Adequate Responsiveness to Scope in Contingent Valuation

by
William Desvousges
Kristy Mathews
Kenneth Train*

January 23, 2012

Abstract

The standard test for scope sensitivity in contingent valuation studies determines whether the response to changes in scope is statistically significant; it does not address whether the *magnitude* of response is plausible. We examine contingent valuation studies that implemented scope tests to determine what they imply about the adequacy of response to scope. We find that in the vast majority of studies, the magnitude of response cannot be assessed. Only three studies permit such an assessment: Samples and Hollyer (1990), Diamond et al. (1993) and Chapman et al. (2009). The first two papers find that responses to their surveys did not vary plausibly with scope. The third study passed the standard scope test, but we show that the magnitude of response in this study is inadequate by straightforward methods of assessment and cannot be explained by diminishing marginal utility or substitution. More research is needed on this issue, including wider application of adding-up tests on incremental parts, as well as the development of other methods that permit an assessment of the magnitude of response.

1. Introduction

Contingent valuation (CV) is a survey method designed to elicit information about respondents' willingness to pay (WTP) for environmental goods; see Carson and Hanneman (2005), e.g., for a review. A central question in the debate about CV is the issue of scope: whether the results from CV studies vary as required for economic rationality with the quantity, extent, or, more generally, the scope of the environmental good. Several early studies, including Kahneman and Knetsch (1992), Diamond et al. (1993), and Boyle et al. (1994), found, in their applications, that CV does not satisfy this criterion. However, as emphasized by, e.g., Carson (1997), these studies' results can constitute evidence of a problem in the design of these particular CV studies rather than in the CV method itself.

In response to this and other issues related to CV, the National Oceanic and Atmospheric Administration (NOAA) convened an expert panel to investigate and make recommendations regarding the use of CV methods for environmental damage assessment. The panel developed a list of guidelines for CV studies (Arrow et al, 1993) and stated that "…the burden of proof of reliability must rest on the survey designers. They must show through pretesting or other

* Corresponding author. Email: train@econ.berkeley.edu. Address: Department of Economics, 530 Evans Hall, University of California, Berkeley, CA 94720-3880. The research was supported in part by BP. The analysis and conclusions in this paper reflect the views of the authors and should not be attributed to anyone else.

experiments that their survey does not suffer from the problems that these guidelines are intended to avoid. Specifically, if a CV survey suffered from any of the following maladies, we would judge its findings 'unreliable'." Among the list that followed is "Inadequate responsiveness to the scope of the environmental insult."

Since the issuance of the NOAA panel guidelines, it has become fairly standard practice for CV studies to perform a scope test by specifying a reduced level of the environmental good[1] and soliciting WTP information for both the reduced and original levels. If the results for the two levels are statistically different in the expected direction, the study is said to have passed the scope test. To our knowledge, ninety CV studies, cited below, have conducted scope tests of this form on environmental goods.

This type of scope test necessarily reduces the "scope" of the scope issue. The NOAA panel described the potential malady as "inadequate responsiveness to the scope of the environmental insult." The standard test examines whether there is any statistically significant response to scope, not whether the amount of responsiveness is adequate.[2] And insofar as the test is applied to a severely reduced level of the environmental good, it does not examine responsiveness (either its existence or adequacy) to less severe, but economically meaningful, changes in scope.

In order to determine what the literature has said about adequacy of response to scope, we examined all the CV studies that we have been able to find that address scope empirically. In the following section, we discuss the implications of these studies in relation to adequacy of response, finding that practically none of them provides information about adequacy. In section 3, we describe the only study that passed the standard scope test and also permits an evaluation of adequacy of response to scope. We show that the response to scope in this study is greatly inadequate under straightforward methods of assessment. In the final section we discuss implications for future work.

2. Review of Previous Studies with Scope Tests

To evaluate the extent to which implemented scope tests allow an assessment of adequacy, we undertook an extensive literature review of empirical scope studies. To find these studies, we used standard research strategies such as internet key word searches, on-line databases of nonmarket valuation studies, and reference lists in published studies and in textbooks on contingent valuation. We made an effort to include unpublished study reports as well as published papers in peer-reviewed journals and books.

We reviewed more than 250 studies that appeared to contain scope tests. We applied several criteria for including studies in our detailed review for adequacy. For example, we focused only on studies that are consistent with a traditional CV approach. Studies based on a choice experiment design are not included. Because our interest is in the use of scope tests in natural

[1] Or a raised level of the good. We use "reduced" for linguistic convenience.

[2] Some members of the NOAA panel even stated (Arrow *et al.* 1994) that a statistically significant response is not what they meant in their guideline to test for an "adequate" response. They indicate that a "plausible" response to scope is a more consistent interpretation of their recommendation.

resource damage assessments, we focused on studies that evaluate commodities that are natural resource services. Studies based on market goods or services, or health states independent of environmental changes, are not included. Additionally, we included studies that may not have been designed explicitly to test for scope effects if other researchers could interpret the design as a scope test.[3] Earlier, unpublished studies are included if they contain scope tests that are different from a subsequent published version.[4] Studies that include a re-analysis of earlier data and report scope test results that differ from the original results are also included.[5] Table 1 contains 90 studies that meet these criteria and form the group of studies on which our detailed review for adequacy is based.[6]

For nearly all of the studies in Table 1, the authors report on the results of more than one scope test. For example, some authors evaluate sensitivity to scope based on both parametric and non-parametric criteria for the same data (e. g., Berrens et al. 2000). Others consider more than two levels of a good (e. g., Whitehead and Cherry 2007). Some scope studies report on scope tests for different surveys conducted on different commodities (e. g., Kahneman and Knetsch 1992). In light of the multiple tests within a study, the third column of Table 1 summarizes the results of each study's scope tests. "Pass" and "Fail" mean that all tests reported by the authors demonstrated sensitivity and insensitivity to scope, respectively. "Mixed" means that some of the tests passed while others did not.

