Available online at www.sciencedirect.com

SCIENCE DIRECT®

Ecological Economics 55 (2005) 47–60

ECOLOGICAL ECONOMICS

www.elsevier.com/locate/ecolecon

ANALYSIS

# Using contingent valuation to value a noxious weeds control program: the effects of including an unsure response category

Patricia A. Champ[a,*], Anna Alberini[b], Ignacio Correas[c]

[a]USDA Forest Service, Rocky Mountain Research Station, 2150 Centre Ave., Bldg A, Fort Collins, CO 80526, United States
[b]University of Maryland, Department of Agricultural and Resource Economics, 2200 Symons Hall, College Park, MD 20742, United States
[c]University of Colorado, Department of Economics, Campus Box 256, Boulder, CO 80309-0256, United States

Received 17 September 2002; received in revised form 5 July 2004; accepted 24 October 2004
Available online 13 February 2005

**Abstract**

The National Oceanic and Atmospheric Administration Contingent Valuation panel recommended that a response category be included along with the vote in favor and vote against options associated with a referendum contingent valuation question that allows individuals to express lack of a well defined opinion. However, the recommendation did not include guidance on how to analyze such responses. In this article, we describe the results of a carefully designed split sample experiment that allowed for comparison of a standard dichotomous-choice referendum treatment to a treatment that included the option of responding unsure to the contingent valuation question in addition to the standard vote in favor and vote against response categories. We are able to examine several options for dealing with the unsure responses and conclude that the unsure responses should be included in the value estimation as the respondents who choose this response category are distinct from respondents who choose the vote in favor and vote against response categories.


*Keywords:* Nonmarket valuation; Contingent valuation; Willingness to pay; Dichotomous-choice format; Unsure responses; Referendum

## 1. Introduction

The National Oceanic and Atmospheric Administration (NOAA) Contingent Valuation Panel recommended the inclusion of an explicit would not vote response category in addition to the vote in favor and vote against response categories of a referendum contingent valuation (CV) question (Arrow et al., 1993). The implications of this recommendation have not been extensively investigated and the Panel did not provide guidelines for interpreting the would not vote option.

Subsequent to the NOAA panel recommendation, researchers have experimented with response formats for closed ended CV questions that, in addition to the vote in favor and vote against response categories, allow for refraining from voting altogether, or allow

* Corresponding author. Tel.: +1 970 282 0084; fax: +1 970 295 5959.
*E-mail address:* pchamp@fs.fed.us (P.A. Champ).

0921-8009/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolecon.2004.10.011

for expressions of uncertainty. The research has taken one of two approaches: comparisons of contingent values to actual behavior (Carson et al., 1986; Champ et al., 1997; Vossler and Kerkvliet, 2003; Vossler et al., 2003) or comparisons within CV treatments (Wang, 1997; Carson et al., 1998; Haener and Adamowicz, 1998).

The common finding across these studies is that when respondents are explicitly given the option of expressing uncertainty about the CV question, many respondents choose this response option. Yet each of the studies comes to a different conclusion about the best way to treat these expressions of uncertainty.

In this paper we describe a split sample experiment involving two CV treatments parallel in all aspects except one treatment uses a standard dichotomous-choice response format while the other treatment uses a standard dichotomous-choice response format with the addition of an unsure response category. These data allow us to explore three related topics: (1) what effect does inclusion of the unsure response category have on data quality? (2) How should the unsure responses be treated for purposes of analyses and value estimation? (3) Why do respondents choose the unsure response category?

The remainder of the paper is organized as follows. Section 2 reviews previous studies. In Section 3, we describe our study design and questionnaire. In Section 4, we examine the responses to the referendum question in our questionnaire and data quality issues. In Section 5, we examine the effect of alternative interpretations of the unsure responses on the WTP estimates. We investigate whether these interpretations are reasonable using multinomial logit models, sample frequencies, and cross tabulations of the answers to attitudinal and debriefing questions. Section 6 offers our concluding remarks.

## 2. Previous research

Three studies have looked at how, in the CV setting, offering a would not vote or don't response option affects responses to the willingness to pay question. Carson et al.'s (1998) working hypothesis was that inclusion of a would not vote option and recoding of those responses as vote against responses results in a distribution of response similar to that of a standard dichotomous-choice CV question. Two independent samples of respondents were administered versions of a CV survey about willingness to pay to prevent oil spills and the related damages to natural resources in Alaska. Approximately 18% of the survey respondents chose the would not vote response category when the option was explicitly offered by the interviewer. The distribution of responses to the CV question was statistically similar between the standard dichotomous-choice version and the version which offered the would not vote option, if the would not vote responses were conservatively recoded as would vote against.

Moreover, this recoding convention produced an estimate of median willingness to pay that was not statistically different from that based on the standard dichotomous-choice CV treatment. Carson et al. concluded that with a conservative interpretation of the responses, inclusion of a would not vote option does not reduce estimates of willingness to pay relative to a standard dichotomous-choice CV response format.

Haener and Adamowicz (1998) investigated various options for treating don't know responses to a referendum CV question about preservation of old growth forests in Alberta, Canada. In their study, approximately 18% of the respondents chose the don't know response category. Analysis of a multinomial logit model revealed that demographic characteristics were the main determinants of individuals choosing the don't know response category. Specifically, income, gender, and hiking activity were significant predictors of the probability of choosing the don't know response category.

Haener and Adamowicz included a question after the referendum question about how much their household would be willing to give up (in dollars) each year to preserve old growth forests. This information was used to recode the don't know responses to the referendum question into yes or no responses. Comparisons of three binary logit models, one that excludes the don't know responses, a second that recodes the don't know responses as no responses and a third that recodes the don't know responses based on the debriefing question about how much the household would be willing to give up for old growth forest preservation, suggest that recoding based on the debriefing question is superior to the other two approaches.

Estimates of mean willingness to pay were $121 for the model that excluded the don't know responses, $69 for the model that recoded don't know to no, and $89 for the model that recoded the don't know responses based on the debriefing question. The Haener and Adamowicz approach uses information from the full sample—a more appealing option than completely deleting the don't know responses.

