Comments on Chapman et al. (2009): Inadequate Response to Scope while Passing the Scope Test

by
Kenneth Train
Department of Economics
University of California, Berkeley
January 24, 2012

Abstract

In contingent valuation studies that pass the scope test, the specification of the environmental goods usually does not permit an evaluation of the *magnitude* of response to scope – just its direction. A study by Chapman *et al.* (2009) is the only exception that I have been able to find. In that study, the magnitude of the response is inadequate under straightforward methods of comparison and cannot be explained by diminishing marginal utility or substitution. Moreover, the survey responses evidence an actual violation of scope, as the estimated willingness to pay is lower for a greater amount of environmental benefits. The implication of these findings is that a CV study can pass the standard scope test while providing implausible results for response to changes in scope.

1. Introduction

Contingent valuation (CV) has been used to elicit information about survey respondents' willingness to pay (WTP) for environmental improvement or to avoid environmental degradation (see, e.g., Carson and Hanemann, 2005.) A central concern with CV is whether the estimated value of an environmental amenity rises with the size, extent, or more generally "scope" of the amenity. A scope test is often applied by specifying different levels of the good in question, conducting a CV survey for each, and testing whether responses differ. Examples include: Natural Resource Damage Assessment, Inc. (1994), Bateman, *et al.*(2005), Banzhaf, *et al.* (2006), and Chapman, *et al.* (2009).

The scope test examines whether there is a statistically significant response to scope, but it does not examine whether the magnitude of response is plausible or reasonable. The difficulty, of course, is that there is little guidance for assessing the magnitude of response. In most studies, any difference in response, no matter how small, can conceivably result from diminishing marginal utility and/or substitution. Since there is no possibility of finding that the magnitude is too small, the reasonableness of the estimated magnitudes cannot be evaluated.

A study by Chapman *et al*. (2009), however, is an exception. For this study, the two goods were specified in a way that facilitates an evaluation of the magnitude of the difference in response. The study passed its scope test, evidencing a statistically significant lower willingness to pay for a program with reduced scope. However, as I describe in detail in the sections below, the difference in response to the two specified levels of scope is implausibly small under several

straightforward methods of assessing adequacy. When combined with an adding-up identity, the responses even constitute an actual violation of the scope criterion by giving a lower value to a greater amount of environmental amenity.

In the next section I describe the pertinent aspects of the Chapman *et al*. study (hereafter "the Study" and its "Report"), including the environmental good that is being valued, the reduced good that is used for the scope test, and the results as presented in the Report. I compare the mean WTP's for the original and reduced goods, showing that their relative magnitudes are implausible and, when considered in relation to an adding-up identity, actually imply a violation of the scope criterion. In section 3, I show that diminishing marginal utility and substitution, which in other contexts can potentially explain small differences in response to large differences in scope, does not explain the Study's findings.

2. The Study's Design and Findings

The goal of the Study was "to measure that natural resource damages associated with excess phosphorus from poultry waste and other sources entering the Illinois River system [within Oklahoma state] and Tenkiller Lake." The phosphorus creates excess algae that depletes the oxygen in the water, which is needed by other species to survive. Respondents were informed about the problem of excess phosphorus and that the state has requested a ban on the spreading of poultry litter but that the ban would not deal with the phosphorus that is already in the land and water. Respondents were told that the excess phosphorus could be removed by putting alum on the land and in the water, which binds to the phosphorus, rendering it harmless. Respondents were told that similar alum treatments have been used successfully in many other states. In a referendum-type question, respondents were asked about their willingness to pay for a program of alum treatment.

Two scenarios were specified: a "base scenario" that was meant to capture the value of the damaged resource, and a "scope scenario" with reduced benefits from alum treatment. For the base scenario, respondents were told that, without alum treatment, the rivers[1] would recover in 50 years and the lake would recover in 60 years as a result of the ban on spreading of poultry litter. With alum treatments, the rivers would recover in 10 years instead of 50, which is 40 years earlier, and the lake would recover in 20 years instead of 60, which is also 40 years earlier. Figure 1 depicts the impact of the alum treatments as shown to the respondents who faced the base scenario.

For the scope scenario, the natural recovery and the impact of the program were specified differently than in the base scenario. Respondents were told that the rivers would recover quickly on their own without treatment because of the ban. The alum treatment was for the lake, making its recovery "somewhat faster." In particular, respondents were told that, without alum treatment, the rivers would recover in 10 years and the lake would recover in 60 years, as a result of the ban. With alum treatment of the lake, the lake would recover in 50 years instead of 60 years, which is 10 years earlier. Figure 2 depicts the impact of alum treatments as shown to the respondents who faced the scope scenario.