We examined all the studies in Table 1 to determine what, if anything, they reveal about adequacy of response to scope. Amazingly, there is little information to be obtained. In most studies, the original and reduced environmental goods that are specified for the scope test differ in ways that prevent an assessment of the magnitude of response. For example, Berrens et al. (2000) find that the average value of saving one fish species in one river is $57 while the value of saving 11 fish species in four rivers is $74, which might appear to be inadequate response (i.e., "look fishy") but might also be the result of highly diminishing marginal utility of saving fish. Similarly, Carson et al. (2004) find that the value of saving two fish species over a period of fifteen years is worth $30 on average, while the value of saving two bird and two fish species for fifty years is $56; the low value for the extra species and years might reflect substitution across species and/or diminishing marginal utility.[7] The design of the scope tests in these studies makes it impossible to find that the response to scope is inadequate, which means that the studies do not

[3] See McClelland et al. (1992) as an example.

[4] For example, Desvousges et al. (1992) report on results from a separate dataset that Boyle et al. (1994) do not include in the published version.

[5] For example, Giraud, Loomis, and Johnson (1999) report original results and Poe, Giraud, and Loomis (2005) later re-analyze the same data.

[6] Although the NOAA panel was explicit in its recommendation for using split-sample or external scope tests, the literature includes many examples of internal scope tests. Thus, we include both in this table.

[7] Reasons have also been proposed for scope test failure. For example: Bateman et al. (2004) demonstrate that scope test failures may be attributed to the lack of advance notice of the entire valuation task. Powe and Bateman (2004) indicate that some scope test failures occur when respondents perceive that the provision of the higher level of the commodities is unrealistic. Heberlein et al. (2005) hypothesize that psychological and emotional beliefs held by respondents can explain some scope test failures. These reasons can also cause an inadequate response when the test is passed.

provide information on whether the response is adequate.[8] By the scientific method, a hypothesis can be tested only with a method that has a theoretical possibility of rejecting it.

Seventeen studies (marked as such in Table 1) have designs that allow testing of an adding-up criterion within CV. The concept of an adding-up test, as suggested by Diamond et al. (1993), is based on a definitional identity that must hold for preferences to be meaningful. In particular, consider two environmental goods A and B, where B is specified as being incremental to A (that is, B is provided as an extra benefit once A is already provided.) By definition, WTP for A plus WTP for B is equal to the WTP for A and B combined. This identity provides a test that is implemented as follows: A CV study is conducted for A, for B (stated explicitly as being incremental to A), and for A and B combined. Then the hypothesis is tested of whether the sum of the first two WTP's is equal to the third. The issue of adequacy is addressed by this procedure, since the response to scope is inadequate if the sum of the parts exceeds the whole.

The adding-up test is only applicable when the parts are specified incrementally. Several authors, e.g., Nunes and Schokkaert (2004), point out that, due to diminishing marginal utility and/or substitution, economically rational values of non-incremental parts need not sum to the value of their combination. This result implies that neither failing nor passing an adding-up test on non-incremental parts is informative about adequacy of response: (1) Failing an adding-up test on non-incremental parts does not indicate that responses are inadequate, since rational values need not sum as tested. The impossibility of rejecting adequacy means that adequacy is not actually being tested. (2) Passing an adding-up test with non-incremental parts does not indicate that response to scope is adequate, for the following reason. If the sum of the true values of the non-incremental parts differs from the true value of their combination, as is rationally possible, then a finding that the sum of CV values for the parts equals the CV value of the whole indicates that the CV responses are *incorrect*. Since we do not know the true values, we cannot know whether passing an adding-up test on non-incremental parts indicates that CV is accurate or inaccurate.

Unfortunately, of the seventeen studies that have used adding-up designs, only three specified the parts as increments: Samples and Hollyer (1990), Diamond et al. (1993), and Binger, Copple, and Hoffman (1995).[9,10] Of these three, only the first two give results of their adding-up tests. Binger, Copple, and Hoffman focus on other issues related to consistency of responses and, while their design would allow a type of adding-up test on incremental parts, they did not perform this test or report the information needed for us to perform it.

[8] Bateman (2011) expresses the same concern when he states (p. 326) that testing for response to changes in scope "has a major flaw in that we have no clear prior expectation about what a 'correct' answer should be. … All outcomes are eminently feasible." Bateman's suggestions for addressing this issue enter our "Discussion" section below.

[9] Even Diamond et al. did not specify parts that were *exactly* incremental. They estimated the WTP to not develop certain named wilderness areas. For their adding up test, the number of wilderness areas already saved from development was raised incrementally when asking about WTP not to develop an additional named area. However, the respondent was not told the *names* of the areas that were already saved.

[10] There is also an issue of income effects even with incremental parts. A person's WTP for B as an increment over A can differ depending on whether, or how much, the person had to pay for A, since the payment reduces the person's remaining income. None of the studies that we reviewed explicitly addresses this issue. However, in the context of environmental goods where any payment would be small relative to income, it is doubtful that the difference matters.

We have found only three studies that allow an assessment of the adequacy of response to scope: the two studies with incremental adding-up tests, i.e., Samples and Hollyer (1990) and Diamond et al. (1993), plus a recent study by Chapman et al. (2009) that did not implement an adding-up test but nevertheless, as we describe below, permits an assessment of adequacy. Samples and Hollyer found that the sum of estimated values for incremental parts greatly exceeded the estimated value of the combination of parts, which implies inadequate responsiveness to scope. Diamond et al. concluded, based on the mixed results of their various scope and adding-up tests, that the response to scope in their CV was not plausible. Chapman et al. applied a scope test, which was passed, and did not explicitly address the issue of adequacy of response; however, the design of their scope test, unlike those of other studies, permits such an assessment. We discuss the findings of this study in the next section.

3. Analysis of the Chapman et al. Study

The purpose of the study by Chapman et al. was to measure the natural resource damages to a river system and a lake in Oklahoma from excess phosphorus caused by poultry litter and other sources. An in-person clustered, stratified random sample was recruited from most of the state of Oklahoma, with a split-sample design for the scope test. The survey design was developed through an extensive pretesting process including the use of focus groups, pretests, and a large scale pilot study. The study was conducted by a team of experts in economics, survey design, and sampling methodology.

The study's scope test involves three incremental parts, as illustrated in Figure 1:

A. Speed recovery of specified rivers[11] by 40 years (starting in 10 years)
B. Speed recovery of a specified lake by 10 years (starting in 50 years), conditional on A already being obtained
C. Speed recovery of the same lake by another 30 years (starting in 20 years), conditional on A and B already being obtained

Their base program consists of A, B, and C combined and was found to be valued at $184 per household. Their scope program consists of B (with its conditioning on A explicitly stated to respondents) and was found to be valued at $138. The responses passed the scope test, in that the hypothesis of no difference was rejected. However, the magnitude of response is demonstrably inadequate. By the standard calculation of present value, B is worth at most one-quarter of B and C combined. However, it is estimated at 75 percent of A, B, and C combined. Any positive discount rate and any positive value for A makes the comparison even more implausible.

