Journal of Agricultural and Resource Economics 37(2):289–300

A Method for Improving Welfare Estimates from Multiple-Referendum Surveys

GwanSeon Kim, Daniel R. Petrolia, and Matthew G. Interis

This paper introduces an alternative non-market value elicitation method—the "quasi-double-referendum" (QDR)—applied to barrier island restoration in Mississippi. It is appropriate for surveys that elicit willingness-to-pay responses to multiple projects differing in scale only and can be used to increase efficiency while mitigating bias. We compare results to the more commonly used single-referendum (SR) method under two admissible ranges of willingness to pay: negative to positive infinity, and zero to income. The confidence intervals of the QDR models were narrower. We argue that the QDR approach should be less subject to bias than the commonly used double-referendum approach.

Key words: quasi double-bound method, single-bound method, willingness to pay

Introduction

The single-referendum (SR) dichotomous-choice question, initially proposed by Bishop and Heberlein (1979), is the simplest and most frequently used approach for eliciting willingness to pay (WTP) in a contingent valuation setting. The major shortcoming of the SR method is that it collects relatively little information from each respondent. All that is revealed is that the respondent's true WTP is either greater than or less than a single offered bid. Hanemann (1985); Carson (1985); and Carson, Hanemann, and Mitchell (1986) proposed adding a follow-up question as a remedy. In this approach, developed by Hanemann, Loomis, and Kanninen (1991) and now known as the double-referendum (DR) method,[1] a respondent is offered a higher bid if he votes "yes" to the initial bid amount, or is offered a lower bid if he votes "no" to the initial bid. More information is collected regarding the location of WTP, resulting in tighter welfare estimates, all else equal (Hanemann, Loomis, and Kanninen, 1991; Kanninen, 1993; Carson, Wilks, and Imber, 1994; Alberini, 1995; McLeod and Bergland, 1999; Carson and Groves, 2007).

Nevertheless, the DR method has been criticized because it suffers from various forms of response bias (Cameron and Quiggin, 1994; Clarke, 2000; Cooper, Hanemann, and Signorello, 2002; Haab and McConnell, 2002), including starting-point bias, in which responses to the follow-up question depend on the initial bid amount offered (Mitchell and Carson, 1993; Flachaire and Hollard, 2006; Herriges and Shogren, 1996); shifting-effect bias, in which the respondent interprets a change

GwanSeon Kim is a Ph.D. Student, Department of Agricultural & Applied Economics, University of Georgia; Daniel R. Petrolia and Matthew G. Interis are Assistant Professors, Department of Agricultural Economics, Mississippi State University.
The authors wish to thank Dr. Greg Carter and Gabe Blossom at the Gulf Coast Geospatial Center for providing the maps used in the survey and for providing feedback on survey design. This research was conducted under award NA06OAR4320264 06111039 to the Northern Gulf Institute by the NOAA Office of Ocean and Atmospheric Research, U.S. Department of Commerce, and by the USDA Cooperative State Research, Education & Extension Service, Hatch project MIS-012030, "Valuation of Environmental Goods and Natural Resources."

Review coordinated by Gary Brester.

[1] The two approaches are also commonly referred to as "single-bounded" and "double-bounded." Later in the paper, however, we discuss the possibility of imposing bounds of zero and annual income on WTP. In order to avoid confusion between these bounds based on an intuitive argument, and bounds created from asking the WTP preference questions, we refer to the two approaches as single and double "referendum."

in the offered price to be a signal of altered quality of the project (Carson et al., 1992; Alberini, Kanninen, and Carson, 1997; Watson and Ryan, 2007); and strategic bias, in which respondents see the new bid amount as a signal that they can bargain over the price (Altaf and DeShazo, 1994; McLeod and Bergland, 1999; Haab and McConnell, 2002; Cooper, Hanemann, and Signorello, 2002; Carson and Groves, 2007).

This paper introduces an alternative method called the "quasi-double-referendum" (QDR) method. The QDR method asks respondents to evaluate three or more different versions of a single policy in the same survey instrument; these versions differ only in the quantity (scale) and price of the good provided. Estimates of WTP for intermediate-scale programs are improved by incorporating data from the questions pertaining to smaller- and larger-scale programs.

The objective of this paper is to introduce the QDR method and report the results of its application to valuation of a proposed barrier-island restoration program for the state of Mississippi, which was originally analyzed by Petrolia and Kim (2009). We test the hypotheses that 1) median WTP obtained using the QDR method is equal to that of the SR method, and 2) the variance of WTP obtained using the QDR method is smaller than that of the SR method.

We find that median WTP estimates using the QDR method are significantly higher than those using the SR method in unbounded models, but slightly lower in bounded models. We find that the QDR approach yields significantly smaller variances (i.e., narrower confidence intervals).[2] We believe that the proposed method can be applied to any survey that elicits willingness-to-pay responses to multiple programs that differ only in scale and can be used to increase estimate efficiency for a variety of restoration programs evaluated in other surveys while mitigating the bias introduced by a more traditional DR survey.