Wang (1997) developed a model for estimating mean willingness to pay that uses information provided by the don't know respondents and applied it to CV data collected via a mail survey. The model assumes that individuals have a distribution of willingness to pay rather than a single value. Wang argues it is straightforward for people to answer yes (no) to a dichotomous-choice CV question when the offer amount assigned to the respondent is sufficiently low (high) relative to her true mean willingness to pay amount. By contrast, respondents answer don't know when the offer amount is close to their true mean willingness to pay. The corresponding statistical model is thus a variant of the ordinal probit (or logit), which identifies by how much willingness to pay must exceed (be less than) the offer amount for the respondent to say yes (no).

Wang found that for the four offer amounts used in the survey, a relatively large percentage (30%) of the respondents chose the don't know response category. Treating the don't know responses as vote against responses results in the lowest estimate of mean willingness to pay ($2.65). The ordinal logit model proposed by Wang used the information from the don't know responses and produced an estimate of mean willingness to pay equal to $11.86, an estimate very close to the estimate of mean willingness to pay obtained from a standard logit model that removes the don't know responses from the willingness to pay estimates ($10.23). The standard error around mean willingness to pay is a bit less for the Wang model (1.527) than that from the model in which the don't know responses are removed (1.703). Wang concluded that the NOAA panel plea for including a don't know response category is appropriate, and recommends using information from the don't know respondents as described in his paper.

Four studies investigated the effects of offering a don't know response category in a referendum CV by making comparisons with actual voting situations. Carson et al. (1986) compared a CV phone survey question to the actual vote on a referendum in California to issue bonds to construct sewage treatment plants. In that study, 24.6% of the CV respondents chose the don't know option. They find the CV result to be a good predictor of the actual referendum vote when 50–70% of the don't know responses are recoded as no responses.

Champ and Brown (1997) conducted a similar experiment in which a CV phone survey was administered prior to a referendum in Fort Collins, CO on retaining surplus tax funds to use for road maintenance. They found that 10% of the survey respondents were uncertain and the CV results were a good predictor of the actual referendum results when these uncertain responses were recoded as no responses.

Vossler et al. (2003) and Vossler and Kerkvliet (2003) conducted studies in Corvallis, OR that allowed for comparisons of a contingent referendum with an undecided response category to actual referenda results. Vossler et al. (2003) found empirical evidence supporting the treatment of the undecided votes as no votes, but Vossler and Kerkvliet (2003) concluded that recoding the undecided responses as no responses is inappropriate. Instead, they recommended recoding 54.6% of the undecided votes as no and 45.6% as yes votes.

The lack of a consistent recommendation for dealing with unsure responses suggests that the issue is complex and that a one-size-fits-all recommendation may not be in order. Our study adds to this literature by investigating the addition of an unsure response category to a referendum CV similar to the Carson et al. (1998), Haener and Adamowicz (1998) and the Wang (1997) studies. We are also able to conduct a more comprehensive statistical investigation of the reasons why people choose the unsure response option.

## 3. Study design

In the split sample design of this study, the two treatments are parallel in all aspects except the response format to the CV question. The final survey instrument was developed after conducting seven focus groups and a small pretest. The data were collected via mail surveys. As this is a methodological

study and we do not generalize the results, we used a convenience sample.

We recruited participants via ads placed in the general news section of three Sunday newspapers from June 27 to July 18, 1999.[1] A total of 891 Colorado residents responded to the ads and were mailed our questionnaires. Out of these, 743 questionnaires were returned, for a response rate of 84%.

The topic of the survey was controlling invasive plant species or noxious weeds on National Forests. Despite the seriousness of the threat posed to biodiversity by noxious weeds and the considerable coverage of this topic in the media, the focus groups we conducted revealed that many people were unaware of the noxious weeds problem. It was therefore necessary to provide study participants with a substantial amount of information prior to asking them about their willingness to pay for a program to control noxious weeds in National Forests. After providing information about National Forests and noxious weeds, study participants were asked a series of questions designed to measure their previous experience and attitudes toward National Forests, noxious weeds and the environment in general. Study participants were told that the Noxious Weeds Control Program would be financed with revenue from a special one-time tax. As it would take many years of treatment to control the noxious weeds, the revenue from the tax would be placed in an interest earning trust fund, and the funds would be used over the next 10 years to implement the Noxious Weeds Control Program. The CV question read as follows:

If the *Noxious Weeds Control Program* is implemented, the cost to your household would be $ (offer amount). Would you vote in favor or against the program?

In Treatment 1, the standard dichotomous-choice response format (vote in favor/vote against) was administered. In Treatment 2, the response format included an unsure response category in addition to the vote in favor and vote against categories, for a total of three response options. In both treatments, one of five offer amounts ($5, $10, $25, $50, $75) was randomly assigned to the CV question.

After the referendum question, we asked people to tell us the reasons for their answers. Possible answers to this question include, among other things, whether they felt uncertain about their future income and/or the benefits of the program, and whether they would want more information about various aspects of the program.

## 4. The data

### 4.1. Responses to the referendum question

Our first order of business is to compare the distributions of the responses to the survey questions prior to the CV question and the demographic questions to assess whether the respondents in the two treatments are drawn from similar populations. As shown in Table 1, the respondents in Treatment 1 (vote in favor/vote against) are very similar to the respondents in Treatment 2 (vote in favor/vote against/unsure) in virtually all respects (demographics, rates of visitation of National Forests, attitudes towards forests and environment, and prior knowledge of the weeds problem). This suggests that any differences across the two subsamples in the responses to the payment questions should be attributed to the treatment.

In Treatment 1, 76% of the valid responses were in favor of the program, compared to 62% of the valid responses in Treatment 2. In Treatment 2, 13% of the valid responses were vote against and 25% were unsure. This result is consistent with previous studies that found that when an explicit opportunity to express uncertainty was provided, a non-trivial percentage of respondents chose the response category rather than providing a definite response to the willingness to pay question.[2]

[1] The ads ran in the Denver Post and the Rocky Mountain News (two Denver based daily newspapers) and The Gazette (a Colorado Springs daily newspaper).