The adding-up identity can be used in this study, since the parts are specified as increments to each other. In particular: if the value of A, B, and C combined is $184, and the value of B is $138, then the combined value of A and C is $46. This result constitutes an actual violation of the scope criterion: Parts A and C provide more service than B in each dimension: more types of resources (the rivers and the lake *versus* just the lake), more years of service (40 years of additional service for the rivers and 30 years for the lake *versus* 10 years for the lake), and a closer time period (recovery occurring 10 years in the future for the rivers and 20 years in the

[11] For linguistic convenience, we use "rivers" to denote "river and creeks," which is the term used in the study.

future for the lake *versus* 50 years in the future for the lake). Yet these two parts are found by the study to be valued less than part B -- a third as much.[12]

As stated above, most CV studies' results cannot be evaluated since their base and scope programs are specified such that any difference in WTP can reflect diminishing marginal utility and/or substitution. However, these factors cannot explain the results in Chapman et al. A salient feature of their design is that the scope program is incremental to some of the benefits provided in the base program. In particular, the base program includes 40 years of faster recovery of the rivers (part A). In the scope scenario, this faster river recovery is said to occur without a program, and so the respondent obtains this environmental benefit without the scope program. The scope program provides 10 years of faster lake recovery as an increment to the river recovery (i.e., part B). Diminishing marginal utility might cause B to be worth less than A, since B is incremental to A. But, instead, the study finds that B is worth *more* than A (and even more than A and C combined.) Similarly, substitution between lakes and rivers might cause the lake recovery to be less valuable after the rivers have already recovered. But the study found the opposite: the lake recovery is more valuable than the river recovery to which it is a possible substitute and was already provided.

In the appendix, we discuss another form of diminishing returns, namely, that marginal utility diminishes as the number of years of recovered lake service rises. This type of diminishing returns has the unusual implication that an extra year of faster lake recovery in the distant future is more valuable than one closer in time, since the latter is incremental to the former. For example: If natural recovery would occur in 2050, then the first year of faster recovery occurs in 2049, the second year occurs in 2048, and so on. If the second year of faster recover is less valuable than the first year, as suggested by this form of diminishing utility, then the year of recovered lake service in 2049 is more valuable than the year of service in 2048 – which is equivalent to a negative discount rate over time. However, even if this unusual implication is accepted, the analysis in the appendix shows that this form of diminishing marginal utility does not plausibly rationalize the study's results.

4. Discussion

What can we conclude from all this? One thing seems clear to us. The NOAA panel was concerned with the possibility of "inadequate responsiveness to the scope of the environmental insult." The standard scope test does not address this concern, since it tests for statistical significance rather than adequacy of magnitude. In nearly all past applications of the test, the design of the study does not permit an evaluation of adequacy. And in the one study that permits such an evaluation while passing the standard scope test (i.e., Chapman et al.), the response to scope is demonstrably inadequate -- which shows that the standard scope test does not provide a reliable indication of adequacy.

Heberlein et al. (2005) propose that the standard scope test should perhaps be abandoned for efficacy and cost-effectiveness reasons. Their argument is that numerous cognitive and

[12] The hypothesis that the combined value of the A and C equals or exceeds the value of the B can be rejected at a confidence level of over 99.9 percent, using the confidence intervals reported on page 7-5 of Chapman et al.

attitudinal factors can induce a lack of scope response and even "negative scope," by which respondents give a smaller value to a larger enviromental good. These factors are important to investigate from a scientific perspective, in order to understand how respondents process information and respond to CV questions. Indeed, there must be *some* explanation for any observed response, and discovering the various causes is useful. Our analysis confirms Heberlein et al.'s suggestion that the standard scope test is not very informative, but the implication we draw is different, namely, that some other form of evidence of plausible response to scope is needed in order to inform the practical uses to which the damage assessments are usually applied.

Bateman (2011) also argues that standard scope tests are uninformative and recommends more extensive testing of the reliability of CV and stated-preference methods in general. He notes that the rich literature in experimental economics suggests types of validity tests that could be incorporated into study designs. He implemented several of these tests in the context of stated choice experiments, and alternative forms of these tests can perhaps be developed for standard CV. Our conclusions are consistent with his recommendations and, in particular, with his encouragement to practitioners to think creatively about addressing the adequacy of CV responses to changes in scope.

Our examination of the literature indicates that environmental economics as a field has barely addressed the issue of adequate response, despite the large number of studies on scope in CV. To remedy this situation, CV practitioners need to specify and implement procedures under which a conclusion of inadequate response is possible: otherwise, analyses, such as the standard scope test, will not actually test for adequacy. As stated above, adding-up tests applied to incremental parts provide one means of assessing adequacy. Important directions for future work include more applications of the incremental adding-up test (beyond the two by Samples and Hollyer, and Diamond et al.) and the development of other methods to assess adequacy of response.

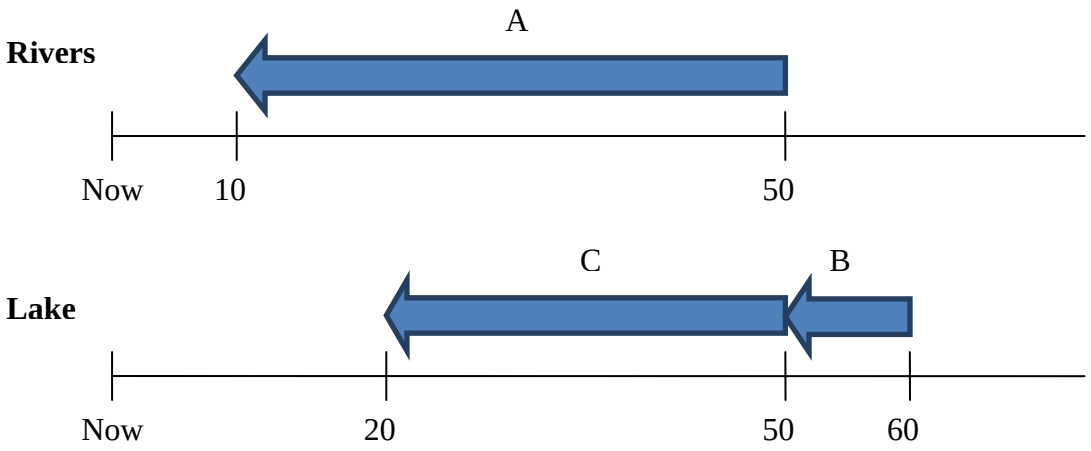
Table 1: CV Studies That Test for Scope Sensitivity
(P = Pass, F = Fail, M = Mixed, NR = Not Reported, X = Yes)

