
6 Testing the robustness of contingent valuation estimates of WTP to survey mode and treatment of protest responses

John Loomis, Armando González-Cabán and Joseph Champ¹

INTRODUCTION

Over the past four decades the contingent valuation method (CVM) has become a technique frequently used by economists to estimate willingness-to-pay (WTP) for improvements in environmental quality and protection of natural resources. The CVM was originally applied to estimate recreation use values (Davis, 1963; Hammack and Brown, 1974) and air quality (Brookshire et al., 1982; Randall et al., 1974). In the second decade the CVM was extended to valuing the general public's option and existence values of environmental improvements (Walsh et al., 1984). As part of this evolution, the overall design of CVM studies now attempts to construct a market for the public good (Carson, 1991). As such, a typical CVM survey describes the public good to be valued, how the good will be paid for (that is, payment vehicle), the WTP question format (that is, open ended or closed ended) and a rule for deciding whether the good will be supplied (for example, majority rule in a referendum or total benefits exceed total cost). Each of these design elements has received substantial testing to determine whether the CVM WTP estimates are sensitive to the payment vehicle or WTP question format (Boyle, 2003). The key concern of CVM regarding the validity of the WTP estimates has also received extensive testing from the early days of CVM (Bishop and Heberlein, 1979) to the more recent (Murphy et al., 2005).

Survey Mode

The CVM relies more heavily on survey research than many other economic valuation techniques such as the hedonic property method. Thus, it is important to understand the sensitivity of CVM WTP estimates to various aspects of survey design such as mode (for example, phone, mail, in person). One of the conditions needed for mainstream economists and

policy-makers to take the CVM estimates of WTP seriously is that they should be robust for different survey modes. The most commonly used survey mode is mail, owing to it being relatively inexpensive and allowing presentation of graphics and photos. Phone surveys are used when visual aids are not critical. In-person interviews are considered by some as the 'gold' standard (Arrow et al., 1993; Mitchell and Carson, 1989). Their expense has resulted in such interviews primarily being used in large natural resource damage cases (Carson et al., 2003) or other high-profile public policy studies. Internet surveys have slowly been gaining ground as an approach for conducting CVM surveys. With commercial companies providing consumer panels with representative demographics and hosting the surveys, CVM surveys using the Internet, especially for large nationwide surveys, have become more common (Banzhaf et al., 2006; Berrens et al., 2004). Basic research is now being conducted to compare WTP responses of Internet surveys with more traditional surveys such as mail (Banzhaf et al., 2006).

Because no one survey mode is ideal for CVM, mixed mode surveys emerged in the late 1980's as a means of utilizing the strengths of each method (Loomis and King, 1994). Hanemann et al. (1991) conducted one of the first mixed mode CVM surveys. They used random digit dialling to initiate contact with a random sample of respondents. Then a survey booklet with visual aids and questions were sent to respondents prior to the phone interviews. Finally, the phone interviews were conducted by leading the respondent through the booklet and obtaining their answers. This was argued to combine the best features of the phone survey such as its use of a random digit dialling sample frame, a live interviewer to motivate respondents to complete each question and answer questions for the respondent, along with the advantages of mail, that is, providing the respondent a booklet with visual aids. While more expensive than mail surveys alone, the approach was shown to yield lower item non-response than mail surveys (Loomis and King, 1994). In this chapter this approach will be referred to as a phone-mail booklet-phone interview approach.

This chapter reports on a comparison of this phone-mail booklet-phone interview approach with a new video survey mode. Both of these approaches attempt to mimic the preferred but expensive survey administration format, in-person interviews (Arrow et al., 1993; Mitchell and Carson, 1989). An alternative to the phone-mail booklet-phone approach and to Internet surveys is to send the respondent a videotape or DVD that utilizes an on-camera interviewer who explains the elements of the constructed market and asks the survey questions. The respondent records their answers on a sheet that is mailed back. This videotape method mimics an in-person interview, may appeal to a 'video' generation and can

be inexpensively applied to large samples. In contrast, large samples in the phone-mail booklet-phone can quickly become expensive due to repeated phone contacts and in-depth (10–15 minute) phone interviews.

Protest Responses

In a constructed market, economists rely on visitors' or households' intended behaviour as manifested in their statements of value towards the natural resource of interest. In order for a survey to have face validity, these respondents must be valuing the same public good as intended by the researcher. Thus, the design elements of the constructed market (for example, description of the good, payment vehicle) should be credible or realistic enough to facilitate rather than interfere with eliciting a household's true WTP. However, some respondents object to the particular method of paying for the public good (for example, rejecting the fairness of the payment vehicle – Morrison et al., 2000), the method of provision (for example, a particular government agency) or doubt whether the proposed public programme will solve the environmental problem described in the survey. In these cases, their WTP responses reflect a rejection of one or more premise of the constructed market, rather than their value of the public good. These responses have been labelled protest responses (Boyle, 2003; Mitchell and Carson, 1989).

