CONTINGENT VALUATION SURVEYS

AND TESTS OF INSENSITIVITY TO SCOPE*

Richard T. Carson
Department of Economics
University of California, San Diego
La Jolla, CA 92093

Draft

January 1995

*An Early Version of This Paper Was Presented at the International Conference on Determining the Value of Non-marketed Goods: Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods, Bad Homburg, Germany, July 1994.

Introduction

The central claim made by critics of contingent valuation (CV) is that respondents to a CV survey are inherently unresponsive to the characteristics of the good being valued.[1] In particular, it is alleged that respondents are not willing to pay more for more of a particular good; in economic jargon, respondents are insensitive to the scope of the good offered in the valuation exercise.[2] The focus of this debate is not whether CV results showing insensitivity to scope are obtainable, a point readily conceded by CV practitioners, but rather whether this result is generally avoidable with appropriate survey design, pretesting, and administration.

In the first section of the paper, I provide a historical sketch of the scope insensitivity debate. In particular, I discuss the claim of generic scope insensitivity in CV surveys and the evidence used to support this claim. I examine some of the tangential claims about respondents' behavior in CV surveys which can be seen to provide interpretations as to why the phenomenon of scope insensitivity may be observed. I also look at some of the subtle interplay between cognitive psychology and economics which underlies the scope insensitivity debate.

In the second section, I advance arguments and indirect evidence which suggest that the hypothesis that CV survey results will typically exhibit scope insensitivity is unlikely to be true. First, drawing upon the literature on surveys of public policy questions, I investigate whether

---

[1]Examples of such criticisms can be found in *Contingent Valuation: A Critical Response*, the collection of Exxon-sponsored research on CV (1993; J.A. Hausman ed.) and testimony on behalf of several oil and chemical companies by Peter Diamond, William Desvousges, Jerry Hausman, Daniel Kahneman, Daniel McFadden, Steven Shavell and others before the NOAA Blue Ribbon Panel on Contingent Valuation co-chaired by Kenneth Arrow and Robert Solow.

[2]We use the term "scope" rather than "embedding" because scope is the term used in the recommendation section of the NOAA Blue Ribbon Panel (Arrow *et al.*, 1993) and is a well-defined economic concept. In contrast, embedding (Kahneman and Knetsch, 1992a) has been used to describe a variety of different phenomena, some of which represent violations of economic theory and some of which are predicted by it (Smith, 1992; Carson and Mitchell, forthcoming).

there is evidence that respondents are insensitive to differences in the way policies are described. Next, turning to the CV literature, I argue that there are various ways in which scope insensitivity would likely be reflected if one looks across and within CV studies.

In the third section, I examine the formal CV tests of the scope insensitivity hypothesis which have been conducted using two or more independent samples. In contrast to assertions by Diamond and Hausman (1994) that few tests of scope insensitivity which use independent samples exist, other than those conducted by Exxon and Kahneman, I show that there are over 30 such tests which provide evidence on this point. Because some CV critics, such as Desvousges *et al.* (1993) draw a distinction between goods with principally direct use values and those with substantial passive use values, while others do not (Kahneman and Knetsch, 1992a), I address whether the results from this literature suggest that a distinction should be drawn.

1.0 Scope Insensitivity: Origins and Frequently Cited Evidence

The notion that CV survey results are likely to be insensitive to the scope of the good being valued originated with Kahneman (1986).[4] Kahneman argued (on the basis of a graph he presented) that respondents' willingness to pay (WTP) for cleaning-up all of the lakes in the Canadian Province of Ontario was not much larger than cleaning-up all of the lakes in Muskoka, a small part of the Province. He also offered an explanation for the phenomena: respondents were expressing ideological rather than economic values. Further, Kahneman argued that CV researchers had not observed the result because the experiments which varied the scope of the good being valued were with-in subject designs.

[4]Kahneman's original presentation was at a 1984 U.S. Environmental Protection Agency workshop. A written account can be found in Cummings, Brookshire, and Schulze (1986).

Kahneman's argument received considerable attention for four reasons. First, he claimed to have presented empirical evidence that CV results violate the simple economic maxim that people should be willing to pay more for more of something they want. Second, he provided an explanation for this occurrence which sounded, to many CV critics, like an argument that people were not really willing to pay the amount they stated. Third, he provided a simple explanation of why CV researchers had missed the scope insensitivity effect. Fourth, Kahneman was well respected in the research community and known for his interesting work on human behavior and response effects.

Contingent valuation researchers took the possibility of scope insensitivity seriously but attributed it to survey design and administration problems rather than to fundamental cognitive difficulties. Mitchell and Carson (1989) devoted considerable attention to different survey problems which might produce an apparent insensitivity to the characteristics of the good being valued.[5] They emphasized that respondents may confuse the good being valued with either a larger or smaller good (part-whole bias), that a poorly described good may be perceived as symbolic of a larger good (symbolic bias), that the researcher could be defining the good in a different metric than that used by respondents (metric bias), and that respondents might be skeptical that the good would actually be provided (probability of provision bias). In all of these cases, the solution was to present a more understandable description of the good and a more plausible provision context. Fischoff and Furby (1988) interpreted the problem similarly by

[5]Kahneman's results had come from a short telephone survey with very little description of the good or how it would be provided. Mitchell and Carson (1989) informally compared estimates of willingness to pay from two very similar surveys, one valuing national water quality changes and the other valuing regional water quality changes, noting the large difference in estimates. A formal statistical test of this comparison is provided in Carson and Mitchell (1993).

emphasizing that respondents would tend to fill in any of the key aspects of the good or provision context not explicitly provided by the CV scenario.

Another problem, which can give the appearance of scope insensitivity, occurs when the researcher believes one good encompasses another, but respondents find the two goods offered indistinguishable. For example, suppose an ecosystem which provides habitat for five endangered species is at risk. As a scope test, the researcher informs one sample that the habitat will be purchased to protect the five species and then informs another sample that the habitat will be purchased to protect only two species. Respondents in the second sample may reason that protecting the habitat will provide protection for all five species and are therefore paying for the same good as that offered to the first sample. Economists call this joint production. From a CV perspective, it is simply a case where the survey designer is either unaware of the problem (and therefore, a form of part-whole bias) or a case where respondents refuse to accept the clear counter-factual in formulating their response.[6]

As a rule, economists do not traffic in motives with this principal enshrined as respect for consumer sovereignty. As a result, Kahneman's notion of ideological values did not go very far on its own. Unexpected findings often result in a questioning of the adequacy of the assumed economic theory and provoke a search for alternative explanations.[7] Kahneman and Knetsch (1992a) interpret contingent valuation responses as follows: "Contingent valuation responses reflect the willingness to pay for the moral satisfaction of contributing to public goods, not the

---

[6]CV researchers became aware of this problem in the context of valuation of air quality improvements where the U.S. Clean Air Act does not allow the U.S. Environmental Protection Agency to assess health benefits in dollar terms but allows non-health benefits such as visibility improvements to be assessed in dollar terms. See Carson, Mitchell and Ruud (1990), Fischoff and Furby (1986), and McClelland *et al*, (1991) for discussions and empirical applications.

[7]See, for instance, Hanemann's (1991) theoretical demonstration that the tight link between willingness to pay and willingness to accept for price changes need not hold for imposed quantity changes.

economic value of these goods." However, from an economic perspective, motives are of little importance as long as the agent would actually pay for the good rather than do without it.[8] The fact that the good is a public good further imposes an obligation on the government (or the survey designer) that people clearly understand what they are paying for. Insensitivity to scope suggests that one of these two conditions is being violated.

As a postscript on the origins of the scope insensitivity hypothesis, it is interesting to note that Kahneman and Knetsch (1992a) repeat the earlier Kahneman (1986) assertion:

> "A finding that we obtained some time ago illustrates the embedding effect: the expressed willingness of Toronto residents to pay increased taxes to prevent the drop in fish populations in all Ontario lakes was only slightly higher than the willingness to pay to preserve the fish stocks in only a small area of the province [Muskoka]."