[2] This finding also confirms that numerous unsure responses can be observed in any survey that explicitly allows for such response category, and not just with in-person surveys, where, it has been suggested, respondents may opt for the unsure response when they are truly against the plan, but are reluctant to say so for fear of offending the interviewer or appearing to be socially or politically incorrect.

Table 1
Respondent experience and demographic characteristics by treatment

| | Treatment 1 (vote in favor/vote against) | Treatment 2 (vote in favor/vote against/unsure) |
|---|---|---|
| *Ever visited or seen a National Forest?* | | |
| Yes | 92% | 95% |
| No | 5% | 3% |
| Unsure | 3% | 2% |
| *Prior to this survey, had you heard about noxious weeds?* | | |
| Yes | 52% | 56% |
| No | 48% | 44% |
| *In the last year have you contributed money to an environmental organization?* | | |
| Yes | 25% | 22% |
| No | 75% | 78% |
| *Demographic measures* | | |
| Percent female | 39% | 43% |
| Mean age | 48 years | 46 years |
| Mean years in CO | 24 years | 24 years |
| Education: | | |
| Eighth or less | 0% | 0% |
| Some high school | 4% | 3% |
| High school graduate | 11% | 10% |
| Some college or technical school | 31% | 27% |
| Technical or trade school graduate | 9% | 9% |
| College graduate | 24% | 26% |
| Some graduate work | 7% | 10% |
| Advanced degree | 14% | 14% |
| Household income: | | |
| Less than $10,000 | 9% | 8% |
| $10,000–19,999 | 12% | 14% |
| $20,000–$29,999 | 15% | 15% |
| $30,000–$39,999 | 13% | 12% |
| $40,000–$49,999 | 12% | 13% |
| $50,000–$59,999 | 13% | 11% |
| $60,000–$69,999 | 8% | 5% |
| $70,000–$79,999 | 4% | 5% |
| $80,000–$89,999 | 4% | 6% |
| $90,000–$99,999 | 5% | 4% |
| $100,000–$149,999 | 4% | 5% |
| Over $150,000 | 2% | 1% |

### 4.2. Data quality

In this paper, we consider two aspects of data quality: (i) the number of missing responses, and (ii) the responsiveness to the offer amount. The former is an example of item non-response, in that a respondent fails to provide an answer to the referendum question, but answers most of the other questions and returns the questionnaire to us. The latter is a test of construct validity which checks whether the data are consistent with predictions from economic theory.

Regarding the frequency of missing responses to the payment question, it is sometimes speculated that when a study participant feels conflicted or uncertain about how to answer a dichotomous-choice CV question, he may skip it. If this is the case, explicit inclusion of an unsure response category should reduce item non-response.

We find that in Treatment 1 (vote in favor/vote against) 5% of the study participants did not answer the CV question. In Treatment 2 (vote in favor/vote against/unsure), the item non-response on the CV question is only 2% (Table 2). The difference in item non-response for the two treatments is significant ($\chi^2$=4.125, $p$=0.039).

This result has two important implications. First, given the relatively large fraction of unsure responses, it would appear that explicit inclusion of an unsure response category does not just attract people who would have skipped the question anyway. Second, whether or not including an explicit unsure response category is advantageous, in terms of reducing item non-response and in turn providing more usable observations, depends on how the unsure responses are treated in the statistical modeling of the data. To further elaborate on this latter point, if information from the unsure respondents is used in deriving estimates of mean willingness to pay (as in Wang, 1997), decreasing item non-response is very important. If unsure respondents are excluded from the sample of usable observations, including an unsure response category will inevitably reduce the usable sample size (the loss of observations in our case would be of 25% of the original number of

Table 2
Response to willingness to pay question by treatment

| | Treatment 1 (vote in favor/vote against), $n$=379 | Treatment 2 (vote in favor/vote against/unsure), $n$=345 |
|---|---|---|
| Vote in favor | 72% | 61% |
| Vote against | 23% | 12% |
| Unsure | | 25% |
| No response | 5% | 2% |

respondents). We investigate these issues in Section 5 below.

Regarding the issue of sensitivity of the vote responses to the bid amount, Table 3 shows that the percentage of unsure responses is non-decreasing in the bid, ranging from 16% at $5 to 32% at $75. Since the percentage of no votes is also increasing in the bid amount, this raises the question whether the unsure responses should be interpreted as equivalent to no votes, and should be treated as such in the statistical models of the data. This decision can be contrasted, for example, with the decision to remove the unsure responses from the usable sample, or with the decision to use a statistical model of WTP that treats the unsure responses as distinct from the others.

Table 3
Response to willingness to pay question by offer amount and treatment

| | Treatment 1 (vote in favor/vote against) | Treatment 2 (vote in favor/vote against/unsure) | | |
|---|---|---|---|---|
| | | No recoding | Unsure dropped | Unsure recoded as no |
| $5 | $n$=72 | $n$=65 | $n$=55 | $n$=65 |
| Vote in favor | 89% | 78% | 93% | 78% |
| Vote against | 11% | 6% | 7% | 22% |
| Unsure | | 16% | $\chi^2$=0.537[a]; $p$=0.464 | $\chi^2$=2.76[b]; $p$=0.097 |
| $10 | $n$=74 | $n$=66 | $n$=50 | $n$=66 |
| Vote in favor | 86% | 67% | 88% | 67% |
| Vote against | 14% | 9% | 12% | 33% |
| Unsure | | 24% | $\chi^2$=0.061[a]; $p$=0.805 | $\chi^2$=7.77[b]; $p$=0.005 |
| $25 | $n$=75 | $n$=64 | $n$=49 | $n$=64 |
| Vote in favor | 79% | 70% | 92% | 70% |
| Vote against | 21% | 6% | 8% | 30% |
| Unsure | | 24% | $\chi^2$=3.80[a]; $p$=0.051 | $\chi^2$=1.28[b]; $p$=0.258 |
| $50 | $n$=65 | $n$=73 | $n$=51 | $n$=73 |
| Vote in favor | 74% | 51% | 72% | 51% |
| Vote against | 26% | 19% | 28% | 49% |
| Unsure | | 30% | $\chi^2$=0.025[a]; $p$=0.875 | $\chi^2$=7.80[b]; $p$=0.005 |
| $75 | $n$=73 | $n$=69 | $n$=47 | $n$=69 |
| Vote in favor | 49% | 46% | 68% | 46% |
| Vote against | 51% | 22% | 32% | 54% |
| Unsure | | 32% | $\chi^2$=4.10[a]; $p$=0.043 | $\chi^2$=0.123[b]; $p$=0.726 |
| Overall | $n$=359 | $n$=337 | $n$=252 | $n$=337 |
| Vote in favor | 76% | 62% | 83% | 62% |
| Vote against | 24% | 13% | 17% | 38% |
| Unsure | | 25% | $\chi^2$=4.88[a]; $p$=0.027 | $\chi^2$=14.745[b]; $p$=0.00 |

[a] The $\chi^2$ statistic is based on a comparison of the distribution of response at each offer amount between Treatment 2 with the unsure responses dropped and Treatment 1.

