduct their study on its premises. The able assistance of Paul Grant and Karen E. Jacowitz in conducting the study is gratefully acknowledged. The authors wish to thank Gary McClelland and William Schulze for their comments and suggestions. This research was supported by grants from the Institution for Social and Policy Studies at Yale and a Chevron grant for risk assessment research to the University of California.

Public Opinion Quarterly Volume 58:49-67 © 1994 by the American Association for Public Opinion Research
All rights reserved. 0033-362X/94/5801-0007S0.50
20 years, a technique of survey research known as the contingent valuation method (CVM) has been developed in order to measure the public’s aggregate willingness to pay (WTP). To date, several hundred studies have employed contingent valuation (see literature reviews in Cummings, Brookshire, and Schulze 1986; Mitchell and Carson 1989).

Because contingent valuation has been assigned presumptive validity in legal proceedings, a great deal is at stake in the current debate over the quality of CVM assessment. The social significance of CVM is amplified by the fact that this method often yields impressive estimates of social value that ultimately guide policy decisions and damages awards. Mead (1993), for example, cites estimates of $32 billion per year for the preservation of the whooping crane and $244 million for the prevention of one offshore oil spill in which forty thousand seabirds might die and 100–200 miles of seashore might be affected. Not surprisingly, perhaps, the method that yields such estimates has attracted critical attention (e.g., Diamond et al. 1993; Fischhoff 1991; McFadden and Leonard 1993; Mead 1993; Milgrom 1993; Philips and Zeckhauser 1989; Shrader-Frechette, Freeman, and Freeman 1988). A striking feature of this intense debate, however, is that discussions of contingent valuation have been carried out mainly in government sponsored studies, privately commissioned reports, and journals devoted to the study of environmental economics (cf. Mitchell and Carson 1989). This is unfortunate, in our view, because the dispute over contingent valuation involves questions of potential interest to those who study mass opinion and survey methodology.

**Measuring Values**

At the heart of CVM is the notion that individuals would be willing to purchase public goods if an appropriate market existed. This purchase model assumes that people can determine how much a specified public good is “worth” to them by comparing two states of affairs: the status quo versus a situation in which the public good is provided and the individual’s wealth is reduced by a certain amount. The value of the good is measured by the amount at which the individual is indifferent between the two states, the so-called reservation price. In order to assess the economic value of the improvement to society as a whole, survey researchers draw a representative sample and ask respondents about their willingness to pay. The theory is that this estimate of value can then be used as a basis for benefit-cost analysis of a proposed public action or for a legal claim for restitution when public goods are damaged by the actions of individuals or firms. An alternative view of the responses to contingent valuation surveys (Kahneman et al. 1993)
is that they are better understood as hypothetical contributions to good causes than as hypothetical purchases of goods, and as measures of attitudes rather than reservation prices.

The economic theory of value and the logic of the purchase model impose special constraints on the variables that affect CVM results. In particular, WTP should increase appropriately with the amount of the good provided: much criticism of the method has focused on failures of this requirement (Diamond and Hausman 1993; Kahneman and Knetsch 1992; NOAA Panel 1993). In addition, a measure of economic values should not be affected by minor variations of the survey method, so long as these remain within the boundaries of accepted technique. A brief look at the CVM literature, however, indicates that subtleties matter. Recent studies (Fischhoff and Furby 1988; Lazo et al. 1992) have shown that the amount of background provided about a particular commodity (e.g., groundwater) can affect the valuation process. Seemingly irrelevant changes in the response format used to measure WTP in surveys, such as the starting point in bidding, may alter the overall assessment of value (Thayer 1981). As Tolley et al. (1983) and Green (1992) have shown, question sequence also affects willingness to pay.

The present study addresses another facet of the measurement problem, related to the normative environment in which willingness to pay is elicited. This aspect of value measurement has received attention from scholars who study donations to charitable causes (Reingen 1978; Schwarzwald, Bizman, and Raz 1983) but has been overlooked by researchers performing contingent valuation. How do subjects respond when, in the process of evaluating a proposed program, they are reminded that fundraising constitutes a large-scale collective undertaking? We show that a seemingly inconsequential message reminding respondents of the number of potential contributors to the public good can have a very substantial effect on stated WTP.