[1] I use the term "the rivers" for linguistic convenience, while the Study used "the river and creeks."

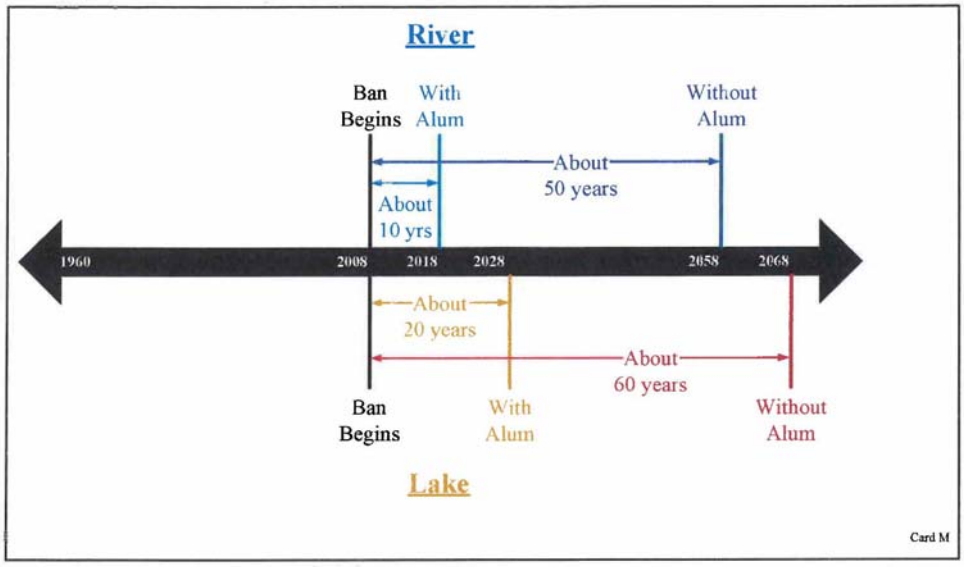

Figure 1: Base Program

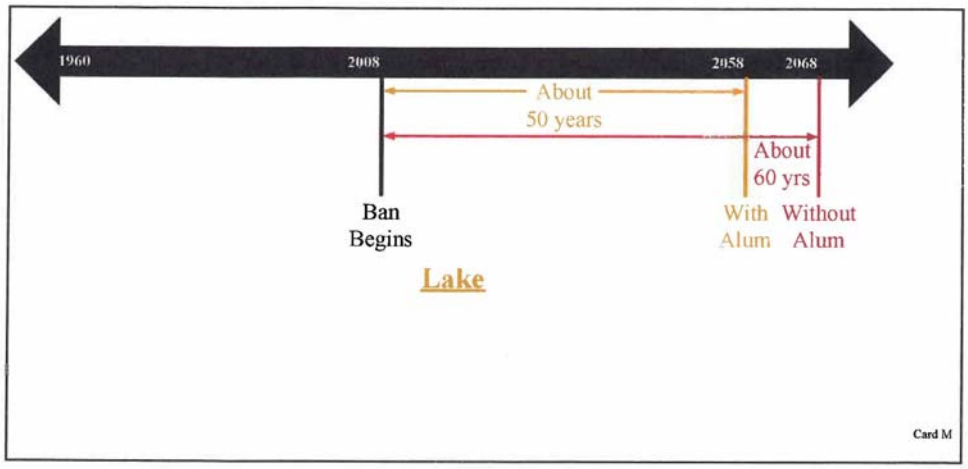

Figure 2: Scope Program

A referendum-type question was asked under both scenarios: whether the respondent would vote in favor or against a proposition under which the state implements the program of alum treatment

and each household is levied a one-time tax of $X to pay for the treatment, with X varied over respondents from $10 to $405. Based on the survey responses, the Report states that its estimate of the mean WTP per household is $184.55 for the base program and $138.51 for the scope program.

To repeat in summary form: The base program, which speeds recovery of the rivers and the lake by 40 years each, is estimated to be valued at $184.55. The scope program, which provides no benefit for the rivers and speeds the recovery of the lake by only 10 years, with these 10 years starting 50 years in the future, is estimated to be valued at $138.51, which is 75 percent as much as the base program. This difference is implausibly small, for the following reasons.