[b] The $\chi^2$ statistic is based on a comparison of the distribution of response at each offer amount between Treatment 2 with the unsure responses coded as vote against responses and Treatment 1.

## 5. Statistical treatment of the unsure responses

In this section we compare statistical models and value estimates to provide insight into the most appropriate way to deal with the unsure responses. There are five options, each corresponding to a different interpretation of the reasons why a respondent chose the unsure response category: (i) remove the unsure responses from the data set, (ii) recode the unsure responses as vote against responses, (iii) recode the unsure responses as vote in favor responses, (iv) recode some unsure responses as vote in favor and some as vote against, and (v) keep the unsure responses and treat them as such in modeling and value estimation. We examine four of these options. We do not investigate recoding some of the unsure responses as vote against and others as vote in favor (i.e., option (iv)) because we did not include a follow-up question of the type used by Haener and Adamowicz (1998) to implement this approach.

### 5.1. Dropping the unsure responses

If individuals truly do not have an opinion about an issue and are not given an opportunity to express their uncertainty, they may vote randomly or use a decision heuristic that reduces the sensitivity of the CV responses to the offer amount (Krosnick et al., 2002). This suggests that we should examine how the percentage of unsure respondents varies with the bid amount, and what is the effect of removing the not sure responses from the Treatment 2 data.

As shown in Table 2, inclusion of the unsure response category reduces the percent of both the vote

Table 4
Mean WTP by treatment

| | Treatment 1 (vote in favor/vote against) | Treatment 2 (vote in favor/vote against/unsure) | | |
|---|---|---|---|---|
| | | Unsure recoded as yes | Unsure dropped | Unsure recoded as no |
| Mean WTP (standard error) | $78.15 (8.23) | $140.61 (30.97) | $103.61 (17.82) | $61.65 (8.71) |
| *t*-statistic for testing difference in mean WTP relative to Treatment 1 | | −1.96 | −1.30 | −1.38 |
| Statistically different from Treatment 1 mean WTP? | | Yes at 5% level | No | No |

in favor and vote against responses relative to the standard dichotomous-choice response format. Carson et al. (1998) report similar findings, and note that, when the would not vote response are excluded, the split between the yes and no in the remainder of the sample is similar to that observed when only two response categories are offered. In this study, we find that dropping the unsure responses from the data provides statistically similar distributions of responses to the willingness to pay question between the two treatments at the $5, $10 and $50 offer amounts (see Table 3). However, at the $25 and $75 offer amounts, we notice that there are significantly more vote in favor responses in Treatment 2 with the unsure responses dropped relative to Treatment 1.

Based on a probit model of the responses to the payment question from Treatment 1 and on a procedure detailed in Cameron and James (1987),[3] we estimate mean WTP for Treatment 1 to be $78.15 (standard error is $8.23) (see Table 4). Both mean willingness to pay and its standard error are larger when we use Treatment 2 data and the unsure responses are dropped. As shown in Table 4, mean willingness to pay is now $103.61, an estimate that is not statistically distinguishable from the $78.15 implied by the Treatment 1 data.

We conclude that exclusion of the unsure responses does not improve the value estimates in the sense of providing an estimate similar to the dichotomous-choice estimate or providing an estimate of mean willingness to pay with a smaller standard error. Furthermore, when we examine possible reasons why individuals choose the unsure response category, as we do in the next section, we find that unsure respondents have different attitudes toward National Forests and concerns about noxious weeds. This suggests that consideration should be given to statistical models where the unsure responses are not removed from the sample.

### 5.2. Recoding unsure responses to vote against

For both versions of the survey, the percentage of vote in favor responses to the willingness to pay question is highest at the lowest offer amount, declining as the offer amount increases (Table 3). In Treatment 2, the percentage of unsure and vote against responses increase with the offer amount, reaching 32% and 22%, respectively, at the top offer amount of $75. This result might suggest that the unsure responses could be interpreted and reclassified as if they were vote against responses. However, as previously mentioned, when we do so the split between vote in favor and vote against at the various offer amount levels reproduces the distribution of responses from the standard dichotomous-choice in Treatment 1 for only three of the five offer amounts ($5, $25 and $75).

In spite of these differences across the two distributions, the estimates of mean willingness to pay are not significantly different across the treatment groups. As shown in Table 4, mean willingness to pay is $78.15 based on the dichotomous-choice Treatment

[3] The mean willingness to pay is computed using a fully parametric approach. Specifically, we fit a probit model where the dependent variable is a dummy indicator that takes on a value of one if the respondent voted in favor of the Noxious Weeds Control Program at the stated offer amount, and zero otherwise. The right-hand side of the model includes the intercept and the offer amount. This procedure assumes that the latent WTP variable is normally distributed, and recovers mean/median WTP as minus the intercept, divided by the coefficient on the bid (Cameron and James, 1987). The standard errors are calculated from the covariance matrix of the probit estimates using the delta method (see Cameron, 1991).

1 group data, and $61.65 when the unsure responses from Treatment 2 are recoded as votes against, but these estimates are not statistically different at the 5% significance level.

To explore whether it is appropriate to recode the unsure responses as votes against the program, we use a multinomial logit (MNL) model which predicts the likelihood of selecting each of the three possible response categories as a function of respondent characteristics, acceptance of the scenario and environmental priorities to explore how these variables relate to the response to the WTP question.[4] If the coefficients on the regressors are similar across the vote against or unsure responses to the willingness to pay question, this can be viewed as support for the approach of recoding the unsure responses as vote against.

