University of Massachusetts Amherst

Department of Resource Economics

Working Paper No. 2003-11

A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation

Mihail Samnaliev[1], Thomas Stevens[2], and Thomas More[3]

Abstract:

A field test of cheap talk and two types of certainty calibration in contingent valuation of public lands indicated that cheap talk does not reduce WTP estimates. Use of a ten point certainty calibration scale reduces WTP estimates by about half. However, adjusting for uncertainty using a 'Not Sure' option does not reduce WTP estimates but increases the variance in responses. There may be a conceptual difference between these two ways of accounting for respondents' uncertainty, which may suggest why they provide different WTP value estimates and variances.

Keywords: CVM, Hypothetical bias, Cheap talk, Certainty Calibration

JEL Classification: Q20, Q26

[1] Mihail Samnaliev, Department of Resource Economics
University of Massachusetts, 202 Stockbridge Hall, 80 Campus Center Way
Amherst, MA 01003
E. samnaliev@resecon.umass.edu P. 413-545-6651 F: 413-545-5853

[2] Thomas Stevens, Department of Resource Economics
University of Massachusetts, 202 Stockbridge Hall, 80 Campus Center Way
Amherst, MA 01003
E. tstevens@resecon.umass.edu P. 413-545-5714 F. 413-545-5853

[3] Thomas More, Forest Service
Northeastern Research Station, P.O. Box 968
Burlington, VT 05402
E. tmore@fs.fed

A Comparison of Cheap Talk and Alternative Certainty Calibration Techniques in Contingent Valuation[1]

Mihail Samnaliev*

Thomas Stevens**

Thomas More***

Abstract

A field test of cheap talk and two types of certainty calibration in contingent valuation of public lands indicated that cheap talk does not reduce WTP estimates. Use of a ten point certainty calibration scale reduces WTP estimates by about half. However, adjusting for uncertainty using a 'Not Sure' option does not reduce WTP estimates but increases the variance in responses. There may be a conceptual difference between these two ways of accounting for respondents' uncertainty, which may suggest why they provide different WTP value estimates and variances.

***Keywords*: CVM, Hypothetical bias, Cheap talk, Certainty Calibration**

---

[1] This research was supported in part by funds provided by the USDA, Forest Service, Northeastern Research Station

* Corresponding author
Department of Resource Economics
University of Massachusetts
80 Campus Way,
Amherst, MA 01003 fax: +1-413-545-5853 e-mail: samnaliev@resecon.umass.edu

** Department of Resource Economics
University of Massachusetts
80 Campus Way,
Amherst, MA 01003 e-mail: tstevens@resecon.umass.edu

*** USDA, Forest Service, Northeastern
Research Station, P.O. Box 968.
Burlington, VT 05402

Introduction

Hypothetical bias, whereby respondents to the contingent valuation method, CVM, state that they would pay more than they actually do, is often perceived as a major problem for the valuation of environmental and natural resources (see Harrison and Rustrom, 2002; List and Gallet, 2001). Two calibration instruments, cheap talk and uncertainty adjustment, are now being widely used to correct for hypothetical bias. However, recent research suggests that cheap talk often fails to fully eliminate hypothetical bias (List 2001; Murphy et al. 2003; Aadland and Caplan 2003), and agreement about the appropriate method for uncertainty adjustment is far from universal. For example, Wang (1997), Carson et al. (1994), and Alberini et al. (2003) provide alternative views about calibration for uncertainty.

This study compares the effectiveness of cheap talk and two types of certainty calibration (certainty scale and a Not Sure option) in reducing WTP estimates in a mail survey. We find that cheap talk fails to reduce WTP estimates and we suggest that the two ways to adjust for uncertainty are conceptually different and may produce very different results.

Previous studies

Previous study of hypothetical bias in contingent valuation has established that bias is a problem in many contingent valuation applications. For example Harrison and Rustrom (2002) summarized 39 studies of hypothetical bias and found that bias was significant in 31 (79%) of them. The magnitude of bias varied widely across studies but in most it was within the range of 25% and 160%. List and Gallet (2001) analyzed 29 studies and concluded that hypothetical valuations were on average three times larger than actual payments, and the well known NOAA panel recommended that hypothetical values should be divided by two.

Several methods to correct for hypothetical bias have been suggested, but “Cheap Talk” and uncertainty scale adjustment have received the most attention. Cheap talk attempts to eliminate hypothetical bias through an explicit discussion of the problem. This includes a script attached to the CV survey which describes the hypothetical bias phenomena, and asks respondents to bear it in mind and answer as if they were in a real situation. Evidence of the ability of cheap talk to actually correct for hypothetical bias has been mixed. The cheap talk script proposed by Cummings and Taylor (1999) eliminated hypothetical bias. However, List (2001), who tested cheap talk using a different commodity and research design found that the cheap talk reduced hypothetical bias to a lesser extent than reported by Cummings and Taylor and only among consumers without experience in the market for the commodity being valued[2]. Brown and Ajzen (2002) reported that respondents who faced higher prices were more influenced by a cheap talk script as compared to respondents who faced lower prices. Other researchers including Loomis et al. (1996), Poe et al.(1997), Murphy et al. (2003), and Aadland and Caplan (2003) have also tested the cheap talk approach in experimental settings, but with mixed results.