For comparison, suppose (i) a year of faster recovery of the rivers is as valuable as a year of faster recovery of the lake, and (ii) the standard form of calculating present values of future streams of environmental services applies, such that 40 years of service starting 20 years from now has at least 4 times the present value of 10 years of additional service starting 50 years from now. The value of the scope program in this case would be less than one-eighth of the value of the base program. But in this Study, the relative value is estimated to be six times larger: 75% of the base program's value. If river recovery is valued differently than lake recovery, then the scope program would still be worth at most one-fourth as much as the base program, no matter how little the river recovery is valued -- but its estimated value is three times larger at 75%. Any strictly positive discount rate makes the comparison even more implausible.

We can also examine the decomposition of benefits that is implied by the scope program. The base program consists of two sets of amenities: those included in the scope program and those not included in the scope program. For convenience, I call the latter the "non-scope elements." These non-scope elements speed the recovery of the rivers by 40 years and induce the lake to recover 30 years sooner than under the scope program. If the base program is valued at $184.55 and the scope program at $138.51, then the non-scope elements are valued at the difference: $46.04. This constitutes an actual violation of the scope criterion. The non-scope elements provide more service than the scope program in each dimension: more types of resources (the rivers and the lake *versus* just the lake), more years of service (40 years of recovered river service and 30 years of recovered lake service *versus* 10 years of recovered lake service), and a closer time period (recovery starting 10 years in the future for the rivers and 20 years in the future for the lake *versus* 50 years in the future for the lake). Yet the non-scope elements are found by the Study to be valued less than the scope program -- a third as much.[2]

3. Diminishing marginal utility and substitution

The most prominent reason that is given for WTP not to be somewhat proportional to the amount of the environmental good is that people have diminishing marginal utility of environmental goods, especially with respect to existence value (see, e.g., Rollins and Lyke, 1998). However, diminishing marginal utility cannot meaningfully explain the results in the Study.

[2] The hypothesis that the value of the non-scope elements equals or exceeds the value of the scope program can be rejected at a confidence level of over 99.9 percent, using the confidence intervals reported on page 7-5 of the Report.

A salient feature of the Study's design is that the scope program is incremental to (i.e., provides benefits in addition to) some of the benefits provided in the base program. In particular, the base program includes 40 years of faster recovery of the rivers. In the scope scenario, this faster river recovery is said to occur naturally, 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. Diminishing marginal utility might cause the scope program to be worth less than the river portion of the base program, since the scope program is incremental to this river recovery. But, instead, the Study finds that the scope program is worth more than the river recovery to which it is incremental (as well as more than the remaining part of the lake recovery).

The relation between the two programs can be described more explicitly as follows. The base and scope scenarios, when considered in relation to each other, represent four successive increments of environmental improvement:

A. Recovery of the rivers starting in 50 years and recovery of the lake starting in 60 years.
B. Recovery of the rivers 40 years faster (so that the rivers recover in 10 years instead of 50.)
C. Recovery of the lake 10 years faster (so that the lake recovers in 50 years instead of 60.)
D. Recovery of the lake 30 years faster (so that the lake recovers in 20 years instead of 50.)

Under the base scenario, A is obtained as a result of the ban on spreading of poultry litter, and the program of alum treatment provides B, C and D environmental benefits in addition to (i.e., as an increment over) A. Under the scope scenario, A and B are obtained as a result of the ban, and the program of alum treatment provides C as a benefit in addition to (i.e., an increment over) A and B. Let *WTP(x)* denote the WTP for *x* as an increment over the previous amounts; e.g., *WTP(C)* is the WTP for C given that A and B are already consumed. The Study found that $WTP(B)+WTP(C)+WTP(D)=184.55$ and $WTP(C)=138.46$, which implies that $WTP(B)+WTP(D)=46.09$.

Suppose, as a first attempt at using diminishing returns as an explanation, that respondents think in terms of one environmental good that is consumed in incrementally larger quantities such that diminishing marginal utility reduces the WTP for each extra unit. Diminishing returns might cause WTP to decrease for successive increments, e.g., $WTP(B)>WTP(C)$ and/or $WTP(C)>WTP(D)$. However, the Study found that $WTP(C)>WTP(B)+WTP(D)$, which implies that $WTP(C)>WTP(B)$ -- the opposite relation from that implied by diminishing marginal utility. Stated in words: the scope program provides benefits that are incremental to recovery of the rivers; yet the responses indicate that WTP for the scope program is greater than for the river recovery to which it is incremental.