**Motivation for This Study**

Our investigation was prompted by the surprisingly large estimate of the total value of preventing the deaths of birds in offshore oil spills (see also Rowe 1992; Rowe et al. 1991). This estimate, obtained through contingent valuation techniques, can be interpreted to mean that the value to the public of saving a single bird might be in the range of thousands of dollars. Although Rowe (1992) has defended this estimate, we find it extraordinarily high and believe that most members of the public would agree with us. In a preliminary study, using a scenario adapted from the Nestucca study, we found that the median
respondent in a group of 62 subjects recruited at the San Francisco Exploratorium was willing to support a public expenditure of no more than $500,000 to save fifty thousand birds from death in an oil spill. This translates into an expenditure of $10 per bird. When asked how large a tax increase they would be willing to accept for the purpose of saving these birds, the median respondent in this same sample gave a figure of $10. This translates into a total of $100 million or $2,000 per bird if such an increase were levied on 10 million taxpayers in the western United States.

How might the purchase model explain this apparent inconsistency? One might argue that the estimate of $2,000 per bird represents the value of a seabird to the population, whereas the $10 expenditure represents a judgment about the costs of saving the birds. The $1,990 difference would then represent an estimate of the approximate "surplus" citizens derive from an inexpensive yet highly valued environmental initiative. This interpretation supposes individual WTP to be unaffected by considerations of the total number of contributors to the activity, by the total amount contributed, or by any estimates of the cost of providing the public good.

We sought to test this hypothesis by using an innocuous variation in the framing of the WTP question. The experimental manipulation was a reminder of the number of households in the western United States that would be asked to contribute. We use the term "reminder" because we assume that the respondents in the control condition knew implicitly that many other households would be asked to contribute, even if they did not know the exact number. Our expectation was that the reminder would increase the salience of this fact rather than provide significant new information.

Respondents were asked to state WTP for two separate issues: a version of the seabird problem, modeled closely after the question used by Rowe et al. (1991; but lacking the introductory discussion that they used) and a program of teaching English to immigrant workers. Our principal reason for using two dissimilar programs was to produce a more stringent test of our framing hypothesis. The fact that we collected two sets of willingness to pay information also enables us to explore subsidiary hypotheses of interest. In developing this particular pair of proposals, we speculated that respondents might have some ideas about the value of saving a seabird but would be unlikely to have a clear view of the value of teaching English to one immigrant. If the reminder prompts respondents to think of the total amount that might be available to pay for the proposed activity, we expected it to affect responses to the seabird problem, but not to the English question. Our research design also enables us to consider the possible effects of question order: would a question about a social problem make saving
seabirds appear relatively insignificant, and reduce WTP? Although not central to our interest here, effects of question order clearly pose a threat to the economic interpretation of WTP as an indicator of value.

The third variable in the design was the payment vehicle. Half the respondents stated maximal WTP in the context of a tax increase and half in the context of a voluntary individual contribution. Once again, the purchase model of WTP predicts that payment vehicle should have no effect, because the respondent is assumed in this model to state a reservation price for the public good. On the other hand, if WTP is affected by considerations of the overall cost of the program, we should expect it to be higher for voluntary contributions than for taxes since the latter is likely to generate much more revenue. Considerations of cost, to the extent that they affect WTP, should make the payment vehicle especially influential when subjects are reminded of the broad scope of the proposed appeal for funds.

In summary, the design of the study included factorial variation of three variables: (1) inclusion or omission of a reminder about the number of affected households; (2) order of the questions about seabirds and about English teaching; and (3) payment vehicle, whether tax or voluntary contribution. After reviewing the results of this experiment, we discuss two supplementary studies designed to refine our understanding of the reminder effect.

An important qualification is in order before we describe our results: the present study is an experimental investigation of willingness to pay; it is not a full-blown contingent valuation. Probably the most important difference is that our respondents were not given the elaborate introductions to each problem that are customary in contingent valuation studies. We see no obvious reason why this difference should affect the impact of the variables that were manipulated in the present study, but caution is surely warranted in extending the conclusions to standard contingent valuations. More generally, we advocate a two-stage process for methodological investigations of CVM. Simplified procedures are appropriate to initial experimental studies of the impact of different variables on WTP. If these experiments show WTP to be susceptible to survey artifacts, the burden of persuasion reasonably shifts to the proponents of contingent valuation, who must then demonstrate that these anomalies disappear when more elaborate techniques are used.

Experimental Design

During the autumn of 1991, we conducted a $2 \times 2 \times 2$ experiment in which the three factors were (1) whether the subjects were reminded
that others would be asked to contribute, (2) the order in which the two issues, saving seabirds and teaching English, were presented, and (3) the nature of the payment vehicle (taxes or donations). The wording of the two willingness to pay questions, with and without reminders, appears in Appendix A.1

Subjects were drawn from visitors to a science museum, the San Francisco Exploratorium. A stand was set up with a poster offering $2 and a donation to the Exploratorium for the task of filling out a brief questionnaire concerned with environmental issues. Adult participants were arbitrarily assigned to one of the eight experimental conditions. The subjects' median age was 32 years; 55 percent were women; 54 percent had at least a bachelor's degree. The breakdown of partisan affiliation, based on a standard Michigan party identification measure, was 36 percent Democrat, 23 percent Republican, and 27 percent independent, with 14 percent declining to state.