If the WTP question format is open ended (what is the maximum you would pay?) or payment card, then some respondent's statements of zero WTP may reflect protest responses rather than lack of value for the public good or inability to pay for the public good. In the closed-ended dichotomous choice or voter referendum format, where respondents are asked whether they would pay a specific monetary amount for the public good, detecting protest responses is more difficult (Boyle, 2003). If a respondent is asked whether they would pay a particularly high monetary amount (US\$450), they may respond 'No', but still have a positive value for the good that is less than \$450. However, the respondent's 'No' response may also be a protest response, in which case they are rejecting some element of the constructed market rather than indicating no value for the public good. Much of the past research has evaluated how to detect or deal with zero WTP responses, but not identification or handling of protest responses in dichotomous choice CVM (Collins and Rosenberger 2007).

In this chapter we compare two different protocols for detecting and handling protest responses in voter referendum CVM WTP responses. The first method relies on written statements of respondents as to why they would not pay. The second approach utilizes these written statements in

combination with a follow-up WTP question asking whether the respondent would pay \$1 for the programme. This allows us to better identify those who completely reject the CVM constructed market and those who have some positive value for public program, but just less than the bid amount they are asked to pay. We then test whether the WTP amounts from these two different approaches to excluding protest responses are statistically different from one another, and different from the WTP estimate from including all responses, whether protest or not. These tests allow us to investigate whether the WTP estimates are robust for the treatment of protest responses.

The empirical application is to the use of prescribed burning to reduce wildland forest fires. This problem is a policy relevant issue throughout the world, as wildfires are becoming increasing large and destructive in countries such as the USA, Australia, Spain and Greece, to name a few.

DESCRIPTION OF PHONE-MAIL BOOKLET-PHONE INTERVIEW AND VIDEO SURVEY MODES

There are many features of a CVM survey that can potentially influence the validity of respondents' WTP (Barro et al., 1996). An important area that has seen less investigation is that of the convergent validity of traditional and new survey administration modes (Loomis and King, 1994). While in-person surveys are preferred, the most frequently used survey administration form is mail surveys. However, mail surveys require significant reading ability on the part of the respondent. Arrow et al. (1993) suggests that phone interviews may be able to mimic some of the strong features of in-person interviews. However, pure phone surveys are limited due to their inability to provide the respondent with visual aids. A combination of phone recruitment-mailed survey booklet-phone interview with the aid of the survey booklet (hereafter called phone-mail booklet-phone) has proved an effective combination in a number of CVM surveys (Hanemann et al., 1991; Loomis et al., 2002). The respondent has the questions and visual aids in front of them while the phone interviewer asks the questions. However, these phone-mail booklet-phone interviews can also be quite expensive, ranging from \$50 to \$100 per completed interview owing to the multiple contacts required and the cost of the in-depth interviews. Internet surveys may be the way of the future, but inconsistent web browsers and incomplete coverage of low-income households make them less than ideal at present, despite the efforts of Internet panel companies.

However, according to the US Census Bureau, in 1999 approximately 85 per cent of American households with televisions also had video

cassette recorders (VCRs). We suspect with the introduction of low-cost digital versatile disk (DVD) players that nearly every household in America with a television has either a VCR or DVD player, or both. The technology of DVDs and VCRs is easy to use, and the medium offers many of the advantages of the Internet, but few of the drawbacks. In addition, using a videotape with an on-camera narrator and visual aids has the potential to mimic an in-person interview, at a fraction of the cost. Once produced, videotape and especially DVD distribution is relatively inexpensive. To our knowledge, no one has yet taken advantage of even the basic features that videotapes offer for conducting a contingent valuation method survey. We hypothesize that the videotape/DVD medium is as effective as the phone-mail booklet-phone interview method at obtaining an adequate response rate, reducing protest refusals to pay and estimating willingness-to-pay (WTP). Thus, the objective of this study is to evaluate the convergent validity of a new video approach with the more traditional phone-mail booklet-phone survey method with respect to response rates, protest refusals to pay versus non-protest refusals to pay responses, and WTP estimates. We emphasize this is a test of convergent, not criterion validity with respect to survey mode, since phone surveys are not perfect in sample coverage and have their own biases.

SCREENING FOR PROTESTS IN THE DICHOTOMOUS CHOICE WTP QUESTION FORMAT

To investigate whether zero bids or 'No' votes reflect a true expression of respondent values or rejection of some feature of the constructed market, follow-up questions to the WTP question are often asked, particularly of those stating zero WTP or voting against the program. These follow-up questions can be open-ended like 'Why did you answer zero?' or 'Why did you vote against?' These responses are then postcoded into two categories, protests and non-protests following criteria described below. Alternatively, some surveys (Loomis, 1987; Walsh et al., 1984) use a checklist approach which prompts the user to select the main reason or all the reasons why they would not pay. These reasons include what are typically considered non-protest such as 'I do not receive any benefits from the public program' or 'I cannot afford to pay for the public program'. It is in fact heartening to obtain these responses, as it indicates respondents took the obligation for payment seriously even though they were not required to actually pay.

Typically protest responses include one of several factors. As summarized by Jorgensen and Syme (2000), these include (a) it is unfair to be

asked to pay additional money for this public good; (b) existing taxes or funding should be used; (c) government wastes money; (d) this public good is a right that I should not have to pay for; (e) money collected may not actually be used to provide the good; (f) taxes already too high; (g) only users of the good should pay. In our traditional postcoding of respondent's open-ended reasons for not paying (described in more detail below), we were guided by these categories in determining what was a protest response.

