Journal of Agricultural and Resource Economics 39(2):201–216 ISSN 1068-5502

The Effects of Consequentiality in Binary- and Multinomial-Choice Surveys

Matthew G. Interis and Daniel R. Petrolia

We examine the effect of respondent perceptions of consequentiality on a split-sample binary-choice/multinomial-choice stated-preference survey. We fail to observe the knife-edge results predicted in the consequentiality literature in the binary-choice data but do observe them in the multinomial-choice data. In the multinomial-choice data, only respondents who believed the survey was at least somewhat likely to affect future policy behaved consistently with theoretical predictions. Overall, we conclude that failing to control for perceived consequentiality can lead to false conclusions regarding marginal effects and welfare estimates. This is true in both the binary- and multinomial-choice contexts.

Key words: binary choice, consequentiality, knife-edge results, Louisiana coastal wetlands, multinomial choice, willingness to pay

Introduction

Carson and Groves (2007, 2011) brought the issue of consequentiality to the forefront of stated-preference survey research, arguing that as long as a survey is perceived by the respondent as consequential, researchers should be able to predict how he will respond given the incentive structure. A survey question is consequential if the agent believes his response will affect some outcome that he cares about. From such questions we can expect "useful information" (Carson and Groves, 2007, p. 183). Theory, however, says nothing about how those for whom a survey is not perceived as consequential should behave. A natural empirical question is then: do respondents who believe the survey to be inconsequential behave differently from those who believe the survey to be consequential?

This question has to some extent been answered for the binary-choice (BC) setting. In a BC stated-preference survey, respondents choose between only two alternatives, such as whether they are for or against a proposed policy. In lab and field experiments, Vossler and Evans (2009) and Landry and List (2007) find that respondents in inconsequential treatments behave statistically differently from respondents in consequential treatments. Furthermore, they find evidence to support the prediction that respondents in treatments in which the probability of affecting the outcome is strictly between zero and one behave statistically the same as those in treatments with a probability of one. Bulte et al. (2005), who compare the effects of three treatments in a field survey, find that respondents to a treatment in which hypothetical bias (i.e., people's statements of willingness to pay in a hypothetical question differ from the amount they would be willing to pay were the question not hypothetical) was directly discussed behave the same as respondents to a treatment in which it was

Matthew G. Interis is an assistant professor and Daniel R. Petrolia is an associate professor in the Department of Agricultural Economics at Mississippi State University.
The authors thank the editor, two anonymous reviewers, and participants in the W3133 USDA multistate research project for comments which improved the manuscript. 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 supported by the USDA National Institute of Food and Agriculture, Multistate Project W-3133 "Benefits and Costs of Natural Resources Policies Affecting Ecosystem Services on Public and Private Lands" (Hatch Project MIS-033130).

Review coordinated by Gregmar I. Galinato.

directly stated that survey responses would be seen by policy makers. They find that respondents in a purely hypothetical treatment behave differently from the aforementioned groups.

Consistent with these findings, Herriges et al. (2010) find that, in a field setting, inconsequential respondents behave differently from other respondents with varying degrees of perceived consequentiality. Specifically, they find similar willingness-to-pay distributions among respondents who find the survey to be at least minimally consequential, whereas respondents who believed the survey to be inconsequential have statistically lower willingness-to-pay distributions. Vossler and Watson (2013) also find in their field survey of a proposed restoration program in Massachusetts that respondents who believed that survey results would "Not [be] taken into account" by policy makers were statistically less likely to vote in favor of the program than respondents who believed there was at least some chance of the survey results being taken into account.

The results of these studies are consistent with the knife-edge[1] result implied by Carson and Groves (2007): respondents who perceive there to be at least some strictly positive probability of responses affecting actual policy should behave similarly to each other because they face the same incentive structure, whereas theory can make no predictions about how respondents who do not perceive the survey to be consequential should behave. The knife-edge that separates respondents by incentive structure occurs at a probability of exactly zero. However, Vossler, Doyon, and Rondeau (2012) conduct a repeated BC field experiment that allows for choice attribute effects; they fail to observe knife-edge results. Although they find that the behavior of respondents who perceived "more than a weak" degree of consequentiality was statistically indistinguishable from participants in a real-choice (i.e., not hypothetical) treatment; they also find that willingness to pay for attribute increments decreases with the degree of consequentiality. Under knife-edge results, one would expect the preferences of all participants who believed their choices to have a strictly positive probability of affecting the outcome to be the same.

To our knowledge, the issue of consequentiality in a multinomial-choice (MC) setting—when survey respondents choose among more than two alternatives (e.g., three or more alternative proposed policies)—has not been addressed. The popular choice-experiment approach to valuation often uses a MC format. Despite the fact that consequentiality is unlikely to aid incentive compatibility in a MC setting as it does in a BC setting (because MC questions are incentive compatible only under extremely restrictive conditions), it is nevertheless a necessary condition for theory to be capable of making predictions in a MC setting.[2] Furthermore, the popularity of the MC setting stems from the fact that it allows for a more direct means of estimating attribute effects. Perceptions of consequentiality may have effects on attribute estimates in the MC setting not heretofore examined in the literature. Therefore, it is important to examine the implications of consequentiality in a MC setting.

As in the BC setting, the proposition that any response to an inconsequential MC survey gives respondents the same expected level of utility still holds. At the same time, the empirical evidence on BC surveys indicates that respondents for whom the survey is perceived to be inconsequential are more likely to vote against the action alternative relative to respondents who indicate that the survey is consequential. Given this evidence, we might expect to find a higher proportion of no-action votes among inconsequential respondents in the MC setting as well.

Furthermore, in the MC setting, the rationale that any response gives inconsequential respondents the same expected level of utility should extend to the issue of attribute effects. In other words, if—in the eyes of an inconsequential respondent—any alternative is as good as any other, we should expect to find undesirable results (i.e., lack of significance or inconsistent signs for such respondents). For example, Scarpa, Thiene, and Hensher (2010) find evidence of attribute

[1] The term "knife-edge" to describe possible behavioral differences between respondents who believe the survey to be consequential and those who do not was first used in Herriges et al. (2010).