Certainty scale calibration is another technique used to correct for hypothetical bias in dichotomous choice (DC) CV questions where respondents may generally be less certain of their answer to the offer price as compared to an open ended or polychotomous choice format (Ready et al. 1999). In the certainty scale calibration the WTP question is followed by a question which asks how sure respondents are about their response to the valuation question. This is usually accompanied by a 10 point scale for levels of certainty. A common application of the certainty scale is to treat positive answers as 'Yes' only when certainty levels are at least 8 on a 10 point scale with 10 indicating 'very certain' (for example see Champ et al. 1997). The effectiveness of this method has been established by comparing hypothetical payments to actual donations (Champ et al. 1997; Polasky et al. 1996). These, as well as other recent studies suggest that

[2] List also pointed to previous research that found that willingness to pay estimates for environmental goods are only reliable if respondents have some degree of experience with the commodity being valued.

uncertainty scale calibration often reduces hypothetical bias. However, Ekstrand and Loomis (1997) reported that the ability of this approach to reduce hypothetical bias depends on how the scale is used. Bias reduction was reported when certainty levels of at least 8 were used to calibrate only 'Yes' answers, but reduction of bias was questionable when 'No' answers were also calibrated. In addition, the authors found that certainty calibration reduced the goodness of fit (of the logit WTP model) and increased the variance in responses. But Welsh and Bishop (1993) reported that certainty calibration reduced the variance in responses[3].

On the other hand, the NOAA panel suggested that a 'Don't Know' option should be added to the DC CVM format. Alberini et al. (2003) identify three interpretations of responses to this option. One possibility is that 'Don't Know' respondents are not in the market for the good being valued. A second interpretation is that 'Don't Know' respondents have not yet made up their mind. The third possibility is that these responses reflect uncertainty. Moreover, Alberini et al. define two types of uncertainty: (a) "true" uncertainty wherein respondents have insufficient experience, and, (b) "false" uncertainty wherein respondents do not want to spend time thinking about the valuation question or would like to indicate some support for the item being valued, but would not pay the amount asked. However, Wang (1997) presents an alternative approach for treating uncertainty. He argued that Don't Know (or Not Sure) answers represent the point of indifference to the offered bid. As the price of the commodity increases, a typical respondent would switch her answer from Yes to Don't Know and from Don't Know to No. Wang included the Don't Know answers in a multinomial probit model estimation and concluded that they provide useful information about preferences.

Taken together these arguments suggest that responses to valuation questions that include an uncertainty scale may differ from responses to valuation questions that include either a 'No

[3] Several other studies have also applied certainty scales to calibrate both 'Yes' and 'No' responses (See for example, Li and Mattson [1995] and Ekstrand and Loomis [1997]).

answer' or 'Don't Know' option, or formats containing a strict DC (Yes/No) option. Since value estimates can differ by how "uncertain responses" are collected and treated in analysis, understanding the motivation underlying these responses is critical.

This paper contributes to previous literature in two ways. To our knowledge, it is the first comparison of cheap talk and alternative uncertainty adjustments derived from a mail survey. Second, we explore some of the factors that might result in different responses to a quantitative uncertainty scale as compared to a 'Not Sure' category.

Methodology

A mail survey was used to elicit attitudes towards management and user fees to access public lands in the context of the current US Fee Demonstration Program (FDP). The FDP was experimentally implemented for some public lands and allows the Forest Service, the Bureau of Land Management and the U.S. Fish and Wildlife Service to impose access fees for public use of these lands. The purpose of the FDP is to test the appropriateness of entrance fees as a mechanism to raise additional money to maintain public natural resources and recreation sites.

The survey was pretested with a pilot survey in June 2002 and was then mailed in October to 2500 randomly selected households in New Hampshire and Idaho.

Within each state a two-stage cluster sampling was applied in order to distinguish between the urban and rural population. Urban and rural clusters were randomly selected for each state, using the official census classification of urban and rural areas. Households from each cluster were then randomly sampled. The sample size within each cluster was proportional to the population size of the respective cluster. In an effort to increase response rates and reduce non-response bias, we followed the four-step procedure proposed by Dillman (2000). The overall response rate was 33%, giving a total of 807 observations.

The hypothetical part of the survey consisted of a description of a recreation area (a hypothetical public site with a scenic overview) which had become part of the FDP. The willingness to pay (WTP) question was presented in a referendum format in which respondents were asked to make hypothetical payments of randomly assigned prices ($3, $5 or $10) for access to this site. Three versions of the questionnaire were used. The first was a baseline version consisting of 'Yes' and 'No' options followed by a standard 10-point certainty scale. The second version (version NS) included a 'Not Sure' category for the WTP response (Appendix I). The third version (version CT) contained a cheap talk script which preceded the referendum question and explicitly warned respondents about the nature of hypothetical bias (see Appendix I).

The theoretical utility model and the derivation of willingness to pay follow well established procedures, outlined in Appendix II. Mean WTP was calculated by integrating under a logit function where price was truncated at $25 and bounded to be positive[4]:

$$\text{mean WTP} = \int_0^{25} [1 - Gwtp]dW \qquad (1)$$

where W is the dollar amount individuals are asked to pay and G_{wtp} is the distribution function of the true willingness to pay.

Results

(i) Description of variables

Variables included in the analysis are described by survey version in Table 1. Two-sample t-tests for difference in means and proportions showed that the distribution of these variables

[4] As noted by Haab and McConnell this approach only bounds calculated WTP, however the estimated logit function is unbounded. A bound logit model, as proposed by the same authors, in which the WTP function is bound between zero and income as well as a Turnbull estimation was also considered. The results from these two estimations are not reported here but are consistent with the results reported later in this paper.

between survey versions was very similar and statistically indistinguishable, as expected, since survey versions were mailed randomly. This allows evaluation of the effects on WTP that arise due to different treatments and eliminates the possibility of confounding effects due to differences between "average" respondents to each survey version.

Respondents had an average income of between 45 to 60 thousand per year per household and most had at least a college degree. The size of most households varied between 1 and 4, with an average of 2.6 per household. Average age was 56.5 years, skewed towards the upper tail of the population distribution. Fifty six percent of the respondents reported visiting public lands at least three times a year over the last three years[5]. The mean number of visits in the past three years was 11 visits per person.

The distribution of 'Yes', 'No' and 'Not Sure' responses for each survey version is shown in Table 2. About 62% rejected the fee offered in the baseline version and about half of the respondents rejected the fee in the NS and the CT version. Approximately one in 6 chose the 'Not Sure' option for the NS and the CT versions.