Authors	Publication Year	Result of Scope Test	Adding Up Design	Incremental Adding Up Design	Result of Adding Up Test	Possible to Assess Adequacy	Result of Adequacy Assessment
Ahern, Boyle and Hellerstein	2006	M					
Arana and Leon	2008	P					
Bandara and Tisdell	2005	M					
Banzhaf et al.	2006	P					
Bateman Cameron and Tsoumas	2006	M	X		F		
Bateman et al.	2005	M					
Bateman et al.	2004	M					
Bell Huber and Viscusi	2009	P					
Bennett Morrison and Blamey	1998	M					
Bergstrom Dillman and Stoll	1985	P					
Berrens et al.	2000	P					
Berrens Ganderton and Silva	1996	P					
Binger Copple and Hoffman	1995	P	X	X	NR		
Blomquist and Whitehead	1998	M					
Bowker and Didychuk	1994	P					
Boyle et al.	1994	F					
Brookshire Eubanks and Randall	1983	M					
Brown et al.	1995	M					
Brown and Duffield	1995	M					
Carson	1997	P					
Carson et al.	1994	P					
Carson Wilkes and Imber	1994	M					
Chapman et al.	2009	P				X	F
Choe Whittington and Lauria	1996	M					
Christie	2001	M	X		F		
Cooper Poe and Bateman	2004	P					
Desvousges et al.	1992	F					

Authors	Publication Year	Result of Scope Test	Adding Up Design	Incremental Adding Up Design	Result of Adding Up Test	Possible to Assess Adequacy	Result of Adequacy Assessment
Diamond et al.	1993	M	X	X	M	X	M
Douglas and Taylor	1999	M					
DuPont	2003	M					
Eom and Larson	2006	P					
Fischhoff et al.	1993	F					
Gerrrans	1994	F					
Giraud Loomis and Johnson	1999	M					
Goodman Seabrooke and Jaffry	1998	M					
Halstead et al.	2004	P					
Hanemann	2005	P					
Hanemann Loomis and Kanninen	1991	P					
Hanley, Schlapfer and Spurgeon	2003	F					
Heberlein et al.	2005	M					
Hite et al.	2002	P					
Hoevenagel	1996	P	X		NR		
Holmes et al.	2004	M					
Huang Haab and Whitehead	1997	P					
Kahneman	1986	NR					
Kahneman and Ritov	1994	M					
Kahneman and Knetsch	1992	M					
Kemp and Maxwell	1993	M					
Loomis and Gonzalez-Caban	1998	P	X		NR		
Loomis and Ekstrand	1997	P					
Loomis and Larson	1994	P					
Loomis Lockwood and DeLacy	1993	M					
Macmillian and Duff	1998	F	X		F		
Macmillan Hanley and Buckland	1996	M					
Magnussen	1992	M					
McClelland et al.	1992	M					
McDaniels et al.	2003	P					
McFadden	1994	F					
Navrud	1989	M					

Authors	Publication Year	Result of Scope Test	Adding Up Design	Incremental Adding Up Design	Result of Adding Up Test	Possible to Assess Adequacy	Result of Adequacy Assessment
Nunes and Schokkaert	2003	M	X		M		
Ojea and Loureiro	2009	F					
Poe Giraud and Loomis	2005	P					
Pope and Jones	1990	M					
Pouta	2005	M					
Powe and Bateman	2004	M	X		F		
Ready Berger and Blomquist	1997	P					
Riddel and Loomis	1998	P	X		NR		
Rollins and Lyke	1998	M	X		NR		
Romer Pommerehne and Feld	1998	M					
Rowe Shaw and Schulze	1992	M					
Samples and Hollyer	1990	F	X	X	F	X	F
Schkade and Payne	1994	F					
Schulze et al.	1998	M					
Smith Zhang and Palmquist	1997	P					
Stanley	2005	P					
Smith et al.	2005	M					
Stevens et al.	1997	F					
Stevens Benin and Larson	1995	P	X		F		
Streever et al.	1998	NR	X		NR		
Svedsater	2000	F					
Veisten et al.	2004	M	X		F		
Walsh Loomis and Gillman	1984	P					
White et al.	1997	M	X		F		
Whitehead	1992	P					
Whitehead et al.	2009	F					
Whitehead and Cherry	2007	M					
Whitehead and Finney	2002	F					
Whitehead Haab and Huang	1998	P					
Wilson	2000	M					
Wu	1993	NR	X		F		

Figure 1: Incremental Parts of Accelerated Restoration



Estimated values: A,B,C = $184, B=$138, thus A + C = $46

References

Ahearn, M., K. Boyle, and Daniel Hellerstein. 2006. "Designing a Contingent Valuation Study to Estimate the Benefits of the Conservation Reserve Program." In *Handbook of Contingent Valuation*, A. Alberini and J. Kahn, eds. Northampton: Edward Elgar Publishing.

Arana, Jorge E. and Carmelo J. Leon. 2008. "Do Emotions Matter? Coherent Preferences under Anchoring and Emotional Effects." *Ecological Economics* 66:700-711.

Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman. 1993. *Report of the NOAA Panel on Contingent Valuation,* http://www.darrp.noaa.gov/library/pdf/cvblue.pdf

Arrow, Kenneth, Edward E. Leamer, Howard Schuman, and Robert Solow, 1994, "Comments of Proposed NOAA Scope Test," Appendix D of *Comments of Proposed NOAA/DOI Regulations on Natural Resource Damage Assessment*, U.S. Environmental Protection Agency.

Bandara, Ranjith and Clem Tisdell. 2005. "Changing Abundance of Elephants and Willingness to Pay for Their Conservation." *Journal of Environmental Management* 76: 47-59.