HYPOTHESES REGARDING SURVEY MODE RESPONSE RATES AND PROTEST RESPONSES

Test of Response Rate Differences by Survey Mode

There are two contacts in both the phone-mail booklet-phone approach and the videotape/DVD approach. In the phone-mail booklet-phone approach there is an initial random digit dialling phone call with a short initial interview. The address to mail a survey booklet is verified, and a time is scheduled for an in-depth (20-minute) interview. The videotape/DVD survey proceeds in a similar fashion with an initial contact, except a videotape is mailed and the respondent mails back the answer booklet. Thus, the first test of convergent validity is whether households respond equally to the initial phone call, and then whether they follow through on answering the in-depth CVM questionnaire with the two survey administration modes.

$$H_0: \text{RESPONSE}_{\text{video}} = \text{RESPONSE}_{\text{phone-booklet}} \quad (6.1)$$

This will be tested using separate contingency tables and χ^2 tests for both the first and second interviews.

Test of Differences in WTP by Survey Mode

Comparisons of mean WTP estimates across survey administration mode will be tested by equality of the mean WTP estimates from each survey mode. The null hypothesis is:

$$H_0: \text{WTP}_{\text{video}} = \text{WTP}_{\text{phone-booklet}} \quad (6.2)$$

The results are determined by whether the confidence intervals overlap or not.

Test of Different Protest Rates by Survey Mode

The response to the WTP questions elicited during the in-depth interview is the main focus of our analysis. First, the phone-mail booklet-phone and video survey responses are compared in terms of the reasons given for refusing to pay anything for the public program. As noted above, some refusals are valid expressions of zero WTP since they reflect lack of value for the good or low income (that is, inability to pay), while others may be considered protest responses.

Our second survey mode related hypothesis is that different survey administration modes may result in different protest rates. Comparing the overall protest reasons given, we will test the null hypothesis of no difference between the two survey administration modes in terms of proportions of protests and non-protest refusals to pay. The null hypothesis is that the distribution of refusals to pay and protest responses to the CVM survey is independent of survey administration mode:

$$H_0: \text{PROTEST}_{\text{video}} = \text{PROTEST}_{\text{phone-booklet}} \quad (6.3)$$

This will be tested using a contingency table and the test performed using a χ^2 statistic.

Comparison of Traditional Protest Classification and 'Increased Information' Protests

Within the video survey we compare the traditional approach to screening for protest responses that uses only what the respondent stated to the open-ended question of why they would not pay their monetary bid amount. The reasons were content-analysed by the authors to classify answers by similar reasons given by the respondent, according to the classifications of Jorgensen and Syme (2000).

However, unlike an open-ended WTP question or payment card where the respondent can express a zero WTP, the standard binary dichotomous choice or voter referendum approach does not directly allow for a zero WTP response. Rather, individuals indicate whether they would pay the random bid amount assigned to them. This range of bid amounts respondents are asked to pay is usually quite wide, that is, the monetary amounts some respondents are asked to pay can be quite high. Thus, it is possible that some respondents who rejected paying a very high bid amount might still have a small positive WTP, just one much smaller than the bid amount they were asked to pay. To ascertain if protesting respondents completely rejected the public program (for

example, ineffectiveness) or the means of payment, or felt government wasted money, we asked respondents who voted against the program at their initial bid amount, whether they would pay \$1. This approach is in the spirit of Collins and Rosenberger (2007) who indicate that some protest respondents in their payment card CVM did have small positive values. In our study, if respondents said 'No' to the \$1 and wrote down a reason suggesting a protest, they were classified as protests in this second approach. Conversely if they would pay the \$1, this suggests the respondent did have a positive value for the program, but just less than their bid amount, and their rejection may have been related to the amount of money being asked to pay for the program. Thus adding the follow-up question on whether those responding 'No' to their bid would pay \$1 helps the researcher to determine if the 'No' votes at the random bid indicates no value and complete rejection of the constructed market or whether they do hold positive values for the public good.

To test whether the traditional versus new increased information treatment of protests has a statistically significant effect on median WTP, we test the null hypothesis of:

$$H_0: WTP_{\text{traditional protest}} = WTP_{\text{increased info protest}} \quad (6.4)$$

WTP MODEL

Hanemann (1984) views a person responding to a dichotomous choice or voter referendum format WTP question as evaluating the difference in utility associated with the status quo versus paying some amount (\$X) to have the program. If the difference in utility is positive for the program, the individual would respond 'Yes'. If the difference in utility is distributed logistically, a logit model can be used to estimate the parameters and allow for calculation of WTP.

$$\ln \{ [\text{Prob}(Y = 1)] / [1 - \text{Prob}(Y = 1)] \} = \beta_0 + \beta_1(\$X_1) + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (6.5)$$

where β_1 is the coefficient on the dollar amount households were asked to pay, and $\beta_2 \dots \beta_n$ are coefficients on the other explanatory variables.

The log of the odds ratio is linear in the coefficients and the independent variables. Two goodness-of-fit measures are typically used for evaluating the logit model: the McFadden R-squared and the likelihood ratio (LR) statistic. For a given logit regression equation, the LR ratio statistic tests

the null hypothesis of whether collectively all the coefficients in the logistic regression are, as a group, statistically different from zero.