(ii) Model building

Three groups of variables were included in the willingness to pay model, based on the theoretical expectations from classical economics and regardless of their statistical significance: (i) dollar amount requested and income, (ii) individual tastes and preferences, and (iii) social characteristics, respectively. These were represented by the variables price, income, (previous) visits to recreation lands, age and household size.

We hypothesize that residents of Idaho and New Hampshire differ culturally in their preferences, and that residents of rural and urban areas differ in their lifestyle regarding outdoor

[5] This criteria was chosen arbitrarily, we assume that people who visit public lands at least 3 times a year are regular visitors, who have well-formed preferences for public lands, while respondents who visit public lands occasionally, say once a year, may not have well established preferences.

activities. The effect of these two factors was represented by the variables 'state' and 'urban'. In order to adjust for non-response bias typical for surveys of this sort, we also included a variable accounting for the round in which the surveys were returned. Linearity of age and income was examined visually by plotting these variables on a logit scale. The inclusion of the variables state, round and urban was then assessed on the basis of three criteria: (1) significance in a univariate model as a main effect variable, (2) likelihood ratio test after inclusion in the main effects model and, (3) the effect of the variable as a modifier on the other variables (percentage change of the estimated coefficients). Interaction terms were considered on the bases of plausibility and statistical significance.

(iii) Model estimates

For each of the above versions a logit model was estimated where willingness to pay (Yes, No) was regressed on the variables listed in Table 1. A likelihood ratio test for difference of estimated coefficients between the two states showed that the estimates were not statistically different which allowed the data to be pooled. The model estimates for the baseline, the NS (where Not Sure responses are treated as missing) and the CT versions are shown in Table 3.

The effect of price and income and number of household members were as expected: positive effect of income and negative of permit price and number of household members (except for the CT version where the effect was not significant). The income and price variables were statistically significant. We did not have prior expectations of the effect of the variable 'visits'. On one hand visitors of public lands can be expected to be more likely to pay, since they are the users of the commodity that is being valued. However, in this particular study users may be less likely to pay on the basis of principal or strategic objections to user fees. Our estimates show a negative effect for this variable. The effect of variable 'round' can be positive or negative. People who are less interested in public lands can be expected to have a lower response rate, which means that the effect of round would follow the same logic as the effect of previous visits. However, we might expect non-respondents to be mainly working people with busier schedules, which might suggest a positive effect of the variable round.

(iv) Tests

In order to test for the relative effect of cheap talk and each type of certainty calibration we compared the estimates of mean WTP obtained from the baseline version to the estimates for mean WTP derived from versions NS, Certainty8, Certainty10[6] and CT. Mean WTP for all versions are presented in Table 4 and plotted in Figure 1. In addition, we compared the probabilities of a 'Yes' response for each fee level in order to test whether the effect of cheap talk or certainty calibration may change as fee asked for increases. The latter are presented in Table 5.

Logit models, as described earlier, were fit to the data from each survey version. In the NS version, 'Not sure' responses were treated as 'No', 'Yes' or 'missing' and a logit model was estimated for each of these treatments. Mean WTP values were calculated using equation 1. Previous research has suggested two main methods for confidence interval estimation. Park, Loomis and Creel (1991) proposed a simulation method based on the Krinsky and Robb (1986) technique where a Gauss distribution is simulated around each estimated coefficient using its estimate and variance. The second approach, proposed by Duffield and Patterson (1991), is based on bootstrapping (with replacement) from the original sample. We use the second approach , which, as pointed out by Cooper (1994), does not impose normality on the distribution of the coefficients. Bootstrapping was done in SAS. Mean WTP was calculated through integration using MATHEMATICA. Empirical confidence intervals around the point estimate of mean WTP

[6] As noted above in the Certainty8 and Certainty10 versions all 'Yes' responses in the baseline version followed by certainty of less than 8 or 10 respectively were recoded as 'No'.

(see Table 4 and Figure 1) and around the point estimates of the probability of a 'Yes' response by fee levels (see Table 5 and Figures 2 and 3) were constructed by generating 1000 bootstraps with replacement for each version.

I. Effect of cheap talk

A comparison of the empirical (bootstrap) distribution of the mean WTP of the CT and the baseline version showed that the mean bootstrap values of the WTP were similar but the 95% CI of the CT version were larger as compared to the baseline version (See Table 4 and Figure 1).

As noted earlier, in previous studies the effect of cheap talk has been mixed. In this study it did not reduce value estimates and the variance of responses was greater as compared to the baseline version. The ineffectiveness of the cheap talk script may be due to several factors. Initially we assume that hypothetical bias actually exists in our data and that respondents read the CT script but failed to calibrate their WTP despite of the warning information about hypothetical bias contained in the script. Under these conditions it is possible that respondents' reaction to cheap talk was dismissive, because they did not believe that their own responses were biased or because they wanted to appear to be 'good' citizens by saying that they would pay. Another possibility is that people simply did not read the cheap talk script because of lack of time, lack of commitment, or interest.

In the light of previous studies that have related cheap talk to price levels (Brown and Ajzen 2002) and experience (List 2001), we compared the effect of cheap talk for each price level, and examined the effect of cheap talk among visitors and non visitors of public lands (assuming that visitors would have well formed preferences as compared to non visitors).

Cheap talk was ineffective for all price levels, and did not reduce the probability of Yes to a greater extent among respondents faced with $10 as compared to a $5 or $3 fee. The 95% CI between the CT and the baseline version were almost identical for each fee level. This result differed from the findings by Brown and Ajzen who reported that respondents faced with a higher price were more influenced by cheap talk.

Regarding List's findings that respondents with well formed preferences are less influenced by cheap talk, our data did not indicate this to be the case. When the version containing cheap talk was split into two groups: visitors and non visitors of public lands, estimates of the mean WTP were not statistically lower for non users of public lands as List's findings would suggest.