Banzhaf, H. Spencer, Dallas Burtraw, David Evans, and Alan Krupnick. 2006. "Valuation of Natural Resource Improvements in the Adirondacks." *Land Economics* 82(3): 445-464. August.

Bateman, Ian J., 2011, "Valid Value Estimates and Value Estimate Validation: Better Methods and Better Testing for Stated Preference Research." In *The International Handbook on Non-Market Environmental Valuation*, Jeff Bennett, editor, Cheltenham UK: Edward Elgar Publishing.

Bateman, Ian J. M.P. Cameron, and A. Tsoumas. 2006. "Investigating the Characteristics of Expressed Preferences for Reducing the Impacts of Air Pollution: A Contingent Valuation Experiment." Department of Economics. University of Waikato Working Paper 08/06. May.

Bateman, Ian J., Philip Cooper, Stavros Georgiou, Ståle Navrud, Gregory L. Poe, Richard C. Ready, Pere Riera, Mandy Ryan, and Christian A. Vossler. 2005. "Economic Valuation of Policies for Managing Acidity in Remote Mountain Lakes: Examining Validity Through Scope Sensitivity Testing." *Aquatic Sciences* 67: 274-291.

Bateman, Ian J, Matthew Cole, Philip Cooper, Stavros Georgiou, David Hadley, and Gregory L. Poe. 2004. "On Visible Choice Sets and Scope Sensitivity." *Journal of Environmental Economics and Management* 47: 71-93.

Bell, Jason, Joel Huber, and W. Kip Viscusi. 2009. "Voter-Weighted Environmental Preferences." *Journal of Policy Analysis and Management* 28(4): 655-671.

Bennett, Jeff, Mark Morrison, and Russell Blamey. 1998. "Testing the Validity of Responses to Contingent Valuation Questioning." *The Australian Journal of Agricultural and Resource Economics* 42(2): 131-148.

Bergstrom, John C., B. L. Dillman, and John R. Stoll. 1985. "Public Environmental Amenity Benefits of Private Land: The Case of Prime Agricultural Land." *Southern Journal of Agricultural Economics* 17(1): 139-149. July.

Berrens, R.P., A. K. Bohara, C. L. Silva, D. Brookshire and M. McKee. 2000. "Contingent Values for New Mexico Instream Flows: With tests of Scope, Group-Size Reminder and Temporal Reliability." *Journal of Environmental Management* 58: 73-90.

Berrens, Robert P., Philip Ganderton, and Carol L. Silva. 1996. "Valuing the Protection of Minimum Instream Flows in New Mexico." *Journal of Agricultural and Resource Economics* 21(2): 294-309.

Binger, Brian R., Robert F. Copple, and Elizabeth Hoffman. 1995. "The Use of Contingent Valuation Methodology in Natural Resource Damage Assessments: Legal Fact and Economic Fiction." *Northwestern University Law Review* 89(3): 1029-1053.

Blomquist, Glenn C. and John C. Whitehead. 1998. "Resource Quality information and Validity of Willingness to Pay in Contingent Valuation." *Resource and Energy Economics* 20(2): 179-196.

Bowker, J.M. and D.D. Didychuk. 1994. "Estimation of the Nonmarket Benefits of Agricultural Land Retention in Eastern Canada." *Agricultural and Resource Economics Review* 23(2): 218-225.

Boyle, Kevin J., William H. Desvousges, F. Reed Johnson, Richard W. Dunford, and Sara P. Hudson. 1994. "An Investigation of Part-Whole Biases in Contingent Valuation Studies." *Journal of Environmental Economics and Management* 27: 64-83.

Brookshire, David S., Larry S. Eubanks, and Alan Randall. 1983. "Estimating Option Prices and Existence Values for Wildlife Resources." *Land Economics* 59(1): 1-15.

Brown, T.C., S.C. Barro, M.J. Manfredo, and G.L. Peterson. 1995. "Does Better Information about the Good Avoid Embedding Effect?" *Journal of Environmental Management* 44: 1-10.

Brown, Thomas C. and John W. Duffield. 1995. "Testing Part-Whole Valuation Effects in Contingent Valuation of Instream Flow Protection." *Water Resources Research* 31(9): 2341-2351.

Carson, Richard T. 1997. "Contingent Valuation and Tests of Insensitivity to Scope." In *Determining the Value of Non-Marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods.* R.J. Kopp, W. Pommerhene, and N. Schwartz, eds. Boston: Kluwer.

Carson, Richard T., W. Michael Hanemann, Raymond J. Kopp, and Jon A. Krosnick, Robert C. Mitchell, Stanley Presser, Paul A. Ruud, and V. Kerry Smith. 1994. "Prospective Interim Lost Use Value Due to PCB and DDT Contamination in the Southern California Bight." Report to National Oceanic and Atmospheric Administration.

Carson, Richard., and W. Michael Hanneman, 2005, "Contingent Valuation," in *Handbook of Environmental Economics, Vol. 2*, K.-G. Mäler and J. R. Vincent, eds., Chapter 17, Amsterdam: Elsevier.

Carson, Richard T., Leanne Wilks, and David Imber. 1994. "Valuing the Preservation of Australia's Kakadu Conservation Zone." *Oxford Economic Papers* 46: 727-749.

Chapman, David J., Richard C. Bishop, W. Michael Hanemann, Barbara J. Kanninen, Jon A. Krosnick, Edward R. Morey, and Roger Tourangeau. 2009. "Natural Resource Damages Associated with Aesthetic and Ecosystem Injuries to Oklahoma's Illinois River System and Tenkiller Lake." Expert Report for State of Oklahoma. Case No. 05-CV-0329-GKF-SAJ. *State of Oklahoma v. Tyson Foods, et al.* In the United States District Court for the Northern District of Oklahoma. Volume I.

Choe, KyeongAe, Dale Whittington, and Donald T. Lauria. 1996. "The Economic Benefits of Surface Water Quality Improvements in Developing Countries: A Case Study of Davao, Philippines." *Land Economics* 72(4): 519-537.

Christie, Michael. 2001. "A Comparison of Alternative Contingent Valuation Elicitation Treatments for the Evaluation of Complex Environmental Policy." *Journal of Environmental Management* 62(3): 255-269.

Cooper, Philip, Gregory L. Poe, and Ian J. Bateman. 2004. "The Structure of Motivation for Contingent Values: A Case Study of Lake Water Quality Improvement." *Ecological Economics* 50: 69-82.