Respondents might not consider there to be one environmental good. Perhaps river and lake recovery are considered by respondents to be two different environmental goods. Suppose further that respondents care so much less about rivers than the lake that their $WTP(C)>WTP(B)$ even though B provides 40 extra years of river recovery while C provides only 10 extra years of lake recovery and is incremental to B. In this case, the Study's finding that $WTP(C)>WTP(B)+WTP(D)$ is not necessarily inconsistent with diminishing marginal utility

for differentiated environmental goods. E.g., if $WTP(B) = 0$, then the Study's results imply that $WTP(C) > WTP(D)$, which is in the direction expected for diminishing marginal utility since D is incremental to C.

However, even in this case, while the direction might be plausible, the magnitude of the difference is not. Again, assuming that $WTP(B) = 0$ (with values over 0 exacerbating the issue), the Study's results imply that $WTP(D) = 46.04$ while $WTP(C) = 138.51$, despite the fact that D 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[3], 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 the alum program and let y^0 be the years of recovered lake service that the consumer obtains without alum treatment, with $y^0 > 0$ since recovery occurs eventually from just the ban. The Study found that 10 extra years of lake recovery is worth \$138.51 and 40 extra years is worth \$184.55, such that $U(x^0 - 184.55, y^0 + 40) = U(x^0, y^0) = U(x^0 - 138.51, y^0 + 10)$. The value of α that equates the two sides is $\alpha = -0.1376$. Using this value of α in either of the two equations gives $U(x^0, y^0) = x^0 - 185.30$ such that $U(x, y) = x - 185.30 * \exp(-0.1376(y - y^0))$.

Table 1 gives the utility for different numbers of years of recovered lake service.

Table 1: 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 – 2905.90	
y^0 - 10	x^0 - 733.81	2172.09
y^0	x^0 - 185.30	548.50
y^0 + 10	x^0 - 46.79	138.51
y^0 + 20	x^0 - 11.82	34.98
y^0 + 30	x^0 - 2.98	8.83
y^0 + 40	x^0 - 0.75	2.23

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 138.51, and the WTP for an extra 40 years given y^0 years is \$184.55 (i.e., the sum of

[3] For the calculations to follow, I 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.

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 $2,172. If the Study had told respondents that the lake would recover without alum treatment 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 $2,894 (2172.09+548.50+138.51+34.98) 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 $2894) arose simply from specifying that the ban causes lake recovery at different arbitrary years in the distant future, without any change in the number of extra years of lake recovery that is obtained (40 more years in each case.) As I show in the appendix, 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.

For similar reasons, substitution between rivers and the lake cannot rationalize the Study's results. The scope program (i.e., C) is incremental to the river recovery (B). If rivers and lakes are substitutes, the value of C would be less when provided as an increment to B (as specified in the Study) than if B were not provided. Substitution reduces the value of C, and hence might explain a low value for C. Yet the problem with the Study's results is that its estimated value for C is too *high* relative to B and D.

References

Banzhaf, H. S., D. Burtraw, D. Evans, and A. Krupnik, 2006, "Valuation of Natural Resource Improvements in the Adirondacks," *Land Economics*, Vol. 82, No. 3, pp. 445-464.

Bateman, I., P. Cooper, S. Georgiou, S. Navrud, G. Poe, R. Ready, P. Riera, M. Ryan, and C. Vossler, 2005, "Economic Valuation of Policies for Managing Acidity in Remote Mountain Lakes: Examining Validity Through Scope Sensitivity Testing," *Aquatic Sciences*, Vol, 67, pp.274-291.

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

Chapman, D., R. Bishop, W. Hanemann, B. Kanninen, J. Krosnick, E. Morey, T. 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, in 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.)

Natural Resource Damage Assessment, Inc., 1994, *Prospective Interim Lost Use Value Due to PCB and DDT Contamination in the Southern California Bight,* report to the National Oceanic and Atmospheric Administration under task order 56-DGNC-3-50070.

Rollins, K., and A. Lyke, 1998, "The Case for Diminishing Marginal Existence Values," *Journal of Environmental Economics and Management*, Vol. 36, No. 3, pp. 324-344.

Appendix

This appendix addresses two issues that readers might consider relevant to the discussion in the text, namely: alternative specifications of utility, and differences in bounds.