Results. Table 1 presents the pattern of willingness to pay across two of the experimental conditions, payment vehicle and reminder about the scope of the request for funds.2 Looking first at the top panel of table 1, which describes the pattern of WTP for seabirds, we find that the mean valuation dropped precipitously when subjects were reminded that others will be asked to pay. When the payment vehicle was donation, average WTP dropped from $38.82 to $16.85; when the money was to be collected through taxes, mean WTP declined from $54.39 to $31.33. A similar pattern holds for the valuation of teaching English to immigrants. Mean willingness to donate fell from $58.62 to $17.57, and mean willingness to pay taxes dropped from $55.06 to $23.16. Since CVM studies devote special attention to estimates of the mean—recall that the aggregate social value is simply the mean times the number of potential contributors—this pattern of results is damaging to the notion that surveys can be used to value nonmarket goods accurately.3

1. The willingness to pay item we use requests a dollar amount to be supplied by the subject. This format, while frequently used, departs from the "referendum" method recommended by the Panel on Contingent Valuation sponsored by the National Oceanic and Atmospheric Administration of the Department of Commerce (NOAA Panel 1993). We take issue with this NOAA Panel's recommendation because we believe that the referendum format, which presents different subsamples with a yes/no choice concerning a single price (see Green 1992), tends to give upwardly biased estimates of value. Furthermore, the referendum methodology has the practical limitation of demanding vastly greater numbers of experimental subjects.

2. For ease of presentation, we have collapsed the results across different question sequences. We discuss the effects of sequence below.

3. Since these data, like other willingness to pay data (Cameron 1988; Green 1992; Irwin et al. 1990; McFadden and Leonard 1993), were skewed to the right, we took care to check the robustness of the reminder effect using a log transformation of willingness to pay. (To be more precise, we took the natural log of WTP plus $1, because the log of
Table 1. Willingness to Pay, by Experimental Condition

<table>
<thead>
<tr>
<th></th>
<th>Donation</th>
<th></th>
<th>Taxes</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Reminder</td>
<td>Reminder</td>
<td>No Reminder</td>
<td>Reminder</td>
</tr>
<tr>
<td>Seabirds: Mean WTP ($)</td>
<td>38.82</td>
<td>16.85</td>
<td>54.39</td>
<td>31.33</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>107.98</td>
<td>51.64</td>
<td>178.28</td>
<td>125.42</td>
</tr>
<tr>
<td>% zero WTP</td>
<td>13.6</td>
<td>13.6</td>
<td>15.0</td>
<td>11.3</td>
</tr>
<tr>
<td>N</td>
<td>221</td>
<td>221</td>
<td>214</td>
<td>212</td>
</tr>
<tr>
<td>English: Mean WTP ($)</td>
<td>58.62</td>
<td>17.57</td>
<td>55.06</td>
<td>23.16</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>194.60</td>
<td>69.83</td>
<td>144.79</td>
<td>81.60</td>
</tr>
<tr>
<td>% zero WTP</td>
<td>18.2</td>
<td>12.2</td>
<td>21.2</td>
<td>15.6</td>
</tr>
<tr>
<td>N</td>
<td>221</td>
<td>221</td>
<td>214</td>
<td>212</td>
</tr>
</tbody>
</table>

Interestingly, the reminder reduced the proportion of subjects who refused to contribute in three of the four comparisons, with the level of refusals in the other case (seabirds/donations) remaining the same. For example, 21.2 percent of the subjects asked to pay more in taxes to support English instruction stated that they were unwilling to pay anything; but when they were reminded of the scope of the fundraising effort, this figure declined to 15.6 percent. Thus, the effect of the reminder on the mean contribution is not attributable to an increase in refusals.

Table 1 offers less support for the hypothesis that WTP was shaped by the manner in which funds are to be raised. In general, mean WTP was higher for taxes than for voluntary contributions, although this effect was inconsistent in direction and magnitude. The pattern of results for question sequence, not reported in table 1, was similarly unimpressive. In the no reminder condition, for example, mean willingness to pay higher taxes for seabirds declined from $47.65 to $45.55 as the seabirds question moved from the first to the second asked. Question sequence seemed to have limited influence in other treatment conditions as well, whether the WTP item involved saving seabirds or teaching English.