II. Effect of certainty scale

The effect of certainty scale calibration was tested using the same approach as for CT. The results indicated that certainty scale reduced WTP by a factor of 2 in the Certainty8 version and by a factor of 3 in the Certainty10 version. In addition, our estimates of mean WTP, had smaller variances for the two certainty versions as compared to the NS and the CT versions (see Table 4 and Figure 1).

Although we do not have data from actual payments, the twofold reduction in value due to the calibration of Yes responses in the Certainty8 version can be expected to give a reasonable approximation of real willingness to pay, if we rely on the recommendations of the NOAA panel (1994) of dividing hypothetical estimates by two. However, it may reduce but not completely eliminate bias, if we rely on the conclusions of List and Gallet (2001), who report an average three-fold hypothetical bias. The Certainty10 version on the other hand, may eliminate bias, if we rely on the List and Gallet conclusions, but may underestimate WTP if hypothetical bias is twofold as suggested by the NOAA panel. Therefore, the decision about the cut point for the certainty calibration is an important issue.

III. Effect of a 'Not Sure' option

To assess the effect of a 'Not Sure' option, we compared the bootstrap distributions of the NS version to the baseline version . The value estimates of the three measures of WTP when the 'Not Sure' option was missing, treated as 'Yes' or as 'No' remained unchanged (accounting for the 95% intervals) as the values for the baseline version. However, the variances were much larger (see Table 4). This implies that the inclusion of a Not Sure category as recommended by the NOAA panel does not have an effect on mean WTP estimates, at least in this case study.

IV. 'Not Sure' versus certainty calibration

An important question raised in previous research is what information about individual responses does certainty reveal. While there is a fair amount of literature on the motivation behind 'Not Sure' responses, the motivation behind uncertain responses when a scale is used, has not been discussed. An important point is the difference, if any, between what a 'Not Sure' response represents as compared to a low level of certainty (when the certainty scale is used).

Regarding a 'Not Sure' (or a Don't Know) option, Wang (1997) and (Carson et al. 1994) provide evidence for two opposing arguments of why people are unsure of their responses. Carson et al. recommend that 'Not Sure' responses be treated as No, because respondents who choose the 'Not sure' option would say 'No' if actually forced to choose. Wang argued that uncertainty is informative about true WTP: respondents would switch their answers from 'Yes' to 'Not Sure' and from 'Not Sure' to 'No' as price increases. In addition Champ et al. (2003), find that respondents may choose the 'Not Sure' option because they have uncertainty about their income, ability to commit to spending money and about the benefits of the program. Other hypotheses suggest that uncertainty may arise because of lack of knowledge, interest, or inability to make a quick decision. The arguments above imply that 'Not Sure' responses may be caused by a wide range of reasons, which may be the cause of an increased variation in responses when a 'Not sure' option is included, as compared to settings in which a 'Not Sure' option is missing. In this study, examination of the explanatory comments following the 'Not Sure' responses in the NS option, showed that some respondents chose this category instead of saying 'No'. Others chose the 'Not Sure' option because of income considerations, or because they didn't understand the question very well, or because they were unable to picture a specific situation and were, as a result, unsure whether they would actually visit and pay for such a site.

Regarding certainty scales, previous studies that have applied the certainty scale calibration to 'Yes' responses have essentially considered responses with certainty of 8 and less (or 10 and less) for both, 'Yes' and 'No' responses to be equivalent to 'Not sure' responses, and have treated them as 'No'. Such calibration of certainty, however, provides different results as compared to a 'Not Sure' calibration when 'Not Sure' responses are treated as 'No'. For our sample, when Yes responses followed by a certainty level of less than 8 and less than 10 were treated as 'No', mean WTP was $2.65, and $1.68 respectively. However, when 'Not Sure' answers were treated as 'No', mean WTP was $4.87. In addition the variation of WTP values was much smaller for the first two estimates (See Table 4 and Figure 1). To the best of our knowledge previous research has not addressed the issue of why these two ways of expressing and calibrating uncertainty might provide different results.

We hypothesize that there is a conceptual difference between the two: the motivation for a 'Not Sure' choice can differ from the motivation for marking a low level of certainty on a 10 point scale. A 'Not Sure' or 'Don't Know' option is a separate, and independent category (in addition to 'Yes or 'No') which either represents a qualitative judgment of WTP or may be associated with factors, other than price, as pointed out earlier. Certainty scale levels on the other hand may involve greater internal commitment, and are quantitative judgments about the magnitude of a 'Yes' or a 'No' response that individuals have taken. Certainty enhances a positive or a negative response with increasing magnitude along the certainty scale; however, a 'Yes' is still a positive response, regardless of the certainty level. Consequently, models for the two ways of incorporating uncertainty are conceptually different. A model that explains 'Not Sure' responses would predict the probability of selecting 'Not Sure' as opposed to a Yes/No category, however a model that explains a low level of certainty when a scale is used, would predict the probability of being uncertain to a Yes/No response.

The implicit assumptions when certainty scales are being used to calibrate 'Yes' responses (for example as in Champ et al. 1997) can be summarized in Hypothesis I below. Hypothesis II is based on Wang's argument. However, other reasons for uncertainty are also plausible. Two other factors that may play an important role in uncertainty adjustment of a DC CV question are suggested in hypothesis III and IV.

Hypothesis I. Self-reported certainty to a 'Yes' response provides information about the individual's true utility-maximizing price. A respondent who overstates his WTP (in a DC format) calibrates his response, using the certainty scale, until he reaches the optimal price. Certainty to 'No' responses doesn't yield any relevant information about one's WTP.

Hypothesis II. Certainty is lowest at the price that is the true willingness to pay (Wang's argument). In this study since the mean WTP is about $5, we can expect that at $5, certainty levels would be smaller as compared to $3 and $10.