The Quasi-Double-Referendum Method

The use of the QDR method requires two conditions. The first is that the only fundamental difference between the scenarios perceived by the respondent is quantity (scale). The second is that utility is non-decreasing in the quantity provided. Provided these assumptions hold in a three-scenario survey, and letting t be the offered price for the intermediate-scale program, "no" responses to the larger scale program provided at price $t'(>t)$ can be interpreted simultaneously as "no" responses to the intermediate program provided at price t', since respondents who are not willing to pay t' for the large scale project should also not be willing to pay the same amount for less of the public good. Similarly, "yes" responses to the small scale project provided at price $t''(<t)$ can be simultaneously interpreted as "yes" responses to the intermediate scale project provided at price t''. In this way, we can narrow the WTP ranges on the intermediate scale program for respondents who vote "yes-no" on the intermediate and large scale programs or "no-yes" on the intermediate and small scale programs to $[t,t')$ and $(t'',t]$ respectively. For "yes-yes" and "no-no" respondents, no additional information on WTP for the intermediate-scale program can be inferred, and so the range of WTP remains $[t,+\infty]$ and $[-\infty,t]$, respectively. As this discussion implies, the strength of this approach hinges on there being a substantial number of respondents who switch from "yes" to "no" or "no" to "yes." Figure 1 summarizes the logic of the QDR method.

The QDR method should mitigate the potential for the biases associated with the DR method. Similar to the approach taken by Cooper, Hanemann, and Signorello (2002), we inform respondents up front that they will be asked to evaluate the program at three distinct and explicitly described scales. Therefore, the likelihood of a respondent substituting his own perceived quantity (or quality) contrary to the researcher's intention (i.e., shifting effect) and the likelihood of creating a perceived price bargaining situation (i.e., strategic bias) should be mitigated. Because the QDR method requires three valuation questions for programs at different quantities, a new bias is potentially

[2] Note that we do not compare the results of the QDR method to the DR method directly because of the logistical difficulty in implementing both the QDR and DR methods in the same survey. Thus our metric is improvements relative to the SR method.

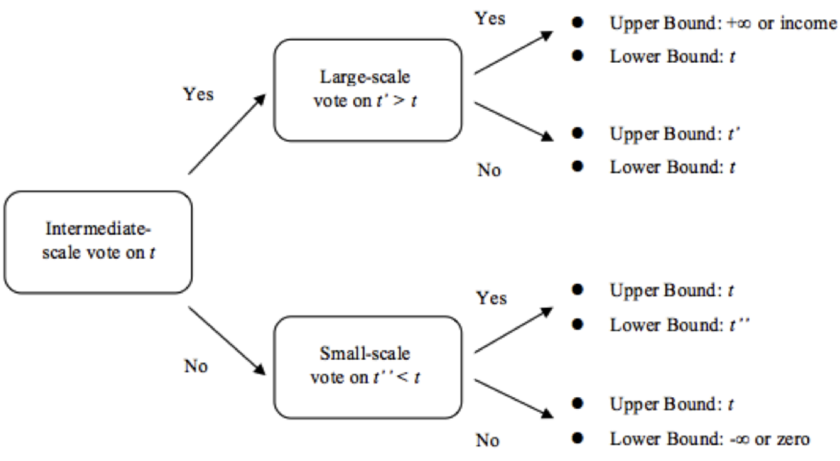

Figure 1: WTP Intervals for the Intermediate-Scale Option, Based on Responses for All Three Programs

introduced: ordering effect (Mitchell and Carson, 1993; Powe and Bateman, 2003; Clark and Friesen, 2008). However, this effect is simply a variant of starting point bias, already present in the DR method, and can be accounted for using question-order treatments in the administration of the survey instrument. Given these measures, there is no obvious reason why a respondent's answers to the three WTP questions in a QDR-based survey instrument should be interdependent, and the method should reduce the potential for bias relative to the DR method.[3]

It might appear that there is something inherently inefficient about asking respondents about three scenarios to improve estimates of one of them. We acknowledge this concern and therefore do not suggest undertaking this approach *for its own sake*. However, there are situations in which researchers may want to ask about multiple projects in the same survey for other research purposes (e.g., Boyle, Welsh, and Bishop, 1993; Poe, Welsh, and Champ, 1997; Powe and Bateman, 2003; Andersson and Svensson, 2008; Nielsen and Kjær, 2011). In these cases, the conditions needed for implementing the QDR approach are already in place, and the approach offers a simple and straightforward way to improve WTP estimates.