Our MNL model includes the bid amount, individual characteristics such as income and education, and various indicators of their attitudes towards noxious weeds, National Forests, and the weeds program. Descriptive statistics for these variables are reported in Table 5.

Estimation results from the MNL model are reported in Table 5. The model shows clearly that the offer amount is one of the strongest determinants of the unsure and vote against responses. The positive coefficients on this variable indicate that, all else the same, as the offer amount increases the likelihood of selecting the unsure and vote against response categories, instead of a vote in favor, increase.

Even more important, the coefficients of the offer amount are virtually the same for the unsure and vote against response options: the appropriate Wald statistic is 1.42, failing to reject the null hypothesis of no difference at all conventional significance levels.[5] This result is very similar to that previously obtained by Carson et al. (1998).

Similar results–in the sense that the coefficients associated with the unsure response are statistically indistinguishable from the corresponding coefficients for the vote against response–are seen with DEFKNOW, DEFSIDE, HARMIMP and DONAT. The coefficients of all of these variables are negative and significant, implying that persons who state they know their future income (DEFKNOW=1), would like to know more about the potential side effects of weed control techniques (DEFSIDE=1), are more seriously concerned about the harm caused by noxious weeds to wildlife (HARMIMP=1) and contribute money to environmental organizations (DONAT=1) are less likely to respond unsure or vote against the CV referendum than vote in favor.

The MNL model also indicates that respondents with higher incomes are less likely to select the unsure option, and that dissatisfaction with the available information about how the program would be funded (DEFPROG) leads people to vote against the Noxious Weeds Control Program, or to say that they are unsure.

In sum, the results of the MNL analysis suggest that the variables related to responding either unsure or vote against to the willingness to pay question are very similar. We conclude that while conditional analyses (i.e., the MNL model) support the recoding of unsure responses to vote against, the unconditional analyses based on the percentage of respondents in favor and against the program (after unsure responses in Treatment 2 are recoded conservatively as votes against the plan) do not support this approach. Therefore, we do not find strong evidence to support the recommendation that all unsure responses be recoded as vote against responses.

[4] The MNL model assumes that each response is associated with a level of utility $V_{ij}=\mathbf{x}_i\beta_j+\epsilon_{ij}$, where $V$ is indirect utility, $\mathbf{x}$ is a vector of individual characteristics or attitudes, $\beta$ is a vector of alternative-specific coefficients, and $\epsilon$ is a vector of i.i.d. error terms that follow the type I extreme value distribution. The subscripts $i$ and $j$ denote the individual and the response category, respectively. It can be shown that the probability that response $k$ is selected by respondent i is $Pr(k) = \exp(\mathbf{x}_i\mathbf{b}_k)/\sum_{j=1}^{3}\exp(\mathbf{x}_i\mathbf{b}_j)$. This model allows one to identify what kind of individuals are more likely to select each of the possible response category, but has the disadvantage that it is not possible to recover estimates of mean willingness to pay. The MNL was one of the tools that led Carson et al. to conclude that persons who declined to vote in one of their two versions of the Alaska oil spill contingent valuation survey should be interpreted as having meant a vote against the proposed program.

[5] The statistic is distributed as a chi square with one degree of freedom under the null hypothesis of no difference and for large sample size. The 5% critical level is 3.84.

Table 5
Multinomial logit model for Treatment 2 (omitted category is vote in favor), $n$=339

| Variable | Mean (standard deviation) | Against the program | Unsure | LR test of the null that the against and unsure coefficients are both equal to zero |
|---|---|---|---|---|
| | | Coefficient ($t$-statistic) | Coefficient ($t$-statistic) | |
| Constant | | 1.244 (0.94) | 2.892 (2.80) | 7.864 (0.020) |
| Offer amount ($5, $10, $25, $50, $75) | 33.761 (22.229) | 0.033 (4.23) | 0.024 (3.87) | 23.373 (<0.0001) |
| Household income (thousand dollars) | 46.253 (33.471) | 0.008 (1.28) | −0.014 (−2.35) | 10.119 (0.006) |
| DEFKNOW (1=know future income; 0=otherwise) | 0.770 (0.421) | −1.022 (−2.16) | −0.963 (−2.67) | 8.561 (0.014) |
| DEFWEED (1=need more information about the problems caused by weeds; 0=otherwise) | 0.519 (0.500) | −0.138 (−0.29) | 1.196 (3.03) | 11.265 (0.004) |
| DEFSIDE (1=need more information about the side effects of weed control techniques; 0=otherwise) | 0.832 (0.374) | −3.428 (−5.14) | −2.030 (−3.42) | 28.275 (<0.0001) |
| DEFPROG (1=need more information about how the program would be funded; 0=otherwise) | 0.723 (0.448) | 1.845 (2.96) | 0.734 (1.53) | 9.180 (0.010) |
| DONAT (1=contribute to environmental organizations; 0=otherwise) | 0.224 (0.417) | −0.960 (−1.65) | −0.978 (−2.13) | 5.970 (0.051) |
| SOILIMP (1=extremely concerned about the soil stability impacts of noxious weeds; 0=otherwise) | 0.448 (0.497) | 0.244 (0.40) | −0.352 (−0.74) | 0.918 (0.632) |
| WATERIMP (1=extremely concerned about the water quality impacts of noxious weeds; 0=otherwise) | 0.640 (0.480) | 0.003 (0.01) | 0.233 (0.48) | 0.268 (0.875) |
| HARMIMP (1=extremely concerned that noxious weeds will harm wildlife habitat; 0=otherwise) | 0.667 (0.471) | −1.545 (−2.58) | −2.032 (−4.02) | 17.586 (<0.0001) |
| HABPLANT (1=strongly agree that National Forests provide habitat for plants; 0=otherwise) | 0.711 (0.453) | −0.200 (−0.40) | −0.702 (−1.86) | 3.5213 (0.172) |
| HABLIFE (1=strongly agree that National Forests provide habitat for fish and wildlife; 0=otherwise) | 0.841 (0.366) | 0.362 (0.57) | −0.357 (−0.77) | 1.5179 (0.468) |
| LIVECO (years lived in Colorado) | 26.077 (20.329) | −0.026 (−2.31) | −0.008 (−1.03) | 5.5162 (0.063) |
| SEEN1 (respondent has seen the noxious weeds) | 0.835 (0.371) | 0.738 (1.12) | −0.189 (−0.47) | 1.965 (0.374) |
| EDUC (years of schooling) | 16.177 (3.005) | −0.068 (−0.93) | −0.025 (−0.45) | 0.884 (0.643) |

### 5.3. Recoding unsure responses as vote in favor

Recoding the unsure responses in the Treatment 2 subsample to votes in favor of the program results in an estimate of mean willingness to pay of about $140 (Table 4). This is (marginally) significantly higher than the estimate of $78 from Treatment 1.

