Journal of Agricultural and Resource Economics, 19(1): 115–128

Differences in Contingent Valuation Estimates from Referendum and Checklist Questions

Jeffrey L. Jordan and Abdelmoneim H. Elnagheeb

This article compares willingness-to-pay (WTP) estimates from an actual survey using a checklist question regarding WTP for groundwater quality improvements to WTP estimates that would have been obtained had a single-bounded referendum (SBR) or a double-bounded referendum (DBR) question been asked. Results indicate differences among estimates from the three types of question formats. There was a loss of statistical efficiency of parameter and WTP estimates when moving from the checklist and DBR formats to the SBR format. WTP estimates from the SBR question were more sensitive to sample size and model specification than the others.

Key words: contingent valuation, Monte Carlo, water quality, willingness to pay.

Introduction

Contingent valuation methods (CVM) have proved to be a valuable approach for assessing the value of nonmarket resources and public goods. Although extensive literature exists on the theoretical and empirical aspects of the CVM (Bishop and Heberlein; Cummings, Brookshire, and Schulze; Hanemann 1984, 1985; Mitchell and Carson; Sellar, Stoll, and Chavas), work continues on refining the method. An important aspect of the CVM is how to elicit an individual's willingness to pay (WTP) for a resource. Researchers have utilized various methods of questioning. The direct question that asks respondents to state the specific dollar amount they are willing to pay has been criticized as being difficult to answer (Hanemann 1985). A questioning format that has become popular in the literature is the referendum question (also known as the dichotomous choice, closed-ended, or take-it-or-leave-it question). The respondent is offered an amount and asked whether she or he is willing to pay that amount. The yes and no responses are then used in a regression model to calculate the average and/or median WTP. Although this method is easy for respondents, the information produced by the responses is diffuse. All that is known is whether the respondent's true WTP is more or less than the offered amount (bid). Hence, a large sample size and a well specified empirical model may be required to obtain a precise estimate of the mean or median WTP from a referendum approach.

Another questioning format is the checklist (also known as the payment-card). Here, the respondent is offered a range of values and is asked to circle the highest amount she or he would be willing to pay. The information obtained from this method is that the respondent's WTP is equal to or greater than the circled value but less than the next higher value. This method also has the advantage of being easy for respondents because they can visually scan a set of value intervals quickly (Cameron and Huppert). The type of information obtained by this method is less diffuse than with the referendum method. In addition to finding that someone's WTP is higher (or less) than a specified value, we also

---

Jeffrey L. Jordan is professor and research leader and Abdelmoneim H. Elnagheeb is a post-doctoral research associate, both in the Department of Agricultural and Applied Economics, University of Georgia, Georgia Experiment Station, Griffin. Seniority of authorship is equally shared.
Helpful comments by anonymous reviewers are highly appreciated.

can determine in which range that WTP lies. Therefore, the single-bounded referendum model requires a larger sample to obtain estimates with a level of accuracy comparable to the checklist model.

In a recent article, Cameron and Huppert compared WTP estimates from an actual checklist format survey to estimates that would have been obtained had the respondents been asked a referendum question. Using Monte Carlo experiments, Cameron and Huppert constructed 200 samples by randomizing the offered amounts (bids) among respondents. The method controlled for behavior bias by making responses to the bids (referendum) consistent with the actual checklist responses. If the randomly assigned bid was less than what the respondent had circled in the checklist, the expected referendum response would be "yes." Cameron and Huppert used the lower bounds of the intervals as the offered bids and in another set of the experiments they used the upper bounds. Conclusions from both sets did not vary significantly. Cameron and Huppert compared the average WTP from the actual checklist survey to the average WTP from the simulated samples. The latter is an average of averages; i.e., for each sample (of the 200 samples), an average WTP was obtained and then these averages were summed and divided by 200. The conclusion from these experiments was that referendum questions can easily lead to a wide range of WTP estimates (the regression parameter estimates will vary as well). Cameron and Huppert also conducted experiments using the double-bounded referendum (DBR) approach as explained and applied by Hanemann, Loomis, and Kanninen. Their results indicated an improvement in the accuracy of estimates over the single-bounded referendum (SBR) approach.

The present study follows Cameron and Huppert's method. However, two sample sizes are used in this study, allowing us to compare results across sample sizes. Although Cameron and Huppert's study alludes to the danger of getting misleading results from dichotomous-choice data when sample size is small, no empirical evidence on the effect of sample size on estimates was provided. Our study will shed some light on this issue. Moreover, because Cameron and Huppert's study was empirical, more empirical work is needed to support or negate their conclusions.

WTP Estimation

Previous empirical studies have indicated that valuations distribution is frequently skewed (Cameron and James). This suggests the use of lognormal distribution as a first approximation for WTP distribution. The ith respondent's true WTP, Y_i, is unobservable and was expressed as

$$\ln(Y_i) = X_i'\beta + e_i, \tag{1}$$

where X_i is a vector of explanatory variables, β is a parameter vector, e_i is independently normally distributed with mean zero and variance σ^2, and $\ln(Y_i)$ is the natural logarithm of Y_i.