Two other alternatives which attempt to increase efficiency over the SR approach while mitigating bias of the DR approach are the One-and-One-Half-Bound (OOHB) approach proposed by Cooper, Hanemann, and Signorello (2002) and a simpler and more straightforward approach proposed by Haab and McConnell (2002). In the OOHB approach, respondents are told ahead of time that the exact cost of the program is unknown but is known to lie within a stated range. In addition to mitigating the potential for bias due to the unexpected nature of the follow-up question, they found that the OOHB method yields more efficient estimates than the SR and DR models. However, it is an open question as to whether the approach mitigates any incentives for strategic behavior. In the Haab and McConnell approach, they argue that reasonable bounds should be placed on the admissible range of WTP during estimation: a lower bound on WTP of zero (non-negativity), and an upper bound on WTP of respondent income. In spite of its reasonableness, it appears that

[3] Carson and Groves (2007) note that in a DR setting, correlation between the WTP distributions associated with the first and second responses should equal one as an indication of no shift in WTP. In our case, however, this is not possible (nor expected) since each of our scenarios represents a different scale of the good, so we should in fact *expect* a shift in WTP. We argue simply that since the three "different" goods are so closely related, we can use bounds from the larger and smaller scale goods as bounds on the intermediate scale good (where appropriate), so long as the assumption of non-decreasing utility in scale holds. Thus, we claim that what we require is that no sort of strategic (or similar) behavior is taking place that would render one or more of the three responses "untruthful." We admit that we cannot guarantee this, but we argue that the likelihood of this is reduced.

this recommendation has not been widely adopted. The obvious advantage of this approach is that it requires no additional questioning, other than eliciting respondent income, which is typically done anyway. The disadvantage is that it has the potential to introduce some error into the data if WTP is negative.

Survey and Data

Motivated by the issue of barrier-island restoration that was touted by Mississippi's state administration following Hurricane Katrina, a contingent valuation survey instrument was designed to measure Mississippi residents' WTP for maintenance and restoration of the Mississippi barrier island chain off the coast of Mississippi in the Gulf of Mexico at three different scales: their current condition (Status-Quo), their condition prior to Hurricane Camille in 1969 (Pre-Camille), and their condition prior to 1900 (Pre-1900).[4] These three programs differ only in quantity of acres restored and correspond to the small-scale project, the intermediate-scale project, and the large-scale project, respectively (because the islands have been disappearing steadily over time); our method uses responses from the Status Quo and Pre-1900 options to increase the efficiency of our estimates of WTP for the Pre-Camille option.[5] Respondents were given the following introduction: "The Mississippi barrier islands are continuously changing shape, size, and location. Overall, though, total land area has decreased by 36% since the 1850s, falling from a combined 10,290 acres to 6,545 acres. Although no exact predictions can be made, it is expected that they will continue to lose more land in the future." Respondents were informed that the islands are part of the Gulf Islands National Seashore and contain designated wilderness areas and beaches which are accessible to the public for recreation; are home to a wide variety of wildlife and separate the coastal waters from the open Gulf waters which help marine life reproduction and growth, provide calmer waters for commercial and recreational fishing and navigation, and reduce storm surge and wave energy during tropical storms, which can reduce storm damage on the mainland.

The Pre-Camille restoration project was described as follows:

> The Pre-Camille option would maintain existing land and restore the land area that has disappeared since Hurricane Camille made landfall in 1969. A total of 2,338 acres would be added to the island under this restoration plan, a 36% increase from current acreage.

Similar descriptions were given for the Status Quo and Pre-1900 options. Each of the three proposed projects was accompanied by a map showing the size of the barrier islands after project completion. Figure 2 presents these maps and illustrates the change in land area over time.

Each restoration option stated that the program would entail restoration and periodic maintenance for the next thirty years and would cost each Mississippi taxpaying household a one-time payment of $\$t$. One of five bids was assigned to each respondent for each restoration option: \$7, \$13, \$20, \$26, and \$33 for the Status-Quo option; \$77, \$153, \$230, \$306, and \$383 for the Pre-Camille Option; and \$195, \$391, \$586, \$782, and \$977 for the Pre-1900 option.[6] Bids corresponded such that, for example, a respondent who received the lowest bid for the Status-Quo scenario (\$7) also received the lowest bid for the other two scenarios (\$77 and \$195, respectively) to ensure that $t'' < t < t'$. The referendum question for each scenario read as follows:

> Suppose a State referendum were held today on the restoration option. A majority vote would be necessary to implement the project and, if passed, the payment would be

[4] The project was proposed as a joint venture between the State of Mississippi and U.S. Army Corps of Engineers.

[5] Note that we could also, in a similar spirit, use responses to the Pre-Camille question to narrow the WTP intervals on the Status-Quo and Pre-1900 scenarios.

[6] Project bids were based on barrier-island restoration cost estimates taken from T. Baker Smith & Sons (1997). Bids varied across surveys and were set as 50%, 100%, 150%, 200%, and 250% of baseline cost estimates.