*Zero is undefined. The log transformation has the virtue of making the distribution of willingness to pay resemble a normal density. Here, as before, we find that the reminder significantly reduced WTP in each of the four comparisons.*
Statistical Analysis

In order to assess the effects of the three experimental conditions more rigorously while taking into account the skewness of the data, maximum likelihood estimation was used. Following McFadden and Leonard (1993), we posited that willingness to pay was zero for those subjects who are indifferent about the proposed policy change (WTP = 0) and lognormally distributed for those who derived some utility from saving seabirds or fostering English education (WTP > 0). Specifically, we assumed that WTP was drawn from a mixed distribution such that observed WTP (x) was positive or zero with distribution

$$F_x(x; 1, \psi) = \begin{cases} 
1 - \pi, & x = 0; \\
1 - \pi + \pi \Phi \left( \frac{\log(x) - \beta}{\sigma} \right), & x > 0.
\end{cases}$$

(1)

The parameters of this distribution are $\psi = (\pi, \beta, \sigma)$. The proportion expressing positive WTP is $\pi$; $\beta$ is the mean log WTP for positive contributions; $\sigma$, the standard deviation of the underlying distribution. Hence, we required that $0 \leq \pi < 1$ and $\sigma > 0$.

The advantage of this estimation approach is that it enabled us to discern three sorts of experimental effects for each of the treatment conditions: effects on the mean of an underlying WTP distribution, on the variance of underlying WTP, and on the proportion of the subjects expressing zero willingness to pay. Since expressing zero WTP was a dichotomous dependent variable, effects here were measured as they would be in a probit analysis. The results are reported in table 2.

The largest and most statistically significant effects to emerge from this analysis were associated with the reminder. The reminder drastically reduced the mean WTP for both seabirds and English instruction ($p < .001$). At the same time, the reminder reduced the proportion of subjects expressing zero WTP, although this effect was significant only in the case of English instruction. The reminder had no apparent effect on the variance of underlying willingness to pay.

The estimates in table 2, while providing an indication of the statistical significance of the reminder estimates, do not give a clear indication of the magnitude of these effects. In order to show what these parameter estimates imply for observed willingness to pay, we use the following formula (Rothschild and Logothetis 1986) for expected mean willingness to pay:

---

4. Previous studies (Cameron 1988; Green 1992; McFadden and Leonard 1993) have found the lognormal distribution to be a close approximation to the empirical distribution of WTP for a variety of different public goods.
Table 2. Effects of Reminder, Payment Vehicle, and Question Sequence on the Distribution of Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th>Birds</th>
<th></th>
<th>English</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Standard Error</td>
<td>Estimate</td>
<td>Standard Error</td>
</tr>
<tr>
<td>Effects on the mean WTP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question order</td>
<td>.19</td>
<td>.12</td>
<td>-.09</td>
<td>.12</td>
</tr>
<tr>
<td>Reminder</td>
<td>-.98*</td>
<td>.12</td>
<td>-1.17*</td>
<td>.12</td>
</tr>
<tr>
<td>Vehicle</td>
<td>.15</td>
<td>.12</td>
<td>.20</td>
<td>.12</td>
</tr>
<tr>
<td>Constant</td>
<td>2.53</td>
<td>.11</td>
<td>2.93</td>
<td>.11</td>
</tr>
<tr>
<td>Effects on the variance of WTP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order</td>
<td>.35</td>
<td>.27</td>
<td>.80*</td>
<td>.28</td>
</tr>
<tr>
<td>Reminder</td>
<td>-.09</td>
<td>.26</td>
<td>-.11</td>
<td>.26</td>
</tr>
<tr>
<td>Vehicle</td>
<td>.37</td>
<td>.27</td>
<td>-.09</td>
<td>.26</td>
</tr>
<tr>
<td>Constant</td>
<td>2.26</td>
<td>.24</td>
<td>2.25</td>
<td>.23</td>
</tr>
<tr>
<td>Effects on the probit of giving zero WTP:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order</td>
<td>-.52*</td>
<td>.12</td>
<td>-.06</td>
<td>.10</td>
</tr>
<tr>
<td>Reminder</td>
<td>-.08</td>
<td>.11</td>
<td>-.23*</td>
<td>.10</td>
</tr>
<tr>
<td>Vehicle</td>
<td>-.30</td>
<td>.11</td>
<td>.13</td>
<td>.10</td>
</tr>
<tr>
<td>Constant</td>
<td>-.87</td>
<td>.10</td>
<td>- .89</td>
<td>.10</td>
</tr>
</tbody>
</table>

Note.—The contrasts were scored as follows: English first to seabirds first; no reminder to reminder; donations to taxes. Thus, the constant refers to the condition in which subjects received no reminder, were asked to make a contribution, and received the questions in the order English first and then seabirds.