The usual practice with referendum[1] data is to fit a logit or probit model and integrate the area bounded by the curve (Bishop and Heberlein; Hanemann 1984). In this article, a full information maximum likelihood (ML) estimation method is used, as suggested by Cameron and James. To describe the method, let the variable S_i take the value one if the ith respondent says "yes" to a bid t_i, and zero if not. The probability of a yes response, $\Pr[S_i = 1]$, is given by

$$\begin{aligned} \Pr[S_i = 1] &= \Pr[\ln(Y_i) \geq \ln(t_i)] = \Pr[(e_i/\sigma) \geq (\ln(t_i) - X_i'\beta)/\sigma)] \\ &= 1 - \phi[(\ln(t_i) - X_i'\beta)/\sigma], \end{aligned} \tag{2}$$

where $\phi[\cdot]$ is the standard normal distribution function. The log-likelihood function is given by

$$\ln(L) = \sum_{i=1}^{n} \{S_i\ln(1 - \phi[z_i]) + (1 - S_i)\ln(\phi[z_i])\}, \tag{3}$$

where

(4) $$z_i = (\ln(t_i) - X_i'\beta)/\sigma.$$

Maximization of (3) allows both β and σ to be estimated due to the presence of $\ln(t_i)$ in contrast to the conventional probit and logit models (Cameron and James). The double-bounded referendum (or dichotomous-choice) approach involves a follow-up question in response to the yes/no response to the single-bounded referendum question. In the follow-up question, the respondent is offered another bid contingent upon the response to the first bid. If the response to the first bid is a "yes," the second bid (denoted t^u) is set greater than the first bid ($t^u > t$). If the first response is a "no," the second bid (t^l) is set lower than the first bid ($t^l < t$).

Let ij ($i, j = y, n$) denote the sequence of responses; e.g., $i = y$ and $j = n$ means the response to the first bid was a "yes" and to the second bid was a "no." The probabilities of the four possible outcomes are given by:

(5a) $$\Pi^{yy} = \Pr(d^{yy} = 1) = 1 - \phi[(\ln(t^u) - X\beta)/\sigma];$$

(5b) $$\Pi^{nn} = \Pr(d^{nn} = 1) = \phi[(\ln(t^l) - X\beta)/\sigma];$$

(5c) $$\Pi^{yn} = \Pr(d^{yn} = 1) = \phi[(\ln(t^u) - X\beta)/\sigma] - \phi[(\ln(t) - X\beta)/\sigma]; \quad \text{and}$$

(5d) $$\Pi^{ny} = \Pr(d^{ny} = 1) = \phi[(\ln(t) - X\beta)/\sigma] - \phi[(\ln(t^l) - X\beta)/\sigma].$$

The corresponding log-likelihood function is given by (see Hanemann, Loomis, and Kanninen for more details):

(6) $$\ln(L) = \sum_i \{d_i^{yy}\ln(\Pi_i^{yy}) + d_i^{nn}\ln(\Pi_i^{nn}) + d_i^{yn}\ln(\Pi_i^{yn}) + d_i^{ny}\ln(\Pi_i^{ny})\},$$

where d^{hj}, $h, j = y, n$, are binary-valued indicator variables; ln stands for logarithm; and Π^{hj} are given by equations (5). The ML method again can be used to estimate the unknown parameters.

In the case of the checklist, the true WTP, Y_i, lies between the circled value (t_{il}, lower bound of WTP interval) and the next higher value (t_{iu}, upper bound). Expressing Y_i as in (1), we can estimate the probability that Y_i lies between t_{il} and t_{iu}, $\Pr(t_{il} \le Y_i < t_{iu})$:

(7) $$\Pr(t_{il} \le Y_i < t_{iu}) = \phi[(\ln(t_{iu}) - X_i'\beta)/\sigma] - \phi[(\ln(t_{il}) - X_i'\beta)/\sigma].$$

The corresponding log-likelihood function is given by

(8) $$\ln(L) = \sum_{i=1}^{n} \{\ln(\phi[z_{iu}] - \phi[z_{il}])\},$$

where

$$z_{iu} = (\ln(t_{iu}) - X_i'\beta)/\sigma, \quad \text{and}$$

$$z_{il} = (\ln(t_{il}) - X_i'\beta/\sigma.$$

Equation (8) is similar to an ordered-probit model (OPM) with known threshold values and can be estimated by LIMDEP's grouped data procedure (Greene). Unlike the OPM, both β and σ can be estimated.

The parameter estimates from the single- and double-bounded and the checklist models will be used to obtain estimates of the median and mean (expected) WTP. The median WTP is given by $\exp(X_i\beta)$, while the mean or expected WTP is $E[Y_i] = \exp(X_i'\beta + \sigma^2/2)$.

The Actual Data

The analysis in this study is based on a survey of Georgia residents which was conducted in February 1991. The objective of the survey was to obtain data on variables measuring people's perception of groundwater contamination and their willingness to pay for improved groundwater quality. Questionnaires were mailed to a random consumer panel of 567 people selected through a telephone survey of Georgia residents conducted by the University of Georgia Survey Research Center. The survey resulted in 192 completed questionnaires with a response rate of about 35%, excluding 14 questionnaires returned due to wrong addresses. Because of budgetary constraints, we did not send reminder cards or follow-up questionnaires. This may partially explain why the response rate was low. However, such a response rate is not uncommon in contingent valuation studies based on mail surveys (Schulze et al.; Brookshire, Coursey, and Schulze; Stoll and Johnson; Randall et al.; Walsh, Sanders, and Loomis).