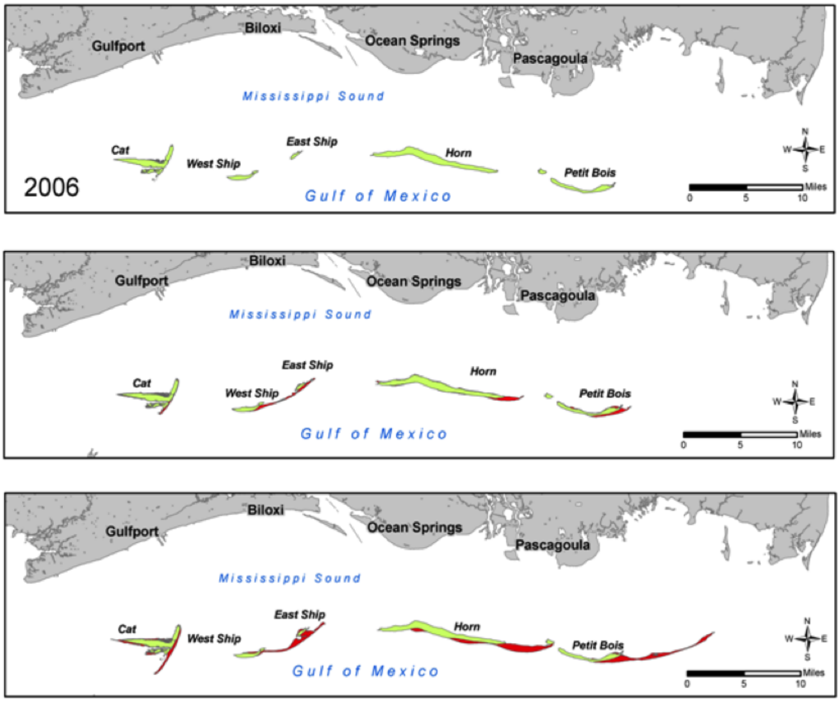

Figure 2: Restoration Option for (from Top to Bottom) the Status-Quo, Pre-Camille, and Pre-1900 Options (Carter and Blossom, 2007)

> collected on your 2008 state income tax return. Would you support the restoration option and therefore be willing to make a <u>one-time</u> payment of $\$t$ to implement it?
>
> **☐Yes ☐No**

Table 1 lists descriptive statistics of the independent variables with their expected signs. The sample was split in order to test the effect of question order of restoration options on responses and a question order variable included in the model. Scenarios were presented in order of ascending quantity (scale) to half the sample and in order of descending quantity to the other half. We have no expectations regarding sign, but significance is an indicator of a shifting effect in preferences from one question to the next.

A survey question was included to ascertain the relative importance of protection of the Mississippi coast among a set of general policy priorities including jobs, healthcare, and education. This was accomplished by asking respondents to allocate a hypothetical $100 to various policy priorities. The dollar amount allocated to protecting the coast and its residents from hurricanes is interpreted as an index of the importance of this priority and is included as the variable “priority.” We should expect the sign on this variable to be positive.[7]

The variables “cultural benefits,” “recreational benefits,” “ecosystem benefits,” and “storm surge benefits” indicate how respondents rate the importance of the barrier islands’ contribution to each of these key benefit types. The “residency” variable captures the effects of respondents who live in one of the three coastal Mississippi counties, and can be interpreted as a rough proxy for differences due to use value (versus non-use value for non-residents). The “no confidence” variable is a binary

---

[7] Although this question involves the “allocation” of dollars, it should not be interpreted as a welfare measure.

Table 1: Independent Variables and Their Descriptions

Variable	Type	Description	Mean	Std. Dev.	Exp.Sign
Residency	Binary	=1 if resident of Hancock, Harrison, and Jackson counties; =0 otherwise	0.68	0.47	+
Priority	Continuous	Budgetary priorities index for protection of the Mississippi coast from hurricanes and other hazards (range 0-100)	23.71	20.22	+
No Confidence	Binary	=1 if No confidence; =0 otherwise	0.55	0.50	-
Cultural Benefit	Binary	=1 if they responded "Extremely Important;" =0 otherwise.	0.63	0.48	+
Recreational Benefit	Binary	=1 if they responded "Extremely Important;" =0 otherwise.	0.63	0.48	+
Ecosystem Benefit	Binary	=1 if they responded "Extremely Important;" =0 otherwise.	0.82	0.39	+
Storm Surge Benefit	Binary	=1 if they responded "Extremely Important;" =0 otherwise.	0.77	0.42	+
Male	Binary	=1 if male; =0 female	0.61	0.49	+/-
White	Binary	=1 if white; =0 otherwise	0.89	0.32	+/-
Age	Continuous	Continuous between 18 - 94	56.13	13.88	+/-
Republican	Binary	=1 if Republican; =0 otherwise	0.53	0.50	-
Income	Ordered Categorical	=1 if < \$20K; =2 if \$20K -\$40K; =3 if \$40K-\$60K; =4 if \$60K-\$80K; =5 if \$80K-\$100K; =6 if \$100K or more	3.45	1.61	+
Comfort	Binary	=1 if comfortable answering survey questions; =0 otherwise	0.98	0.14	+
Question Order	Binary	=1 if order of restoration question is descending; =0 if ascending	0.46	0.50	+/-

indicator of whether the respondent has no confidence in government agencies to carry out the proposed project; its hypothesized sign is negative.