This finding is now typically attributed to Kahneman and Knetsch (1992a), rather than Kahneman (1986), where it was actually introduced. The new results contained in Kahneman and Knetsch (1992a) which pertain to scope are inconsequential and have received little attention. However, the Ontario lakes example is cited throughout the recent benefit-cost literature (*e.g.*, Zerbe and Dively, 1994) and in the decision science literature (Kleindorfer, Kunreuther, and Schoemaker, 1993), both as a "fact" and as the most telling criticism of contingent valuation. What is unfortunate about this turn of events is that no statistical test of this claim was ever presented in either Kahneman (1986) or Kahneman and Knetsch (1992a). The graph in Kahneman (1986), however, can be used to obtain an approximate estimate of willingness to pay for all Ontario lakes and the Muskoka lakes. Performing this exercise shows that WTP for all

---

[8]Harrison (1992), in a now famous line, responds with the standard economic position: "I call my utility 'jolly.' What you choose to call your utility is, as far as I am concerned, your business. What Kahneman and Knetsch want to call it is, of course their business. To be blunt, this hypothesis is 'meaningless' in the standard methodological sense. To paraphrase Samuelson (1947, p. 92), it places no definite restrictions on observable behavior, even under ideal observational conditions."

Ontario lakes is 50% higher than for the Muskoka lakes. While it is possible to argue about what the magnitude of the difference should be, the difference is clearly not "only slightly higher," and experiments incapable of detecting a 50% difference seem pointless to conduct.

The actual story in Kahneman and Knetsch (1992a) is, for the most part, quite different from that of scope insensitivity, and instead, involves the sequence in which public goods are offered. Because this issue continues to caused a great deal of confusion, I now turn to it.

1.1 Economic Theory Behind the Tests of Sequence and Scope

Kahneman's argument was quickly embraced by economists advocating industry's position with regard to the use of CV, as well as those generally opposed to benefit-cost analysis. However, many economists have been reluctant to accept Kahneman's story for two primary reasons. One reason is the *ad hoc* nature of Kahneman's story; more pointedly, his story is not based on a behavioral model but rather on an isolated explanation. For example, in his comment in *Valuing Environmental Goods* (Cummings, Brookshire, and Schulze, 1986), Kahneman draws the distinction between CV studies of use values and CV studies of passive use values. The former appear to be for economic values and the latter for ideological values. From a theoretical perspective, Kahneman's story is very much like Robert Louis Stevenson's Dr. Jekyll and Mr. Hyde. For goods with use values, agents exhibit all of the selfish virtues of *homo economicus*. However, for goods with substantial passive use values, agents are transformed from *homo economicus* into idealogues whose responses are suddenly stripped of any economic content.[9] Kahneman's story fails to provide any link between these two beasts. No appeal to any special or unusual conditions is necessary in order to offer an economic

[9] Kahneman has since reversed his position and now condemns almost all contingent valuation values.

explanation for much of Kahneman's empirical evidence. The ability and willingness to grasp the economic story, however, is influenced by the filters through which one views the evidence. For the economist, the economic story does not offer a surprise ending. However, for the cognitive psychologist looking for evidence against economic theory, the story resembles a black comedy for which missing the ending comes as no great loss.

The recent articulation of the Kahneman critique in Kahneman and Knetsch (1992a) presents two lines of criticism. The first of which is not relevant to scope, but rather argues that economic values may be subject to a "regular embedding" effect.[10] Although the quality of the Kahneman and Knetsch study which provided the evidence in support of this criticism has been severely criticized (Smith, 1992), the claim is one with which Smith and most other economists would not disagree, as economic theory predicts this result.[11]

The second line of criticism is a repackaging of the ideological values criticism found in the Kahneman (1986) comment and is directly related to scope. The name is changed from ideological values to a catchier phrase, the "purchase of moral satisfaction." Kahneman and Knetsch use another term, "warm glow" borrowed from Andreoni (1990), to describe what they believe individuals actually gain when responding to contingent valuation surveys.[12] According

[10] In their original paper, Kahneman and Knetsch refer only to embedding. They draw a distinction between regular embedding and perfect embedding in their reply (1992b) to Smith's (1992) comment.

[11] Indeed, the recent Flores (1994) analysis shows that because one is effectively dealing with an inverted quantity constrained demand system, large sequence effects are likely to be the rule not the exception.

[12] Andreoni's (1990) theory of warm glow offers an explanation of the confounding empirical evidence against the neutrality hypothesis. This hypothesis states that government expenditures on public goods will perfectly crowd out charitable donations for the provision of these same goods. The theory of warm glow presents a model in which government provision is an imperfect substitute for private donations because individuals receive a warm glow from the act of giving not available when the same level is provided through lump sum taxes. Two qualifications of the adoption of the "warm glow" term by CV critics are in order. First, Andreoni's model does not specify a warm glow for anonymous payment of taxes, which is perhaps the most common payment mechanism used in contingent valuation surveys. Second, with multiple public goods, the warm glow must come from the simple act of paying taxes irrespective

to Kahneman and Knetsch, individuals are only basking in their warm glow and will pay basically the same for any good for which they are asked to value. That is to say, they are economically insensitive to the scope of the good they are asked to value.[13] In their reply to Smith, Kahneman and Knetsch refer to this phenomenon as "perfect embedding."

1.2 Predictions From Economic Theory: Nested Goods[14]

In this subsection, I look at valuing sequences of public goods and the theoretical structure of the tests involving sequencing and nesting of such goods. To help make the discussion more concrete, I define three different valuation sequences in Table I, each of which is to be given to a different but equivalent (in a statistical sense) subsample. The valuation sequences (and subsamples) are denoted I, II, and III. Three different goods are offered: **A**, **B**, and **C**. The order in which the goods are offered in the sequence is denoted by a subscript, 1, 2, or 3. It is important to note that economic theory says little about the relationship between the elements of the three valuation sequences without specifying more about the nature of the three goods including their relationships to each other. I will make two assumptions: (1) all of the goods being considered are normal goods, a reasonable assumption for most environmental goods, and (2) all goods in the set of interest are individually and in any group Hicksian substitutes for each other. This latter assumption, while frequently made, is probably

---

of whether the government provides anything in return.

[13]Kahneman and Knetsch do advance an argument that differences in more satisfaction, not economic motivations, may drive differences in WTP. However, it is unclear what Kahneman and Knetsch believe is and is not a legitimate economic motive, as they fail to provide a list of acceptable economic motivations. See Samuelson (1993) for a recent restatement of the neoclassical economic position which rejects the notion that some motives are not legitimate determinants of economic behavior.

[14]This section is largely taken from Carson and Mitchell (forthcoming).

less generally true than the former normality assumption. These two assumptions cause the income and substitution effects to work in the same direction in a WTP valuation sequence.[15]

TABLE I: STRUCTURE OF VALUATION SEQUENCES		
Subsample I	Subsample II	Subsample III
A_I^1		
B_I^2	B_{II}^1	
C_I^3	C_{II}^2	C_{III}^1

Under these conditions, it can be shown that two external WTP tests of economic theory are possible: (1a) $B_I^2 \leq B_{II}^1$ and (1b) $C_I^3 \leq C_{II}^2 \leq C_{III}^1$. Tests of the hypotheses represented by equations (1a) and (1b) will be termed *embedding effect tests*, in that they look at the effect of embedding a good farther down in a specific valuation sequence. If we make the further assumption that the respondent will always get positive utility from possessing any of these goods regardless of whether the respondent already possesses any or all of the other goods at issue, the weak inequalities above can be replaced with strong inequalities. Economic theory does not say anything *a priori* about comparisons between different goods, e.g., (B_j^1, C_j^1).

Nested goods, where B is a proper subset of A and C is a proper subset of B, are a special case of conventional goods so that (1a) and (1b) hold if we make the same assumptions. To get the strong inequality in the nested goods case, one must assume for (1a) that respondents still get positive utility from B after they possess its complement B* and, for (1b), that respondents still get positive utility for C after they possess C*. Testing a null hypothesis of

[15]These two assumptions work in the opposite direction in a willingness to accept compensation valuation sequence (Carson, Flores, and Hanemann, 1994).

equality in an embedding effect test is essentially testing for consistency with a flat utility function with respect to further increases in the good (*e.g.*, WTP(**B**), conditional on having **B***, is zero). Note that, altl ,n a flat utility function with respect to **B** (or **C**) conditional on possessing **B*** (**C***) might not be considered well behaved, it is not ruled out by economic theory assuming local nonsatiation does rule out such behavior.