Out of the 192 responding individuals, 150 obtained their drinking water from city/county water systems (78% of the sample). The remaining 42 individuals obtained water from their own private wells. This ratio is close to estimates from other sources. Approximately 78% of Georgia's population was served by public water supplies in 1985 (Bachtel).

Respondents were first asked to read the following statement in the questionnaire:

> The Environmental Protection Agency (EPA) has ranked the State of Georgia as second in the nation for potential contamination of underground water. At the same time, underground water is a source of drinking water for almost 50% of the U.S. population.
>
> Results from the EPA's five-year study of wells in different states showed that over half of U.S. drinking water wells contain nitrates. Nitrates are chemical substances hazardous to human health if taken in large quantities. Most of the wells surveyed currently have nitrate levels below hazardous levels.
>
> As farmers continue to apply more fertilizers to increase yields, the underground water may become contaminated with nitrates. Adoption of different agricultural practices can reduce the amount of nitrates in the groundwater *but* may increase food prices. On the other hand, if agricultural practices did not change, the amount of nitrates in the groundwater would increase. So the costs of cleaning water from nitrates will go up.
>
> The local water supply companies have to clean pumped water to make sure it is safe for drinking. Since the costs of cleaning water from nitrates will increase, the consumers may have to pay higher water bills.

After reading the above statement, the respondents were asked if they received their water from their own wells or from a city/county public source. If they checked "own well," they were asked to read the following:

> Suppose you found that the amount of nitrates in your well water exceeds the safe level. Suppose also that a local water supplier offers to install *and* maintain new equipment on your well. This equipment will clean your water from nitrates, but the water supplier will charge you for the use of its equipment. If you do not want to pay to the water supplier, the equipment will *not* be installed and you have to bear the risk of increasing nitrates in your drinking water.

If the respondents received public water, they were asked to read the following:

> *Imagine* that the amount of nitrates in underground water will increase. This will increase the costs of cleaning water. Imagine that the local water supply company will make sure that your water is safe for drinking *but* will increase your monthly water bill.

The respondents were then asked to circle, from a set of predetermined values, the most they would be willing to pay *above* their current monthly water bill. The question was stated as follows: "To avoid the risk of increasing nitrates in my drinking water, the *most* I would permanently pay to the water supplier *above* my current monthly water bill is (please circle *one* answer): $0, $1, $5, $10, $25, $50, $100." Respondents who chose zero were not asked why. Cross-tabulations of income by education and perception of water quality did not reveal any inconsistency. Therefore, we chose to keep all zeros.

Respondents who obtained water from public water sources and those who were on private wells could represent two distinct markets. Thus, rather than one commodity (water), there are two separate commodities (publicly treated water from surface or ground versus groundwater from private wells). Further, because of the wording of the WTP questions presented to each group, WTP might be different, as well as the effect of the explanatory variables on WTP. Therefore, two WTP equations were estimated: one for the city/county water users and one for the private well water users.

Table 1. Distributions of WTP: Checklist and Bid Values

Bid Value Interval[a] ($)	Checklist Freq.	Checklist Cum. Freq.	Lower Bound Freq.	Lower Bound Cum. Freq.	Upper Bound Freq.	Upper Bound Cum. Freq.
City/County Water Users:						
0–1	13	13	0[b]	0	13	13
1–5	46	59	59	59	46	59
5–10	51	110	51	110	51	110
10–25	18	128	18	128	18	128
25–50	8	136	8	136	8	136
50–100	1	137	1	137	4	140
100+	3	140	3	140	0[c]	140
Private Well Water Users:						
0–1	7	7	0[b]	0	7	7
1–5	5	12	12	12	5	12
5–10	10	22	10	22	10	22
10–25	16	38	16	38	16	38
25–50	0	38	0	38	0	38
50–100	0	38	0	38	2	40
100+	2	40	2	40	0[c]	40

[a] The interval *i–j* denotes that *i* is the lower bound and *j* is the upper bound which were used in the referendum Monte Carlo experiments. The respondent circled *i*.
[b] Because zero cannot be assigned as a bid in a referendum question, the zeros were replaced with the $1 bid.
[c] No upper bound is defined for this interval and $100 was used as an upper bound.

The survey also obtained information on the respondents' socioeconomic and demographic characteristics and perceptions of water quality. The checklist interval choice frequencies are presented in table 1. Descriptive statistics on variables used in the analysis are presented in table 2. To restore population representativeness, appropriate weights

Table 2. Definitions and Weighted Averages of Variables Used in Analysis

Variable Name	Description	Average: City/County Water Users	Average: Private Well Water Users
INCOME	Household's total income, before tax, for 1990 using midpoints of reported intervals	28,683 (16,613)[a]	22,008 (13,340)[a]
Ln(*INCOME*)	Logarithm of *INCOME*	10.02 (.80)[a]	9.77 (.77)[a]
MALE	One if male, zero otherwise	.45	.51
BLACK	One if Black, zero otherwise	.30	.13
AGE	Age in years	49 (15)[a]	53 (16)[a]
EDUCATION	One if more than high school, zero otherwise	.76	.66
FARM	One if resides on a farm or a ranch, zero otherwise	.01	.25
RISK[b]	One if rated current water quality as poor, zero otherwise	.27	.13
UNCERTAIN[b]	One if uncertain of current water quality, zero otherwise	.23	.14
n	Sample size	140	40

[a] Numbers in parentheses are the sample standard deviations.
[b] Excluded category is comprised of those respondents who rated current water quality as very safe, safe, or fair.

were used (Sonquist and Dunkelberg). Weights were derived from cross-tabulations of race by education and sex. These weights were based on the distributions of respondents' socioeconomic and demographic characteristics from the whole sample and the corresponding distributions from census data (Bachtel; Wetrogan). The weights attempt to give each stratum the same relative importance in the sample that it has in the population, thus minimizing the sample nonresponse bias. Mathematically, the weight for the jth stratum, w_j, is the ratio of the population proportion for the jth stratum divided by the sample proportion for the same stratum ($w_j = P_j/s_j$). It would have been more appropriate had the weights been based on the distributions of characteristics of two populations, city/county water users and private well water users. However, lack of census information on the distributions of socioeconomic and demographic characteristics in the two populations separately impeded this approach.