The “comfort” variable is a binary indicator of whether the respondent was comfortable taking the survey. Although theory says nothing regarding the sign of this variable, we hypothesize that those not comfortable are more likely to vote “no,” and thus the expected sign on this variable is positive. Finally, table 1 also describes several demographic variables.

Members of the staff at Mississippi State University participated in focus groups in order to improve survey quality and clarity. The survey instrument was mailed to 3,000 Mississippi households in February 2008, with half sent to a random sample of coastal residents (Hancock, Harrison, and Jackson counties) and half sent to a random sample across the remaining 79 non-coastal counties. Reminder postcards were mailed two weeks later. A total of 444 surveys were returned, for an overall response rate of 20%.[8] As reported in Petrolia and Kim (2009), the sample has a higher proportion of higher-income, more-educated, white respondents relative to the target population of Mississippi households. Further details regarding the survey can be found in Petrolia and Kim (2009).

Econometric Models

Following Cameron and James (1987), we model the unobserved WTP (Y_i) as a continuous random variable, which is a function of a vector of observed exogenous variables $\boldsymbol{x}_i$ and an error term ε_i, or

[8] Kruskal-Wallis equality-of-population rank tests and t-tests were used to test for treatment bias. No statistical differences in demographic indicators were found across treatments.

Table 2: Bounds on WTP for the SB and QDB Interval-Censored Models

Bounds on WTP using SB method

	Lower Bound		Upper Bound	
Bid	**N**	**Proportion**	**N**	**Proportion**
Left-Censored ($-\infty$)*	166	0.49		
\$77	43	0.13	24	0.07
\$153	42	0.13	29	0.09
\$230	31	0.09	39	0.12
\$306	29	0.09	27	0.08
\$383	25	0.07	47	0.14
Right-Censored ($+\infty$)*			170	0.51
Total	336		336	

Bounds on WTP using QDB method

	Lower Bound		Upper Bound	
Bid	**N**	**Proportion**	**N**	**Proportion**
Left-Censored ($-\infty$)*	56	0.17		
\$7	17	0.05		
\$13	16	0.05		
\$20	26	0.08		
\$26	20	0.06		
\$33	31	0.09		
\$77	43	0.13	24	0.07
\$153	42	0.13	29	0.09
\$195			11	0.03
\$230	31	0.09	39	0.12
\$306	29	0.09	27	0.08
\$383	25	0.07	47	0.14
\$391			17	0.05
\$586			9	0.03
\$782			11	0.03
\$977			12	0.04
Right-Censored ($+\infty$)*			110	0.33
Total	336		336	

Notes: Single asterisk (*) indicates left- and right-censored observations replaced with \$0 and respondent income, respectively, in bounded models.

$Y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i$. $\boldsymbol{\beta}$ is a vector of parameters to be estimated. From the responses to the WTP questions, we infer whether the respondent's true WTP is greater than or less than the offered price t. In the QDR model, the additional information collected from the small- and large-scale scenarios is incorporated directly into existing observations by updating bounds on WTP for the intermediate-scale scenario. This approach results in an interval-censored model (Hanemann, Loomis, and Kanninen, 1991) in which the dependent variable is composed of two data points: a lower and upper bound on WTP. For example, suppose a respondent voted "yes" to a \$77 bid at the intermediate scale and voted "no" to a \$195 bid at the large scale. These two observations would be combined into a single interval-censored observation on the WTP for the intermediate-scale project with \$77 as its lower bound and \$195 as its upper bound. In general, the contribution to the likelihood function of a respondent whose unobserved WTP is bounded on both sides (by t_{1i} on the left and t_{2i} on the right) is represented by the probability $\Pr(t_{1i} \leq Y_i \leq t_{2i})$.

Observations in which no useful information is gained from the small- and large-scale WTP questions simply remain left-censored ($\Pr(Y_i \le t_{Li})$) for "no" votes and right-censored ($\Pr(Y_i \ge t_{Ri})$) for "yes" votes, as they would in a SR model. Assuming normally distributed errors, the log-likelihood function then takes the form:

$$\begin{aligned}\log L \quad = \quad & \sum_{i\in L}\log\Phi\left(\frac{t_{Li}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)+\sum_{i\in R}\log\left\{1-\Phi\left(\frac{t_{Ri}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)\right\}+ \\ (1) \qquad & \sum_{i\in I}\log\left\{\Phi\left(\frac{t_{2i}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)-\Phi\left(\frac{t_{1i}-\boldsymbol{x}_i'\boldsymbol{\beta}}{\sigma}\right)\right\},\end{aligned}$$

where observations in L are left-censored, observations in R are right-censored, observations in I are interval-censored, and $\Phi(\cdot)$ is the standard cumulative normal distribution. Note that the traditional SR method is a special case of the interval-censored model in which there are no observations in I.