[2] Carson and Groves (2007, 2011) argue that any response in an inconsequential survey gives the respondent the same expected utility, so there can be no predictions made based on economic theory. This holds true for either a BC or MC question format.

nonattendance in MC settings and note that it is likely to be a function of personal relevance of the survey topic. They stop short, however, of identifying the source of attribute nonattendance and leave it as a topic for future research. We hypothesize here that perceived inconsequentiality of the survey results in attribute nonattendance.[3]

We also specifically examine knife-edge results in an MC setting, which has not been done before. Understanding whether knife-edge results hold in an MC setting has implications for the format of questions eliciting respondent consequentiality perceptions; if knife-edge results hold in an MC setting, then a simple yes/no consequentiality question suffices. If knife-edge results do not hold and, for example, value estimates vary by consequentiality perceptions, then a more sophisticated question eliciting consequentiality perceptions, which perhaps offers some kind of range in confidence of the survey's consequentiality, is required. The presence or absence of knife-edge results in an MC setting also has implications for decisions about handling respondents who perceive the survey to be relatively less consequential or not consequential at all—for example, whether to weight their responses or to exclude them from data analysis altogether.

We test these hypotheses using a stated-preference survey for Louisiana coastal wetland and barrier island restoration which includes both BC and MC treatments, the differences between which were designed to be minimal. The results of the BC treatment are inconsistent with the knife-edge prediction of Carson and Groves (2007) and instead show that willingness to pay increases with perceived consequentiality. In the MC treatment, we find that failing to control for perceptions of consequentiality lowers the apparent construct validity of the instrument; respondents who believe the survey to be consequential are more sensitive to project attributes and behave consistently with scope predictions whereas respondents who do not believe the survey to be consequential exhibit behavior inconsistent with theoretical predictions. We also check for knife-edge results in the MC treatment. We find knife-edge results regarding sensitivity to the offered bid and choice attributes but also find that the value of the intercept increases with perceived consequentiality.

Experimental Design

We designed a split-sample binary-choice/multinomial-choice survey instrument to estimate welfare associated with coastal wetland and barrier island restoration in Louisiana's Barataria-Terrebonne National Estuary.[4] Louisiana has been one of the states most affected by wetland loss in the United States. Couvillion et al. (2011) estimate that coastal Louisiana has undergone a net change in land area of approximately 1,883 square miles from 1932 to 2010, representing a 25% decrease, which is an area of land about the size of the state of Delaware. These losses in Louisiana account for about 90% of total wetland loss in the lower forty-eight states (Couvillion et al., 2011). Of Louisiana's nine major basins, the Barataria and Terrebonne basins have the highest and second highest land loss rates, respectively. Furthermore, these two basins alone account for 49% of total persistent losses across all basins (Couvillion et al., 2011).

The survey instrument included language that explained to respondents that wetlands and barrier islands in the estuary were being lost due to "natural erosion, sea-level rise, sinking of land, winds, tides, currents, and major storms," as well as human development such as the construction of river channels and levees. The project under consideration was a large-scale land restoration project which included "wetland building, barrier island restoration, freshwater and sediment diversions, and the movement of large amounts of soil on barges and via pipelines." The survey focused on three main benefits of restoration: improved wildlife habitat, measured as the percentage of created land generally suitable for wildlife habitat; storm surge protection, measured as the percentage of

[3] Vossler, Doyon, and Rondeau (2012), who capture attribute effects via repeated BC format, find that marginal WTP for attribute increments decreases as consequentiality increases. While this result does not directly address attribute attendance, it is some indication that respondents' reactions to attributes are a function of consequentiality.

[4] The full survey is available as a supplement on the journal's website (http://www.waeaonline.org/publications/jare). Also, see Petrolia, Interis, and Hwang (2014) for more background on land loss in the estuary.

	With Project: 50% of lost land restored	**Without Project (No Action):** Land loss expected to continue at 4,500 to 7,100 acres per year
Wildlife habitat	50% of restored land suitable as habitat	No additional habitat and current habitat expected to decline
Storm surge protection	Improved protection for 30% of residents	No improvement and current protection expected to decline
Commercial fish harvest	15% higher harvest levels	No improvement and current harvest levels expected to decline
Share of total cost to your household (one-time tax)	$925	$0
I prefer:	☐	☐

I prefer not to vote: ☐

Figure 1. Example Binary-Choice Valuation Question

	Project A: 50% of lost land restored	**Project B: 50% of lost land restored**	**No Action: Land loss expected to continue at 4,500 to 7,100 acres per year**
Wildlife habitat	25% of restored land suitable as habitat	50% of restored land suitable as habitat	No additional habitat and current habitat expected to decline
Storm surge protection	Improved protection for 5% of residents	Improved protection for 30% of residents	No improvement and current protection expected to decline
Commercial fish harvest	Maintains current harvest levels	15% higher harvest levels	No improvement and current harvest levels expected to decline
Share of total cost to your household (one-time payment)	$155	$285	$0
I prefer:	☐	☐	☐

☐ **I prefer not to vote**

Figure 2. Example Multinomial-Choice Set

residents in the area that would have improved storm surge protection; and improved commercial fish harvest, measured as the percentage improvement in harvest levels of major commercial (Gulf of Mexico) fish such as oysters and shrimp. Respondents were randomly assigned to either the BC or the MC treatment. An example choice question for each treatment is given in figures 1 and 2.[5] Each respondent in the MC treatment was assigned to one of twelve blocks and presented with two alternative restoration programs with attributes set at one of the levels shown in table 1 and

[5] Hwang, Petrolia, and Interis (2014) examine the respondent decision to opt out of answering the choice questions.

Table 1. Attribute Levels and Descriptions

	Action Alternatives: *50% of lost land restored*			**No Action Alternative (SQ):** *Land loss expected to continue at 4,500 to 7,100 acres per year*
	Low	*Medium*	*High*	
Wildlife habitat: x% of restored land suitable as habitat	25%	50%	75%	No additional habitat and current habitat expected to decline
Storm surge protection: improved protection for x% of residents	5%	30%	50%	No improvement and current habitat expected to decline
Commercial fisheries harvest: x% higher harvest levels	Maintains current harvest levels	15%	30%	No improvement and current harvest levels expected to decline
Price: $x one-time tax	$25, $90, $155, $285, $545, $925			$0

a status quo alternative.[6] Respondents in the BC treatment were presented with a single project at the medium attribute levels and with a price in the set {$25, 90, 155, 285, 545, 925, 1,305, 2,065, 2,825}, which were determined after focus group feedback and pretesting. To avoid any confounding effects related to repeated choices, each respondent in both treatments was presented with exactly one choice set (see for example McNair, Bennett, and Hensher, 2011; Day et al., 2012).