Another source of possible bias is the unit or item nonresponse (Mitchell and Carson). Among variables used in the analysis, income has the highest number of missing observations: six for the city/county water users and five for private well water users. Following the recommendation by Mitchell and Carson, we imputed income for missing observations. This task was accomplished by regressing the logarithm of income on other variables having complete observations and using the predicted income to fill the missing observations. The independent variables used in the regression were race, education, employment, and marriage status. These imputations raised the number of complete observations on all variables used in the analysis from 163 to 180, 140 for city/county water users and 40 for private well water users.[2] The 180 observations make up the basis for this analysis.

The Referendum Monte Carlo Experiments

Single-Bounded Referendum

In the actual survey, there were seven values from which respondents could choose. Hence, these seven values could have been assigned (as bids) to respondents in a referendum question. Following Cameron and Huppert, the threshold values were generated from frequency distributions similar to the observed frequencies from the actual survey. Two sets of experiments were conducted. In the first set of experiments, the threshold distributions mimic the distributions of the lower bounds of the actual intervals, while in the second set, the upper bounds were used as the thresholds. The frequency distribution from the actual survey was used to generate 100 samples. The actual interval values and their corresponding frequencies are presented in table 1.

The 100 simulated samples were generated as follows. For the city/county water users, random integers ranging from 1 to 140 were generated (without repetition; this is the actual sample size). Each random integer corresponded to a respondent. The actual frequency distribution (table 1) was then followed to assign the threshold values. This point is illustrated for the case where the lower bounds were used as bids. For respondents who were randomly assigned the integers 1 to 59, the assigned threshold was \$1, those assigned the integers 60 to 110 were assigned the threshold \$5, those assigned the integers 111 to 128 were assigned the threshold \$10, and so forth. In this way, the distribution of the thresholds will mimic the actual frequency distribution of the lower bounds. This procedure was followed for the case where upper bounds were used as bids. Each time we generated a set of 140 random integers, one simulated sample was generated. The same steps were followed for the private well water users.

The second step was to determine the responses to the assigned threshold values. These are responses one would have expected had the referendum question been asked. The yes/no responses were generated by comparing the assigned threshold value to the actual value circled by the respondent in the survey. For example, if the respondent had circled \$10 from the checklist, a yes response to a \$5 threshold value would be assigned. These responses were then used to estimate the β and σ parameters in equation (1).

Table 3. Maximum Likelihood Estimates: Checklist versus Single-Bounded Referendum, City/County Water Users

	Checklist	100 Single-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Asymp. t-Ratio)	Mean (t-Ratio)[a]	Max./Min.	Mean (t-Ratio)[a]	Max./Min.
Ln(*INCOME*)	.11928	.12716	.31540	.09011	.29700
	(2.175)	(1.779)	−.07726	(1.415)	−.02492
MALE	−.51665	−.51611	−.02736	−.49981	.09220
	(−2.420)	(−2.451)	−1.1790	(−2.210)	.99760
AGE	−.00410	−.00307	.01447	−.00098	.01421
	(−.605)	(−.468)	−.01649	(−.139)	−.01685
BLACK	.67025	.59043	1.3980	.51893	1.3250
	(2.831)	(2.195)	−.1160	(1.834)	−.2006
EDUCATION	.63665	.43892	1.2610	.63663	1.7440
	(2.054)	(1.248)	−.57260	(1.728)	−.3678
FARM	.03523	−1.7446	4.291	.68803	4.6950
	(.038)	(−.791)	−4.474	(.290)	−4.3440
RISK	.05349	.17459	.76780	.14378	.72800
	(.214)	(.743)	−.45890	(.539)	−.61130
UNCERTAIN	.66736	.75707	1.6040	.70738	1.8550
	(2.599)	(3.094)	.08835	(2.483)	−.19160
σ	1.13394	1.0517	1.8490	1.15960	1.8200
	(14.396)	(4.536)	.62310	(5.593)	.60940
Max. ln(L)		−62.958		−60.123	
		(5.835)		(6.123)	

[a] Mean divided by across-sample standard deviation of estimates.

Double-Bounded Referendum

The experiments for the double-bounded referendum (DBR) were similar to those for the single-bounded referendum (SBR) except for a follow-up question in the case of the DBR. If the simulated response to the SBR question was a "yes," the next highest bid (which was about twice the assigned first bid) was assigned as a second bid. A yes or no response was then simulated to this second bid. If, on the other hand, the simulated first response was a "no," the next lowest bid (which was about half the assigned first bid) was assigned as a second bid and a yes or no response was simulated. To illustrate, suppose a respondent circled \$25 as a WTP from the checklist survey and was assigned \$5 as a first bid. The simulated first response would be "yes" because \$25 > \$5. The second bid to assign would be \$10 (highest next to \$5) and the simulated second response would be "yes" because \$25 > \$10.