Hypothesis III. Certainty represents consistency between answers. People tend to avoid personal contradictions, and once they choose a 'Yes' or a 'No' response they would tend to back it with a high level of certainty

Hypothesis IV. Certainty represents attitudes about the program being valued. By indicating high levels of certainty to a 'Yes' response respondents may be expressing their support of the program being valued and by marking high levels of certainty to a 'No' response they may express objection on principal

Our data did not seem to support Wang's hypothesis. The proportion of respondents who were certain, was not lowest at the $5 price. The consistency hypothesis (hypothesis III) is based on the theory of stability (For example, see Schwarz and Sudman 1996) according to which people tend to reduce mental pressure by avoiding personal contradictions. When asked a question, a respondent may hesitate about whether to say 'Yes' or No'. However, when asked later about her certainty, the respondent would say that she is certain, in order to avoid self-contradiction. We may expect patterns of consistency during in person interviews when respondents feel watched. However, consistency can also be an internalized norm of behavior that will appear regardless of social settings.

We hypothesize that certainty levels may represent consistency between answers regardless of true WTP. If consistency were a factor, an intuitive consequence would be that the distribution of certainty levels for the whole sample would be skewed towards 10. Our data did show an uneven distribution of certainty levels. Along the certainty scale, certainty levels were strongly skewed towards 10, with very few certainty levels less than 5. About half of the responses were followed by a certainty level of 10 indicating 'very sure'. This was the case for all price levels and for both 'Yes' and 'No' responses (see Table 5).

Hypothesis IV argues that certainty of a WTP response may be a manifestation of attitude, rather than true willingness to pay. In order to test this hypothesis we explored the association between certainty levels and attitudes towards user fees. Respondents who objected to fees in principle were more certain in rejecting the price asked. Males tended to be more certain in their answers than women but this result is mixed; the significance of gender depended on how certainty was coded. The associations of certainty levels to gender and attitudes about user fees are summarized in Table 6. Certainty to ‘No’ responses is greater among those who object to fees, implying that high certainty to a ‘No’ response is a way to assert objection. Certainty to ‘Yes’ responses was not correlated with price, or attitudes.

Table 7 presents logistic regression estimates in which certainty is regressed on objection to fees, gender and whether price was $5(which is the average WTP) or not. Among ‘No’ responses, negative attitudes towards user fees have a highly significant effect on certainty levels. Respondents who objected to user fees in principle were on average 2.7 times more likely to indicate certainty of 10. The effect of gender was insignificant as was the effect of price.

Conclusions

This study conducted a mail contingent valuation survey about access to public lands in the context of the Fee Demonstration Program. Cheap talk was not an effective instrument to reduce WTP estimates. A certainty scale, however, reduced estimates by about two to three times, which may suggest elimination of hypothetical bias.

The two common ways of expressing uncertainty: a certainty scale and a 'Not Sure' option produced different results and are possibly motivated by different reasons. The inclusion of a 'Not Sure' option increased the variance in responses as compared to a strictly DC format but estimated mean WTP was not reduced. This is possibly because of the wide range of factors that may underlie a 'Not Sure' response.

However, when a certainty scale is used, high levels of certainty may be an indication of consistency between answers and, also, may represent expressions of attitudes; thus, certainty levels may not be informative about the true values respondents place on the commodity being valued. This implies that in field applications of contingent valuation, certainty calibration may not consistently correct for hypothetical bias across studies.

In the context of the current debate about the appropriate way to adjust for respondents' uncertainty, further research on the relationship between certainty scale levels and individual characteristics is needed in order to verify the validity of this technique and its ability to consistently reduce or eliminate hypothetical bias. In addition, much more work is needed in order to determine the optimal cut off point for recoding 'Yes' answers as 'No' when respondents are uncertain.

References

Aadland, D. and Caplan, A.J., 2003. "Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias." American Journal of Agricultural Economics. 85(2) 492-502.

Alberini, A., Boyle, K. and Welsh, M., 2003. "Analysis of Contingent Valuation Data With Multiple Bids and Response Options Allowing Respondents to Express Uncertainty." Journal of Environmental Economics and Management. 45, 40-62.

Brown, T.C. and Ajzen, I., 2002. "Further Tests Of Entreaties To Avoid Hypothetical Bias In Referendum Contingent Valuation". Journal of Environmental Economics and Management, forthcoming.

Carson, R. T. , Hanemann, W. M., Kopp, R.J., Krosnick, J.A., Mitchell, R. C., Ruud, P.A., and Smith, V. K., 1994. "Prospective Interim Lost Use due to DDT and PCB Contamination in the Southern California Bight." A Report of Natural Resource Damage Assessment, Inc. and Industrial Economics, Inc. to US. NOAA.

Champ, P., Bishop, R., Brown, T. and McCollum, D., 1997. "Using Donation Mechanisms to Value Nonuse Benefits From Public Goods." Journal of Environmental Economics and Management. 33(2), 151-63.

Champ, P., Alberini, A. and Correas, I., 2003. "Using Contingent Valuation to Value a Noxious Weeds Control Program: The Effects of Including an "Unsure" Response Category." Working paper.

Cummings, R.G. and Taylor,L.O., 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." American Economic Review. 89(3): 649-65.

Cooper, J. C., 1994. "A Comparison of Approaches to Calculating Confidence Intervals for Benefit Measures from Dichotomous Choice Contingent Valuation Surveys." Land Economics. February, 70(1): 111-22.

Dillman , D. A., 2001. "Mail and Internet Surveys: The Tailored Design Method." New York: John Wiley & Sons, Inc.

Duffeld, J. W. and Patterson, D. A., 1991. "Inference and Optimal Design for a Welfare Measure in Dichotomous Choice Contingent Valuation." Land Economics. 67(2) 225-39.

Ekstrand, E. R. and Loomis, J., 1997. "Estimating Willingness to Pay For Protecting Critical Habitat For Threatened And Endangered Fish With Respondent Uncertainty" Unpublished Manuscript.

Haab, T.C. and McConnell, K.E., 2002. "Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation: New York: Edward Elgar.