Most models of WTP assume an unbounded WTP range between $-\infty$ and $+\infty$. As discussed earlier, Haab and McConnell (2002) note that this can sometimes yield either negative or unreasonably large WTP estimates. They argue that reasonable bounds be placed on the admissible range of WTP: a lower bound on WTP of zero (non-negativity) and an upper bound on WTP of respondent income. Given the apparent lack of consensus on this issue, we estimate models under both assumptions: either the WTP range can be unbounded or bounded between zero and income (which we impose manually by updating the interval observations). Bounding WTP in this manner places all observations in I.

For the interval-censored models estimated, two dependent variables are specified for each respondent: a lower and an upper bound on WTP, represented by the offered bid. Table 2 reports the distribution of lower and upper bounds on WTP for the intermediate-scale Pre-Camille scenario between the SR and the QDR methods. Note that these imply WTP responses: a lower-bound represents a "yes" vote to the offered bid, and an upper-bound represents a "no" vote to the offered bid. The key difference with the QDR method relative to the SR method is the reduced incidence of infinite bounds (i.e., left- and right-censored observations) because these observations are "updated" with tighter bounds implied by responses to the smaller- and larger-scale scenarios. For the zero-income bounded models presented subsequently, the infinite bounds are replaced with \$0 and respondent income, respectively. Altogether, we compare four models: the unbounded and zero-income bounded SR and QDR models.

Results

Tables 3 and 4 show the coefficient estimates for the SR and QDR methods for the unbounded and bounded models, respectively. The estimated coefficients of the interval-censored model can be interpreted as the marginal contributions to WTP of the independent variables. In each of the four models, the likelihood ratio test of joint significance of the independent variables indicates that the models fit better than models with no independent variables.

For the unbounded models in table 3, all the coefficient signs in both the SR and QDR methods have the same sign as predicted in table 1, but only the priority, no confidence (QDR only), ecosystem benefit, storm surge benefit, white (SR only), and income variables are significant. Although parameter coefficients are not strictly comparable across models because they are being regressed on different dependent variables (statistically speaking), the explanatory variables in the QDR model generally have smaller estimated coefficients (i.e., smaller marginal contributions to WTP).

The zero-income bounded models presented in table 4 are generally similar to those in table 3 in terms of sign and significance, the only exceptions being that the no confidence variable is also significant in the SR model, whereas the white variable is not. The differences between the bounded SR and QDR models are much less pronounced, however, and both of these models are strikingly

Table 3: Estimated Coefficients for the Unbounded SR and QDR Models

	SR		QDR	
Variables	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
Residency	58.89	61.95	7.17	20.13
Priority	2.94*	1.51	1.19**	0.47
No Confidence	−56.90	54.30	−37.06**	18.53
Cultural Benefit	39.58	65.47	8.16	22.28
Recreational Benefit	40.73	63.24	21.82	21.36
Ecosystem Benefit	199.47**	92.05	55.46**	25.83
Storm Surge Benefit	154.73**	74.91	63.32***	22.71
Male	34.95	57.45	7.67	19.40
White	228.28*	121.49	42.91	31.83
Age	0.58	2.02	0.54	0.70
Republican	−71.34	62.02	−20.36	20.19
Income	74.47***	24.26	27.84***	6.19
Comfort	312.05	251.36	79.67	69.27
Question Order	87.10	60.30	4.28	18.18
Constant	−1,004.91	423.38	−129.02	86.58
Sigma	350.35	89.72	128.15	8.31
Observations		336		336
Log Likelihood		−191.47		−257.61
LR $\chi^2(13)$		70.00		68.05
Prob > χ^2		0.00		0.00

Notes: Single, double, and triple asterisks (*, **, ***) represent significance at the 10%, 5% and 1% level.

similar to the unbounded QDR model. In other words, applying Haab and McConnell bounds to the SR model produces results consistent with the unbounded QDR model.

Median WTP estimates and confidence intervals were obtained using the Krinsky and Robb (1986) procedure as outlined in Haab and McConnell (2002): a new parameter vector from a multivariate normal distribution is calculated as $\boldsymbol{\beta}_d = \hat{\boldsymbol{\beta}} + \boldsymbol{C}\boldsymbol{x}_K$, where $\hat{\boldsymbol{\beta}}$ is the vector of estimated parameters, $\boldsymbol{C}$ is the Cholesky decomposition of the variance-covariance matrix, and $\boldsymbol{x}_K$ is a vector of random variables drawn from the standard normal distribution. This procedure is repeated $N = 10{,}000$ times to generate an $N \times K$ matrix of simulated parameter estimates $\boldsymbol{\beta}_d^N$, from which median WTP is calculated N times. Mean and variance of median WTP are then calculated over these simulated observations.[9]

Table 5 reports means and variances of simulated median WTP. Differences in means between the SR and QDR methods were tested using t-tests (not assuming equal variance), and differences in variances were tested using ANOVA (i.e., F-test). Both means and variances were found to be statistically different at the 1% significance level under both the unbounded and bounded models. Nevertheless, the differences in means are not drastic: an \$8 difference across the unbounded models and a \$10 difference across bounded models. However, the differences in variances are quite pronounced between the unbounded models, indicating the ability of the QDR method to reduce variance with only minor impact on the mean. For the bounded models, the reduction in variance is marginal, but we should expect this since the ranges of individual WTP values will be relatively closer in the bounded QDR and SR models than in the unbounded QDR and SR models due to the imposition of zero and income bounds on WTP.