The nesting property also allows predictions about other possible comparisons the first of which are internal tests using top-down or bottom-up sequences.[16] Here, since a top-down valuation sequence is used, it can be shown that: (2a) $\mathbf{A}_{\mathrm{I}}^{1} \geq \mathbf{B}_{\mathrm{I}}^{2} \geq \mathbf{C}_{\mathrm{I}}^{3}$ and (2b) $\mathbf{B}_{\mathrm{II}}^{1} \geq \mathbf{C}_{\mathrm{II}}^{2}$. A variant on (2a) and (2b) using a bottom-up approach can be written as: (2c) $\mathbf{C}_{\mathrm{I}}^{1} \leq \mathbf{B}_{\mathrm{I}}^{2} \leq \mathbf{A}_{\mathrm{I}}^{3}$, and (2d) $\mathbf{C}_{\mathrm{II}}^{1} \leq \mathbf{B}_{\mathrm{II}}^{2}$. It is important to note that the willingness-to-pay amounts for the goods in the top-down and bottom-up valuation sequences will, in general, not be the same. This is because of the differences respondents perceive in the choice set implied to be available (initially large or broad depending on the type of nesting in the top-down sequence, and initially small or narrow in the bottom-up sequence) and because of differences in whether the respondent believes he or she already possesses the complement of the good being valued (the top-down case) or has just purchased it (the bottom-up case).[17] The (2a) and (2b) results follow from monotonicity assumptions. Again, the weak inequalities may be replaced by strong inequalities if it is

[16]We define a top-down sequence as essentially asking, "How much less are you willing to pay for **A** if **B** is not part of **A**?" and a bottom-down as, "How much more are you willing to pay for **A** given that you have just paid for **B**?"

[17]It is of course possible to define top-down and bottom-up sequences which, from a theoretical perspective, ask the same thing. However, the implementation of one of the two sequences will almost always involve asking respondents to "forget" the choice information they have previously been given, something which survey respondents generally do not do (Schwarz and Strack, 1991). This problem is avoided with the test described in equation (3) below.

assumed that respondents get utility from the complement of the element being valued. Tests of the hypotheses represented by these equations will be called *nested sequence tests*.

Another comparison which can be made is along the diagonal of Table I: (3) $A_I^I \geq B_{II}^I \geq C_{III}^I$. I will term this comparison a *test of component sensitivity*. It is the type of *scope insensitivity hypothesis test* generally implemented in recent CV surveys. This test looks at whether respondents are sensitive to differences in levels of the good (quantitative nesting) or the level of inclusion (categorical nesting) in a way that should change its value. It is the external test version of equation (2) and has the same theoretical properties as in (2).[18]

It is important to note that this test does not say anything about the magnitude of the difference that should be observed. Some CV critics, such as Diamond (1993), have proposed theoretical tests which they claim allow one to test not only the direction of the difference but also the magnitude. Thus far, the validity of such tests depends crucially on a set of strong auxiliary assumptions which seem unlikely to be satisfied for the types of goods usually valued with contingent valuation surveys.[19]

---

[18]Note that this test has quite different statistical properties because it is a between sample test rather than a within sample test. Between sample tests will generally require substantially larger sample sizes for the same level of precision in the estimate of the difference between WTP estimates for different goods. Three other factors can also substantially reduce the statistical power of tests of (3). The first is to ignore the known relationship between the goods being valued if there are more than two goods. The second is to conduct the test in a situation where there are a large number of respondents who do not value any level of the good. The third is to use a two-sided hypothesis test when (3) clearly calls for a one-sided test.

[19]For instance, Smith and Osborne (1994) and Hanemann (1994) show that Diamond's assumption in the example he uses from the Desvousges *et al.* (1992) bird study is equivalent to assuming that people are indifferent between a government program which would keep some number of birds from being killed by oil and one which would simply increase the bird population by the same number. More generally, Smith and Osborne (1994) show by moving from Diamond's quasi-linear utility function to something as simple as a CES utility function, the adding-up condition proposed by Diamond need no longer hold.

1.3 Kahneman and Knetsch Scope and Sequence Test

Table II displays the results from Kahneman and Knetsch's (1992a) embedding test which was implemented using a short telephone survey. Drawing on Table I and the discussion of it, going across the rows provides a test of the embedding effect and the results are in accord with economic theory. Looking down the columns, one finds the nested sequence effect tests. These results are also in accord with economic theory. Down the diagonal, one finds the test of component sensitivity which is the hypothesis test for scope insensitivity. A comparison of the first two diagonal elements suggests scope insensitivity; however, there was little reason for Group I respondents to think that the long description of improved disaster preparedness that was read to the Group II respondents was actually contained in their description of the environmental services that would be delivered (Smith, 1992; Harrison, 1993).[20] Looking at the second two diagonal elements, one observes again the expected result.

TABLE II: WILLINGNESS TO PAY FOR SELECTED CLASSES OF GOODS AND ALLOCATIONS OF TOTALS TO LESS INCLUSIVE GROUPS

Public Good		Group I (N=66)	Group II (N=78)	Group III (N=74)
Environmental Services	Mean Median	$135.91 $50.00		
Improve Disaster Preparedness	Mean Median	$29.06 $10.00	$151.6 $50.00	
Improve Rescue Equipment Personnel	Mean Median	$14.12 $1.00	$74.65 $16.00	$122.64 $25.00

[20]It should also be noted that respondents are usually suspicious that much will actually be provided by paying for vaguely defined improvements in a broad aggregate commodity. This probability of provision effect with large goods (Fishchoff *et al.*, 1993) can lead to situations where the less inclusive good is actually valued more highly.

1.4 The Exxon Scope Experiments

Along with the Kahneman studies, the Exxon-sponsored scope experiments in the Hausman volume (1993)—Diamond *et al.*, Desvousges *et al.*. Schkade and Payne—are usually cited as the principal evidence supporting the contention that contingent valuation estimates are not responsive to the nature of the good being valued.[21] Indeed, it is often asserted that these are basically the only studies which have looked at the issue. For instance, Diamond and Hausman in their 1994 *Journal of Economics Perspectives* symposium paper on contingent valuation make the following remarks with respect to the embedding effect (insensitivity to scope) which they call the "main contingent valuation anomaly":[22]

> "With a pattern of results that are inconsistent with the usual economic assumptions, two interpretations are always possible: the surveys were defective or the contingent valuation method as currently practiced does not measure preferences with accuracy. One should consider all the surveys that attempt to test for consistency in order to judge which interpretation is likely to be correct. The studies we have described [Kahneman and Exxon's] have been criticized as not done well enough to be an adequate test. However, they are the only quantitative tests we are aware of. No comparable comparison tests have been done by proponents of the accuracy of contingent valuation, although the embedding effect has long been recognized." [bracketed material added for clarity]

1.4.1 Diamond *et al.*

Diamond *et al.* (1993) use a number of split-samples to ask respondents about programs involving different wilderness areas. Here I will only address their scope test which is labeled

---

[21]The fourth reported Exxon experiment, Kemp and Maxwell (1993), looked at different valuation sequences but not at scope, and as a result, is not discussed here.

[22]Diamond and Hausman say "An example of embedding would be a willingness-to-pay to clean up one lake roughly equal to that for cleaning up five lakes, including the one asked about individually. The embedding effect is usually thought to arise from the non-existence of individual preferences for the good in question and from the failure of survey respondents, in the hypothetical circumstances of the survey, to consider the effects of their budget constraints." Much of Diamond and Hausman's actual discussion concerns sequencing effects and here, after dismissing the importance of income and substitution effects, they conclude: "If a survey question reveals a true valuation, it should not matter whether the question is asked by itself or with other questions, nor if asked with any other questions, what the order of questioning is."

as Hypothesis I and operationalized as a test of whether willingness to pay for each of three different wilderness areas was equivalent.[23] They hypothesized further that WTP should vary by the size of the wilderness area: Selway Bitteroot (1.3 million acres), Bob Marshall (1.0 million acres), and Washakie (0.7 million acres).