Median willingness to pay is calculated as:

$$\text{Median WTP} = ((\beta_0 + \beta_2 X_2 + \beta_3 X_3 + \dots \beta_n X_n) / |\beta_1|) \quad (6.6)$$

where β_0 is the constant term, and $\beta_2 \dots \beta_n$ are the other non bid coefficients.

DATA

Phone Survey-Booklet and Video Design

The public program used to compare video and phone-mail booklet-phone interviews was forest fire prevention in California. The survey booklet and videotape were developed in conjunction with forestry professionals in California and described the acreage that is burned by wildfires in an average year as well as the typical number of houses lost to wildfire each year. The current situation (the without program) was illustrated with a colour drawing refined through several focus groups. Next, a program increasing the use of prescribed fire or controlled burning in California was described. Specifically, respondents were told that the prescribed burning fuel reduction program would reduce potential wildfire fuels through periodic controlled burning. It was acknowledged that prescribed burning does create some smoke, although far less than a wildfire. A similar drawing showing that prescribed burning has lower flame length, slower rate of fire spread along with fire crews standing by was used to illustrate the public prescribed burning program. Then the respondent was provided additional information and drawings contrasting wildfire and prescribed fire. The cost of financing this program of prescribed burning was described as a cost-share program between the state of California and the county of residence of the individual. Respondents were told the new program would be implemented only if a majority of county residents voted in favour of the program.

The WTP elicitation wording was:

California is considering using some state revenue as matching funds to help counties finance fire prevention programs. If a majority of residents vote to pay the county share of this program, the Expanded California Prescribed Burning program would be implemented in your county on federal, state, and private forest and rangelands. Funding the Program would require that all users of California's forest and rangelands pay the additional costs of this program . . . If the Program was undertaken it is expected to reduce the number of acres of wildfires from the current average of 362,000

acres each year to about 272,500 acres, for a 25 per cent reduction. The number of houses destroyed by wildfires is expected to be reduced from an average of 30 a year to about 12. Your share of the Expanded California Prescribed Burning program would cost your household \$X a year. If the Expanded Prescribed Burning Program were on the next ballot would you vote ___In favour ___Against?

The \$X was replaced with one of 10 different bid amounts developed from previous fire prevention surveys in Florida (Loomis et al., 2002) and ranged from \$10 to \$470. The exact bid amounts were \$10, 20, 40, 60, 90, 120, 160, 250, 350 and 470 having been selected based on pre-testing and prior fire surveys.

The basic format of the survey booklet and script had previously been through several focus groups in two different states. In the first treatment, the survey was conducted through a phone-mail booklet-phone process. To obtain a representative sample of households, random digit dialling of the households living in a sample of California counties was performed. The counties were selected so there was a mix of counties that frequently experience wildfires, counties that occasionally experience wildfires, and counties that almost never experience wildfires. Once initial contact was established, we elicited initial attitude and knowledge of wild and prescribed fire, followed by the scheduling of appointments with individuals for detailed follow-up interviews. During the interim time period, a colour survey booklet was mailed to the household.

The 15-minute videotape was designed to follow closely the layout of the booklet, the same two colour drawings, and the question order of the phone interview. First, a script was created by adhering to the exact wording of the survey booklet and interviewer script used in the first phone interview. The video was simple and included only a headshot of the narrator, the same two still graphics in the booklet and occasional written text on the screen including the wording of the questions. In order to focus solely on the survey mode effects, the video did not deviate from the booklet or phone script. Thus, we did not take full advantage of the dynamic nature of video to be consistent with the visuals in the booklet. The video, like the booklet, begins by defining important fire management terms like 'prescribed fire' and 'wildfire'. Then, the narrator continues to describe the current problem and suggested solution in detail. Ultimately, respondents were asked questions about whether or not they agree with the proposed solution and whether or not they would be willing to pay a certain dollar amount for the solution to be implemented. The script and video were edited and revised slightly following two focus groups.

The initial contact of potential households for the videotape was much like the phone-interview process. To obtain a representative sample,

Table 6.1 *Response rates for video and phone-mail-phone surveys*

	Video	Phone	Total
<i>First wave – RDD screener</i>			
Total initial sample contacted	1261	794	2055
Completed initial	588	328	916
1st wave response rate	46.6%	41.31%	
Chi-square			2.158
<i>Second wave – In-depth interview/returned</i>			
Net sample for second wave	588	257	845
Completed	174	187	361
2nd wave response rate	29.6%	72.76%	
Chi-square			49.73***

random digit dialling of the households was used in the same counties that were used in the phone survey. A videotape or DVD, answer sheet, and postage-paid self-addressed envelope were mailed to individuals who agreed to participate in the survey. Follow-up contact was made with non-respondents, including sending a replacement videotape if necessary.

The relative cost of the two approaches should be noted. The upfront cost of the video production is fairly expensive (about \$15000) but the cost per unit is pretty small (\$7 per unit-videotape and mailing costs). The phone-mail booklet-phone survey was contracted via a survey research centre and averaged about \$75 per completed interview. Thus, if more than 200 surveys are planned it may be cheaper to use the video.