[9] To facilitate testing means, we report simulated median WTP rather that that obtained from estimated coefficients directly using $\bar{\boldsymbol{x}}'\boldsymbol{\beta}$ where $\bar{\boldsymbol{x}}$ is the vector of means. However, differences are trivial.

Table 4: Estimated Coefficients for the Bounded SR and QDR Models

	SR		QDR	
Variables	**Coeff.**	**Std. Error**	**Coeff.**	**Std. Error**
Residency	10.45	20.33	4.74	18.47
Priority	1.27***	0.48	1.10**	0.43
No Confidence	−34.01*	18.71	−39.00**	16.94
Cultural Benefit	4.54	22.58	3.05	20.43
Recreational Benefit	20.36	21.68	21.18	19.54
Ecosystem Benefit	56.29**	25.60	49.15**	23.36
Storm Surge Benefit	60.93***	22.90	58.06***	20.67
Male	6.90	19.56	3.51	17.69
White	35.85	31.35	31.43	28.65
Age	0.81	0.70	0.66	0.64
Republican	−19.53	20.48	−17.65	18.52
Income	29.47***	6.31	26.20***	5.64
Comfort	44.56	58.29	41.91	54.79
Question Order	10.86	18.45	5.69	16.60
Constant	−100.16	76.16	−61.80	70.08
Sigma	127.34	8.10	118.74	6.93
Observations		336		336
Log Likelihood		−244.43		−266.61
LR $\chi^2(13)$		66.38		66.76
Prob > χ^2		0.00		0.00

Notes: Single, double, and triple asterisks (*, **, ***) represent significance at the 10%, 5% and 1% level.

Table 5: Means and Variances of Median WTP.[a]

	Unbounded		Bounded	
	Mean	**Variance**	**Mean**	**Variance**
SR	$226.52	$695.60	$248.05	$93.19
QDR	$234.75	$82.69	$237.46	$71.28

Notes: [a] Means and Variances between SR and QDR method were statistically different at 1% significance level.

Conclusions

This study proposes a variant of the double-referendum (DR) contingent valuation method, referred to as the “quasi-double-referendum” (QDR) method, and we apply the method to a WTP study of barrier-island restoration in Mississippi. This study made two assumptions needed to apply the QDR method. The first assumption is that the only fundamental difference among the three scenarios is quantity. This assumption is likely to hold in our study, in which the main difference highlighted between the scenarios was the amount of land restored. The second assumption is that utility is non-decreasing in quantity. Although this assumption could be violated, it is difficult to think of a situation in which one would prefer less restoration, *ceteris paribus*.

We compared the QDR approach with the traditional single referendum (SR) approach, allowing the range of willingness to pay to be in the unbounded range $[-\infty, +\infty]$ and also restricting the range of willingness to pay to be bounded between zero and income. In each case, the QDR approach yielded smaller variance. Median WTP estimates using the QDR method were slightly higher than those of the SR method in the unbounded models, but slightly lower in the bounded models. Overall, though, the result was substantially reduced variance with only slight change in mean.

We believe that the method can be applied to any survey that elicits willingness-to-pay responses to multiple programs differing only in scale and can be used to increase the efficiency of welfare estimates while mitigating biases introduced by a more traditional DR survey. Thus, the proposed methodology should be a valuable alternative to obtaining welfare estimates for a variety of restoration issues involving mining, reforestation, riparian areas, or drought/flood impacted areas.

This study suggested that use of the QDR method should mitigate the issues of starting point, shifting effect, and strategic bias that can occur in the DR method. By using three scenarios that differ in scale, a variant of starting-point bias is introduced: bias due to the order in which the three programs are presented. Results indicate, however, that no significant question order effect existed. Shifting effect bias, in which respondents view a change in the offered price as a signal that the quality or quantity of the program has been changed, should not be of concern in the QDR method since three unique, explicitly characterized, programs are evaluated.

Lastly, a DR survey will suffer from strategic bias if respondents view a change in the offered price as a signal to bargain. We believe that our approach mitigates this possibility by informing respondents ahead of time that they will be evaluating three unique programs at different scales and prices. At the same time, however, another form of strategic bias could be introduced. When respondents are presented with multiple competing programs, they may, for example, state that they are not willing to pay for a given program despite actually being willing to do so, in order to increase the chances of one of the other, more preferred, programs being implemented. This is a common issue with any multiple program survey, particularly the commonly used choice experiment survey approach. Research in this area is ongoing.

[Received December 2011; final revision received March 2012.]

References

Alberini, A. "Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single- bound, Double-Bound, and Bivariate Model." *Journal of Environmental Economics and Management* 28(1995):287–306.

Alberini, A., B. Kanninen, and R. T. Carson. "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data." *Land Economics* 73(1997):309–324.