Diamond *et al.* (1993) tested this hypothesis on a cleaned data set using a non-parametric Kruskall-Wallis test which ignores the size ordering of the three wilderness areas. They report a p-value of 0.42 for the hypothesis test and consider this result as evidence that respondents are insensitive to the scope of the good being valued.

Carson and Flores (1993) fit the following model in order to test Hypothesis I:

$$Y = \alpha + \beta \text{ ACRES} + \epsilon ,$$

where the p-value reported for β in Table III is from the one-sided test indicated by (1a).

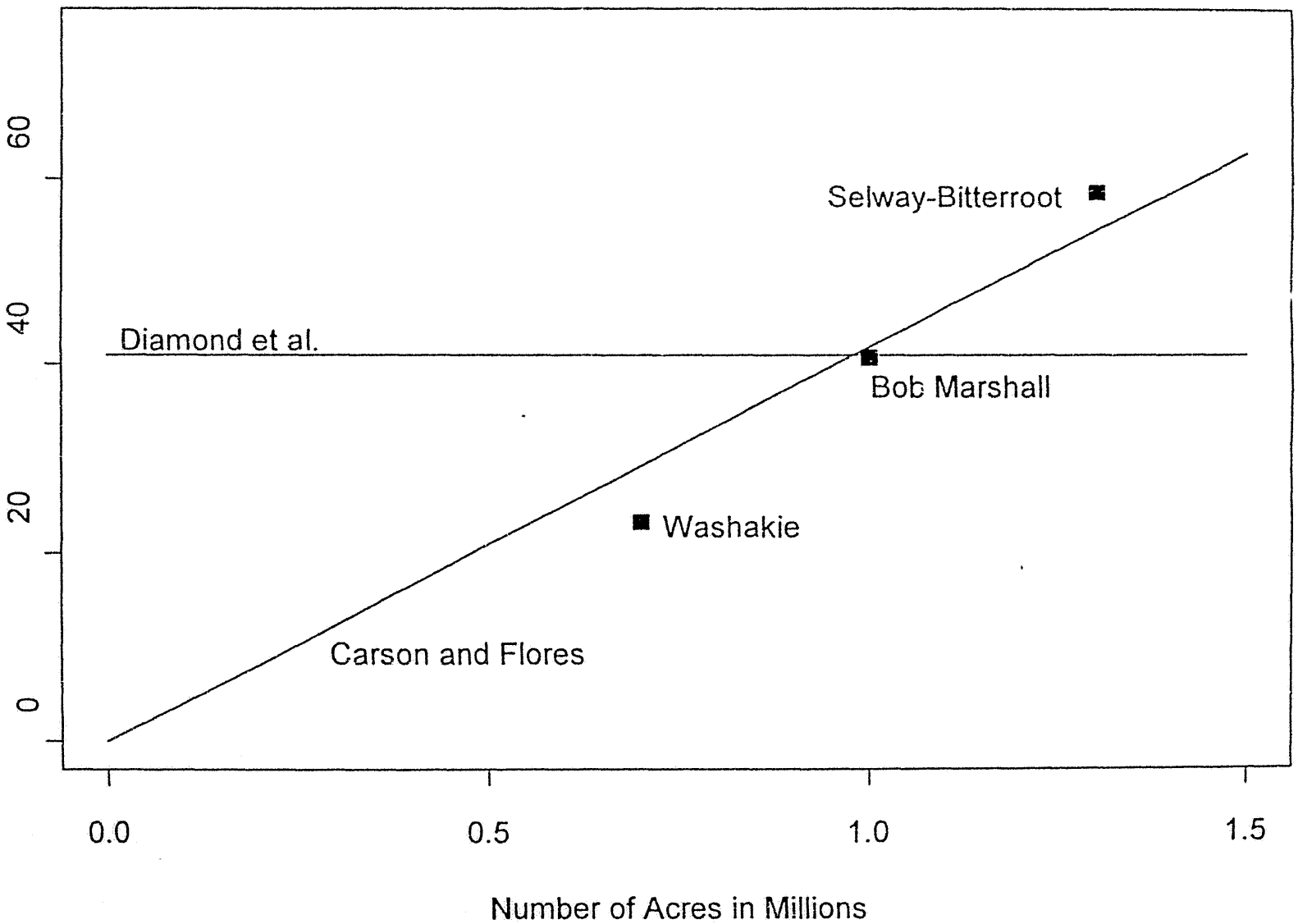
TABLE III: LINEAR REGRESSION ESTIMATE USING DIAMOND *ET AL.* DATA

Parameter	Estimate	t-statistic	p-value
α	-17.834	-0.643	0.52
β	58.782	2.187	0.01*

Using a likelihood ratio test one can further look at the hypothesis, $\alpha=0$. This null hypothesis is accepted ($\chi_1^2=0.421$, p-value=0.52). Following this result, Carson and Flores (1993) then fit the model without a constant ($\beta=41.866$; $t_\beta=6.60$). Figure 1 displays the difference between the regression line fit by Carson and Flores and the null hypothesis accepted by Diamond *et al.*

[23]Carson and Flores (1993) discuss the Diamond *et al.* experiments in detail. In particular, they show that finding Hypothesis I (stated WTP to Protect Each of the Three Areas [Selway Bitterroot, Washakie, and Bob Marshall] is the same) was true is crucial to the validity of each of the additional hypotheses that Diamond *et al.* attempt to test.

FIGURE 1: DIAMOND ET AL. [1993] WILDERNESS STUDY



A further test of the scope insensitivity hypothesis is available by noting that Diamond *et al.* also asked about willingness to pay for Selway with no wilderness areas developed and about developing all 57 wilderness areas. This comparison clearly rejects the scope insensitivity hypothesis at $p < 0.001$.[24]

1.4.2 Desvousges *et al.*

Desvousges *et al.* (1992; 1993) performed two scope experiments, one involving preventing birds from being killed in oil ponds in the Rocky mountain states and the other involved protecting coastal areas from oil spills, which provide results suggesting insensitivity to scope. Both were self-administered surveys in Atlanta shopping malls and had a high fraction of young respondents and both show little ability to explain the WTP responses. The oil spill experiment is marred by the fact that the larger good invoked a different (and lower) probability of success of preventing a large spill than had been used in the second treatment, thereby providing a significant confounding factor. The bird study used three treatments: saving 2000 birds, 20,000 birds, and 200,000 birds. These treatments were also labeled "much less than 1% of the population," "less than 1% of the population" and "about 2% of the population." The mean estimates for the three treatments were $80, $78, and $88, respectively, with the median for all three treatments equal to $25. A careful look at the data suggests a number of outliers in the right tail. Table IV shows a regression using the 10% trimmed mean data displayed in the appendix to their 1992 monograph. This result suggests that there is a significant

[24]Selway Bitterroot, which was one of the largest wilderness areas with approximately 10% of the total acreage, has a mean WTP of $28.54 with a standard error of the mean of $4.84 (N=286) while the 57 wilderness areas as a group have a mean of $78.90 with a standard error of $13.18 (N=297). This hypothesis test (t=3.53) turns out to be the only one in the Diamond *et al.* paper which does not require an assumption about which wilderness areas had already been slated for development (Carson and Flores, 1993).

relationship between the treatments. However, even the trimmed data is not monotonically increasing in the number of birds saved and, if one is willing to accept evidence from a quick shopping mall intercept survey, Desvousges *et al.* should probably be counted as supporting the hypothesis that CV results are insensitive to scope.

TABLE IV: DESVOUSGES *et al.* (1992) BIRD STUDY			
10% Trimmed Data Regression Analysis — WTP = $\alpha + \beta$ (% BIRDS SAVED) + ϵ			
Parameter	Estimate	t-Statistic	p-value
α	47.476	13.088	
β	8.484	3.073	0.001

1.4.3 Schkade and Payne

Schkade and Payne (1993; 1994) use the Desvousges *et al.* bird questionnaire with the same three treatments in the third Exxon scope experiment. This experiment uses a "think aloud" technique from psychology called verbal protocol analysis. In it, Schkade and Payne attempt to delve into what people are thinking and what motivates their responses. I do not address those issues here; rather, I note that the slower pace of the think-aloud technique and the presence of an interviewer in the room at the marketing research firm provide a situation more like that of an in-person interview than did those in the Desvousges *et al.* self-administered mall-intercept surveys.[25]

In the Exxon conference version of the paper (1993), Schkade and Payne report a p-value of 0.42 for an analysis of variance test of the scope insensitivity hypothesis using willingness to

[25]Indeed Schkade and Payne, Desvousges *et al.*, and Diamond and Hausman all argue that since the Schkade and Payne and the Desvousges *et al.* willingness to pay estimates are similar in magnitude, and both show insensitivity to scope, that the mode of survey administration makes no difference to their claim of scope insensitivity.

pay as the response variable. The journal version of the paper (1994) presents an additional analysis of variance test of the hypothesis using the log of willingness to pay as the response variable and reports a p-value of 0.18 for the test and notes that the WTP data looks log-normal.