Comparison of Survey Response Rates

Because the survey was conducted in two waves, we compare the response rates from the initial random digit dial phone survey and the follow-up in-depth interviews separately in Table 6.1. We obtained 46.6 per cent in the initial phone contact in the video survey and 41.3 per cent with the initial phone contact in the phone-mail booklet-phone interview, a response rate not statistically different at the 5 per cent level using a chi-square test (calculated χ^2 of 2.158 versus critical of 3.84 with one degree of freedom). However, response rates to the follow-up were higher for the phone-mail booklet-phone at nearly 73 per cent as compared with the 30 per cent for the video. The direction of the difference is surprising as one would have expected the more novel video survey would have yielded a higher response rate. The difference is statistically significant at the 1 per cent level (calculated χ^2 is 49.73). Perhaps, having a preset appointment with

Table 6.2 Reasons why respondent would not pay for the program by survey mode

Valid reasons	Video	Phone	Total
No value/No benefits	7	1	8
Cannot afford	4	3	7
Total	11	4	15
Protest reasons			
<i>Taxes already too high</i>	<i>1</i>	<i>2</i>	<i>3</i>
<i>Should be paid for with existing taxes</i>	<i>1</i>	<i>4</i>	<i>5</i>
<i>Government wastes money</i>	<i>2</i>	<i>0</i>	<i>2</i>
<i>Those that live in WUI/forest should pay</i>	<i>1</i>	<i>0</i>	<i>1</i>
<i>Don't trust government</i>	<i>4</i>	<i>0</i>	<i>4</i>
<i>Program won't work</i>	<i>2</i>	<i>0</i>	<i>2</i>
<i>Government should pay</i>	<i>2</i>	<i>0</i>	<i>2</i>
<i>Other</i>	<i>8</i>	<i>1</i>	<i>9</i>
Total	21	7	28
Per cent of 'No' responses that are protests	66%	64%	
Overall survey protest rate	12%	4%	

Note: Italicized considered protest responses for purposes of the chi-square analysis.

the phone interviewer calling back and recording responses, with nothing for the respondent to have to initiate or mail back is an advantage to the phone-survey booklet approach over the video.

Comparing Reasons Why Households Would Not Pay for the Program by Survey Mode

Table 6.2 presents the analysis of refusals to pay, that is, individuals who indicated they would neither pay their initial bid amount nor the \$1 in the follow-up willingness to pay question (our increased information screening). Table 6.2 lists the reasons why a person would not pay using Jorgensen and Syme's (2000) classification. The first two reasons listed in Table 6.2 are not considered protest responses because having no value for the program or receiving no benefits from the program, as well as not being able to afford to pay, are valid reasons for zero WTP. However, the other categories of responses (italicized in Table 6.2) are considered protests because they were frequently prefaced with, 'I am in favour of program' or 'I'm all for it, but I think the program should be paid for by those living in the forests or with existing taxes'.

The overall protest rate for the phone-mail booklet-phone is 4 per cent. The video protest rate is 12 per cent, while three times larger than the phone-mail booklet-phone, is still low compared to protest rates reported in the literature (see Collins and Rosenberger 2007). The calculated chi-square of protest refusals to pay versus non-protest refusals to pay for video versus phone is 0.014 so we accept the null hypothesis that there is no statistical difference between the two survey modes in terms of non-protest and protest reasons for not paying. As can be seen at the bottom of Table 6.2, the percentage of respondents giving protest 'No' responses is nearly identical between the video and the phone-mail booklet-phone interview.

LOGIT REGRESSIONS RESULTS

Comparison of Phone-Mail Booklet-Phone and Video Logit Regression Results

Tables 6.3a and 6.3b presents the logit regression result for the phone-mail-phone survey mode and the video survey mode, respectively. For comparison of survey mode, both logit regression samples include all observations, including about seven protest responses in the phone-mail-phone and 21 protest responses in the video. As can be seen in Tables 6.3a and 6.3b, the bid slope coefficients are statistically different from zero at conventional levels (1 per cent for the phone-mail-phone and 5 per cent for the videotape) for both types of survey administration. The sign on the bid coefficient is negative, indicating the higher the cost to the household, the less likely a household would agree to pay for the program. The likelihood ratio statistic indicates that both overall logit models are statistically significant at the 1 per cent level.

WTP RESULTS

Comparing Median WTP for Video and Phone-Mail Booklet-Phone

Median willingness to pay for the video survey was \$323 while for the phone-mail-phone was \$423, a 24 per cent difference between the two survey modes. Confidence intervals (CI) were calculated using a technique developed by Park et al. (1991). The 90 per cent CI for the phone-booklet-phone is \$337 to \$598. For the video survey the confidence interval is larger, spanning from \$234 to \$510. Given the overlapping CI's there is no

Table 6.3a Phone-mail-phone interview results and no protests dropped

Dependent variable: VOTE RX BURN

Observations: 263

Variable	Coefficient	Std. error	z-statistic	Prob.
Constant	1.7442	0.2966	5.8800	0.0000
BIDRX	-0.0041	0.0010	-4.0822	0.0000
RXHEALTHPROB	-0.0822	0.4493	-0.1829	0.8548
WITNESSFIRE	0.0252	0.2985	0.0846	0.9326
Mean dependent var	0.7452	S.D. dependent var		0.4365
Log likelihood	-140.59	Restr. log likelihood		-149.251
LR statistic (3 df)	17.312	Probability(LR stat)		0.00061
		McFadden R-squared		0.0579
N with dependent = 0	67	Total N		263
N with dependent = 1	196			