Is it reasonable, however, to interpret the unsure responses as votes in favor of the program? The MNL model of Table 5 suggests otherwise, since it finds that the unsure and favorable response options relate to the independent variables in the models in different ways. To further explore this issue, in Table 6 we compare responses of unsure respondents to vote in favor respondents for a series of Likert scale items developed to measure the reasons for the response to the willingness to pay question.

Table 6 reveals an interesting pattern. On four of the five items, contingency table analysis suggests that the distributions of response for the two groups (vote in favor and unsure respondents) are statistically different. The consistent pattern is that more vote in favor respondents chose the extreme point on the scale that corresponds to definitely true. Relative to the unsure respondents, more vote in favor respondents said it was definitely true that the program was worth the stated amount, that they wanted to show their support for the environment in general, that the goals of the program were an important consideration when deciding how to vote, and that preserving the health of National Forest was very important to them. These results seem intuitive and suggest that unsure respondents are distinct from respondents who vote in favor. We conclude that there is no evidence that in the absence of an explicit unsure response category, the unsure respondents would vote in favor or that the

Table 6
Comparing unsure respondents to vote in favor respondents

| How true is each statement? | | Definitely true | Somewhat true | Somewhat false | Definitely false |
|---|---|---|---|---|---|
| I felt the Noxious Weeds Control Program would be worth the amount I was asked to pay. ($\chi^2$=110.85; $p$=0.000) | Unsure | 10% | 64% | 23% | 3% |
| | Vote in favor | 75% | 23% | 1% | 1% |
| I would vote for the program to show my general support for the environment ($\chi^2$=119.25; $p$=0.000) | Unsure | 10% | 69% | 16% | 5% |
| | Vote in favor | 80% | 17% | 1% | 1% |
| The goals of the Noxious Weeds Control Program were an important consideration when deciding how to vote ($\chi^2$=37.01; $p$=0.000) | Unsure | 27% | 60% | 12% | 1% |
| | Vote in favor | 65% | 32% | 2% | 1% |
| The use of herbicides was an important factor when deciding how to vote ($\chi^2$=2.078; $p$=0.556) | Unsure | 32% | 48% | 14% | 5% |
| | Vote in favor | 31% | 42% | 18% | 9% |
| Preserving the health of the National Forests is very important to me ($\chi^2$=14.681; $p$=0.002) | Unsure | 74% | 23% | 3% | 0% |
| | Vote in favor | 90% | 9% | 0% | 1% |

unsure respondents are similar to the vote in favor respondents.

### 5.4. Including the unsure responses

The model proposed by Wang (1997) allows for retaining the information provided by the unsure respondents. In this model, the unsure responses are treated as distinct from both the votes in favor and vote against responses. The model assumes that individuals have distributions of willingness to pay rather than single values of willingness to pay. Wang argues unsure responses are informative about the underlying distribution of willingness to pay because they signal that the respondent's mean willingness to pay is very close to the offer amount. This model assumes that respondents vote in favor of the program if their mean willingness to pay amount is sufficiently greater than the offer amount—in fact, if it exceeds the offer amount by more than a certain threshold amount, which we denote as $t_1$.

The model is completed by assuming that people that are against the program have a mean willingness to pay that is sufficiently smaller than the offer amount. These persons will vote against only if their willingness to pay is less than the offer amount, minus an appropriate threshold. For identification purposes, this latter threshold, denoted as $t_2$, is assumed equal to the threshold $t_1$. In other words, the unsure band is symmetric around the bid amount. Finally, if willingness to pay lies between (bid$-t_2$) and (bid$+t_1$), the respondent chooses the unsure response category.

The contributions to the resulting likelihood function are thus:

$$\begin{aligned} Pr(\text{yes} \mid B_i, \mathbf{x}_i) &= Pr(\text{WTP}_i > B_i + t_1) \\ &= Pr(x_i\beta + \varepsilon_i > B_i + t_1) \\ &= Pr(\varepsilon_i/\sigma > -x_i\beta/\sigma + B_i/\sigma + t_i/\sigma), \end{aligned} \tag{1}$$

$$\begin{aligned} Pr(\text{no} \mid B_i, \mathbf{x}_i) &= Pr(\text{WTP}_i < B_i - t_1) \\ &= Pr(\varepsilon_i/\sigma < -x_i\beta/\sigma + B_i/\sigma - t_1/\sigma) \end{aligned} \tag{2}$$

and

$$\begin{aligned} Pr(\text{not sure} \mid B_i, \mathbf{x}_i) &= Pr(B_i - t_1 < \text{WTP}_i < B_i + t_1) \\ &= Pr(\varepsilon_i/\sigma < -x_i\beta/\sigma + B_i/\sigma + t_i/\sigma \\ &\quad - (Pr(\varepsilon_i/\sigma < -x_i\beta/\sigma + B_i/\sigma \\ &\quad - t_1/\sigma). \end{aligned} \tag{3}$$

If one assumes that willingness to pay follows the normal distribution, the three contributions become:

$$Pr(\text{yes} \mid b_i, \mathbf{x}_i) = \Phi(x_i\beta/\sigma - B_i/\sigma - t_1/\sigma), \tag{4}$$

$$Pr(\text{no} \mid B_i, \mathbf{x}_i) = \Phi(-x_i\beta/\sigma + B_i/\sigma - t_1/\sigma) \tag{5}$$

and

$$\begin{aligned} &Pr(\text{not sure} \mid B_i, \mathbf{x}_i) \\ &= \Phi(-x_i\beta/\sigma + B_i/\sigma + t_1/\sigma) \\ &\quad - \Phi(-x_i\beta/\sigma + B_i/\sigma - t_1/\sigma), \end{aligned} \tag{6}$$

where $\Phi$ denotes the standard normal cdf.