* Estimates significant at the .05 level, using a two-tailed test. Estimates obtained by maximum likelihood.
Expected mean WTP = \( \hat{\pi} e^{\hat{\mu} + \hat{\sigma}^2/2} \).  

Here \( \hat{\pi} \) is the estimated proportion of subjects expressing positive WTP, \( \hat{\mu} \) is the estimated mean of the underlying WTP distribution, and \( \hat{\sigma}^2 \) is the variance of the underlying distribution. Consider the following example illustrating the impact of a reminder on the seabird/donation group who were first asked the questions related to English instruction. For the control group given no reminder, \( \hat{\mu} \) for the seabirds proposal is 2.53; \( \hat{\sigma}^2 \) is 2.26; and \( \hat{\pi} \) is 0.808.\(^5\) Therefore, the expected mean WTP for the control group was

\[ \hat{X}_c = 0.808 e^{2.53 + 2.26/2} = $31.39. \]  

In order to obtain the corresponding mean for the reminder condition, we added the reminder coefficients to the control group estimates. Now \( \hat{\mu} \) for the seabirds proposal was 2.53 + -0.98; \( \hat{\sigma}^2 \) was 2.26 + -0.09; and \( \hat{\pi} \) was 0.829. Thus,

\[ \hat{X}_r = 0.829 e^{1.55 + 2.17/2} = $11.56. \]  

In other words, the insertion of a reminder about the scope of the finding effort lowered mean willingness to pay for seabirds from $31.39 to $11.56, a reduction of 63 percent. The change in expected means was slightly larger in the case of English instruction. Here the estimate for the control group was $46.91 while the corresponding figure for the reminder condition was $14.72, a reduction of 69 percent.\(^6\) There is no support, in other words, for our earlier speculation that the reminder’s effects are larger when subjects have a firmer sense of the per-unit value of the public good in question.

The other experimental treatments had much more limited effects. None of the estimates of payment vehicle effects turned up significant (\( p > 0.05 \), two-tailed test). If anything, the results suggest that taxes inspired greater WTP, which runs counter to the idea that subjects gauge their WTP according to an image of the appropriate collective expenditure for a particular good. There was no general effect of question sequence. However, the propensity to offer a positive WTP for seabirds was significantly diminished when the English item was asked first. This result is readily interpretable: although the seabird issue can be quite compelling when viewed on its own, it becomes less significant in the context of a social problem that affects the welfare of humans.

\(^5\) Each of these figures is taken from table 2, except that the calculation of \( \hat{\pi} \) requires a conversion from logits to probabilities. A Z-score of -0.87 implies that 81 percent of the subjects derive some positive value from the seabirds proposal.

\(^6\) Using observed sample means instead of the expected means derived from the analysis reported in table 2, the reminder effects parallel those just reported. For seabirds, WTP declines from $34.29 to $11.59 (a drop of 66 percent); for English, WTP drops from $60.35 to $11.34 (a reduction of 81 percent).
Kahneman and Ritov (in press) find related "perference reversals" in which the importance of other species appeared to diminish when they were placed in the context of human problems. For example, a majority of respondents stated higher WTP to "provide refuges for the American elk" than to "increase the number of safety inspectors in workplaces," but this preference was strongly reversed in a direct choice between these options.

Finally, we note that expanding our analysis to include interactions between our experimental treatments turns up no significant effects. For example, there is no evidence of an interaction between the reminder and the type of payment vehicle in question ($p > .05$, using a likelihood-ratio test). Nor were there any indications of higher-order interactions using the same significance criterion. Just one finding stands out: altering the question text so as to alert respondents to the fact that they are being asked to participate in a collective endeavor undercuts WTP. The magnitude of the reminder effect is not only large in absolute terms: it is large relative to the narrow range of values that contingent valuation surveys ordinarily obtain for goods as different from each other as saving seabirds and teaching immigrants. (See D'Arge [1989], Kahneman and Knetsch [1992], Kahneman and Ritov [in press], and Kahneman et al. [1993] for a discussion of the limited variance of CVM estimates across a wide array of proposed interventions.)

Interpreting the Reminder Effect

When we first considered the effect of reminding respondents that others would be asked to share the costs of the public good in question, we suspected that people might have a sense of how much they would like to see spent by the collectivity on a particular public good. We reasoned that upon hearing that 10 million households were to be taxed in order to save fifty thousand seabirds, respondents would reflect on what their contribution would mean for the total funding for the rescue effort if others were forced to follow suit. A $10 tax would imply a collective expenditure of $2,000 per bird, presumably more than most people would think appropriate. Thus, we expected respondents to moderate their willingness to pay in order to bring the collective expenditure into line with what they thought society as a whole ought to commit to this problem.