Results

The set of independent variables used to explain variations in WTP is similar to what has been used in other studies (Shultz and Lindsay). Table 2 contains descriptive statistics of the independent variables (X) used in equation (1). The variables *RISK* and *UNCERTAIN* measure the respondent's perception of water quality. These variables were derived from responses to the question, "Overall, how would you rate your drinking water quality?" The respondent was requested to choose one of five answers: very safe, safe, fair, poor, or don't know.

Parameters for both the checklist and referendum (single- and double-bounded) models were estimated by "weighted" maximum likelihood (ML) using the LIMDEP software (Greene). Appropriate weights were used to restore sample representativeness, as discussed above.

Table 4. Maximum Likelihood Estimates: Checklist versus Single-Bounded Referendum, Private Well Water Users

		100 Single-Bounded Referendum Samples			
	Checklist	Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Asymp. t-Ratio)	Mean (t-Ratio)[a]	Max./Min.	Mean (t-Ratio)[a]	Max./Min.
Ln(*INCOME*)	.12571	.4995	18.170	.15650	.86580
	(1.526)	(.143)	−22.400	(.924)	−.11820
MALE	−.82210	−6.2376	.45220	−.90501	.09245
	(−2.145)	(−.289)	−159.200	(−1.422)	−3.1820
AGE	−.00877	−.03136	1.3630	−.01297	.05290
	(−.750)	(−.114)	−1.4760	(−.566)	−.10220
BLACK	1.26447	7.2838	265.60	2.28100	8.8000
	(2.245)	(.241)	−31.780	(1.161)	−.15950
EDUCATION	1.00902	8.8178	309.50	1.17570	23.310
	(2.073)	(.229)	−8.9780	(.505)	−.22290
FARM	1.23931	1.3331	61.410	1.7098	10.830
	(2.805)	(.150)	−29.450	(1.009)	.09427
RISK	.00912	−1.1741	31.470	−.00667	3.4610
	(.015)	(−.140)	−48.590	(−.004)	−13.250
UNCERTAIN	.99975	10.9390	322.10	1.67660	19.990
	(1.868)	(.278)	−.36040	(.676)	−1.376
σ	1.06174	.76783	1.8820	1.3699	4.824
	(7.330)	(1.721)	.00892	(.889)	−9.548
Max. ln(L)		−10.399		−16.099	
		(4.743)		(3.790)	

[a] Mean divided by across-sample standard deviation of estimates.

Single-Bounded Referendum

The ML parameter estimates for both the checklist and single-bounded referendum models for the city/county water users are presented in table 3. Corresponding results for the private well water users are presented in table 4.

For each group of water users, 100 samples were generated for the referendum question resulting in 100 vectors of parameter estimates. Averages of these vectors were calculated (tables 3 and 4). Following Cameron and Huppert, the degree of dispersion in these estimates was approximated by the standard deviation across the 100 samples. These standard deviations were used to calculate what Cameron and Huppert called artificial t-statistics. The t-statistics were obtained by dividing the across-sample average of each parameter estimate by its corresponding across-sample standard deviation. For the checklist model, the ML asymptotic standard errors measure the variability in the parameter estimates. The corresponding asymptotic t-ratios are reported in tables 3 and 4.

As shown in table 3, the averages of the parameter estimates from the referendum experiments carry the same sign as their corresponding estimates from the checklist model. One exception is the coefficient on *FARM* for the upper-bound referendum experiments. However, this coefficient was not statistically significant in either the original checklist model or the referendum model. The parameter estimates from the checklist model and the averages of the corresponding parameter estimates from the referendum are comparable in magnitude (table 3). However, the checklist estimates were more efficient, as indicated by higher t-statistics, than the referendum estimates. This loss in statistical efficiency in the referendum estimates was also noted by Cameron and Huppert. However, for most of the significant coefficients in the checklist model, the ratio of their t-statistics to the corresponding t-statistics from the referendum experiments is less than two. This ratio was two or greater for most of the estimates in Cameron and Huppert's study (note that our survey used fewer checklist values than theirs). Hence, this result may indicate

Table 5. WTP Estimates: Checklist versus Single-Bounded Referendum, City/County Water Users

	Checklist	100 Single-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Mean (Std. Error)	Mean (Std. Error)	Max./Min.	Mean (Std. Error)	Max./Min.
Median	5.85[a]	6.76[c]	23.93	6.19[c]	15.81
	(.27)[b]	(3.72)[d]	2.29	(2.25)[d]	3.33
Mean	11.13	11.14	29.06	12.08	21.94
	(.51)	(3.64)	7.25	(3.19)	8.10

[a] Sample mean across respondents.
[b] Standard error of the mean = $s/\sqrt{n}$, where s = standard deviation across respondents, and n = sample size.
[c] Mean of 100 sample-means (across samples).
[d] Standard deviation across samples of means.

that the relative efficiency of the checklist model to the referendum model (both using the same number of discrete response values) is smaller when the checklist is shorter.[3]