The results of the Wang model for normal willingness to pay are reported in Table 7. In this specification, the thresholds are allowed to vary with respondent characteristics and with variables capturing acceptance of the scenario:

$$t_1 = \mathbf{z}_i\delta. \quad (7)$$

The $\delta$ coefficients are identified only if the variables that enter in the determination of the threshold (the $\mathbf{z}_i$'s in Eq. (7)) do not overlap with variables than enter in the expression for mean willingness to pay (the $\mathbf{x}_i$'s).

The results of the Wang model in Table 7 make intuitive sense and confirm some of the insights learned from the MNL model. Mean willingness to pay increases significantly with respondent confidence about his or her future income (by $34), with respondent need for more information about the side effects of weed control (by $97; presumably, this signals seriousness about undertaking the program), and is typically greater among persons who contribute to environmental organizations (by about $33). Concern over wildlife impacts of uncontrolled noxious weeds also tends to increase willingness to pay (by about $55). By contrast, skepticism about the funding of the noxious weeds program reduces willingness to pay by about $49.

The unsure region, i.e., the band around the offer amount within which the respondent is unable to provide a firm vote in favor or vote against response, is made considerably tighter (by about $12) by each year of formal education and by personal experience with the species of weeds. It should be noted, however, that the former effect is only statistically significant at the 10% level, and that the effect of the latter (about $5) is not very pronounced. We also find that males seem to have somewhat tighter uncertainty regions, but this effect is not statistically significant. Finally, we also experimented with including the respondent's voting experience among the determinants of the threshold variables. However, previous voting experience in national or local elections does not have any explanatory power for the thresholds.

The estimated mean willingness to pay based on the Wang model $102.36 and the standard error is $13.94. The asymptotic $t$-test to compare the mean based on the Wang model to the estimate of mean willingness to pay based on the standard dichotomous-choice data and Treatment 1 data (which is equal to $78.15) results in a statistic of −1.4955. The

Table 7
Wang model for treatment 2 ($n$=339)

| Variable | Coefficient ($t$-statistic) |
|---|---|
| Constant | −15.941 (−0.760) |
| Household income (thousand dollars) | −0.020 (−0.074) |
| DEFKNOW (1=know future income; 0=otherwise) | 34.466 (2.369) |
| DEFWEED (1=need more information about the problems caused by weeds; 0=otherwise) | −10.738 (−0.864) |
| DEFSIDE (1=need more information about the side effects of weed control techniques; 0=otherwise) | 96.700 (3.880) |
| DEFPROG (1=need more information about how the program would be funded; 0=otherwise) | −49.891 (−2.684) |
| DONAT (1=contribute to environmental organizations; 0=otherwise) | 32.672 (1.977) |
| SOILIMP (1=extremely concerned about the soil stability impacts of noxious weeds; 0=otherwise) | 5.524 (0.340) |
| WATERIMP (1=extremely concerned about the water quality impacts of noxious weeds;0=otherwise) | −4.243 (−0.263) |
| HARMIMP (1=extremely concerned that noxious weeds will harm wildlife habitat; 0=otherwise) | 54.624 (2.855) |
| PLANTIMP (1=extremely concerned about the effects of noxious weeds on native plants; 0=otherwise) | 12.709 (0.799) |
| HABPLANT (1=strongly agrees that National Forests provide habitat for plants; 0=otherwise) | 13.643 (0.997) |
| HABLIFE (1=strongly agrees that National Forests provide habitat for fish and wildlife; 0=otherwise) | −6.321 (−0.400) |
| Standard deviation of WTP ($\sigma$) | 72.487 (4.870) |
| Coefficients | |
| Constant | 81.022 (3.732) |
| Education | −12.591 (−1.775) |
| MALE (1=male; 0=female) | −16.181 (−1.428) |
| SEEN1 (respondent has seen the noxious weeds) | −4.281 (−1.967) |
| Mean WTP (standard error) | $102.36 (13.94) |

difference between the two means is not statistically different.

Is fitting the Wang model, then, justified by the fact that the unsure responses are different from the other response categories? In an effort to more formally examine the choice of the unsure response category to the CV question and relationships to other variables collected in the survey, we conducted a correlation analysis, which we report in Table 8.

Table 8 suggests that the unsure respondents did not find the Noxious Weeds Control Program to be as important as individuals who voted in favor. The unsure respondents were also less likely to think that the program could be funded without the additional taxes and objected less to paying additional taxes compared to the vote against respondents. Table 8 also confirms that as the offer amount goes up, there are more unsure responses.

Many variables were found to be negatively correlated with unsure responses. Not surprisingly, unsure responses were correlated with needing more information about the problems weeds cause, the effectiveness of control methods, and how the program would be funded. Lower levels of education and income were also associated with unsure responses. Unsure responses were also found to be correlated with not thinking that habitat of plants, fish and wildlife were important reasons for having National Forests. Unsure responses were also correlated with not agreeing that the environment should be protected even if people have to go without some products and not being willing to make personal sacrifices to protect the environment.