At first glance, the data presented here seem to confirm this interpretation. People do reduce their contributions when they are reminded that fundraising is a collective enterprise. At the same time, however, the insensitivity of the value of WTP to payment vehicle suggests that
an individual’s willingness to pay is not affected by his or her preference for aggregate spending. With the number of target households held constant, taxes should generate more revenue than donations. Since the type of payment vehicle did not have a main effect on WTP nor did it interact with the reminder effect, we infer that aggregate spending was not a salient determinant of WTP.

It is possible, of course, that subjects were simply inattentive to the way in which funds were to be raised. For this reason, we set out to test the aggregate spending hypothesis more directly. In another experiment conducted at the San Francisco Exploratorium, we assigned 195 subjects to one of two versions of the reminder condition prior to asking their WTP for seabirds. The question format was identical to that presented in Appendix A, except that 97 subjects were told that 1 million households would be asked to donate, while the remaining 98 were told that donations would be solicited from 10 million households. If anticipated aggregate expenditure is the driving force behind the reminder effect, then the latter condition should yield a smaller mean WTP than the former. No such result materialized. Instead, the means were $27.80 and $27.46, respectively, which is an insignificant difference ($p > .05$, one-tailed) regardless of how the data are transformed or analyzed.\(^7\)

Another interpretation of the reminder effect observed in the first study is that respondents are behaving less like judicious tax collectors and more like free-riders. Mention of the millions of others who will be asked to pay may have diminished respondents’ sense of personal obligation and hence reduced WTP. To be sure, the reminder dampened the exuberance of many subjects. Whereas just 6 percent in the reminder condition offered to contribute more than $100 to rescue seabirds, 14 percent did so in the no reminder condition. The differences were even more pronounced for contributions to English instruction, 6 percent and 18 percent. As noted above, however, the percentage of free rides—expressions of zero willingness to pay—did not conform to the free-rider hypothesis. Instead, the reminder tended to decrease the proportion giving zero bids, although the difference was statistically significant only for the English case. This pattern took us by surprise, and admittedly achieves statistical significance in only one of two issues, although it works in the same direction for both. Nevertheless, it is interesting to speculate about the thought process that might have led subjects to moderate their bids upon hearing that others would participate but not withdraw support altogether.

---

7. Again, this may be due to subjects’ failure to discriminate between revenue-generating proposals that differ by an order of magnitude: 1 million and 10 million may simply be perceived as large numbers.
Our favored interpretation is that the reminder condition causes subjects to adjust their willingness to pay in accordance with a social norm. Those who would otherwise have contributed nothing made positive offers, while those who would otherwise have contributed a large sum reined in their contribution. In order to test this hypothesis more directly, a third study was conducted, again using subjects from the San Francisco Exploratorium. The 104 subjects were divided into two groups, both of which considered the proposal to fund an English program for recent immigrants through voluntary donations. In the control condition, no mention was made of other contributors, while in the treatment group, subjects were reminded that 20 million other households would be asked to make a voluntary contribution. What differentiates this experiment from the previous two is that in place of the willingness to pay question, we substituted a pair of questions. The first asked respondents to evaluate the appropriateness of a person's refusal to contribute; the second, to state what size contribution would be appropriate (see App. B).

Our hypothesis is that the reminder about other contributors changes respondents' assessment of what is socially appropriate behavior in two ways. First, because the scope of the appeal indicates the viability of the collective effort on behalf of this public good, it will be less socially appropriate to turn down a request to contribute. Second, since others are sharing the financial burden, subjects may feel that a person may appropriately reduce the size of his or her contribution.

The data shown in table 3 conform to both hypotheses rather well. As predicted, the reminder increases the extent to which people feel

| Table 3. The Effect of a Reminder on Perceived Appropriateness of Refusing to Contribute (Percentages, with N in Parentheses) |
|---|---|---|
| | No Reminder | Reminder |
| Completely inappropriate to refuse to donate | 6 (3) | 11 (6) |
| Somewhat inappropriate to refuse to donate | 24 (12) | 30 (16) |
| Acceptable | 46 | 44 |
| (23) | (24) |
| Completely appropriate to refuse to donate | 24 (12) | 15 (8) |
| Total | 100 (50) | 100 (54) |
Table 4. The Effects of a Reminder on Socially Appropriate Willingness to Pay. Controlling for the Perceived Appropriateness of a Refusal to Contribute