The survey was subject to a fairly extensive vetting process. First, the researchers met with staff at the Barataria-Terrebonne National Estuary Program center in Thibodaux, Louisiana, to discuss the feasibility and believability of projects like the one proposed in the survey, the relevant project attributes that people would most likely care about, etc. Two focus groups were held in early 2011 using staff from various departments at Mississippi State University,[7] the first of which was used only to narrow down the appropriate attributes for the survey and the second of which focused on a more complete version of the survey to check for clarity, bias, etc. The survey instrument was then pretested through Knowledge Networks, who administered the survey to approximately thirty respondents, the main focus being on checking the sensitivity of respondents to the restoration program prices used in the survey (i.e., that respondents were less likely to favor a restoration program offered at a higher price, all else equal). After the pretest we decided to add higher prices ($1,305, $2,065, and $2,825) to the BC treatment because there were still a high proportion of favorable responses to the then highest price of $925. A pilot test was then administered to approximately 100 respondents, which left us satisfied with our range of restoration program prices. The final version of the survey was administered by Knowledge Networks to a random sample of U.S. households that were part of their Knowledge Panelï£¡ in June and July of 2011. Out of 4,402 people sampled, 3,464 (77%) responded. After eliminating incomplete responses, there were 3,297 usable observations, of which 1,974 were in the MC treatment and the remainder of which were in the BC treatment.

Respondents were informed that the results of the survey were "advisory," meaning that they would "be used to inform policy makers on the opinion and preferences of taxpayers to help them decide if and what projects should be funded. . . " The respondent's perception of the consequentiality of the survey on future policy was elicited with this question, which followed the choice question:

[6] The choice experiment was designed to maximize the D-efficiency of a fractional factorial design, using the %MktEx macro in the statistical program SAS (see Kuhfeld, 2010).

[7] These participants were deliberately chosen not to be "experts" in anything related to the study since our target population was the general U.S. population.

Table 2. Description of Variables and Descriptive Statistics

		BC Data (N = 1,323)		MC Data (N=1,974)	
	Description	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Education	Highest education received. 9 = high school graduate, 10 = some college, no degree, 11 = associate degree	10.32	1.98	10.29	2.04
Conservative	1 = extremely liberal, 7 = extremely conservative	4.23	1.48	4.20	1.49
DWH	= 1 if followed the Deepwater Horizon oil spill somewhat or very closely, 0 if not at all	0.84	0.37	0.85	0.36
Changes	= −1 if made no changes, 0 if minor, 1 if major changes to lifestyle for environmental reasonsa	−0.08	0.60	−0.06	0.60
Income	Income range. 11 = $40-50K, 12 = $50-60K	11.87	4.36	12.05	4.37
Age	Age, in years	48.80	16.74	48.99	16.87
Male	= 1 if male	0.50	0.50	0.49	0.50
Minority	= 1 if not white	0.25	0.44	0.24	0.43

Notes: [a] The counts for no, major, and minor changes were 292, 838, and 193 for the BC treatment and 420, 1,247, and 307 for the MC treatment

How likely do you think it is that the results of this survey will shape the direction of future policy in the Lower Barataria-Terrebonne Estuary?

a) Very likely
b) Somewhat likely
c) Unlikely
d) I don't know

Carson and Groves (2007) argue that the incentive properties of a binary-choice advisory survey hold as long as the respondent perceives some positive probability that an increase in the proportion of survey responses in favor of the project will increase the likelihood that the actual project is implemented. In other words, there is no theoretical difference in the incentive properties facing a respondent who perceives it "very likely" that the survey outcome will affect implementation versus those facing a respondent who perceives it as "somewhat likely". Herriges et al. (2010) find empirical evidence in favor of this position, referring to it as a "knife-edge" relationship. Vossler, Doyon, and Rondeau (2012), however, find that the incentive properties are triggered at a relatively higher perception of consequentiality. Furthermore, they find that the marginal willingness to pay for project attributes decreases with consequentiality. Thus, although our scale is an abbreviated version of that used in the previous literature, it still allows for some gradient of perceived consequentiality to further investigate whether the knife-edge result holds. Our response options allow for "I don't know" responses, which previous literature has not, and which allows us to test for any differences in how such respondents behave. We included this option because we thought it reasonable that some respondents would not have well-defined perceptions about the consequentiality of the survey.

We also asked how closely respondents followed news about the *Deepwater Horizon* explosion and oil spill in the Gulf of Mexico which occurred in April 2010, just over one year before the survey was administered, because we hypothesized that for some respondents willingness to pay for the project might be driven by recent media exposure of the region of Louisiana under consideration. We also asked whether they had made changes in their lifestyle within the past five years to help protect the environment and to rate their political ideology on a seven-point scale. Finally, standard demographic data (age, gender, etc.) were precollected by Knowledge Networks.

Data

Descriptive statistics and descriptions of the individual-specific demographic variables are reported in table 2.[8] One of the advantages of using Knowledge Networks is that they guarantee a

[8] The data are available from the authors upon request.

representative sample for the general United States population (our population of interest). The variables for education, conservative, income, age, male, and minority were precollected by Knowledge Networks. Means are similar across consequentiality types, within treatments, for all variables except whether respondents had made changes to their lifestyle for environmental reasons.[9] A well-known result in the literature is that more environmentally oriented individuals tend to be more likely to vote in favor of environmental programs like the ones presented in our survey. Thus, we are cautious about treating the responses to the consequentiality question as exogenous and we formally test for this (discussed later).