Table 6.3b Video logit results and no protests dropped

Dependent variable: VOTE RX BURN

Observations: 156

Variable	Coefficient	Std. error	z-statistic	Prob.
Constant	-3.5766	1.5290	-2.3391	0.0193
RXBIDAMT	-0.0040	0.0012	-3.2536	0.0011
RXHEALTHPROB	1.9810	0.6607	2.9982	0.0027
WITNESSFIRE	0.7970	0.4032	1.9765	0.0481
Mean dependent var	0.6474	S.D. dependent var		0.4793
Log likelihood	-87.911	Restr. log likelihood		-101.247
LR statistic (3 df)	26.671	Probability(LR stat)		6.90E-06
		McFadden R-squared		0.1317
N with dependent = 0	55	Total N		156
N with dependent = 1	101			

statistical difference between the two mean willingness to pay estimates, despite the 24 per cent difference in mean WTP. Thus, it appears there is no statistical difference between the new videotape/DVD approach and more traditional phone-booklet-phone yield. Therefore, the choice between the two methods can be based on selecting the mode that best matches the complexity of the good to be valued and the cost relative to the target number of surveys.

Comparison of Different Protest Identification Procedures Logit Regression Results

Table 6.2 previously presented the classification of protest 'No' responses in the two surveys using the increased information procedure or what we called 'No-No' protests. These responses are classified as protests if the respondent not only voted against their random bid amount but also voted against paying \$1, and then cited one of the protest reasons listed in Table 6.2. Because this procedure uses information on two dichotomous choice responses to two different bid amounts, it helps identify individuals that have a positive WTP, but a WTP less than the random bid amount. This procedure identified the 21 protest responses shown for the video survey in Table 6.2. These 21 observations were dropped from the logit regression, and the estimation results are shown in Table 6.4a. In this logit regression the bid amount is negative and statistically significant, but the other two variables are not individually significant at conventional levels. However, the likelihood ratio statistic (LR Statistic) indicates that as a group, all three of these variables are statistically different from zero at the 1 per cent level.

In contrast to this newer approach of identifying protests, the standard protest classification in many dichotomous choice studies is to ask those who will not pay their bid amount why they would not. Using the same protest categories in Table 6.2, but ignoring individual's response to the \$1 bid, we find a total of 32 protests in our logit regression data, representing an additional 11 protest responses over the No-No approach. These additional protests are concentrated in Table 6.2 protest categories of 'Taxes already too high', 'Government wastes money', 'Those living in the forest should pay', and a few more in the 'Other' protest category. Dropping these 32 responses from the full data-set yields 124 complete observations and the estimation results presented in Table 6.4b.

In this logit regression, as in the newer approach, the bid amount is negative and statistically significant, but the other two variables are not individually significant at conventional levels. However, the likelihood ratio statistic (LR statistic) indicates that as a group, all three of these variables are statistically different from zero at the 1 per cent level.

Testing the Robustness of Median WTP to Different Treatment of Protests

The second main thrust of this chapter is to compare the effect of three treatments of protest responses on median WTP. The \$323 median WTP from the video survey noted above was estimated including protest responses to the hypothetical voter referendum (as would be done in a

Table 6.4a Increased information 'No-No' protest analysis

Dependent variable: VOTE RX BURN

Observations: 135

Variable	Coefficient	Std. error	z-statistic	Prob.
Constant	0.60716	1.9644	0.3090	0.7573
RXBIDAMT	-0.00481	0.0014	-3.4329	0.0006
RXHEALTHPROB	0.3967	0.8858	0.4455	0.6559
WITNESSFIRE	0.3767	0.4373	0.8615	0.3890
Mean dependent var	0.7462	S.D. dependent var		0.4367
Log likelihood	-68.912	Restr. log likelihood		-75.897
LR statistic (3 df)	13.969	Probability(LR stat)		0.0029
		McFadden R-squared		0.0920
N with dependent = 0	34	Total obs		135
N with dependent = 1	101			

Table 6.4b Dropping standard protest responses

Dependent variable: VOTE RX BURN

Observations: 124

Variable	Coefficient	Std. error	z-statistic	Prob.
Constant	0.6031	2.0138	0.2995	0.7645
RXBIDAMT	-0.0055	0.0015	-3.4489	0.0006
RXHEALTHPROB	0.9575	0.8595	1.1140	0.2652
WITNESSFIRE	0.0390	0.5148	0.0758	0.9395
Mean dependent var	0.813008	S.D. dependent var		0.391500
Log likelihood	-50.89464	Restr. log likelihood		-59.26529
LR statistic (3 df)	16.74130	Probability(LR stat)		0.000799
		McFadden R-squared		0.141240
N with dependent = 0	23	Total obs		124
N with dependent = 1	101			

real election). However, it is common in many CVM surveys, particularly those conducted for benefit-cost analysis, to identify and drop protest responses (Mitchell and Carson 1989).