In my reanalysis of Schkade and Payne's data, it is useful to make two key points. First, the hypothesis Schkade and Payne tested does not correspond to the theoretical argument they claimed to want to test. Schkade and Payne test a more general hypothesis, and as a result of their small samples sizes, they lack almost any statistical power because the models estimated do not incorporate information that the treatments (and the economic hypothesis) are monotonically ordered. It is possible to show that simply regressing the log of willingness to pay on the number of birds results in a rejection of the null hypothesis that WTP is not monotonically increasing in the number of birds valued at $p < 0.05$ using a one-sided t-test.[26] The other problem with the data is that much of the value is driven by a small number of very large outliers. The results of dropping the largest two outliers and the largest six outliers are displayed in Table IV along with the original Schkade and Payne results.[27] This analysis suggests that dropping out just the two highest observations eliminates the apparent violation of the economic restriction that WTP increase monotonically with increases in the level of the good. Dropping these two observations also results in a substantial reduction in the estimates of the standard error of the mean for the first and third treatments. Dropping the next four

[26]The estimate equation is log(WTP)=3.0704 + .0034*(BIRDS/1000) where the t-statistics on the constant and BIRDS variable are 13.72 and 1.80, respectively.

[27]The Schkade and Payne data have a very long right tail with two obvious breaks in the data: 2 observations above $600 (1 at $1000 and 1 at $1200) and 4 additional observations above $200 (3 at $500 and 1 at $600).

largest observations does not change the relative rankings of WTP for the three treatments, but as one would expect, does further reduce the standard errors.

TABLE V: SCHKADE AND PAYNE BIRD EXPERIMENT DATA				
2,000 BIRDS (MUCH LESS THAN 1% OF POPULATION TREATMENT)				
SAMPLE	MEAN	S.E. MEAN	MEDIAN	N
ALL OBS	$84.23	$34.70	$25	28
OBS ≤ $1000	$53.70	$17.06	$25	27
OBS ≤ $500	$38.31	$7.61	$25	26
20,000 BIRDS (LESS THAN 1% OF POPULATION TREATMENT)				
ALL OBS	$62.47	$18.82	$30	29
OBS ≤ $1000	$62.47	$18.82	$30	29
OBS ≤ $500	$45.13	$7.57	$30	28
200,000 BIRDS (ABOUT 2% OF POPULATION TREATMENT)				
ALL OBS	$121.82	$39.51	$50	30
OBS ≤ $1000	$88.13	$21.28	$50	29
OBS ≤ $500	$60.67	$10.03	$50	27

2.0 Scope Sensitivity: Supporting Arguments and Evidence

Before undertaking the task of looking at other direct tests of scope insensitivity, I believe it is useful to ask the question: is there survey evidence from other sources that is likely to shed light on how plausible this hypothesis is likely to be? The answer, of course, is yes. At one level, a contingent valuation survey can simply be thought of as a very in-depth survey which elicits a respondent's views on a specific government policy. As such, there is a large body of evidence to draw upon to ascertain whether responses to survey questions on public policy issues

are generally sensitive to the specifics of the policy question asked and the societal conditions at the time the question is asked. With respect to contingent valuation itself, there are a very large number of studies which may provide indirect evidence which tends to either contradict or support the general proposition of scope insensitivity. This evidence falls into a number of different categories. I first look at whether there is substantial variation in contingent valuation estimates across commodities as the scope insensitivity hypothesis suggests a lack of such variation. Second, I look at the correlation between CV and revealed preference (RP) estimates for the same good when it is possible to use these two different valuation approaches. The scope insensitivity hypothesis suggests that one will not see a strong correlation unless RP estimates are also insensitive to the scope of the good valued. Next, I will look at whether CV estimates tend to systematically vary with respondent characteristics. Then, I will look at whether CV estimates are sensitive to non-scope characteristics of the good, such as the payment vehicle used and the cost stated when the binary discrete choice elicitation method is used. Finally, using an example from a recent study, I will look at internal tests of scope and argue that these tests are not as uninformative as Kahneman contends.

2.1 Survey Research on Policy Issues

Much survey research over the last 50 years has been devoted to showing how, and understanding why, small differences in question wording can lead to substantial differences in the results obtained (Payne, 1951; Schuman and Presser, 1981; Sudman and Bradburn, 1983; Krosnick and Fabrigar, forthcoming). If anything, one gets the impression reading such works, as well as the standard fare of articles in journals such as *Public Opinion Quarterly*, that responses to survey questions are perhaps too sensitive to question wording and that this can

often create problems in the interpretation of policy questions (Schuman and Scott, 1987) by those unaware of the nuances of the question that was asked.

Two interesting examples of this come from split-sample experiments described in Stanley Payne's 1951 classic, *The Art of Asking Questions*. In the first experiment, 82% of the sample receiving the question, "Do you think anything *should* be done to make it easier for people to pay doctor or hospital bills?" responded positively, while 77% of sample receiving the word "*could*" rather than "*should*" responded positively, while only 63% of the sample responded positively when the word "*might*" replaced "*should*." In a second experiment, 92% of the sample responded positively to the statement, "Some people have a kind of insurance for which *they* pay a certain amount each month to cover any hospital care *they* or *their* families may have in the future," but only 66% of the respondents agreed with the statement when it was framed in terms of taking the action personally. None of these differences is at all surprising once it is realized that fundamentally different questions were asked. More importantly, though, these differences suggest respondents *can* pay close attention to subtle differences in question wording.

Payne's experiments are in no sense isolated examples. For instance, large differences in the percentage favoring a government policy toward abortion have long been observed based on whether or not that policy includes federal funding for abortion (NORC, 1991). Support for welfare programs is highly sensitive to whether the program requires able bodied participants to work (Smith, 1987). In more recent times, one sees large differences in the percentage in favor of supporting United Nations peace-keeping efforts in places like Bosnia, Somalia, Rwanda, and Haiti based on whether committing U.S. ground troops is part of the proposal.

Support for environmental protection has been shown to be dependent on whether it simply involves higher costs or substantial unemployment effects (Dunlap and Scarce, 1991).

One can also look at responses to the same survey question over time and ask whether the public and policy makers are sensitive to changes in economic and social conditions over time. Here Page and Shapiro (1990) have provided substantial evidence that public opinion on national expenditures and policy actions tend to track each other fairly closely, sometimes following and sometimes leading as one might expect in a democracy.[28] For environmental expenditures, the National Opinion Research Center's General Social Survey demonstrates that an increase of 1% in the perception that the country is "spending too little on the environment" tends to be followed by a 1.3% increase in the government's rate of spending on the environment. Mueller (1985; 1994) documents how military actions in Vietnam and Kuwait were highly sensitive to responses to public opinion questions.[29] Knowledge on issues such as the greenhouse effect can be shown to steadily increase with the continual flow of information, with the percentage of "don't knows" falling from 36% to 11% over the course of seven years (Dunlap and Scarce, 1991). In another interesting example, the percent saying that oil companies were doing a poor job protecting coastal areas, which had been stable over time, more than doubled when the question was asked six months after the Exxon Valdez spill (Dunlap and Scarce, 1991).

---

[28]See Ferris (1983) for a more economic based approach to this use of survey data.

[29]One of the more interesting examples with respect to military action was the Reagan administration's commission of a survey posing various policy options toward Libya. The most popular option with the public, bombing Qaddafi, was quickly thereafter implemented (Anderson and Van Atta, 1988).