Table 3 contains the breakdown of votes in favor of and opposed to project implementation by consequentiality type. To facilitate a simple comparison of votes for and against project implementation between the BC and MC data, votes in favor of either project A or project B in the MC data are pooled into a single "In Favor" category.[10] Lowercase letters correspond to pairwise proportions tests of equal vote proportions across consequentiality responses within data sets. In the BC data, the proportion of respondents in favor of implementing the project increases with perceived consequentiality and the proportions for the "unlikely" and "I don't know" groups do not statistically differ from one another. In the MC data the proportions for the "very likely" and "somewhat likely" respondents are not statistically different, with a greater proportion in favor of the project relative to the "unlikely" respondents. Also, the proportions for the "unlikely' and "I don't know" respondents statistically differ. Based on this initial examination of the data, the predicted knife-edge results appear to hold for the MC data only.[11]

Regression Analysis

Let the utility for individual i from option j with perceived consequentiality response k be described as

$$U_{ijk} = \alpha_{jk} + \boldsymbol{\beta}_k' \boldsymbol{x}_{ij} + \boldsymbol{\delta}_j' \boldsymbol{z}_i + \varepsilon_{ijk}, \tag{1}$$

where α_{jk} is an alternative-specific constant for perceived consequentiality response k, $\boldsymbol{\beta}_k$ is a vector of coefficients on choice-specific attribute levels $\boldsymbol{x}_{ij}$ (including bid) for perceived consequentiality response k; $\boldsymbol{\delta}_j$ is a vector of coefficients on individual-specific characteristics $\boldsymbol{z}_i$ for option j; and ε_{ijk} is an error term. For the BC model, the vector $\boldsymbol{\beta}_k$ in equation (1) contains bid only and thus reduces to the scalar β_k. This specification allows for identification of the effects of consequentiality via both the intercept and slopes of alternative-specific attributes.

One fundamental implication of Carson and Groves' model of consequentiality is that respondents who believe their responses are inconsequential derive the same expected utility from any response to the willingness-to-pay question (because their responses do not affect anything they care about). On the other hand, respondents who believe their responses are consequential should behave consistently with the incentive structure of the choice question. Carson and Groves (2007) and Carson, Groves, and List (2014) show that respondents in a BC survey who perceive the survey to be consequential,[12] to whatever degree, face the same incentive structure. Herriges et al. (2010)

[9] Demographic descriptive statistics broken down by consequentiality type are available from the authors upon request.

[10] The MC results in table 3 are admittedly an oversimplified picture of the data since there were twelve different choice sets, each with two alternatives that varied by project attributes. We do not wish to imply that any strong conclusions should be drawn from this picture but provide it for comparison to the BC data and to avoid the clutter of breaking down the data by choice set.

[11] Any elicitation of perceptions of the consequentiality of the survey must be subjective, so differences between "unlikely," "somewhat likely," and "very likely" also have only a subjective interpretation. We acknowledge that to test for knife-edge results necessarily implies that our perception scale is meaningful and that we can never be more confident in our results than we are in the meaningfulness of our scale. It is also true that a variety of scales are used in the literature, which increases the difficulty of identifying the true thresholds along this subjective continuum.

[12] Another necessary condition is that respondents believe they will have to pay if the program is implemented. Unfortunately, we did not elicit their beliefs about this, and we must therefore make the assumption that respondents believed the one-time tax would definitely be imposed if the project were to be implemented.

Table 3. Vote Tally (Proportions) by Consequentiality Response

Binary-Choice Data	**In Favor**	**Opposed**
Very Likely[a]	61	24
	(72%)	(28%)
Somewhat Likely[b]	288	218
	(57%)	(43%)
Unlikely[c]	139	295
	(32%)	(68%)
I Don't Know[c]	85	213
	(29%)	(71%)
Total	573	750
	(43%)	(57%)

Multinomial-Choice Data	**In Favor**	**Opposed**
Very Likely[a]	75	22
	(77%)	(23%)
Somewhat Likely[a]	604	187
	(76%)	(24%)
Unlikely[b]	349	280
	(55%)	(45%)
I Don't Know[c]	183	274
	(40%)	(60%)
Total	1,211	763
	(61%)	(39%)

Notes: "In Favor" indicates a vote for either project A or project B. [a], [b], and [c] indicate results of pair-wise proportions tests. Groups that share the same letter are not statistically different from each other but are statistically different from groups with other letters. Comparisons were made within each data set only, not across data sets.

refer to differences in behavior between these two groups of respondents with differing incentive structures as knife-edge results. If we observe knife-edge results, then:

- H01: In the BC treatment, the constant and attribute parameters are equal for respondents who answered "very likely" or "somewhat likely" to the consequentiality question.
- H02: In the BC treatment, these same parameters differ from those respondents who answered "unlikely" to the consequentiality question.

In the BC model, we test these by examining the constant and attribute (which, in the BC model, includes only the bid value because other project attributes do not vary) coefficient estimates for each consequentiality response.

While Carson and Groves (2007) argue that respondents in a BC survey who believe their responses to be consequential have an incentive to choose their more preferred alternative, there is no equivalent theoretical prediction that respondents in a MC choice situation (> two options) should always choose their most preferred alternative. The reason is that under the assumption that the alternative that receives the most respondent votes is most likely to be implemented it is well known that choices in MC situations generally depend on the respondent's perceptions of how others will choose (Myerson and Weber, 1993). In particular, when a respondent believes his most preferred alternative is likely to garner the fewest votes, he has an incentive to choose his second-most preferred alternative. Nevertheless, the underlying assumption in econometric models used to estimate the parameters of MC models is that respondents choose the alternative that gives them the greatest utility (Greene, 2012). Just as in a BC survey, respondents in a MC survey who do not believe their responses to be consequential do not (necessarily) fall within this utility maximization framework. Furthermore, two respondents with the same preferences and the same

Table 4. Likelihood Ratio Tests of Parameter Equivalence across Consequentiality Responses (Binomial Choice Data)

Null Hypothesis	χ^2	p-value
Knife-edge tests		
$H_0 : \beta_{Very} = \beta_{Somewhat}$	0.31	0.58
$H_0 : \beta_{Very} = \beta_{Somewhat}, \alpha_{Very} = \alpha_{Somewhat}$	5.97*	0.05
$H_0 : \beta_{Very} = \beta_{Somewhat} = \beta_{Unlikely}$	0.39	0.82
$H_0 : \beta_{Very} = \beta_{Somewhat} = \beta_{Unlikely}, \alpha_{Very} = \alpha_{Somewhat} = \alpha_{Unlikely}$	61.29***	0.00
"I don't know" equality tests		
$H_0 : \beta_{Very} = \beta_{Don'tKnow}$	0.30	0.58
$H_0 : \beta_{Very} = \beta_{Don'tKnow}, \alpha_{Very} = \alpha_{Don'tKnow}$	33.52***	0.00
$H_0 : \beta_{Somewhat} = \beta_{Don'tKnow}$	0.01	0.94
$H_0 : \beta_{Somewhat} = \beta_{Don'tKnow}, \alpha_{Somewhat} = \alpha_{Don'tKnow}$	35.43***	0.00
$H_0 : \beta_{Unlikely} = \beta_{Don'tKnow}$	0.14	0.70
$H_0 : \beta_{Unlikely} = \beta_{Don'tKnow}, \alpha_{Unlikely} = \alpha_{Don'tKnow}$	0.17	0.92
Joint equality tests		
$H_0 : \beta_{Very} = \beta_{Somewhat} = \beta_{Unlikely} = \beta_{Don'tKnow}$	0.46	0.93
$H_0 : \beta_{Very} = \beta_{Somewhat} = \beta_{Unlikely} = \beta_{Don'tKnow}, \alpha_{Very} = \alpha_{Somewhat} = \alpha_{Unlikely} = \alpha_{Don'tKnow}$	78.48***	0.00