Hanemann, M., 1984. “Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses.” American Journal of Agricultural Economics. 66(3): 332-41.

Harrison, G.W. and Rustrom, E.E., 2002. “Experimental Evidence on the Existence of Hypothetical Bias in Value Elicitation Methods.” Forthcoming in Handbook of Results in Experimental Economics. C. Plott and V.L. Smith, eds. New York: Elsevier Science.

Krinsky, I. and Robb, A. L., 1986. “On Approximating the Statistical Properties of Elasticities.” Review of Economics and Statistics. 68 (Nov): 715-19.

Li, C. and Mattsson,, 1995. “Discrete Choice Under Preference Uncertainty: An Improved Structural Model for Contingent Valuation.” Journal of Environmental Economics and Management. 28: 256-69.

List, J. A., and Gallet, Graig A., 2001. “What Experimental Protocol Influence Disparities between Actual and Hypothetical Stated Values?” Environmental and Resource Economics. 20:241-254.

List, J. A., 2001. “Do Explicit Warnings Eliminate the Hypothetical Bias in Elicitation Procedures? Evidence from Field Auctions of Sportscards.” American Economic Review. 91(5): 1498-1507.

Loomis, J. B., Lucero, T. and Peterson, G., 1996. “Improving Validity Experiments of Contingent Valuation Methods: Results of Efforts to Reduce the Disparity of Hypothetical and Actual Willingness to Pay.” Land Economics. 72(4): 450-61.

Murphy, J. , Stevens,T., Weatherhead, D., 2003. “An Empirical Study of Hypothetical Bias in Voluntary Contribution Contingent Valuation: Does Cheap Talk Matter? Working Paper, University of Massachusetts, Amherst.

National Oceanographic and Atmospheric Administration (NOAA)., 1994. “Proposed Rules for Valuing Environmental Damages.” Federal Register. 59 (Jan): 1062-91.

Park, T., Loomis, J.B., and Creel, M., 1991. “Confidence Intervals for Evaluating Benefits Estimates from Dichotomous Choice Contingent Valuation Studies.” Land Economics. February, 67(1) 64-73.

Polasky, S., Gainutdinova, O. and Kerkvliet, J., 1996. “Comparing CV Responses with Voting Behavior: Open Space Survey and Referendum in Corvallis, Oregon.” in Herriges, J., compiler, W-133 Benefits and Costs Transfer in Natural Resources Planning, 9th Interim Report. Department of Economics, Iowa State University, Ames, IA.

Poe, G., Welsh, M. and Champ, P., 1997. “Measuring the Difference in Mean Willingness to Pay When Dichotomous Choice Valuation Responses Are Not Independent.” Land Economics. 73(2): 255-67.

Ready, R. and Navrud, S., 1999. “How Do Respondents With Uncertain Willingness to Pay Answer Contingent Valuation Questions?” Manuscript, Agricultural University of Norway.

Schwarz,N. and Sudman,S., 1996. “Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research.” San Francisco: Jossey-Bass Publishers.

Wang, H., 1997. “Treatment of Don’t Know Responses in Contingent Valuation Surveys: A Random Valuation Model.” Journal of Environmental Economics and Management. 32(2): 219-32.

Welsh, M. and Bishop, R., 1993. “Multiple Bounded discrete Choice Models.” In W-133 Benefits and Costs Transfers in Natural Resources Planning, 6th interim report. Department of Agricultural Economics, University of Georgia,

Table 1. Descriptive characteristics (N=807)

Variable	Description	Coding	Mean (stdev)			
			All versions	**Baseline version n=285**	**NS version n=261**	**CT version n=253**
Price ($)	Dollar amount asked in the questionnaire.	$3, $5, $10	5.9(2.9)	6.05(2.9)	5.86(2.9)	5.91(2.9)
Income category	1 =less than $10,000 to 10 = above $120,000 per year	1 to 10	5.2(2.2)	5.3(2.3)	5.1(2.1)	5.1(2.3)
Visits	Whether respondents visited public lands more than 3 times a year in the past 3 years	1=Yes 0= otherwise	0.56(0.49)	0.56(0.49)	0.60(0.49)	0.53(0.49)
HH	Number of household members	Continuous	2.6(1.6)	2.64(1.3)	2.82(2.0)	2.54(1.2)
State	Whether resident of NH or Idaho	0= NH 1 = Idaho	NH (n=374) Idaho (n=418)	0.52(0.2)	0.50(0.5)	0.56(0.5)
Age	Age of respondent	Continuous	56.5(15.01)	57.2(14.9)	55.7(14.7)	56.6(15.3)
Round	Whether survey was returned in first round or in 2nd round	0=first round 1=second round	1.27(0.42)	0.27(0.40)	0.28(0.45)	0.26(0.44)
Urban	Whether survey was sent to a urban or a rural cluster	0=rural 1=urban	0.09(0.25)	0.08(0.2)	0.10(0.3)	0.08(0.3)

Table 2. Distribution of willingness to pay

WTP	N
Baseline Version	281
wtp=No	133(62.3%)
wtp=Yes	75(37.7%)
Not Sure (NS) version	259
wtp=No	129(49.8%)
wtp=Yes	83(32.1%)
wtp=Not Sure	47(18.1%)
Cheap Talk (CT) version	250
wtp=No	133(53.2%)
wtp=Yes	75(30.0%)
wtp=Not Sure	42 (16.8%)

Table 3. Logistic estimation of WTP Function

Variable (Expected sign)	Baseline version Estimate (st error)	NS version (Treating Not Sure as missing) Estimate (st error)	CT version Estimate (st error)
Intercept	3.30(2.01)*	2.34(2.47)	0.64(2.04)
Price (-)	-0.26(0.05)***	-0.17(0.05)***	-0.17(0.05)***
Income (+)	0.18(0.08)**	0.15(0.08)*	0.22(0.08)***
Visits (?)	-0.72(0.30)**	-1.20(0.37)***	-0.53(0.36)
HH (-)	-0.23(0.11)**	-0.23(0.16)	0.09(0.15)
State (?)	0.20(0.29)	0.16(0.33)	-0.15(0.35)
Age (?)	-0.13(0.06)**	-0.06(0.07)	-0.09(0.06)
Age2	0.001(0.0005)**	0.0005(0.0006)	0.001(0.0006)*
Round (?)	0.45(0.32)	0.40(0.37)	0.38(0.36)
Urban (?)	0.23(0.46)	0.30(0.58)	0.26(0.59)