The ML results for the private well water users are presented in table 4. The sample size for this group is only 40 in contrast to 140 for the city/county water users. Most of the parameter estimates for the referendum models agree in sign with their counterparts for the checklist model. However, the difference in magnitude is large. This result is more obvious if we examine the fourth and last columns of table 4 where the maximum and minimum parameter estimates are reported. The result is also more obvious for the referendum experiments which used the upper bounds of the checklist as bids than it is for the lower-bound experiments. This result indicates the sensitivity of the assignment of bids in referendum models to sample size. The loss in statistical efficiency of the parameter estimates also is clear. Except for one (or two at most), all of the estimates from the referendum experiments were statistically insignificant at the 10% level. This result is in contrast to six significant parameter estimates for the checklist model. Two factors might contribute to the loss in statistical efficiency. First, the number of bids assigned to the private well water users was less than for the city/county water users. The reason for this is that two values ($25 and $50) were not checked in the actual survey by the private well water users (see table 1). Hence, these values were not used as bids. The second factor could be the small sample size for the private well group.[4]

One objective of estimating a WTP regression model is to obtain an estimate of the average (or median) WTP. Using the estimates of the β and σ parameters of the checklist model, both the median and the expected WTP ($E[Y_i]$) for each respondent were calculated.

Table 6. WTP Estimates: Checklist versus Single-Bounded Referendum, Private Well Water Users

	Checklist	100 Single-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Mean (Std. Error)	Mean (Std. Error)	Max./Min.	Mean (Std. Error)	Max./Min.
Median	7.97[a]	20,105E9[c]	10,950E11	28,079E5[c]	28,070E7
	(1.12)[b]	(12,066E10)[d]	0	(28,074E6)[d]	2.06
Mean	14.01	20,151E9	10,990E11	28,222E5	28,200E7
	(1.96)	(12,100E10)	0	(28,201E6)	0

Refer to table 5 footnotes.

Table 7. Maximum Likelihood Estimates: Checklist versus Double-Bounded Referendum, City/County Water Users

		100 Double-Bounded Referendum Samples			
	Checklist	Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Asymp. t-Ratio)	Mean (t-Ratio)[a]	Max./Min.	Mean (t-Ratio)[a]	Max./Min.
Ln(*INCOME*)	.11928	.12620	.21000	.10284	.16970
	(2.175)	(4.157)	.05267	(3.701)	.03784
MALE	−.51665	−.45361	−.23160	−.54010	−.31330
	(−2.420)	(−4.767)	−.67650	(−5.604)	−.81990
AGE	−.00410	−.00398	.00591	−.00111	.00746
	(−.605)	(−1.150)	−.01382	(−.291)	−.01103
BLACK	.67025	.65341	.96110	.56235	.91610
	(2.831)	(4.850)	.36380	(4.218)	.27440
EDUCATION	.63665	.49067	1.10600	.70544	1.07400
	(2.054)	(2.319)	−.19330	(4.437)	.28970
FARM	.03523	−.21630	1.05200	.02215	1.34400
	(.038)	(−.204)	−4.494	(.042)	−3.92900
RISK	.05349	.11120	.42510	.11801	.38900
	(.214)	(.909)	−.12970	(1.081)	−.15740
UNCERTAIN	.66736	.73693	1.01700	.71593	1.04100
	(2.599)	(6.002)	.34240	(5.812)	.37420
σ	1.13394	1.03690	1.23300	1.11420	1.27000
	(14.396)	(11.130)	.77400	(16.062)	.95970
Max. ln(L)		−142.63		−141.53	
		(57.04)		(6.13)	

[a] Mean divided by across-sample standard deviation of estimates.

These values were then averaged across respondents and their standard deviations were calculated.

In the case of the referendum experiments, the expected WTP and the median WTP were calculated for each individual in each sample. The across-individual averages of expected WTP and median WTP were then obtained for each sample. These values were then averaged across the 100 simulated samples.

The estimates of WTP obtained from the checklist and referendum experiments for the city/county water users are presented in table 5. Corresponding results for the private well water users are presented in table 6. Results in these tables are consistent with those reported in tables 3 and 4. For the city/county group, the WTP estimates from the referendum experiments are close to those from the checklist model (table 5). On the other hand, there are huge differences between the WTP estimates from the referendum experiments and their counterparts from the checklist model for the private well group (table 6). This result is due mainly to the small sample size of private well water users. The result is also in line with the expectations of Cameron and Huppert, who reported that the smaller the sample size, the greater the danger of obtaining misleading results from the referendum model (p. 917). The result is important because researchers are often forced to use small samples due to budgetary constraints. The implication is that researchers should be aware of the danger of obtaining misleading WTP estimates when the dichotomous-choice (or referendum) question is administered to a small sample.

Double-Bounded Referendum

The ML parameter estimates for the checklist and double-bounded referendum (DBR) models for the city/county water users are presented in table 7. Corresponding results for the private well water users are presented in table 8.