Desvousges, William. H., F.Reed Johnson, RichardW. Dunford, Sara .P. Hudson, K.N. Wilson, and Kevin J. Boyle. 1992. *Measuring Nonuse Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy.* Research Triangle Institute Monograph 92-1. Research Triangle Park, NC.

Diamond, Peter A., Jerry A. Hausman, Gregory K. Leonard, and Mike A. Denning. 1993. "Does Contingent Valuation Measure Preferences? Experimental Evidence." In *Contingent Valuation, A Critical Assessment*, J.A. Hausman, ed., pp. 41-89. Amsterdam: Elsevier.

Douglas, Aaron J. and Jonathan G. Taylor. 1999. "The Economic Value of Trinity River Water." *Water Resources Development* 15(3): 309-322.

Dupont, Diane P. 2003. "CVM Embedding Effects When There Are Active, Potentially Active and Passive Users of Environmental Goods." *Environmental and Resource Economics* 25: 319-341.

Eom, Young-Sook and Douglas M. Larson. 2006. "Improving Environmental Valuation Estimates through Consistent Use of Revealed and Stated Preference Information." *Journal of Economics and Environmental Management* 52: 501-516.

Fischhoff, Baruch, Marilyn Jacobs Quadrel, Mark Kamlet, George Loewenstein, Robyn Dawes, Paul Fischbeck, Steven Klepper, Jonathan Leland, and Patrick Stroh. 1993. "Embedding Effects: Stimulus Representation and Response Mode." *Journal of Risk and Uncertainty* 6: 211-234.

Gerrans, P. 1994. An Economic Valuation of the Jandakot Wetlands. Occasional Paper No. 1 Edith Cowan University.

Giraud, K.L., J.B. Loomis, and R.L. Johnson. 1999. "Internal and External Scope in Willingness-to-Pay Estimates for Threatened and Endangered Wildlife." *Journal of Environmental Management* 56: 221-229.

Goodman, S.L., W. Seabrooke, S.A. Jaffry. 1998. "Considering Conservation Value in Economic Appraisals of Coastal Resources." *Journal of Environmental Planning and Management* 41(3): 313-336.

Halstead, John M., Thomas H. Stevens, Wendy Harper, and L. Bruce Hill. 2004. "Electricity Deregulation and the Valuation of Visibility Loss in Wilderness Areas: A Research Note." *The Journal of Regional Policy Analysis* 34(1): 85-95.

Hanemann, Michael. 2005. "The Bird Study Revisited." Presentation. U.C. Berkeley.

Hanemann, Michael, John Loomis, and Barbara Kanninen. 1991. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73(4): 1256-1263.

Hanley, Nick, Felix Schlapfer, and James Spurgeon. 2003. "Aggregating the Benefits of Environmental Improvements: Distance-Decay Functions for Use and Non-use Values." *Journal of Environmental Management* 68: 297-304.

Heberlein, Thomas A., Matthew A. Wilson, Richard C. Bishop, and Nora Cate Schaeffer. 2005. "Rethinking the Scope Test as a Criterion for Validity in Contingent Valuation." *Journal of Environmental Economics and Management* 50: 1-22.

Hite, Diane, Darren Hudson, and Walaiporn Intarapapong. 2002. "Willingness to Pay for Water Quality Improvements: The Case of Precision Application Technology." *Journal of Agricultural and Resource Economics* 27(2):433-449.

Hoevenagel, Ruud. 1996. "The Validity of the Contingent Valuation Method: Perfect and Regular Embedding." *Environmental and Resource Economics* 7: 57-78.

Holmes, Thomas P., John C. Bergstrom, Eric Huszar, Susan B. Kask, and Fritz Orr, III. 2004. "Contingent Valuation, Net Marginal Benefits, and the Scale of Riparian Ecosystem Restoration." *Ecological Economics* 49: 19-30.

Huang, Ju-Chin, Timothy C. Haab, and John C. Whitehead. 1997. "Willingness to Pay for Quality Improvements: Should Revealed and Stated Preference Data Be Combined?" *Journal of Environmental Economics and Management* 34: 240-255.

Kahneman, Daniel. 1986. "Comments by Professor Daniel Kahneman." In *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method.* R. Cummings, D. Brookshire, and W. Schulze, eds. pp. 185-194. Totowa: Rowman and Allanheld.

Kahneman, Daniel and Jack L. Knetsch. 1992. "Valuing Public Goods: The Purchase of Moral Satisfaction." *Journal of Environmental Economics and Management* 22: 57-70.

Kahneman, Daniel and Ilana Ritov. 1994. "Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method." *Journal of Risk and Uncertainty* 9: 5-38.

Kemp, Michael A. and Christopher Maxwell. 1993. "Exploring a Budget Context For Contingent Valuation Estimates." In *Contingent Valuation, A Critical Assessment,* J.A. Hausman, ed., pp. 217-269. Amsterdam: Elsevier.

Loomis, John and Armando Gonzalez-Caban. 1998. "A Willingness-to-Pay Function for Protecting Acres of Spotted Owl Habitat from Fire." *Ecological Economics* 25: 315-322.

Loomis, John and Earl Ekstrand. 1997. "Economic Benefits of Critical Habitat for the Mexican Spotted Owl: A Scope Test Using a Multiple-Bounded Contingent Valuation Survey." *Journal of Agricultural and Resource Economics* 22(2):356-366.

Loomis, John B. and Douglas M. Larson. 1994. "Total Economic Values of Increasing Gray Whale Populations: Results from a Contingent Valuation Survey of Visitors and Households." *Marine Resource Economics* 9: 275-286.

Loomis, John, Michael Lockwood, and Terry DeLacy. 1993. "Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection." *Journal of Environmental Economics and Management* 24: 44-55.

Macmillan, D.C. and E.I. Duff. 1998. "Estimating the Non-Market Costs and Benefits of Native Woodland Restoration Using the Contingent Valuation Method." *Forestry* 71 (3): 247-259.

Macmillan, Douglas, Nick Hanley, and Steve Buckland. 1996. "A Contingent Valuation Study of Uncertain Environmental Gains." *Scottish Journal of Political Economy* 43(5): 519-533.