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively. β_k: bid coefficient. α_k: constant coefficient.

beliefs about how others will choose but who have different, but strictly positive, perceptions about the consequentiality of their responses face the same incentives. We therefore might expect to observe knife-edge results in an MC survey. We therefore test whether:

- H03: In the MC treatment, the constant and attribute parameters are equal for respondents who answered "very likely" or "somewhat likely" to the consequentiality question.
- H04: In the MC treatment, these same parameters differ from those respondents who answered "unlikely" to the consequentiality question.
- H05: In the MC treatment, respondents who answered "unlikely" to the consequentiality question are insensitive to choice attributes.

We test H05 by testing whether choice attribute coefficient estimates are equal to zero.

The binary-choice model is estimated using a logit model (see Haab and McConnell, 2002), which assumes that the respondent chooses alternative j instead of alternative m if $U_{ijk} > U_{imk}$. The multinomial-choice model is estimated using a multinomial logit model (see Greene, 2012) which assumes that the respondent chooses alternative j instead of alternatives m and n if $U_{ijk} > U_{imk}$ and $U_{ijk} > U_{ink}$.[13] Utility associated with the status-quo choice is set to zero in the MC estimation.

We tested for endogeneity of the consequentiality response on the BC data with a two-step instrumental-variable probit model using Newey's 1987 minimum Chi-squared estimator. Following Vossler, Doyon, and Rondeau (2012), we relied on sociodemographic indicators as excluded instruments, including gender, age, minority, and income. We concluded from a test of overidentifying restrictions (Baum et al., 2010) that the null hypothesis that the excluded instruments are jointly uncorrelated with the error term and correctly excluded from the binary-choice model

[13] A variety of alternative model specifications were estimated, including random-parameter specification for all choice-specific nonprice attributes. However, none of these were found to be statistically random. This result was robust across alternative distributional assumptions on the random parameters and models with subsets of attributes specified as random. Given that respondents were presented with only a single choice set, allowing for scale heterogeneity was not warranted (Train and Weeks, 2005).

Table 5. Likelihood Ratio Tests of Parameter Equivalence across Consequentiality Responses (Multinomial-Choice Data)

Null Hypothesis	χ^2	p-value
Knife-edge tests		
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat}$	11.49	0.12
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat}, \alpha_{Very} = \alpha_{Somewhat}$	14.82*	0.10
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Unlikely}$	21.57*	0.09
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Unlikely}, \alpha_{Very} = \alpha_{Somewhat} = \alpha_{Unlikely}$	87.87***	0.00
I don't know equality tests		
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Don'tKnow}$	12.94*	0.07
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Don'tKnow}, \alpha_{Very} = \alpha_{Don'tKnow}$	45.20***	0.00
$H_0: \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Don'tKnow}$	17.25**	0.02
$H_0: \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Don'tKnow}, \alpha_{Somewhat} = \alpha_{Don'tKnow}$	115.30***	0.00
$H_0: \boldsymbol{\beta}_{Unlikely} = \boldsymbol{\beta}_{Don'tKnow}$	15.91**	0.03
$H_0: \boldsymbol{\beta}_{Unlikely} = \boldsymbol{\beta}_{Don'tKnow}, \alpha_{Unlikely} = \alpha_{Don'tKnow}$	24.63***	0.00
Joint equality tests		
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Unlikely} = \boldsymbol{\beta}_{Don'tKnow}$	39.34***	0.01
$H_0: \boldsymbol{\beta}_{Very} = \boldsymbol{\beta}_{Somewhat} = \boldsymbol{\beta}_{Unlikely} = \boldsymbol{\beta}_{Don'tKnow}, \alpha_{Very} = \alpha_{Somewhat} = \alpha_{Unlikely} = \alpha_{Don'tKnow}$	163.61***	0.00

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively. $\boldsymbol{\beta}_k$: vector of bid and attribute coefficients. α_k: constant coefficient.

could not be rejected ($\chi^2(1) = 0.19,\ p = 0.67$). Given this set of valid instruments, we also failed to reject the null hypothesis of exogenous consequentiality responses according to a Wald test ($\chi^2(3) = 3.15,\ p = 0.37$). Given these results, we proceeded under the assumption of exogenous consequentiality responses.[14]

Binary-Choice Regression Results

We tested several restrictions in the BC data, the results of which are displayed in table 4 and are based on the specification in equation (1) using all observations. We test whether model fit under each of the specified restrictions in table 4 is significantly different from that of a completely unrestricted specification in which all constants and choice attribute coefficients vary by response to the consequentiality question. In table 4, β_k is a bid coefficient and α_k is a consequentiality-specific constant term, where k indicates the consequentiality type from the set {very likely, somewhat likely, unlikely, and I don't know}. Our tests fall into three main categories. The first category contains tests for knife-edge results in which we test for coefficient equality between the "very likely" and "somewhat likely" groups (H01) and whether both of these are equal to the coefficients for the "unlikely" groups (H02). The second category contains tests for equality of coefficients between the "I don't know" group and each of the other groups so that we can examine whether respondents who choose "I don't know" behave similarly to other respondents or if they are, in fact, a unique group. The third category contains tests of joint equality of coefficients across all consequentiality types. The tests of the BC data in table 4 show that we cannot reject equality of the bid coefficients across all consequentiality types.[15] Neither can we reject the hypothesis of equal intercept coefficients between

[14] We did not test for endogeneity in the MC model because we know of no such test. Our tests of endogeneity are admittedly weak, given the instrumental variables available, which undermines the validity of the exclusion restrictions. Short of this approach, however, we know of no other way to address this issue directly. So although these results are not fully conclusive, they do offer some support for the assumption of exogeneity moving forward.