Altaf, M. and J. DeShazo. "Bid Elicitation in the Contingent Valuation Method: The Double Referendum Format and Induced Strategic Behavior." Urban Planning Committee Working Paper, Harvard University, 1994.

Andersson, H. and M. Svensson. "Cognitive Ability and Scale Bias in the Contingent Valuation Method: An Analysis of Willingness to Pay to Reduce Mortality Risk." *Environmental and Resource Economics* 39(2008):481–495.

Bishop, R. C. and T. A. Heberlein. "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61(1979):926–930.

Boyle, K. J., M. P. Welsh, and R. C. Bishop. "The Role of Question Order and Respondent Experience in Contingent-Valuation Studies." *Journal of Environmental Economics and Management* 25(1993):S80–99.

Cameron, T. A. and M. D. James. "Estimating Willingness to Pay from Survey Data: An Alternative Pre-Test-Market Evaluation Procedure." *Journal of Marketing Research* 24(1987):389–395.

Cameron, T. A. and J. Quiggin. "Estimation Using Contingent Valuation Data from a 'Dichotomous Choice with Follow-up' Questionnaire." *Journal of Environmental Economics and Management* 27(1994):218–234.

Carson, R. T. *Three Essays on Contingent Valuation*. Ph.D. thesis, University of California, Berkeley, 1985.

Carson, R. T. and T. Groves. "Incentive and Informational Properties of Preference Questions." *Environmental and Resource Economics* 37(2007):181–210.

Carson, R. T., W. M. Hanemann, and R. C. Mitchell. "Determining the Demand for Public Goods by Simulating Referendums at Different Tax Price." Working paper, University of California San Diego, Department of Economics, 1986.

Carson, R. T., R. C. Mitchell, W. M. Hanemann, R. J. Kopp, S. Presser, and P. A. Ruud. "A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill." Report to the Attorney General of Alaska, NRDA, Inc., La Jolla, CA, 1992.

Carson, R. T., L. Wilks, and D. Imber. "Valuing the Preservation of Australia's Kakadu Conservation Zone." *Oxford Economic Papers* 46(1994):727–749.

Carter, G. and G. Blossom. "Unpublished Maps." 2007. Gulf Coast Geospatial Center, Gulf Coast Research Lab, University of Southern Mississippi.

Clark, J. and L. Friesen. "The Causes of Order Effects in Contingent Valuation Surveys: An Experimental Investigation." *Journal of Environmental Economics and Management* 56(2008):195–206.

Clarke, P. M. "Valuing the Benefits of Mobile Mammographic Screening Units Using the Contingent Valuation Method." *Applied Economics* 32(2000):1647–1655.

Cooper, J. C., M. Hanemann, and G. Signorello. "One-and-One-Half-Bound Dichotomous Choice Contingent Valuation." *Review of Economics and Statistics* 84(2002):742–750.

Flachaire, E. and G. Hollard. "Controlling Starting-Point Bias in Double-Bounded Contingent Valuation Surveys." *Land Economics* 82(2006):103–111.

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

Hanemann, M., J. Loomis, and B. Kanninen. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73(1991):1255–1263.

Hanemann, W. M. "Some Issues In Continuous–and Discrete–Response Contingent Valuation Studies." *Northeastern Journal of Agricultural and Resource Economics* 14(1985):5–13.

Herriges, J. A. and J. F. Shogren. "Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning." *Journal of Environmental Economics and Management* 30(1996):112–131.

Kanninen, B. J. "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation." *Land Economics* 69(1993):138–146.

Krinsky, I. and A. L. Robb. "On Approximating the Statistical Properties of Elasticities." *The Review of Economics and Statistics* 68(1986):715–719.

McLeod, D. M. and O. Bergland. "Willingness-to-Pay Estimates Using the Double-Bounded Dichotomous-Choice Contingent Valuation Format: A Test for Validity and Precision in Bayesian Framework." *Land Economics* 75(1999):115–125.

Mitchell, R. C. and R. T. Carson. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, D.C.: Resources for the Future, 1993.

Nielsen, J. S. and T. Kjær. "Does Question Order Influence Sensitivity to Scope? Empirical Findings from a Web-Based Contingent Valuation Study." *Journal of Environmental Planning and Management* 54(2011):369–381.

Petrolia, D. R. and T. Kim. "What are Barrier Islands Worth? Estimates of Willingness to Pay for Restoration." *Marine Resource Economics* 24(2009):131–146.

Poe, G. L., M. P. Welsh, and P. A. Champ. "Measuring the Difference in Mean Willingness to Pay When Dichotomous Choice Contingent Valuation Responses are Not Independent." *Land Economics* 73(1997):255–267.

Powe, N. A. and I. J. Bateman. "Ordering Effects in Nested 'Top-Down' and 'Bottom- Up' Contingent Valuation Designs." *Ecological Economics* 45(2003):255–270.

Watson, V. and M. Ryan. "Exploring Preference Anomalies in Double Bounded Contingent Valuation." *Journal of Health Economics* 26(2007):463–482.