Magnussen, Kristin. 1992. "Valuation of Reduced Water Pollution Using the Contingent Valuation Method - Testing for Amenity Misspecification." In *Pricing the European Environment*, edited by Ståle Navrud. Oxford: Oxford University Press.

McClelland, Gary H., William D. Schulze, Jeffrey K. Lazo, Donald M. Waldman, James K. Doyle, Steven R. Elliott, and Julie R. Irwin. 1992. "Methods for Measuring Non-Use Values: A Contingent Valuation Study of Groundwater Cleanup." Report Prepared Under EPA Cooperative Agreement CR815183. October.

McDaniels, Timothy L., Robin Gregory, Joseph Arvai, and Ratana Chuenpagdee. 2003. "Decision Structuring to Alleviate Embedding in Environmental Valuation." *Ecological Economics* 46: 33-46.

McFadden, Daniel. 1994. "Contingent Valuation and Social Choice." *American Journal of Agricultural Economics* 76(4): 689-708. November.

Navrud, Ståle. 1989. "Estimating Social Benefits of Environmental Improvements from Reduced Acid Rain Deposition: A Contingent Valuation Survey." In *Valuation Methods and Policy Making in Environmental Economics*, edited by Folmer & van lerland. Amsterdam: Elsevier.

Nunes, Paulo A.L.D. and Erik Schokkaert. 2003. "Identifying the Warm Glow Effect in Contingent Valuation." *Journal of Environmental Economics and Management* 45: 231-245.

Ojea, Elena and Maria L. Loureiro. 2009. "Valuation of Wildlife: Revising Some Additional Considerations for Scope Tests." *Contemporary Economic Policy* 27(2): 236-250.

Poe, Gregory, Kelly L. Giraud, and John B. Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal Agricultural Economics* 87(2): 353-365.

Pope, C. Arden, III and Jeffrey W. Jones. 1990. "Value of Wilderness Designation in Utah." *Journal of Environmental Management* 30: 157-174.

Pouta, Eija. 2005. "Sensitivity to Scope of Environmental Regulation in Contingent Valuation of Forest Cutting Practices in Finland." *Forest Policy and Economics* 7: 539-550.

Powe, Neil A. and Ian J. Bateman. 2004. "Investigating Insensitivity to Scope: A Split-Sample Test of Perceived Scheme Realism." *Land Economics* 80(2): 258-271.

Romer, Anselm U., Werner W. Pommerehne, and Lars P. Feld. 1998. "Revealing Preferences for Reductions of Public Risks: An Application of the CV Approach." *Journal of Environmental Planning and Management* 41(4): 477-503.

Ready, Richard C., Mark C. Berger, and Glenn C. Blomquist. 1997. "Measuring Amenity Benefits from Farmland: Hedonic Pricing vs. Contingent Valuation." *Growth and Change* 28: 438-458.

Riddel, Mary and John Loomis. 1998. "Joint Estimation of Multiple CVM Scenarios under a Double Bounded Questioning Format." *Environmental and Resource Economics* 12: 77-98.

Rollins, Kimberly and Audrey Lyke. 1998. "The Case for Diminishing Marginal Existence Values." *Journal of Environmental Economics and Management* 36: 324-366.

Rowe, R.D., W.D. Shaw, and W. Schulze. 1992. "Nestucca Oil Spill." In *Natural Resource Damages: Law and Economics,* K. M. Ward and J. W. Duffield, eds. pp. 527-554. New York: John Wiley.

Samples, K.C and J.R. Hollyer. 1990. "Contingent Valuation of Wildlife Resources in the Presence of Substitutes and Complements." In *Economic Valuation of Natural Resources: Issues, Theory, and Applications,* R.L. Johnson and G.V. Johnson, eds., pp. 177-192. Boulder, CO: Westview Press.

Schkade, David A. and John W. Payne. 1994. "How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis of Willingness to Pay for an Environmental Regulation." *Journal of Environmental Economics and Management* 26: 88-109.

Schulze, William D., Gary H. McClelland, Jeffrey K. Lazo, and Robert D. Rowe. 1998. "Embedding and Calibration in Measuring Non-use Values." *Resource and Energy Economics* 20: 163-178.

Smith, V. Kerry, Xiaolong Zhang, and Raymond B. Palmquist. 1997. "Marine Debris, Beach Quality, and Non-Market Values." *Environmental and Resource Economics* 10: 223-247.

Smith, Anne E., Michael A. Kemp, Timothy H. Savage, and Catherine L. Taylor. 2005. "Methods and Results from a New Survey of Values for Eastern Regional Haze Improvements." *Journal of Air and Waste Management Association* 55: 1767-1779.

Stanley, Denise L. 2005. "Local Perception of Public Goods: Recent Assessments of Willingness-to-Pay for Endangered Species." *Contemporary Economic Policy* 23(2): 165-179.

Stevens, Thomas H., Nichole E. DeCoteau, and Cleve E. Willis. 1997. "Sensitivity of Contingent Valuation to Alternative Payment Schedules. *Land Economics* 73(1): 140-148.

Stevens, T.H., S. Benin, and J.S. Larson. 1995. "Public Attitudes and Values for Wetland Conservation in New England." *Wetlands* 15(3): 226-231.

Streever, W.J., M. Callaghan-Perry, A. Searles, T. Stevens, and P. Svoboda. 1998. "Public Attitudes and Values for Wetland Conservation in New South Wales, Australia." *Journal of Environmental Management* 54(1): 1-14.

Svedsäter, Henrik. 2000. "Contingent Valuation of Global Environmental Resources: Test of Perfect and Regular Embedding." *Journal of Economic Psychology* 21: 605-623.

Train, Kenneth, 2011, "Comments of Chapman et al. (2009): Inadequate Response to Scope while Passing the Scope Test," working paper, Department of Economics, University of California, Berkeley.

Veisten, Knut, Hans Fredrik Hoen, Ståle Navrud, and Jon Strand. 2004. "Scope Insensitivity in Contingent Valuation of Complex Environmental Amenities." *Journal of Environmental Managemen.* 73: 317-331.

Walsh, Richard, John Loomis, and Richard Gillman. 1984. "Valuing Option, Existence, and Bequest Demands for Wilderness." *Land Economics* 60(1): 14-29.

White, Piran C. L., Keith W. Gregory, Patrick J. Lindley and Glenn Richards. 1997. "Economic Values of Threatened Mammals in Britain: A Case Study of the Otter *Lutra Lutra* and the Water Vole *Arvicola Terrestris." Biological Conservation* 82: 345-354.