[15] The results of these tests may partially be driven by the fact that we have relatively small subsamples for certain consequentiality responses, in particular, the "very likely" response.

Table 6. Regression Results for the Binary-Choice Model

	Estimate	S.E.
Education	0.08**	0.03
Conservative	−0.27***	0.04
Followed DWH Oil Spill	0.69***	0.19
Made Environmental Lifestyle Changes	0.34***	0.11
Bid	−0.00***	0.00
Constant_Very Likely	0.93**	0.46
Constant_Somewhat Likely	0.30	0.41
Constant_Unlikely or I Don't Know	−0.67*	0.41

Notes: Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% levels, respectively. N = 1,323. Log L = −782.61.

"unlikely' and "I don't know" respondents. If the BC respondents were consistent with knife-edge properties, we would expect equality of coefficients between the "very likely" and "somewhat likely" respondents, and these would differ from those of "unlikely" respondents. Thus, our tests indicate a lack of knife-edge results, and we observe this in the sensitivity to bid, constant coefficients, as well as overall. Our results are therefore inconsistent with the knife-edge results of Herriges et al. (2010) and more consistent with those of Vossler, Doyon, and Rondeau (2012), in that willingness to pay appears to depend on the degree of perceived consequentiality.

Specification of the BC regression model is based on the results of the aforementioned likelihood ratio tests. The bid coefficients are restricted to be equal across consequentiality types, but intercepts are allowed to vary with the restriction that those of the "unlikely" and "I don't know" respondents are equal. The results are shown in table 6. The parameter on the bid is negative and highly significant. Respondents with more formal education, those who had followed the 2010 oil spill in the Gulf more closely, and those who had made greater changes to their lifestyle for environmental reasons were more likely to vote in favor of the program. Respondents who rated themselves more politically conservative were less likely to vote in favor.

Multinomial-Choice Regression Results

Table 5 shows the restrictions tests for the MC data. As with table 4, we test whether the model fit under each of the specified restrictions significantly differs from that of a completely unrestricted specification in which all constants and choice attribute coefficients are allowed to vary by consequentiality response. $\boldsymbol{\beta_i}$ is a vector of attribute coefficients (including bid). The knife-edge tests indicate equality of bid coefficients for "very likely" and "somewhat likely" respondents but lack of equality for their constant coefficients. Unlike the BC data, however, we reject the hypothesis of joint equality of bid coefficients across all consequentiality types. The constant coefficients do differ statistically. Thus we observe mixed results regarding knife-edge results in the MC data; there appears to be knife-edge behavior regarding sensitivity to the bid and the choice attributes, but the constant coefficient varies by perceived consequentiality. This exercise illustrates that, although knife-edge results may appear to hold based strictly on voting behavior (as seen in the bottom of table 3), there may yet be some underlying differences between respondents with varying, positive perceived degrees of consequentiality.

The estimates from the MC data, in which—based on the results of the aforementioned likelihood ratio tests—the bid and attribute coefficients of the "very likely" and "somewhat likely" respondents are equal, are reported in table 7. The low levels of the project attributes were set as the omitted base levels. Results regarding demographic variables are consistent with those in the BC data. Additionally, given the similarities of coefficients between alternatives "A" and "B" (i.e., the first and second choices, respectively) in each choice set, there is no evidence of order bias.

Table 7. Regression Results for the Multinomial-Choice Model

	Estimate	S.E.
Consequentiality Response: Very likely or Somewhat likely		
Bid	−0.00***	0.00
Wildlife Habitat: Medium	0.22**	0.12
Wildlife Habitat: High	0.31**	0.15
Storm Surge Protection: Medium	0.36***	0.11
Storm Surge Protection: High	0.39**	0.17
Commercial Fisheries Harvest: Medium	0.52***	0.14
Commercial Fisheries Harvest: High	0.53***	0.17
Consequentiality Response: Unlikely		
Bid	−0.00***	0.00
Wildlife: Medium	0.41***	0.16
Wildlife: High	0.16	0.20
Storm Surge Protection: Medium	0.18	0.14
Storm Surge Protection: High	−0.20	0.22
Commercial Fisheries Harvest: Medium	0.97***	0.19
Commercial Fisheries Harvest: High	0.70***	0.23
Consequentiality Response: I Don't Know		
Bid	−0.00***	0.00
Wildlife: Medium	0.31	0.21
Wildlife: High	0.18	0.24
Storm Surge Protection: Medium	0.75***	0.19
Storm Surge Protection: High	−0.07	0.31
Commercial Fisheries Harvest: Medium	0.14	0.26
Commercial Fisheries Harvest: High	0.32	0.29
Choice A		
Education	0.21***	0.03
Conservative	−0.25***	0.04
Followed DWH Oil Spill	0.84***	0.19
Made Environmental Lifestyle Changes	0.45***	0.11
Constant_Very Likely	−1.11**	0.48
Constant_Somewhat Likely	−1.38***	0.42
Constant_Unlikely	−2.50***	0.45
Constant_I Don't Know	−3.12***	0.46
Choice B		
Education	0.18***	0.03
Conservative	−0.22***	0.04
Followed DWH Oil Spill	0.81***	0.18
Made Environmental Lifestyle Changes	0.48***	0.10
Constant_Very Likely	−1.52***	0.49
Constant_Somewhat Likely	−1.32***	0.41
Constant_Unlikely	−2.66***	0.45
Constant_I Don't Know	−2.81***	0.47

Notes: Double and triple asterisks (**, ***) indicate significance at the 5%, and 1% levels, respectively. $N = 1,974$. Log $L = -1,853.19$.