Unless there is something fundamentally different about asking people for their willingness to pay for a particular project, there does not appear to be anything in the survey research literature which suggests that people are inherently unable to recognize and respond to differences in the scope of the good they are asked about. What the survey literature does point out is that one must take great care to ensure that respondents understand what the researcher is describing. It is often difficult and frequently time consuming to ask questions in a manner which the vast majority of the public can comprehend. Respondents will try to answer whatever question is put to them. They often do not realize that the researcher is asking a different question than the one they heard and, even when they are confused about what the question is asking, they frequently do not ask for clarification. This reason is usually found to lie behind most of the aberrant results in the survey research literature, not some fundamental inability of the respondent to deal with the question the researcher really wanted answered.[30]

2.2 Variation in CV Estimates

The most extreme variant of the scope insensitivity argument is that CV respondents are always willing to pay the same amount irrespective of the good being valued. A frequent assertion by CV critics is that all CV studies provide estimates of approximately $30. Other CV critics (*e.g.*, Cummings, 1989; Note, 1992) suggest different amounts. Most are in the $10 to $60 range, but all seem to fall in the range of $10-$120. While a $10-$120 range is broad and

[30]One of the most striking recent examples occurred on a recent Roper survey commissioned by the American Jewish Federation. Twenty-two percent of the respondents appeared to be saying that they believed that the holocaust never happened and 12% were not sure. This was of great concern to the sponsor and at odds with several other survey questions on this issue. The problem was traced back to a confusing double-negative in the question. Roper reworded the question so as to avoid the double-negative and refielded the survey. The percentage then dropped to 1% who believed that the holocaust never happened and 8% were unsure, a response in line with other well-known surveys on the issue. See Morin (1994) for an account.

would cover a large fraction of the non-grocery store purchases made by consumers, even that range can easily be shown to be false. Thompson (1986), interviewing a sample of respondents suffering from rheumatoid arthritis found that they were willing to pay on average $5000 for a particular arthritis treatment. Randall and Kriesel (1990) found American households willing to pay almost $700 on average for large improvements in several national environmental programs. Brookshire *et al.* (1985) found respondents to be willing to pay $4650 more for the same house in the Los Angeles area if it is located outside an earthquake zone. Mitchell and Carson (1993) found respondents willing to pay on average $275 for a very large improvement in national water quality. In contrast, Milon *et al.* (1993) found Florida fishermen willing to pay on average only $0.65 for a management plan which would increase the allowable harvest of the Pompano fish, while Carson *et al.* (1992) found households willing to pay less than $1 to improve 10 days of bad visibility in the Grand Canyon in the winter. Thayer (1981) finds respondents willing to pay about $2.50 more per trip to prevent geothermal development in the Jemez Mountain area of New Mexico while Boyle and Bishop (1987) respondents were willing to pay about $5 to protect an endangered species. While these examples, and any sort of systematic inspection of the estimates from CV studies, suggests that any argument that all CV estimates fall in any sort of narrow range is clearly specious, a formal statistical test (Neil, 1992) of the hypothesis based on collecting datasets from a number of CV survey rejects it.

Perhaps even more telling are two meta-analysis which have been done of CV estimates for specific classes of goods. Walsh, Johnson and McKean (1992) look at CV studies valuing outdoor recreation trips. While they find a mean value of $34 from the 287 studies, the individual studies produced estimates ranging from $4 to $220. Moreover, the differences in

the estimates were systematically related to the characteristics of the trips. Smith and Osborne (1994) look at visibility estimates for national parks and find that they systematically vary with the scope of the change looked at and show that the estimates from the various studies are largely consistent with each other.

2.3 Comparisons Between CV and RP Estimates

For quasi-public goods, it is sometimes possible to use both contingent valuation and revealed preference approaches, such as hedonic pricing and travel cost analysis, to estimate willingness to pay for the good. If respondents in a CV survey for such a good were insensitive to its scope, one would not expect to see much of a correlation between CV and RP estimates for such goods unless actual behavior is also largely insensitive to the scope of the good being valued. Carson *et al.* (1994) have recently performed a meta-analysis looking at 616 comparisons of CV and RP estimates for the same approximate good from 83 studies spanning almost 30 years and 6 continents. They find the correlation between the CV and RP estimates ranges between 0.83 and 0.98 depending upon the whether the full, trimmed or weighted sample is used. This suggests that, at least for quasi-public goods, CV estimates are roughly as sensitive to scope as are estimates obtained from techniques based on observed behavior.

2.4 Systematic Relationships Between CV Estimates and Respondent Characteristics

If one assumes the scope insensitivity hypothesis, then one would expect that willingness to pay *in general* would not vary with respondent characteristics. This proposition is easy to examine because it is common practice to estimate valuation functions in analysis of contingent valuation data. Looking at this large body of evidence, it is quite apparent that CV estimates usually vary in a systematic and expected way with various covariates. Such variation is taken

as a sign of construct validity (Mitchell and Carson, 1989) whereas the lack of such relationships (*e.g.*, in the Desvousges *et al.* bird study) should be taken as a sign of potential problems with the study as a whole or as an indicator of gross outliers in the data.

As it would be possible to write a lengthy book discussing the various valuation functions found in contingent valuation studies, I simply note here some of the more frequently found or interesting relationships between WTP responses and various covariates. With respect to standard demographic variables one often finds that income is a significant predictor of willingness to pay, particularly in studies where the mean WTP estimate is fairly high. Consistent with economic theory, a large number of contingent valuation studies have found a positive and statistically significant relationship between income and WTP (*e.g.*, Cocheba and Langford, 1978; Walsh, Miller and Gilliam, 1983). This relationship tends to be stronger for studies valuing large changes (Thompson, 1986; Randall and Kreisel, 1990; Carson and Mitchell, 1993). Age is often significantly (and usually negatively) related to WTP, particularly for preservation issues (*e.g.*, Hoehn, 1991, Loomis 1987, Whitehead and Blomquist, 1991). Education is often positively related to WTP, particularly for development and risk issues (*e.g.*, Bergstrom, Dillman and Stoll, 1985; John, Walsh and Moore, 1992; Loomis and duVair, 1993) as well as for some of the more exotic forms of outdoor recreation (*e.g.*, Cicchetti and Smith, 1976; Daubert and Young, 1981).

As one might expect, strong environmentalists are usually willing to pay more for an environmental good than other respondents (*e.g.*, Bergstrom *et al.* 1990; Hanley and Craig, 1991; Carson *et al.* 1992). Specific environmental attitudes, particular those closely tied to the

good being valued, are also usually good predictors (*e.g.*, Randall and Kriesel, 1990; Carson, Wilks and Imber (1994); Carson *et al.*, 1992).

Starting with Knetch and Davis' well-known 1966 study of outdoor recreation in the Maine woods, a number of authors have found statistically significant relationships between past recreational experiences and WTP for future recreation (*e.g.*, Johansson, 1990; Karou, 1993; Loomis, Creel and Park, 1991). A number of studies have found a negative relationship between a respondent's distance from the site and his or her willingness to pay for it (*e.g.*, Gramlich, 1977; Sutherland and Walsh, 1985; Bateman *et al.*, 1992). Researchers have typically found that direct users of a resource are willing to pay substantially more than passive users (Barrick and Beazley, 1990; Carson and Mitchell, 1993).