Whitehead, John C. 1992. "Ex Ante Willingness to Pay with Supply and Demand Uncertainty: Implications for valuing a Sea Turtle Protection Programme." *Applied Economics* 24: 981-988.

Whitehead, John C. and Todd L. Cherry. 2007. "Willingness to Pay for a Green Energy Program: A Comparison of Ex-ante and Ex-post Hypothetical Bias Mitigation Approaches." *Resource and Energy Economics* 29: 247-261.

Whitehead, John, Peter Groothuis, Rob Southwick, and Pat Foster-Turley. 2009. "Measuring the Economic Benefits of Saginaw Bay Coastal Marsh with Revealed and Stated Preference Methods." *Journal of Great Lakes Research* 35(3):430-437.

Whitehead, John and Suzanne Finney. 2003. "Willingness to Pay for Submerged Maritime Cultural Resources." *Journal of Cultural Economics* 27(3): 231-240. November.

Whitehead, John C., Timothy C. Haab, and Ju-Chin Huang. 1998. "Part-Whole Bias in Contingent Valuation: Will Scope Effects Be Detected with Inexpensive Survey Methods?" *Southern Economic Journal* 65(1): 160-168.

Wilson, Matthew A. 2000. "Rethinking Scope Sensitivity and Contingent Valuation Surveys: Strong Environmental Attitudes and Contingent Economic Values." Dissertation submitted to the Graduate School of the University of Wisconsin-Madison. UMI Number: 9981890.

Wu, Pei-Ing. 1993. "Substitution and Complementarity in Commodity Space: Benefit Evaluation of Multidimensional Environmental Policy." *Academia Economic Papers* 21(1): 151-182, March.

Appendix: Another Form of Diminishing Marginal Utility with the Chapman et al. Results

As described in the body of the paper, the Chapman et al. study found that WTP for B is $138 and, by the adding-up identity, WTP for A and C combined is $46 (where increments A, B, and C are defined in section 3 above.) Since C is incremental to B, it might be thought that diminishing returns for lake recovery (which are B and C), coupled with a low value for river recovery (A), could rationalize the results. However, while the direction of difference is plausible, the magnitude of the difference is not. We demonstrate this as follows.

Let WTP for A be zero (with values over zero rendering the study's results less plausible than shown below.) The study's results then imply that WTP for C is $46 while WTP for B is $138, despite the fact that C provides 3 times as many years of recovered lake service. The degree of diminishing marginal utility must be very great in order to off-set this difference in number of years. Suppose, for example, that respondents' direct utility takes the form $U(x,y) = x + \frac{k}{\alpha}\exp(\alpha y)$, where x is the quantity of non-environmental goods consumed measured in dollars, y is years of recovered lake service[13], $k>0$, and parameter $\alpha < 0$ captures the degree of diminishing marginal utility for recovered lake service, such that $MU(y) > 0$ and $dMU(y)/dy < 0$. Let x^0 be the consumption of other goods without payment for a program and let y^0 be the years of recovered lake service that the consumer obtains without a program, with $y^0 > 0$ since recovery occurs eventually without a program. The study found that 10 extra years of lake recovery is worth $138.46 and 40 extra years is worth $184.55, where we include the cents to allow for greater accuracy in our calculations. These values imply the following equalities for the utility function: $U(x^0 - 184.55, y^0 + 40) = U(x^0, y^0) = U(x^0 - 138.46, y^0 + 10)$. The value of α that equates the two sides is $\alpha = -0.1375$. Using this value of α in either of the two equations gives $U(x^0, y^0) = x^0 - 185.31$ such that $U(x,y) = x - 185.31 * \exp(-0.1375(y - y^0))$.

Table A gives the utility for different numbers of years of recovered lake service.

Table A: Utility from years of recovered lake service, holding consumption of other goods constant at x^0.

Years of service y	Utility	Incremental utility
y^0 - 20	x^0 - 2899.44	
y^0 - 10	x^0 - 733.00	2166.44
y^0	x^0 - 185.31	547.60
y^0 + 10	x^0 - 46.85	138.46
y^0 + 20	x^0 - 11.84	35.00
y^0 + 30	x^0 - 2.99	8.85
y^0 + 40	x^0 - 0.76	2.23

[13] For the calculations to follow, we assume that years are not discounted for relative distance in the future. Again, this assumption is generous to the Study since any discounting would exacerbate the problem being described.

By construction, the degree of diminishing returns is consistent with the study's results: the incremental utility (in dollars) of an extra 10 years of recovery given y^0 years are already obtained is \$138.46, and the WTP for an extra 40 years given y^0 years is \$184.55 (i.e., the sum of the last four values of incremental utility). However, this degree of diminishing marginal utility is not plausible when viewed for previous increments. The utility function that is consistent with the study's results implies that the WTP to obtain 10 extra years of service given that $y^0 - 10$ are already obtained is \$548, and that the WTP for 10 extra years given that $y^0 - 20$ are already obtained is \$2,166. If the Study had told respondents that the lake would recover without a program in 80 years, and asked about their WTP to speed the recovery from 80 years to 40 years (the same number of years of faster recovery as in the original specification of the base program), the degree of diminishing returns in the current study implies that respondents would have said that they are willing to pay \$2,887 (2166.44+548.60+138.46+35.00) on average for the program. The implausibility of this figure suggests that respondents do not actually posses the degree of diminishing marginal utility that would rationalize the study's findings, even when coupled with a low value for river recovery and no discounting over time. Furthermore, since the study's purpose was to value the resource, its results would be deemed unreliable for this purpose if changes of this magnitude (from \$184.55 to \$2887) arose simply from specifying that the natural lake recovery starts at different arbitrary years in the distant future, without any change in the number of extra years of lake recovery that is obtained from the program (40 more years in each case.) As Train (2011) shows, any alternative representation of marginal utility (i.e., different from the parameterization above) that is consistent with the study's findings has the same implication.[14]

[14] It has been suggested by a reader that respondents' marginal diminishing utility might apply to years of recovered service *that are provided by the program*, rather than to the years of service from the resource itself. If this is case, then respondents are valuing the program rather than the resource, and their responses do not provide information about the value of the resource, which was the stated goal of the study and is the criterion used for damage assessment.