Table 8. WTP by Consequentiality Response and Program Attribute Level: Mean (95% CI)

	Low-Level Program	Medium-Level Program	High-Level Program
Multinomial-Choice Data[a]			
Very Likely	\$270 (57, 506)	\$677 (448, 953)	\$715 (460, 974)
Somewhat Likely	\$168 (64, 285)	\$576 (451, 735)	\$614 (439, 777)
Unlikely	−\$374(−784, −147)	\$476 (273, 749)	−\$51(−703, 329)
I Don't Know	−\$1, 299(−4, 698, −454)	−\$127(−1, 516, 376)	−\$1, 020(−5, 850, 143)
Binary-Choice Data[b]			
Very Likely		\$2,193 (1,236, 3,368)	
Somewhat Likely		\$1,007 (667, 1,399)	
Unlikely or I Don't Know		\$−790(−1, 461, −339)	

Notes: [a] With the exception of the WTP of the "I don't know" respondents for the low and high-level programs, pairwise tests indicate inequality of mean WTP values both (i) across consequentiality types within program and (ii) across programs within consequentiality type. [b] Pairwise tests indicate inequality of mean WTP values across consequentiality types.

The bid coefficients are negative, significant, and similar in magnitude across respondents of varying consequentiality beliefs. Consequential respondents (those who answered "very likely" or "somewhat likely") are highly sensitive to choice attributes; attribute coefficients are all significant and have the expected signs. Post-estimation tests indicate that the medium- and high-level coefficients within each attribute are not statistically different, indicating that results are only weakly consistent with scope predictions. The most striking observation, however, is that respondents who did not believe the survey to be consequential or who did not know whether the survey was consequential were generally less sensitive to variations in price and insensitive to variations in the nonprice choice attributes. That is, we find evidence of attribute nonattendance among these respondents (H05), and there is no evidence that scope effects hold for these respondents.

Willingness-to-Pay Estimates

We are ultimately interested in welfare estimates (willingness to pay) for restoration programs and for increases in program attribute increments. Table 8 reports the estimated willingness-to-pay means and 95% confidence intervals for the proposed program at low, medium, and high attribute levels by consequentiality response. Means and confidence intervals were estimated using the Krinsky and Robb simulation procedure (see Haab and McConnell, 2002) with 5,000 random draws. In the BC model, WTP for the medium-level restoration program increases with perceived degree of consequentiality, with estimated WTP actually negative for "unlikely" and "I don't know" respondents. Based on t-testing, we reject the hypothesis of equality of means between any pair of estimates at the 1% level.

For the MC model, only the very/somewhat likely group is consistently willing to pay more for increased levels of the attributes, and the mean differences are statistically different at the 1% level. Thus, the WTP values of consequential respondents are consistent with scope predictions. As with the BC data, estimated WTP is generally negative for "unlikely" and "I don't know" respondents. Of course, these respondents indicated that they did not believe the survey to be consequential (or were at least unsure), so it is not clear what conclusions can be drawn regarding their true preferences. We tested for knife-edge results within each of the three program levels but found that estimated mean WTP values statistically differ across all consequentiality types.

Table 9 reports the WTP estimates for incremental increases in project attributes based on the MC results. These were also estimated using the Krinsky and Robb procedure. We performed pairwise tests of equality of mean willingness-to-pay values both within each attribute level across consequentiality responses and within each consequentiality response across attribute levels. We found statistical difference except where noted. Once again, only the values of the very/somewhat

Table 9. WTP for Attribute Increments (Relative to Low Level) by Consequentiality Response: Mean (95% CI)

	Medium Attribute Level	High Attribute level
	Wildlife Habitat	
Very or Somewhat Likely	\$82 (1, 165)	\$110[a] (9, 197)
Unlikely	\$224 (58, 439)	\$70 (−172, 246)
I Don't Know	\$310 (−101, 1155)	\$110[a] (−511, 587)
	Storm Protection	
Very or Somewhat Likely	\$133 (58, 218)	\$140 (26, 244)
Unlikely	\$98 (−54, 271)	−\$124[b] (−463, 112)
I Don't Know	\$708 (256, 2,418)	−\$156[b] (−1,434, 409)
	Fisheries Productivity	
Very or Somewhat Likely	\$192 (86, 313)	\$195 (80, 306)
Unlikely	\$529 (285, 929)	\$376 (151, 683)
I Don't Know	\$154 (−480, 797)	\$326 (−335, 1,029)

Notes: We tested for pairwise equality of means (i) within each attribute level across consequentiality responses and (ii) within each consequentiality response across attribute levels.
[a,b] Indicate pairs of mean WTP values for which we could not reject the hypothesis of equality at the 5% level. All other pairwise tests indicate rejection of the same hypothesis.

likely respondents are consistent with theoretical predictions of preferences as they are willing to pay more for greater attribute levels. Respondents who found the survey to be at least somewhat consequential are willing to pay most for increased fisheries productivity and least for restored wildlife habitat. The "unlikely" and "I don't know" respondents are generally willing to pay less for the high levels of the attributes than the medium levels. Although the WTP estimates for the "unlikely" and "I don't know" respondents are generally positive, their overall WTP for a program is nevertheless negative in some cases, as seen in table 8.

Conclusion

Carson and Groves (2007) predict that survey respondents who perceive their responses to be at least somewhat consequential face the same incentives and should respond to the willingness-to-pay question in the same way, all else equal. This implies that the willingness-to-pay estimates of these respondents should be statistically equivalent. The results of our binary-choice survey indicate that willingness to pay for the proposed coastal restoration program increases with the respondents' perceived degree of survey consequentiality.

Our study is the first to examine the effects of consequentiality on responses in a multinomial-choice setting. We find that only respondents who believe it to be at least somewhat likely that the results of the survey will affect future policy behave consistently with theoretical predictions: they are sensitive to project attribute levels and, ceteris paribus, are more likely to vote in favor of a project with medium or high levels of the attributes than projects with low levels of the attributes. On the other hand, based on the preponderance of attribute parameter estimates not significantly different from zero, we cannot reject the hypothesis that the behavior of respondents who either did not perceive the survey to be consequential or did not know is random with specific regard to project attributes. In other words, we find strong evidence of attribute nonattendance among these respondents. We find mixed evidence for knife-edge results in the multinomial-choice data: we cannot reject the hypothesis that respondents who answered "very likely" or "somewhat likely" to the consequentiality question have equal choice attribute coefficients, but their constant

coefficients do differ. We also find that willingness to pay for restoration increases with perceived consequentiality, thus failing to observe knife-edge effects in welfare estimates.