2.5 Sensitivity of CV Estimates to Important Non-Scope Characteristics

If respondents in a CV survey are insensitive to the scope of the good they are asked to value, one might reasonably expect to see insensitivity to other important characteristics of the good such as the payment vehicle used and its price. Sensitivity to the payment vehicle was observed more than a decade ago using split-sample designs quite early on by CV researchers (*e.g.*, Brookshire, Randall, and Stoll, 1980; Greenley, Walsh, and Young, 1981) and continues to be documented (*e.g.*, Duffield, 1992).[31]

Bishop and Heberlein (1979) in their seminal paper on eliciting CV responses using a binary discrete choice framework observed considerable sensitivity to the price of the good the

[31]This sensitivity initially bothered some environmental economists lacking a public choice background. In that tradition, public preferences for particular tax structures has long been an area of research by both economists and political scientists and a subject of interest to bureaucrats and politicians. See the discussion on this point in Cummings, Brookshire and Schulze (1986). Those who tend to see "value" as something which should be context independent are still troubled when two substantially different estimates of willingness to pay are obtained using two different payment vehicles. See Mitchell and Carson (1989) for an extensive discussion of the role of payment vehicles in CV surveys.

respondent was asked about. This result was extremely robust and, given the large number of binary discrete choice CV studies which have been conducted, I have not located a single clear example of insensitivity to price.[32]

2.6 With-in Subject Tests of Scope Insensitivity

Sensitivity to scope can be tested internally within subjects or externally between subjects, with the same restriction from economic theory being tested in both cases. Internal tests of scope, where the same respondent is asked about willingness to pay for different levels of the good of interest, have a long history in contingent valuation (*e.g.*, Randall, Ives, and Eastman, 1974). This should not be surprising, since curves defined over levels of the good of interest—not a single point estimate—are the standard tools of benefit-cost analysis. It is much easier (and cheaper) to ask a single respondent about several levels of a good than to ask several respondents about a single level of the good and one almost always finds that respondents are sensitive to the different levels of the good. From an economic perspective, it is often interesting to observe different patterns of declining marginal valuation (often going to zero) for respondents with different attributes. The information from such questions can be incorporated into a valuation function and, in the statistical estimation of the valuation function, the level of the good is almost always a highly significant WTP predictor (*e.g.*, Carson and Mitchell, 1993).

Having consistently observed internal tests of the scope insensitivity hypothesis being rejected in samples of any size, one might reasonably expect to see this hypothesis rejected in external tests. The typical argument by CV critics against the large number of internal tests

[32]This remark should not be taken to imply that there are no violations of the restriction from economic theory that the percent willing to pay should not increase as the price increases at particular design points. Such violations are, however, generally small in number and well within what one might expect from sampling theory. No doubt there is a demand curve in some CV study which is not downward sloping, the results of which would clearly be questionable.

of scope sensitivity having any bearing on external tests of scope sensitivity is that when a respondent gets asked about multiple levels of a good they try to act in an internally consistent way by simply valuing higher levels higher. Essentially this is an argument that more always costs more and is better, and hence, respondents are still not necessarily paying much attention to the actual level of the good. This need not be the case, as is shown in the following example from a study Michael Hanemann designed for the California Water Resources Control Board on Mono Lake (Jones and Stokes, 1993). In that study, respondents were asked to value three changes in the water level of Mono Lake with the cost of purchasing additional water being the driving force behind the need to determine willingness to pay for the changes. Respondents valued the first change in the water level quite highly, the second change more highly than the first, but the third change much less than either the first or the second; and indeed, the third change appears to be not much preferred to the base condition. What was the cause of this result since the lake level clearly kept rising over the change as did the perceived cost of buying water to fill the lake? What happened was that the first two lake level changes involved clear environmental improvements to the respondents. The third rise in lake level did not; it toppled a significant fraction of the tufa towers the lake is famous for and, while providing substantially more habitat for some species, it displaced others. Obviously, respondents were paying attention to the details and not simply saying more water is better because it cost more.

3.0 Direct Tests of the Hypothesis

Part of what has lent credence to those asserting the validity of the scope insensitivity hypothesis is their claim that only a few tests of this hypothesis have been conducted and that those tests almost uniformly show scope insensitivity. This claim can easily be shown to be

false by searching the CV literature for split-sample comparisons which allow one to test the scope insensitivity hypothesis. In most instances, these split-sample tests were designed to look at different project configurations rather than to formally test the scope insensitivity hypothesis. Often the justification behind the use of such split samples was that the elements of one possible project configuration might influence respondent views about other project configurations with different characteristics. The advantage of these studies is that they tended to receive much more developmental work than did the Kahneman and Exxon surveys and they tended to use larger samples and better modes of survey administration.

3.1 Studies Rejecting Scope Insensitivity Hypothesis

Table VI presents a list of studies, appearing since Kahneman's original 1984 presentation, which contain a rejection of the scope insensitivity hypothesis at $p < 0.10$.[33] Most of the studies contain rejections at $p < 0.05$ and many contain rejections at better than $p < 0.001$. Studies where passive use is thought to be important are marked with a dagger (†).

TABLE VI: STUDIES WITH SIGNIFICANT SCOPE EFFECTS

STUDY	GOOD VALUED
Bowker & Didychuk (1994)†	Preservation of different numbers of acres of agricultural land in Canada
Boyle, Welsh, & Bishop (1993)	Different water flow levels of the Grand Canyon River
Brown, Layton, & Lazo (1994)†	Different options for policies on the Pacific Northwest old growth forests and Northern spotted owls

[33]There are several earlier CV studies which effectively used a split sample by obtaining willingness to pay estimates under a variety of different congestion conditions. For example, McConnell (1977) asks respondents for their WTP for the beach day they have just experienced and finds that the level of beach congestion on the day of the interview is a highly significant predictor of WTP. Cicchetti and Smith (1976) obtained willingness to pay for wilderness hiking days for both the day of the interview and under alternative congestion scenarios pooling these sources to obtain an estimate of the value of congestion effects.

TABLE VI: STUDIES WITH SIGNIFICANT SCOPE EFFECTS	
Buzby, Ready, & Hu (1994)	Reductions in risks associated with pesticides on grapefruits
Carson & Mitchell (1991)	Four scenarios for avoiding urban water shortages
Carson, Mitchell, & Ruud (1989)	Combinations of air pollution visibility & health effects
Carson & Mitchell (1993)†	National versus regional water quality
Carson, Wilks, & Imber (1994)†	Comparison of two different impact scenarios for Australia's Kakadu Conservation Zone
Carson *et al.* (1994)†	Comparison of two different natural resource injury accelerated recovery plans
Diamond *et al.* (1992) as reanalyzed†	Comparisons of different wilderness areas in four Rocky Mountain states
Duffield & Neher (1991)	Comparison of Montana waterfowl hunting trips differing by number of birds hunted
HBRS, Inc. (1994)†	Different impacts on downstream resources from Glenn Canyon Dam operations
Hoevenagel (1994)†	Package of six environmental programs versus specific components (acid rain, greenhouse effect)
Jakus (1992)	Two different Gypsy moth control programs in Pennsylvania and Maryland
Krieger (1994)	Sport fishing toxics information programs
Loomis, Lockwood, & DeLacy (1993)†	Forest area protection program of Southeastern Australia and two smaller portions of the same area
Magnussen (1992)†	Comparisons of various Norwegian pollution control programs with emphasis on water pollution and the North Sea
McDaniels (1988)	Avoidance of different numbers of automobile deaths
Mitchell & Carson (1986)	Comparison of different levels of drinking water risk reduction
Navrud (1989)†	Comparison of different acid rain programs

TABLE VI: STUDIES WITH SIGNIFICANT SCOPE EFFECTS	
Propper (1990)	Different National Health Service waiting times
Ready (1990)†	Preserving different percentages of Kentucky Horse Farms
Römer & Pommerehne (Forthcoming)	Different hazardous waste risk reductions
Rowe *et al.* (1991)†	Multiple versus single Pacific Northwest oil spills
Schkade & Payne (1994) as reanalyzed†	Preventing different numbers of birds from being killed in oil ponds
Smith & Zhang (1994)†	Cleaning-up different levels of marine debris on distant beaches
Tolley & Babcock (1986)	Comparison of different number of days of light health symptoms; comparison of two different variations of Angina
Veisten, *et al.* (1993)†	Environmental programs including biodiversity in Norwegian forests
Whitehead (1992)†	Different probabilities of North Carolina Sea turtle extinction
Whitehead & Blomquist (1991)†	Different Kentucky wetland programs
Wu (1991)†	Different improvement programs at Ohio's Big Darby Creek

It would take a great deal of space to discuss each of these studies in any detail. A few highlights, however, are likely to be useful. First, the scope insensitivity hypothesis is strongly rejected ($p < 0.001$) by two large recent in-person contingent valuation studies, Carson, Wilks and Imber (1994) and Carson *et al.* (1994), which used extensive visual aids and very clean experimental designs to value goods thought to have substantial passive use considerations. Second, other very recent studies of goods thought to have substantial passive use considerations such as increasing the probability of preserving spotted owls (Brown, Layton and Lazo, 1994)

and cleaning-up marine debris on distant beaches (Smith et al., 1994), show respondent sensitivity to very subtle changes in split-sample tests.