Based on this correlation analysis, along with the other analyses described in this article, it would seem that unsure respondents who choose the unsure response category are not completely convinced that the Noxious Weeds Control Program is a worthwhile program. They are also individuals who might be less "environmentally" focused. Finally, there appears

Table 8
Correlations between providing an unsure response to the CV question and other survey variables (1=unsure; 0=vote in favor or vote against)

| | Pearson's correlation statistic ($p$-value) |
|---|---|
| *Variables positively correlated with providing an unsure response* | |
| I felt the Noxious Weeds Control Program would be worth the amount I was asked to pay (1=def. true, 4=def. false)[a] | 0.577 (0.000) |
| I would vote for the program to show my general support for the environment (1=def. true, 4=def. false)[a] | 0.547 (0.000) |
| The goals of the Noxious Weed Control Program were an important consideration when deciding how to vote (1=def. true, 4=def. false)[a] | 0.314 (0.000) |
| Preserving the health of National Forests is very important to me (1=def. true, 4=def. false)[a] | 0.134 (0.022) |
| I think that the project can be funded without additional taxes (1=def. true, 4=def. false)[b] | 282 (0.001) |
| I object to paying more taxes than I already do, regardless of the reason (1=def. true, 4=def. false)[b] | 0.197 (0.025) |
| Economic progress is more important than environmental concerns (1=strongly disagree, 5=strongly agree) | 0.126 (0.022) |
| Offer amount | 0.121 (0.026) |
| *Variables negatively correlated with providing an unsure response* | |
| I need more information about problems that weeds cause (1=def. true, 4=def. false) | −0.221 (0.000) |
| I need more information about the effectiveness of the methods that would be used to control weeds (1=def. true, 4=def. false) | −0.137 (0.014) |
| I need more information about how the program would be funded (1=def. true, 4=def. false) | −0.138 (0.014) |
| Highest year of schooling | −0.123 (0.025) |
| Household income | −0.184 (0.001) |
| Importance of habitat for plants as a reason for having National Forests (1=not at all important, 4=extremely important) | −0.179 (0.001) |
| Importance of habitat for fish and wildlife as a reason for having National Forests (1=not at all important, 4=extremely important) | −0.199 (0.000) |
| The environment should be protected even if people have to go without some products (1=strongly disagree, 5=strongly agree) | −0.122 (0.027) |
| I am willing to make personal sacrifices to protect the environment (1=strongly disagree, 5=strongly agree) | −0.136 (0.013) |

[a] Only the vote in favor and unsure respondents responded to these statements.
[b] Only the vote against and the unsure respondents responded to these statements.

to be a financial aspect as well, in that there are more unsure respondents at the higher offer amounts and respondents with lower incomes were more likely to provide an unsure response to the CV question.

## 6. Discussion and conclusions

The NOAA Panel recommended inclusion of a would not vote response category in addition to the vote in favor and vote against response categories. The Panel also recommended additional research in alternative ways of presenting and interpreting the would not vote option. This recommendation has spurred several studies that have all come to different recommendations on how to handle these responses.

Our research adds to this previous body of literature by looking at a wider variety of options for dealing with the unsure responses. We administered two variants of our CV survey questionnaire to two independent samples. The two variants used the referendum format and were identical in all respects, except for the fact that the version administered to Treatment 2 group allowed three response options (vote in favor, vote against, and unsure), while the version assigned to Treatment 1 respondents employed a dichotomous-choice question (vote in favor, vote against).

Confirming previous studies, we found that a substantial number of survey respondents (25%) chose the unsure response. We employ multinomial logit models and debriefing questions to investigate the reasons why people choose to the unsure response category. While multinomial logit models suggest that there is a considerable degree of similarity between not sure and against responses, the analysis of the debriefing questions suggests otherwise.

While it is possible that some of these respondents did not want to make the cognitive effort to examine their preferences and come up with a definite response, it appears as though there are legitimate explanations for the choice of the unsure response, that they were distinct from both the vote in favor and the vote against responses, and that the causes of uncertainty are complex and likely vary among respondents.

The latter implication is confirmed when we recode the unsure responses as votes in favor or against the program. When we do so, we reproduce the distribution of the responses from the standard dichotomous-choice Treatment 1 group only at three of the five bid amounts.

This result suggests to us that approaches such as the Wang model which allow for interpretation of responses at face value may be one of the most promising approaches. This confirms earlier results by Alberini et al. (2003), although it should be emphasized that our study seeks to identify determinants of uncertain response options to a referendum question, and does not examine the effect of polychotomous choice response options (a broader variant than the vote in favor/vote against/unsure response format), combined with multiple bid amounts.

We conclude by emphasizing that more research is needed on this topic. In addition to statistical modeling issues, we would recommend that attention be devoted to developing appropriate debriefing questions that could be used to sort out issues of uncertainty.

## References

Alberini, A., Boyle, K., 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.

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Shuman, H., 1993. Report of the NOAA panel on contingent valuation. Federal Register 58, 4601–4614.

Cameron, T.A., 1991. Interval estimates of non-market resource values from referendum contingent valuation surveys. Land Economics 67 (4), 413–421.

Cameron, T.A., James, M.D., 1987. Efficient estimation methods for "closed-ended" contingent valuation surveys. The Review of Economics and Statistics 69, 269–276.

Carson, R.T., Hanemann, W.M., Mitchell, R.C., 1986. The use of simulated political markets to value public goods. Unpublished manuscript. Economics Department, University of California, San Diego.

Carson, R.T., Hanemann, W.M., Kopp, R.J., Krosnick, J.A., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K., Conaway, M., Martin, K., 1998. Referendum design and contingent valuation: the NOAA panel's no-vote recommendation. Review of Economics and Statistics 80 (3), 484–487.

Champ, P.A., Brown, T.C., 1997. A comparison of contingent and actual voting behavior. Proceedings from W-133 Benefits and Cost Transfer in Natural Resource Planning, 10th Interim Report, pp. 77–98.

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

Haener, M.K., Adamowicz, W., 1998. Analysis of don't know responses to referendum contingent valuation questions. Agricultural and Resource Economics Review 27 (2), 218–230.

Krosnick, J.A., Holbrook, A.L., Berent, M.K., Carson, R.T., Hanemann, W.M., Kopp, R.J., Mitchell, R.C., Presser, S., Ruud, P.A., Smith, V.K., Moody, W.R., Green, M.C., Conaway, M., 2002. The impact of "no opinion" response options on data quality. Public Opinion Quarterly 66, 371–403.

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

Vossler, C.A., Kerkvliet, J., Polasky, S., Gainutdinova, O., 2003. Externally validating contingent valuation: an open-space survey and referendum in Corvallis, Oregon. Journal of Economic Behavior and Organization 51, 177–261.

Wang, H., 1997. Treatment of "don't-know" responses in contingent valuation surveys: a random valuation model. Journal of Environmental Economics and Management 32, 219–232.