<table>
<thead>
<tr>
<th></th>
<th>Completely Inappropriate to Refuse</th>
<th>Somewhat Inappropriate</th>
<th>Acceptable</th>
<th>All Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>No reminder:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean WTP</td>
<td>200.00</td>
<td>134.83</td>
<td>34.22</td>
<td>79.08</td>
</tr>
<tr>
<td>Median WTP</td>
<td>200.00</td>
<td>37.50</td>
<td>25.00</td>
<td>25.00</td>
</tr>
<tr>
<td>N of cases</td>
<td>3</td>
<td>12</td>
<td>23</td>
<td>38</td>
</tr>
<tr>
<td>Reminder:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean WTP</td>
<td>9.17</td>
<td>481.87*</td>
<td>17.54</td>
<td>174.69</td>
</tr>
<tr>
<td>Median WTP</td>
<td>6.00</td>
<td>10.00</td>
<td>5.00</td>
<td>7.25</td>
</tr>
<tr>
<td>N of cases</td>
<td>6</td>
<td>15</td>
<td>23</td>
<td>44</td>
</tr>
</tbody>
</table>

* Two subjects named amounts of $3,000 and $4,000. Note that the difference in medians between experimental conditions is nevertheless significant at $p < .001$.

that it is inappropriate to refuse to contribute (from 30 percent to 41 percent). Although the effect falls just below conventional levels of statistical significance ($p = .073$ by a one-tailed Mann-Whitney test), its magnitude matches the results presented in table 2, again hinting that the reminder reduces expressions of zero WTP.

Despite the fact that the reminder makes people less comfortable about refusing to contribute, the perceived socially appropriate willingness to pay is markedly reduced in the reminder condition ($N = 102$, $p < .05$, one-tailed Mann-Whitney test). Indeed, when we restrict the analysis to just those respondents who did not previously express the view that a contribution of zero is "completely appropriate," we find that the reminder has even stronger effects on the socially appropriate willingness to pay ($N = 82$, $p < .01$, one-tailed Mann-Whitney). As shown in table 4, the median amount that subjects perceived to be socially appropriate declines from $25.00 to $7.25 when subjects are reminded about the scope of the fundraising effort.

Implications for CVM

Four main findings emerged from our study of the robustness of WTP estimates. These estimates were highly robust to variations of the payment vehicle (taxes versus donations). The willingness to contribute to saving seabirds was diminished somewhat by an implicit comparison
with the issue of teaching English to immigrants. Estimates of WTP proved highly sensitive to a seemingly innocuous reminder about the number of people who would be asked or required to make a payment. Finally, there is no indication that the effect of the reminder varied with the substance of the program or the way in which revenue was to be raised.

These findings can be evaluated in the light of alternative interpretations of willingness to pay. As we noted earlier, the contingent valuation method rests on a purchase model of WTP, where respondents indicate that payment level at which they are indifferent between the status quo and the provision of a public good at a cost. An alternative contribution model has been proposed, based on the idea that respondents assimilate the contingent valuation interview to a more familiar situation, in which they are requested for a charitable contribution for a good cause (Kahneman and Knetsch 1992; Kahneman and Ritov, in press; Kahneman et al. 1993). In the contribution model, WTP is an expression of attitude toward a problem, and questions about support for taxes or about personal donations are interpreted as requests for voluntary contributions.

The contribution model is supported by other studies in which WTP and other measures were obtained for a large set of issues. The repeated finding has been that the ranking of issues by WTP is highly correlated with the ranking of the issues by a measure of the personal satisfaction that is derived from contributing, and also by a simple rating of the importance ascribed to the different problems (Kahneman and Knetsch 1992; Kahneman et al. 1993). Another observation that is compatible with the contribution model but not with a purchase model is the finding that WTP is quite insensitive to the scope of proposed interventions: survey expressions of willingness to pay remain about the same when the objective is stated as cleaning fresh water lakes in Haliburton or in all of Ontario, saving one species or many species, and so on.

For quite different reasons, the purchase model and the contribution model predict the finding that payment vehicle had no reliable effect on WTP. In a purchase model, the value of a good does not depend on the modality of payment. In a contribution model, accepting an extra tax and making donations are evaluated alike. The two models diverge, however, in their predictions about the reminder effect. The economic rationality presupposed by the purchase model is inconsistent with the observation that reminding respondents that millions of potential contributors live in the western states reduced average contributions by 50 percent or more. The reminder effect is much more compatible with the psychology of donations, in which social context and social norms are salient.
The observation that a seemingly uninformative message can halve the typical contribution will come as no surprise to students of attitudes or to specialists in survey research. Schuman and Presser (1981), among others, have shown that subtle changes in question order and wording can affect the distribution of survey responses, particularly when these framing effects involve widely understood norms, such as reciprocity. Our research offers additional evidence of this phenomenon, this time involving a norm of collective responsibility. When alerted to the fact that funding for the nonmarket good was to be a collective effort, subjects apparently recognized that people are under greater obligation to contribute something, but less obligation to contribute something substantial. This finding raises not only a methodological but a conceptual question concerning contingent valuation: Which values are to be elicited from survey respondents? Those values respondents offer when they envision themselves as the sole benefactors of a public program? Or those they offer when they envision themselves participating in a collective undertaking?