In general, as shown in table 7, the average (across simulations) of the DBR model's

Table 8. Maximum Likelihood Estimates: Checklist versus Double-Bounded Referendum, Private Well Water Users

	Checklist	100 Double-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Asymp. t-Ratio)	Mean (t-Ratio)[a]	Max./Min.	Mean (t-Ratio)[a]	Max./Min.
Ln(*INCOME*)	.12571	.13638	.28460	.13465	.31010
	(1.526)	(2.141)	−.12940	(2.419)	.018970
MALE	−.82210	−.70323	−.11000	−.96426	−.30240
	(−2.145)	(−2.550)	−1.52400	(−3.503)	−1.98100
AGE	−.00877	−.00416	.01802	−.01071	.00571
	(−.750)	(−.496)	−.02814	(−1.400)	−.04095
BLACK	1.26447	1.0351	1.96300	1.39850	2.13700
	(2.245)	(2.921)	.31290	(5.908)	.86370
EDUCATION	1.00902	.90238	2.16500	1.44550	2.36000
	(2.073)	(2.500)	.00552	(4.021)	.71760
FARM	1.23931	1.02340	2.07300	1.63240	2.87000
	(2.805)	(3.372)	−.18280	(4.085)	.92120
RISK	.00912	.10632	1.40900	.03917	.59610
	(.015)	(.276)	−1.27900	(.140)	−.48930
UNCERTAIN	.99975	1.04240	2.36400	1.21710	2.87800
	(1.868)	(2.373)	−.025020	(2.388)	.32340
σ	1.06174	1.09510	1.49500	1.36880	2.00100
	(7.330)	(5.985)	.56920	(7.117)	.88750
Max. ln(L)		−35.86		−36.32	
		(3.89)		(2.91)	

[a] Mean divided by across-sample standard deviation of estimates.

parameter estimates compare favorably to their counterparts from the checklist model. In comparison to the single-bounded referendum (SBR) model's results (table 3), there are substantial improvements in the efficiency of the DBR model's parameter estimates, as indicated by higher t-ratios (table 7). This result supports the finding of Cameron and Huppert and shows the effect of the additional information obtained from the follow-up question in the DBR format. More important, and in contrast to Cameron and Huppert, the t-ratios for the DBR model are higher than their counterparts for the checklist model (table 7). In only one instance did the maximum/minimum estimates for the significant coefficients change sign (table 7, column 4). What has been stated about table 7 applies also to table 8. Hence, regardless of the sample size, the DBR model statistically outperforms the SBR model.

The WTP estimates from the checklist and DBR models for the city/county water users are presented in table 9. Corresponding results for the private well water users are presented

Table 9. WTP Estimates: Checklist versus Double-Bounded Referendum, City/County Water Users

	Checklist	100 Double-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Std. Error)	Mean (Std. Error)	Max./Min.	Mean (Std. Error)	Max./Min.
Median	5.85[a]	5.88[c]	6.83	6.10[c]	7.22
	(.27)[b]	(.32)[d]	5.14	(.33)[d]	5.43
Mean	11.13	10.13	12.76	11.43	15.56
	(.51)	(.87)	8.32	(1.31)	9.10

Refer to table 5 footnotes.

Table 10. WTP Estimates: Checklist versus Double-Bounded Referendum, Private Well Water Users

	Checklist	100 Double-Bounded Referendum Samples			
		Upper-Bound Distrib.		Lower-Bound Distrib.	
Variable	Coeff. Estimate (Std. Error)	Mean (Std. Error)	Max./Min.	Mean (Std. Error)	Max./Min.
Median	7.97[a]	10.31[c]	24.92	17.06[c]	45.30
	(1.12)[b]	(2.39)[d]	7.34	(7.30)[d]	9.31
Mean	14.01	19.49	55.02	48.68	192.00
	(1.96)	(6.23)	10.82	(33.49)	16.05

Refer to table 5 footnotes.

in table 10. For the city/county group, the WTP estimates from the DBR experiments were very close to those from the checklist model (table 9). The DBR model shows a substantial increase over the SBR model (see table 5) in the accuracy of WTP estimation, as indicated by the standard deviations and the maximum/minimum columns of table 9. The WTP estimates for the private well water users were not as close to their counterparts from the checklist model as were those for the city/county water users. This result is due mainly to the smaller sample size of the private well water users. However, an improvement is revealed in WTP estimation when moving from the SBR model (table 6) to the DBR model (table 10).

Conclusions

In this study, three WTP regression equations were estimated. The WTP data for the first equation were from an actual survey that used a checklist questioning format to elicit people's WTP for drinking-water quality improvements. The WTP data for the second and third equations were Monte Carlo-generated data representing responses that would have been obtained had a single-bounded or a double-bounded referendum WTP question been asked. For both the single-bounded referendum (SBR) and double-bounded referendum (DBR) models, two sets of experiments were conducted. One set of experiments used the upper bounds of the checklist intervals as the offered bids, while the second set used the lower bounds. For each set, 100 samples were generated.

The actual and simulated WTP data were used in combination with other socioeconomic and demographic data to estimate the parameters of the WTP equations. The parameter estimates then were used to estimate the unobservable WTP for two groups of respondents: those who obtained water from a city/county water system and those who obtained water from private wells. The sample sizes were 140 for the first group and 40 for the second. The study compared the across-sample averages of the parameter and WTP estimates from the SBR and DBR experiments to their counterparts from the checklist model.

Results indicated that parameter and WTP estimates from the SBR experiments were close to their counterparts from the checklist model for the city/county water users. However, the differences between the SBR and checklist estimates were huge for the private well water users. These results indicate that the SBR model was more sensitive to sample size than the checklist model. The study also found that there was a loss in statistical efficiency of parameter estimates when moving from a checklist to an SBR format. However, this loss is less noticeable in the case of the larger sample, city/county water users. The result again supports the previous conclusion about the sensitivity of the SBR model to sample size. These results support recent results obtained by Cameron and

Huppert, who concluded that WTP studies that use SBR surveys can produce quite different estimates, even from the same population.