The lesson from these results and others in the literature appears to confirm the predictions of Carson and Groves: that the consistency of respondent behavior with theoretical predictions is predicated upon the condition that respondents perceive the survey instrument to be consequential. Failing to account for perceptions of consequentiality may therefore lead researchers to wrong conclusions: that willingness to pay for a project is negative, that responses are not strongly based on choice attributes, or that scope effects are absent or contrary to expectations. Our study shows that the risk of false conclusions is present in both binary- and multinomial-choice settings, but our results are only a first cut regarding the role of consequentiality in a multinomial-choice setting. More research is warranted to generalize these findings. Based on our results, we would encourage researchers using stated-preference methods to elicit respondent perceptions of the consequentiality of their responses. Clearly these perceptions have an effect on respondent choices and therefore on inferences made by researchers regarding respondent preferences. Additionally, consequentiality and knife-edge results are still relatively young concepts in the literature, and so far there has not been a preponderance of evidence for or against knife-edge results. As such, no "best practice" can yet be suggested for whether consequentiality perceptions should be elicited on a simple yes/no scale or on a scale with more degrees of variation such as we have implemented in our study.

One topic that our study was not designed to address is best practice for "priming" the respondent regarding the consequentiality of the survey. In our study, we informed respondents that the survey results would be shared with policy makers so that they could better incorporate public opinion into their policy decisions. On the other hand, Herriges et al. (2010) used two treatments, one in which respondents read a letter directly from a public official stating how surveys like the present one had been useful in public policy decisions, and one in which respondents did not receive such a letter. This exogenous treatment helped Herriges et al. to account for possible endogeneity of the consequentiality question response, whereby respondents who are more likely to be in favor of the project are more likely to believe the survey to be consequential. They did find some evidence of endogeneity of consequentiality responses, so researchers might consider implementing a similar procedure in their studies. However, it may not always be feasible to do so. In the absence of such an approach, researchers should make a concerted effort to convey to respondents the consequentiality of the survey. To this end, it seems reasonable that researchers should incorporate an exploration into the survey's perceived consequentiality into the design and testing stage—either as part of a focus group or during pretesting—to identify the best (legitimate) means of maximizing perceived consequentiality rather than taking it for granted.

[Received May 2013; final revision received July 2014.]

References

Baum, C. F., V. Wiggins, S. Stillman, and M. E. Schaffer. "OVERID: Stata module to calculate tests of overidentifying restrictions after ivreg, ivreg2, ivreg29, ivprobit, ivtobit, reg3." 2010. Available online at http://ideas.repec.org/c/boc/bocode/s396802.html.

Bulte, E., S. Gerking, J. A. List, and A. de Zeeuw. "The Effect of Varying the Causes of Environmental Problems on Stated WTP Values: Evidence from a Field Study." *Journal of Environmental Economics and Management* 49(2005):330–342.

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

———. "Incentive and Information Properties of Preference Questions: Commentary and Extensions." In J. Bennett, ed., *International Handbook of Non-Market Environmental Valuation,* Northampton, MA: Edward Elgar, 2011, 300–321.

Carson, R. T., T. Groves, and J. A. List. "Consequentiality: A Theoretical and Experimental Exploration of a Single Binary Choice." *Journal of the Association of Environmental and Resource Economists* 1(2014):171–207.

Couvillion, B. R., J. A. Barras, G. D. Steyer, W. Sleavin, M. Fischer, H. Beck, N. Trahan, B. Griffin, and D. Heckman. "Land Area Change in Coastal Louisiana from 1932 to 2010." Scientific Investigations Map SIM 3164, U.S. Geologic Survey, Washington, DC, 2011.

Day, B., I. J. Bateman, R. T. Carson, D. Dupont, J. J. Louviere, S. Morimoto, R. Scarpa, and P. Wang. "Ordering Effects and Choice Set Awareness in Repeat-Response Stated Preference Studies." *Journal of Environmental Economics and Management* 63(2012):73–91.

Greene, W. H. *Econometric Analysis*. Boston: Prentice Hall, 2012, 7th ed.

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

Herriges, J., C. Kling, C. C. Liu, and J. Tobias. "What Are the Consequences of Consequentiality?" *Journal of Environmental Economics and Management* 59(2010):67–81.

Hwang, J., D. R. Petrolia, and M. G. Interis. "Valuation, Consequentiality, and Opt-Out Responses to Stated Preference Surveys." *Agricultural and Resource Economics Review* Forthcoming(2014).

Kuhfeld, W. F. "Experimental Design, Efficiency, Coding, and Choice Designs." SAS support guide, SAS Institute Inc., Cary, NC, 2010. Available online at http://support.sas.com/techsup/technote/mr2010c.pdf.

Landry, C. E., and J. A. List. "Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Markets: Evidence from the Field." *American Journal of Agricultural Economics* 89(2007):420–429.

McNair, B. J., J. Bennett, and D. A. Hensher. "A Comparison of Responses to Single and Repeated Discrete Choice Questions." *Resource and Energy Economics* 33(2011):554–571.

Myerson, R. B., and R. J. Weber. "A Theory of Voting Equilibria." *American Political Science Review* 87(1993):102–114.

Newey, W. K. "Efficient Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables." *Journal of Econometrics* 36(1987):231–250.

Petrolia, D. R., M. G. Interis, and J. Hwang. "America's Wetland? A National Survey of Willingness to Pay for Restoration of Louisiana's Coastal Wetlands." *Marine Resource Economics* 29(2014):17–37.

Scarpa, R., M. Thiene, and D. A. Hensher. "Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does It Matter?" *Land Economics* 86(2010):817–839.

Train, K., and M. Weeks. "Discrete Choice Models in Preference Space and Willingness-to-Pay Space." In R. Scarpa and A. Alberini, eds., *Applications of Simulation Methods in Environmental and Resource Economics,* No. 6 in The Economics of Non-Market Goods and Resources. Dordrecht: Springer, 2005, 1–16.

Vossler, C. A., M. Doyon, and D. Rondeau. "Truth in Consequentiality: Theory and Field Evidence on Discrete Choice Experiments." *American Economic Journal: Microeconomics* 4(2012):145–171.

Vossler, C. A., and M. F. Evans. "Bridging the Gap between the Field and the Lab: Environmental goods, Policy Maker Input, and Consequentiality." *Journal of Environmental Economics and Management* 58(2009):338–345.

Vossler, C. A., and S. B. Watson. "Understanding the Consequences of Consequentiality: Testing the Validity of Stated Preferences in the Field." *Journal of Economic Behavior & Organization* 86(2013):137–147.