Applying just the standard verbal responses to the question of why they would not pay their bid amount resulted in dropping about 21 per cent of the respondents (32 individuals) who would not pay their bid amount. Removing these 21 per cent of full sample observations increases the video median WTP from \$323 to \$454 for a 29 per cent increase in WTP. Using the double criteria for identifying protest responses works particularly well for overcoming the inability of the standard binary dichotomous

choice question to identify protest zero WTP. This 'increased information' approach combines the traditional open-ended statements along with a follow up WTP question regarding whether respondents would pay \$1 for the program. To be classified as a protest by this method requires that respondents not only give a protest reason for their refusal to pay their bid amount, but also indicate they would **not** vote for the program even at \$1 (what we call 'No-No' protesters). This double protest screening criteria identifies fewer protest responses (12 per cent), and yields a median WTP of \$392. This WTP of \$392 is in between the WTP from dropping all those that gave a written protest response (\$454) and the WTP from including all respondents (\$323). The absolute magnitude of the differences in WTP is 18 per cent (no protests dropped versus the 'No-No' protesters) as compared to a 29 per cent difference between not dropping any protests and dropping all respondents who give written responses objecting to paying for the program at their bid amount. However, given the overlapping confidence intervals around these three median WTP estimates, there appears to be no statistical difference in WTP estimates to treatment of protest responses. Our empirical result of robustness may be due to our relatively low protest rate (12 per cent to 21 per cent, depending on which of the two methods is used to classify protesters). For example, Collins and Rosenberger (2007) had 118 out of their 296 observations or 40 per cent protest zero WTP responses to their payment card. Given this large protest rate, how protests were treated in their survey makes a statistically significant difference in WTP estimates. Dropping protest respondents increased mean WTP from \$60 to \$98, a 39 per cent increase. This high percentage of protest zero responses is not unusual (Jorgenson et al., 1999; Lindsey, 1994) and Stevens et al. (1991) recorded protest rates of 50 per cent. In these cases, treatment of protests can substantially affect WTP.

Thus, four recommendations arise from our findings and the findings of others. The first, and most preferred, is to conduct a sufficient number of focus groups and extensive pretesting to uncover protest responses and redesign the survey to minimize them. While some respondents will always object to realistic payment vehicles and provision mechanisms, there are ways to counter this. For example, if respondents distrust government to provide the good, an independent citizen advisory board to approve government expenditures or even an independent non-profit trust to administer the program can be established. Second, is for social scientists to develop a more theoretical basis for including or rejecting protests (Boyle and Bergstrom, 1999). Third, is to develop professionally agreed standards for how to classify and treat protest responses. This would at least provide comparability in relative WTP values that would enhance the ability to rank public programs by their relative benefits. Finally, and

perhaps the most controversial, is that economists could abandon their sole reliance on obtaining values from individuals answering in isolation of one another. Since economists are often asked to value public programs, several authors have suggested group valuation, sometimes called discursive ethics (Haddad and Howarth, 2006), deliberative valuation or values juries to reach an informed and acceptable group valuation.

CONCLUSION

The overall equivalency of results between the two survey modes is encouraging. The videotape/DVD survey costs more for the initial production, but is less expensive per unit than the follow-up phone interviews. The videotape technology also offers the potential to present actual fire footage and more dynamic images, something we did not undertake in this study to maintain consistency with the original booklet. However, the phone-booklet approach offers a live interviewer, which may result in a more engaged respondent even if the interaction is only audio. Further research investigating the convergent validity of video versus a pure mail and in-person interviews is warranted to evaluate the full potential of using videotape to present information about public goods and eliciting willingness to pay.

The results regarding the robustness of WTP to treatment of protest responses is encouraging, as there were no statistically significant differences in WTP across approaches for dealing with protest responses. Despite the lack of statistical significance the 29 per cent difference in WTP for including all protests versus dropping all suspected protests is undesirably high. However, this difference in WTP is lower than many CVM surveys, we suspect in part due to the other surveys' higher protest rate. Thus, to increase the robustness of CVM WTP results to how protests are treated substantial effort must be expended in the focus group and pretest phases to design CVM surveys that minimize protest responses to begin with.

NOTE

1. We wish to thank Hayley Hessehn, University of Saskatchewan for her assistance in pre-testing the video survey. Funding for this study was provided by the Joint Fire Sciences Program and USDA Forest Service. The Survey Research at University of California-Riverside collected the videotape data collected in this research. The Survey Research Center that was formerly at California State University-Chico collected the phone-mail-phone data used in this research. The analysis of this data was supported by the W2133 Regional Research project of the Colorado Agricultural Experiment Station.