*** Variable is significant at 99% level
** Variable is significant at 95% level
* Variable is significant at 90% level

Table 4. Mean willingness to pay, and 95% bootstrap confidence intervals for 1000 bootstraps

Version	Mean WTP ($) Lower CI – Upper CI	Difference between Upper CI and lower CI
Baseline (DC format) n=281	4.32 3.32 – 5.46	2.14
Certainty8 (WTP=Yes only for certainty ≥8) n=259	2.65 1.72-3.59	1.87
Certainty10 (WTP=Yes only for certainty =10) n=259	1.68 0.79- 2.67	1.88
NS 'Not Sure' = missing n=182	5.43 3.57 - 9.11	5.54
NS 'Not Sure' = Yes n=224	7.28 5.25-11.02	5.77
NS 'Not Sure' = No n=224	4.87 2.69-10.58	7.89
CT (Contains a Cheap Talk script) n=208	5.05 3.10 - 8.59	5.49

Table 5. Distribution of Certainty levels
Number (%)

Certainty	All n=260	$3 n=90	$5 n=90	$10 n=80	Yes* n=108	No* n=160
<8	63 (24%)	23 (25%)	21 (24%)	19 (23%)	33 (35%)	29 (18%)
8 to 9	62 (24%)	23 (26%)	20 (22%)	19 (24%)	23 (23%)	39 (25%)
10	135 (52%)	44 (49%)	49 (54%)	42 (53%)	42 (42%)	92 (57%)

* There was no deviation from this distribution when ‘Yes’ and ‘No’ answers were broken down by price level

Table 6. Association between certainty levels and attitudes and gender

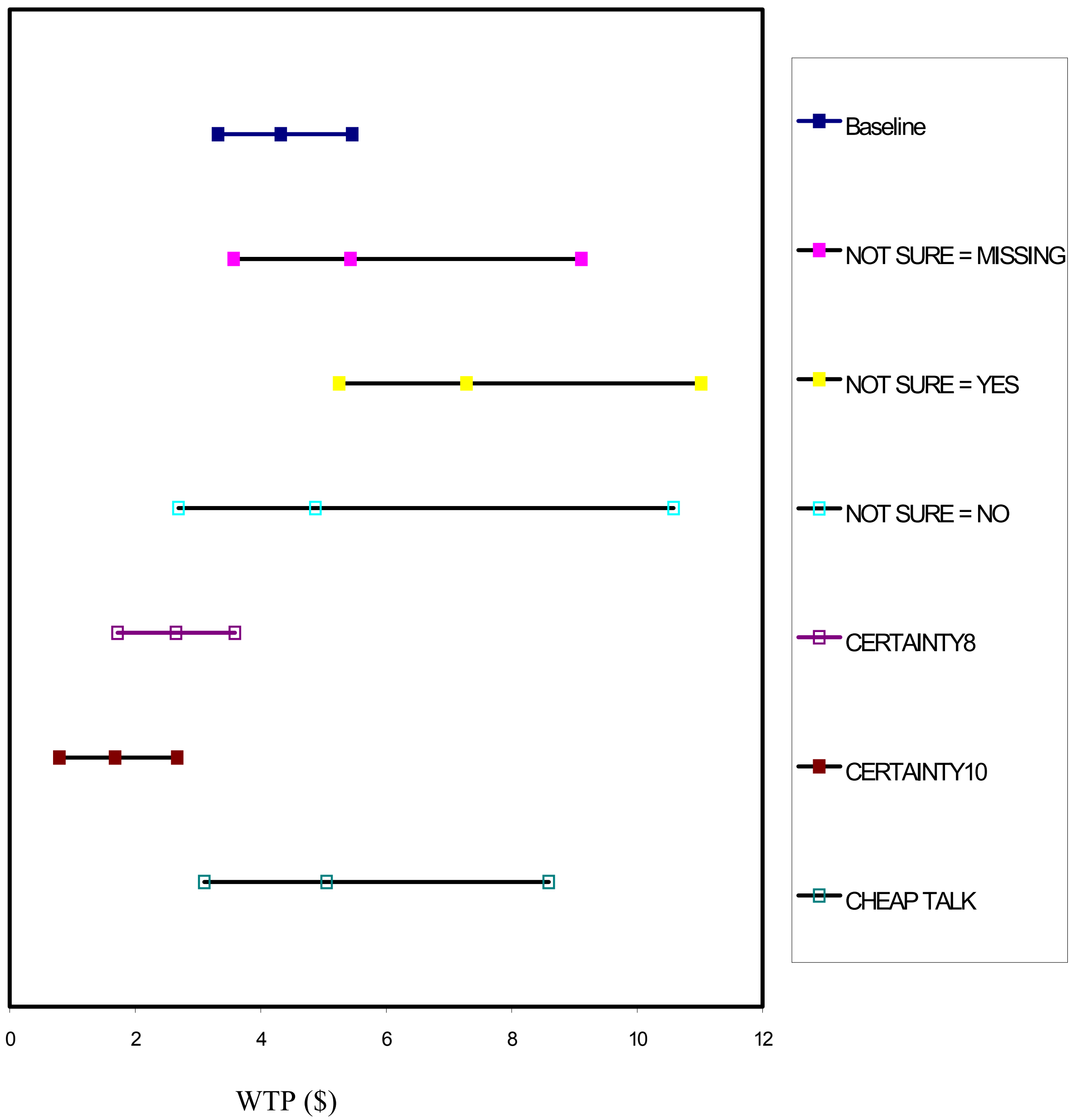
Certainty	Males, %	Females, %	Objected to fees on principal, %	Did not object to fees, %
8 to 10	78.8 (p=0.06) †	67.6	83.3 (p=0.09)†	73.2
10	53.8 (p=0.2) †	45.1	71.2 (p=0.0003) †	45.4