Appendix A

Wordings of Willingness to Pay Questions

Question 1

Over the past 25 years, millions of immigrants have come to the United States from Asia, Latin America, and Eastern Europe. Because many of these new citizens are unable to speak English fluently, there has been a growing concern about the survival of English as our common tongue. As a result, several states, including California and Arizona, have passed laws designed to preserve English as the official language of the land. It turns out, however, that many recent immigrants who wish to learn English are unable to do so because public adult education classes in English are currently over-enrolled and private tutoring is too expensive. It has been suggested that public adult education classes in English should be expanded to accommodate an additional 100,000 people in the western United States. Such classes would enable new citizens to find work for themselves and to participate in mainstream American life.

We are interested in the value that your household would place on expanding public adult education programs so that an additional 100,000 people would learn English each year. [Reminder Condition: Note that there are approximately 20 million taxpaying households in the western states that would be asked to contribute to expand these programs. So, for example, an average contribution of $1 per household would produce a fund of $20 million each year to teach English to 100,000 people; an average contribution of $100 would produce a fund of $2 billion.] If you could be sure that 100,000 people would be taught English each year in the western United States, what is the most
Willingness to Pay for Public Goods

your household would pay, as a voluntary donation on your state or federal income tax form [Tax Condition: In extra federal or state taxes], to support an expanded program?

Question 2

There is a population of several million seabirds living off the Pacific coast, from San Diego to Seattle. The birds spend most of their time many miles away from shore and few people see them. It is estimated that small oil spills kill more than 50,000 seabirds per year, far from shore. Scientists have discussed methods to prevent seabird deaths from oil, but the solutions are expensive and extra funds will be required to implement them. It is usually not possible to identify the tankers that cause small spills and to force the companies to pay. Until this situation changes, public money would have to be spent each year to save the birds. We are interested in the maximal amount that your household would be willing to pay each year as a voluntary donation on your state or federal income tax form, to support an operation to save the seabirds. [Reminder Condition: Note that there are approximately 10 million tax-paying households in the Pacific coastline states that would be asked to contribute to save the birds. So, for example, an average contribution of $1 per household would produce a fund of $10 million each year to save 50,000 birds; an average contribution of $100 would produce a fund of $1 billion.] We are interested in the value your household would place on saving about 50,000 seabirds each year from the effects of offshore oil spills.

If you could be sure that 50,000 seabirds would be saved each year, what is the most your household would pay as a voluntary donation on your state or federal income tax form [Tax Condition: In extra federal or state taxes], to support an operation to save the seabirds? The operation will stop when ways are found to prevent oil spills or to identify the tankers that cause them and make their owners pay for the operation.

Appendix B

Wordings of Items Assessing Reactions to a Refusal to Contribute and Appropriate Willingness to Pay

Over the past 25 years, millions of immigrants have come to the United States from Asia, Latin America, and Eastern Europe. Because many of these new citizens are unable to speak English fluently, there has been a growing concern about the survival of English as our common tongue. As a result, several states, including California and Arizona, have passed laws designed to preserve English as the official language of the land. It turns out, however, that many recent immigrants who wish to learn English are unable to do so because public adult education classes in English are currently over-enrolled and private tutoring is too expensive.
It has been suggested that public adult education classes in English should be expanded to accommodate an additional 100,000 people in the western United States. Such classes would enable new citizens to find work for themselves and to participate in mainstream American life. [Reminder Condition: In order to fund this program, approximately 20 million households in the western states would be asked to contribute. So, for example, an average contribution of $1 per household would produce a fund of $20 million each year to teach English to 100,000 people; an average contribution of $100 would produce a fund of $2 billion.]

Imagine that a person (whose household income is $40,000) were asked to make a voluntary contribution for the purpose of expanding adult education programs to accommodate an additional 100,000 immigrants. Suppose this person refuses to contribute anything to this cause. In your opinion, is this person's unwillingness to contribute (a) completely inappropriate behavior; (b) somewhat inappropriate behavior; (c) acceptable behavior; or (d) completely appropriate behavior? Now suppose that this person had decided to make a donation. How large a donation would be appropriate for this cause?

References


Willingness to Pay for Public Goods