The DBR model showed a substantial improvement over the SBR model in terms of closeness of parameter and WTP estimates to their counterparts from the checklist model and in terms of efficiency of estimates. Moreover, the DBR model's estimates were more efficient than those from the checklist model.

The study also provided empirical evidence on the sensitivity of the SBR model to sample size. This result is important because researchers are often forced to select small samples due to budget constraints. In such cases, the checklist and the DBR formats may be preferable to the referendum format for statistical efficiency for a small sample. This conclusion says nothing about the behavioral distortion or bias introduced by any format, an issue which needs more investigation by CVM researchers. We also believe that increasing the number of offered bids in the referendum question will improve the efficiency of the parameter and WTP estimates.

[*Received December 1992; final revision received June 1993.*]

Notes

[1] Throughout this article, unless otherwise specified, referendum refers to the single-bounded (conventional) referendum approach as developed in Bishop and Heberlein.

[2] Although 40 is a small sample size, a frequently cited CV study (Sellar, Stoll, and Chavas, table 5, p. 171) used samples that varied in size from 15 to 74.

[3] We owe this point to a reviewer.

[4] Following a reviewer's suggestions, we took a random sample of 70 (half the sample) respondents from the city/county water users. Using this subsample, we estimated the checklist and the referendum (single- and double-bounded) models. The Monte Carlo results using this subsample confirmed our conclusion that the single-bounded referendum model is more sensitive to sample size than are the checklist and double-bounded referendum models. Results are available from the authors upon request.

References

Bachtel, D. C., ed. *The Georgia County Guide.* Coop. Ext. Ser., College of Agr., The University of Georgia, Athens, 1989.

Bishop, R. C., and T. A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *Amer. J. Agr. Econ.* 61(1979):926–30.

Brookshire, D. S., D. L. Coursey, and W. D. Schulze. "Experiments in the Solicitation of Private and Public Values: An Overview." In *Advances in Behavioral Economics,* Vol. 2, eds., L. Green and J. Kagel, pp. 173–90. Norwood NJ: Ablex Publishing Corp., 1990.

Cameron, T. A., and D. D. Huppert. "Referendum Contingent Valuation Estimates: Sensitivity to the Assignment of Offered Values." *J. Amer. Statis. Assoc.* 86(1991):910–18.

Cameron, T. A., and M. D. James. "Efficient Estimation Methods for Use with 'Closed-Ended' Contingent Valuation Survey Data." *Rev. Econ. and Statis.* 69(1987):269–76.

Cummings, R. G., D. S. Brookshire, and W. D. Schulze. *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* Totowa NJ: Rowman and Allanheld, 1986.

Greene, W. H. *LIMDEP Software.* New York: Econometric Software, Inc., 1990.

Hanemann, W. H. "Some Issues in Continuous- and Discrete-Response Contingent Valuation Studies." *Northeast. J. Agr. and Resour. Econ.* 14(1985):5–17.

———. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses." *Amer. J. Agr. Econ.* 66(1984):332–41.

Hanemann, W. H., J. Loomis, and B. Kanninen. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *Amer. J. Agr. Econ.* 73(1991):1255–63.

Mitchell, R. C., and R. T. Carson. *Using Surveys to Value Public Goods: The Contingent Valuation Method.* Washington DC: Resources for the Future, 1989.

Randall, A., G. C. Blomquist, J. P. Hoehn, and J. R. Stoll. "National Aggregate Benefits of Air and Water Pollution Control." Interim Rep. to U.S. Environmental Protection Agency, Washington DC, 1985.

Schulze, W. D., R. G. Cummings, D. S. Brookshire, M. A. Thayer, R. Whitworth, and M. Rahmatian. "Methods Development in Measuring Benefits of Environmental Improvements: Experimental Approaches for Valuing Environmental Commodities," Vol. 2. Draft Rep. to Office of Policy Analysis and Resource Management, U.S. Environmental Protection Agency, Washington DC, 1983.

Sellar, C., J. R. Stoll, and J. P. Chavas. "Validation of Empirical Measures of Welfare Change: A Comparison of Nonmarket Techniques." *Land Econ.* 61(1985):156–75.
Shultz, S. D., and B. E. Lindsay. "The Willingness to Pay for Groundwater Protection." *Water Resour. Res.* 26(1990):1869–75.
Sonquist, J. A., and W. C. Dunkelberg. *Survey and Opinion Research: Procedures for Processing and Analysis.* Englewood Cliffs NJ: Prentice-Hall, Inc., 1977.
Stoll, J. R., and L. A. Johnson. "Concepts of Value, Nonmarket Valuation, and the Case of the Whooping Crane." Texas Agr. Exp. Sta. Pub. No. 19360, Dept. Agr. Econ., Texas A&M University, 1985.
Walsh, R. G., L. D. Sanders, and J. B. Loomis. "Wildlife and Scenic River Economics: Recreation Use and Preservation Values." Spec. Rep. to American Wilderness Alliance, prepared through Dept. Agr. and Nat. Resour. Econ., Colorado State University, 1985.
Wetrogan, S. I. "Projections of the Population of States by Age, Sex, and Race: 1989 to 2010." *Current Population Reports,* Ser. P-25, No. 1053. U.S. Department of Commerce, Bureau of the Census, Washington DC, 1990.