† Kruskal-Wallis test. For example, Ho: 78.8 = 67.6

Table 7. Predictors of certainty levels

	Certainty to a 'No' response (n=157)		Certainty to a 'Yes response (n=97)	
Variable (expected sign)	Estimate (st error)	Odds ratio	Estimate (st error)	Odds Ratio
Intercept	0.35(0.54)		0.35(0.54)	
Mean price (-) = 1 if price=$5, 0 otherwise	0.15(0.35)	1.2(0.6-2.3)	0.30(0.43)	1.4(0.6-3.2)
Object (+) = 1 if objected to user fees in principal, 0 otherwise	1.01(0.35)***	2.7(1.4-5.5)***	0.5(1.4)	1.6 (0.1-27.7)
Gender (+) 1=male, 2=female	-0.38(0.37)	0.7(0.3-1.4)	-0.03(0.37)	1.0(0.4-2.6)
	LR=0.01		LR =0.9	

Figure 1. Mean WTP and 95% confidence intervals for each version



APPENDIX I

Hypothetical Settings

Imagine an area with a scenic overlook in a nearby federal or state public forest. In the past, this area was free with only picnic tables and a dirt parking lot. This year the area is the same as always, but it is part of the Fee Demonstration Program (described in the cover letter), so you must buy a permit or face a fine of $100 if caught without a permit. Permits are sold at a visitor's center that you pass on the way to the site.

If a permit to use this area costs $______ per visitor per day, would you buy it, keeping in mind your household income and other financial commitments?

Baseline Version

A. Yes, I would pay this amount.

B. No, I would not pay this amount. (Please explain why

Certainty Calibration

C. How sure are you of your decision about how much you would pay? Please circle one number from 1 to 10, with 1 indicating "very unsure" and 10 indicating "very sure".

1	2	3	4	5	6	7	8	9	10
very unsure									very sure

If you are unsure, please explain why.

Version Certainty8 : All 'Yes' responses followed by certainty less than 8 are recoded as 'No'

Version Certainty10: All 'Yes' responses followed by certainty less than 10 are recoded as 'No'

Version NS

A. Yes, I would pay this amount.

B. No, I would not pay this amount. (Please explain why

C. Not Sure. (Please explain why)

Version CT

In recent studies, several different groups of people were asked to make choices like this. Although they were asked how much they thought they would pay, no one actually paid money. These studies found that on average, people OVERSTATED their actual willingness to pay by as much as 150 percent -- quite a difference. In this question, please decide on how much you would pay exactly as if you were going to pay real money.

APPENDIX II

We follow a simple model where individual utility is a function of Income, Y, a basket of market goods, X, and public land use, Q.

$$U = U(Y, X , Q)$$

Individuals are asked to pay a dollar amount, W for access to a specific public land, Q.

The utility after paying this amount would be

$$U_1 = U_1(Y\text{-}W,X,Q) + \varepsilon_1$$

while the utility if the bid is rejected would be

$$U_0 = U_0(Y,X) + \varepsilon_0$$

Individuals will pay W, if $U_1 \geq U_0$. That is:

$$Pr[Yes] = Pr\ [U_1(Y\text{-}W,X,Q) + \varepsilon_1 \geq U_0(Y,X) + \varepsilon_0$$

which can be rewritten as

$$Pr\ [\varepsilon_0 - \varepsilon_1] \leq U_1(Y\text{-}W,X,Q) - U_0(Y,X)$$

Let

$$\Delta U = U_1(Y\text{-}W,X,Q) - U_0(Y,X),$$

$\eta = \varepsilon_0 - \varepsilon_1$, and

$F\eta$ be the cumulative distribution function of the error.

Then the expression above can be rewritten as

$$Pr\ [(\eta) \leq \Delta U] = F\eta(\Delta U),$$

which if $F\eta(\Delta U)$ is assumed to have a logistic cumulative density function is equal to $(1 + e^{-\Delta U})^{-1}$.

Using the approach described by Hanley et al. (1997) in order to proceed we need to adopt a specific functional form for u(.). assume for example, the following simple form:

$$u = u\ (\alpha + \beta_1\ Y + \beta_2\ X + \beta_3\ Q\ )$$

Then the change in utility would be

$$\Delta U = U_1(Y\text{-}W,X,Q) - U_0(Y,X) =$$

$$[\alpha_1 + \beta_1 *(Y - W)+ \beta_2\ X + \beta_3\ Q\ ] - [\alpha_2 + \beta_1\ Y + \beta_2\ X]=$$

$$(\alpha_1 - \alpha_2) - \beta_1\ W + \beta_3\ Q$$

Then, the probability of a Yes response is:

$$Pr[Yes] = F\eta\ [\ (\alpha_1 - \alpha_2) - \beta_1\ W + \beta_3\ Q]$$

The median WTP is calculated by

$$Pr\ [U_1(Y\text{-}W,X,Q) > U_0(Y,X)] = 0.5$$

We will use the approximation of compensating surplus, using the formula derived by Hanemann (1984)

$$Pr[Yes] = (1 + e^{-\alpha-\beta W}\ )^{-1}$$

then the median WTP = $-\alpha/\beta$

In a binary regression α is the sum of the coefficients of the explanatory variables, multiplied by the mean value of each variable, and β is the coefficient for the variable representing the bid amount.

The mean WTP is calculated by

$$\text{mean WTP} = \int_0^T [1 - Gwtp]dW$$

where G_{wtp} is the distribution function of the true willingness to pay. T is infinite for the true willingness to pay and is truncated at some value for the purpose of estimation.