REFERENCES

- Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner and H. Schuman (1993), 'Report of the NOAA Panel on Contingent Valuation. U.S. Department of Commerce', *Federal Register*, 58 (10), 4602-14.
- Banzhaf, H.S., D. Burtraw, D. Evans and A. Krupnick (2006), 'Valuation of natural resource improvements in the Adirondacks', *Land Economics*, 82 (3), 445-64.
- Barro, S., M. Manfredo, T. Brown and G. Peterson (1996), 'Examination of the predictive validity of CVM using attitude behavior framework', *Society and Natural Resources*, 9 (2), 111-24.
- Berrens, R.P., A.K. Bohara, H.C. Jenkins-Smith, C.L. Silva and D.L. Weimer (2004), 'Information and effort in contingent valuation surveys: application to global climate change using national internet samples', *Journal of Environmental Economics and Management*, 47 (2), 331-63.
- Bishop, R. and T. Heberlein (1979), 'Measuring values of extra-market goods: are indirect measures biased?', *American Journal of Agricultural Economics*, 61 (5), 926-30.
- Boyle, K. (2003), 'Contingent valuation in practice', in P. Champ, K. Boyle and T. Brown (eds), *A Primer on Nonmarket Valuation*, Boston, MA: Kluwer Academic, pp. 111-70.
- Boyle, K. and J. Bergstrom (1999), 'Doubt, doubts, and doubters: the genesis of a new research agenda', in I. Bateman and K. Willis (eds), *Valuing Environmental Preferences*, Oxford: Oxford University Press, pp. 183-206.
- Brookshire, D., W. Schulze, M. Thayer and R. d'Arge (1982), 'Valuing public goods: a comparison of survey and hedonic approaches', *American Economic Review*, 72 (1), 165-77.
- Carson, R. (1991), 'Constructed markets', in J. Braden and C. Kolstad (eds), *Measuring the Demand for Environmental Quality*, Amsterdam: North Holland, pp. 121-62.
- Carson, R.T., R.C. Mitchell, M. Hanemann, R.J. Kopp, S. Presser and P.A. Ruud (2003), 'Contingent valuation and lost passive value: damages from the Exxon Valdez oil spill', *Environmental and Resource Economics*, 25, 257-86.
- Collins, A. and R. Rosenberger (2007), 'Protest adjustments in valuation of watershed restoration using payment card data', *Agricultural and Resource Economics Review*, 36 (2), 321-35.
- Davis, R. (1963), 'Recreation planning as an economic problem', *Natural Resources Journal*, 3 (3), 239-49.
- Haddad, B. and R. Howarth (2006), 'Protest bids, commensurability, and substitution: contingent valuation and ecological economics', in A. Alberini and J. Kahn (eds), *Handbook on Contingent Valuation*, Cheltenham, UK and Northampton, MA: Edward Elgar, pp. 133-52.
- Hammack, J. and G. Brown (1974), *Waterfowl and Wetlands: Toward Bioeconomic Analysis*, Baltimore, MD: Johns Hopkins University Press for Resources for the Future.
- Hanemann, M. (1984), 'Welfare evaluations in contingent valuation experiments with discrete responses', *American Journal of Agricultural Economics*, 66 (3), 332-41.
- Hanemann, M., J. Loomis and B. Kanninen (1991), 'Statistical efficiency of double-bounded dichotomous choice contingent valuation', *American Journal of Agricultural Economics*, 73 (4), 1225-63.
- Jorgensen, B., G. Syme, B. Bishop and B. Nancarrow (1999), 'Protest responses in contingent valuation', *Environmental and Resources Economics*, 14 (1), 131-50.
- Jorgensen, B. and G. Syme (2000), 'Protest responses and willingness to pay: attitude toward paying for stormwater pollution', *Ecological Economics*, 33 (2), 251-65.
- Lindsey, E. (1994), 'Market models, protest bids, and outliers in contingent valuation', *Journal of Water Resources Planning and Management*, 120 (1), 121-9.
- Loomis, J. (1987), 'Balancing public trust resources of Mono Lake and Los Angeles' water right: an economic approach', *Water Resources Research*, 23 (8), 1449-56.
- Loomis, J. and M. King (1994), 'Comparison of mail and telephone-mail contingent valuation surveys', *Journal of Environmental Management*, 4 (4), 309-24.
- Loomis, J., L. Bair and A. González-Cabán (2002), 'Language related differences in a

- contingent valuation study: English versus Spanish', *American Journal of Agricultural Economics*, **84** (4), 1091–102.
- Mitchell, R. and R. Carson (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington, DC: Resources for the Future.
- Morrison, M.D., R.K. Blamey and J.W. Bennett (2000), 'Minimizing payment vehicle bias in contingent valuation studies', *Environmental and Resource Economics*, **16** (4), 407–22.
- Murphy, J., P. Allen, T. Stevens, D. Weatherhead (2005), 'A meta-analysis of hypothetical bias in stated preference valuation', *Environmental and Resource Economics*, **30** (3), 313–25.
- Park, T., J. Loomis and M. Creel (1991), 'Confidence intervals for evaluating benefit estimates from dichotomous choice contingent valuation studies', *Land Economics*, **67** (1), 64–73.
- Randall, A., B. Ives and C. Eastman (1974), 'Bidding games for valuation of aesthetic environmental improvements', *Journal of Environmental Economics and Management*, **1**, 132–49.
- Stevens, T., J. Eschewerria, R. Glass, T. Hager and T. Moore (1991), 'Measuring existence value of wildlife: what do the estimates really show?', *Land Economics*, **67**, 390–400.
- Walsh, R., J. Loomis and R. Gillman, (1984), 'Valuing option, existence and bequest demands for wilderness', *Land Economics*, **60** (1), 14–29.