A1. Alternative specification of marginal utility

For Table 1, I assumed a particular parameterization of utility that is useful to illustrate the issue, but alternative representations are also consistent with the Study's finding. Suppose instead that marginal utility takes a form that avoids the increasingly high values in the top rows of Table 2 and yet is still consistent with the Study's estimates for the base and scope programs. To be concrete, suppose diminishing marginal utility has the extreme form (i.e., most favorable for the Study) shown in Table A1. WTP is constant for speeding recovery from 80 years to 70 years, from 70 years to 60 years, and from 60 years to 50 years; then WTP drops for speeding recovery from 50 years to 40 years; after which WTP drops to zero for any further speeding of recovery. As required, WTP for the scope program (which provides recovery in 50 years instead of 60) is 138.51 and that WTP for the base program (which provides recovery in 20 years instead of 60) is 184.55 (i.e., 138.51+46.04), which are the values found by the Study.

Table A1: Diminishing Marginal Utility of Lake Recovery that Most Favorably Explains The Study's Estimates for Base and Scope Program, assuming no value to river recovery.

	WTP
Recovery in 70 years instead of 80	138.51
Recovery in 60 years instead of 70	138.51
Recovery in 50 years instead of 60	138.51
Recovery in 40 years instead of 50	46.04
Recovery in 30 years instead of 40	0
Recovery in 20 years instead of 30	0

If this pattern of marginal utility were actually the reason for the Study's results, then it would be quite a coincidence that the Study happened to select the amounts of recovery right before and after the point at which the rate of diminishing returns changes drastically: recovery in 50 years instead of 60 for the scope program, which are the last years for which the rate of diminishing marginal is still low (so that previous years do not have implausibly high marginal value), and recovery in 20 years instead of 60 for the base program, during which returns decline very sharply. Moreover, the Study's findings would have been very different if the recovery dates were changed 10 years in either direction. If the ban had been specified to induce recovery in 70 years (10 years later than the Study specified) and the base program had been described as speeding recovery from 70 to 30 years (the same number of years of faster recovery as the Study used), then the estimated WTP would have been 323.06 (138.51+138.51+46.04+0). And if the ban had been specified as inducing recovery in 50 years with the base program speeding recovery from 50 to 10 years, then the estimated WTP would have been 46.04 (46.04+0+0+0). Since the Study's purpose was to value the resource, its results would be deemed unreliable for this purpose if changes in estimated value of this magnitude – a factor of six from 46.04 to

323.06 – arose simply from specifying that the ban causes lake recovery at different arbitrary years in the distant future.

A2. Bounds

The Study's estimate of average WTP for the base and scope program implements the conservative assumption that all people within a WTP interval have a WTP at the lower end of the interval. E.g., the Study found that 81.5% of respondents voted in favor of the referendum on the base program when the cost is given as $10 and that 70.1% voted in favor at a cost of $45, which implies that 11.4% (81.5-70.1) of respondents are willing to pay somewhere between $10 and $45 for the base program. The Study used the lower value of $10 for all 11.4% of them, and similarly for the other ranges. The average WTP for the scope study was calculated in the same way. I have argued that the average WTP for the scope program is implausibly high relative to the average WTP for the base program.

A potential issue with this approach is that relative values of the lower bounds for two statistics do not necessarily reflect the relative values of the statistics themselves. To investigate the potential impact of this issue, we can make opposite assumptions for the two WTP calculations. That is, assume that (i) for the base program all respondents within an interval have WTP at the higher end of the interval, and (ii) for the scope program all respondents within an interval have WTP at the lower end of the interval. This procedure provides the highest possible estimate of WTP for the base program and the lowest possible estimate of WTP for the scope program. If there is still an insufficient difference between these two estimates ("inadequate responsiveness" to scope), then my previous conclusions remain valid.

The comparison just described cannot be completely implemented since the top interval of WTP is "over $405" which has no finite higher end. To account for this fact, I calculated the WTP for people in this top interval that would be required in order to attain a given average WTP for the base program. To attain four times as high a WTP for the base program as the scope program, which is the ratio that would occur if respondents placed no value on river recovery and applied no time discounting to the lake service, the people in the upper interval would have to be willing to pay $1,364 on average for the base program. And to attain eight times the WTP of the scope program, which is the relation that would occur if river service was valued the same as lake service but respondents still did not discount for time, the top-interval people would need to have a WTP of $2,984 on average. The fact that the survey did not ask respondents about costs over $405 suggests that values this high were not considered reasonable even by the Study designers.