The Desvousges *et al.* version of the scope insensitivity hypothesis is stark — one should rarely observe significant differences between the WTP estimates of equivalent subsamples which vary with respect to the scope of the good they are asked to value if that good involves substantial passive use considerations. Table VI, however, shows 19 studies valuing goods thought to have substantial passive use considerations where the hypothesis that respondents are inherently insensitive to differences in scope can be rejected. In contrast, studies where the uniform acceptance of the Desvousges *et al.* hypothesis appears to be indicated are few in number. Kahneman and Knetsch's version of the hypothesis is even more stark and implies that CV results will almost always be insensitive to the scope of the good being valued irrespective of the nature of the good. Here Table VI shows over 30 rejections of Kahneman and Knetsch's hypothesis and very few instances where it is not rejected. From these results I conclude that either version of the hypothesis that scope insensitivity is inevitable in CV surveys is clearly rejected by the large number of available studies which directly address the issue.

This conclusion should not, however, be taken as a statement that there are not potential problems with the results from a particular CV survey. The most plausible alternative hypothesis which is consistent with the data is that there are isolated instances of scope insensitivity which are related to the particular survey design and administration problems noted earlier. One of those deserves special note here because there are a few studies listed in Table VI (*e.g.*, Ready; Magnussen; Loomis; Lockwood and DeLacy) which do not show significant scope effects with respect to the largest program they valued. This is most noticeable in Ready

where preservation of 100% of Kentucky horse farms is actually valued less than 75%. Ready argues, and probably correctly so, that respondents found the 100% preservation plan implausible. Fischoff *et al.*'s experiments suggest the likely reason: the probability of provision of very large programs tends to be seen as less than that of smaller programs.[34] A closely related type of difficulty may arise when respondents already have a clear idea of what a policy will accomplish and refuse to accept the depiction presented in the CV scenario (Carson, Wilks, and Imber, 1994).

It should also be noted that some of the initial CV work on valuing reductions in food safety risks (*e.g.*, Lin and Milon, 1993) does not seem to show responsiveness to the size of the risk being valued. Whether this is due in this particular case to the use of a short telephone survey, general difficulties in conveying small risk reductions, or a lack of an adequate means of conveying the risk reduction and a plausible means of delivering it is unclear. On this last point, it should be noted that Krieger's (1994) original pilot study did not show a difference between a complete and a partial sport fishing toxics risk information program but did show a difference after the survey instrument was redesigned to be more understandable.[35] Finally, for completeness, I should also note that there are also a few instances where using one elicitation method rejects lack of sensitivity to scope while, using another elicitation method, the

[34]What one would obviously like to do is change the scope of the good without changing the perceived probability of provision as that change will generally work against the likelihood that a significant scope effect is observed. It may be possible to improve the power of scope tests by taking steps to help ensure that the probability of provision is held constant across subsamples receiving different goods.

[35]Tolley, Brian, and Fabien (1988) had a similar finding in an early air pollution study where their initial study showed a lack of sensitivity to fairly large differences in air quality improvements whereas changing the presentation of the two programs to make their key elements clearer to respondents resulted in different (plausible) valuations.

scope insensitivity hypothesis cannot be rejected.[36] While there does not seem to be any clear pattern here, there may be interactions between the informational content of elicitation formats and their incentive properties which need to be explored.

4.0 Concluding Remarks

The conclusion that should be drawn from the large body of direct and indirect evidence seems clear: any hypothesis of generic respondent insensitivity to the scope of the good being valued should be rejected. There are over 30 studies with direct split-sample tests of the scope insensitivity hypothesis which reject it. In contrast, there are only a handful of studies in which the hypothesis is not clearly rejected. These studies tend to suffer from (a) small sample sizes, (b) poor survey design (c) shifts in the probability that the good would be provided between subsamples and/or (d) the use of a mode of survey administration, such as the telephone or shopping mall intercepts, which do not encourage respondents to pay close attention to the questions being asked. The large number of available split-sample CV tests do not suggest any difference in scope sensitivity between goods with substantial direct use values and those with substantial passive use values.

In retrospect, it is surprising that the scope insensitivity hypothesis ever gained any currency since the general survey finding is that respondents often pick up on small nuances in question wording. A closer examination of the scope insensitivity hypothesis suggests a number of other patterns which CV data should exhibit if this hypothesis were true. An examination of

[36]Loomis, Lockwood, and DeLacy (1993) find significant differences between all three changes they examined using a dichotomous choice format but only between two of the three changes using an open-ended format. Buzby, Ready, and Hu (1994) find differences using a payment card but do not find a significant difference with a dichotomous choice. Navrud's (1989) results tend to show significant differences between different acid rain programs using a bidding game but not using a payment card.

a very large body evidence shows that such patterns do not generally exist. Once the confusion over the term embedding was sorted out into a test of sequence effects and a test of scope insensitivity (across the rows and down the diagonal, respectively), it became clear that all a test of scope insensitivity required was a split-sample test where one subsample valued a good that was larger than the good valued by the other subsample. Realizing this, it became apparent that there were a fairly large number of existing split-sample CV studies existing in the literature which provided direct evidence rejecting the scope insensitivity hypothesis. Studies conducted subsequent to the Kahneman and Knetsch paper (1992) and the Exxon-sponsored conference (Hausman, 1993) have only tended to confirm this conclusion.

Rejection of generic insensitivity to scope in CV surveys should not be taken to imply that one cannot design a CV questionnaire and administer it in such a way as to find scope insensitivity. Indeed this can be done fairly easily. The remedies for the problem are straightforward in concept but often difficult and expensive in practice to implement.[37] The respondent must (i) clearly understand the characteristics of the good they are asked to value, (ii) find the CV scenario elements related to the good's provision plausible, and (iii) answer the CV questions in a deliberate and meaningful manner.

[37]Robert Mitchell and I discuss different aspects of this issue at length in Mitchell and Carson (1989), Carson and Mitchell (1993), Carson and Mitchell (forthcoming), and Mitchell and Carson (forthcoming).

References

Anderson, Jack, and Dale Van Atta. "The Secret Polls Behind the Libya Raid." *Washington Post*, 29 February 1988.

Andreoni, James. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy* 97, no. 6 (1989): 1447-1458.

------. "Impure Altruism and Donations To Public Goods: A Theory of Warm-Glow Giving." *The Economic Journal* 100 (1990): 464-477.

Arrow, Kenneth, Robert Solow, Paul R. Portney, Edward E. Leamer, Roy Radner, and Howard Schuman. "Contingent Valuation Methodology Report (Report of the NOAA Panel on Contingent Valuation)." *Federal Register* 58(15 January 1993): 4602-4614.

Barrick, Kenneth Andrew, and Ronald I. Beazley. "Magnitude and Distribution of Option Value for the Washakie Wilderness, Northwest, Wyoming, USA." *Environmental Management* 14, no. 3 (1990): 367-380.

Bergstrom, John C., B.L. Dillman, and John Raymond Stoll. "Public Environmental Amenity Benefits of Private Land: The Case of Prime Agricultural Land." *Southern Journal of Agricultural Economics* 17, no. 1 (1985): 139-149.

Bishop, Richard C., and Thomas A. Heberlein. "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" *American Journal of Agricultural Economics* 61, no. 5 (1979): 926-930.

Blomquist, Glenn C., and John C. Whitehead. *The Effect of Alternative and Reclaimed Areas on the Value of Wetlands*. Final report OSM 597 to the U.S. Department of the Interior. Office of Surface Mining Reclamation and Enforcement, 1990.

Bowker, J.M. and D.D. Didychuk. "Estimation of the Nonmarket Benefits of Agricultural Land Retention in Eastern Canada." *Agricultural and Resource Economics Review* 23, no. 2 (1994): 218-225.

Boyle, Kevin J., and Richard C. Bishop. "Valuing Wildlife in Benefit-Cost Analyses: A Case Study Involving Endangered Species." *Water Resources Research* 23, (1987): 943-950.

Boyle, Kevin J., Michael P. Welsh, and Richard C. Bishop. "The Role of Question Order and Respondent Experience in Contingent Valuation Studies." *Journal of Environmental Economics and Management* 25, no. 1 (1993): 45-55.

Brookshire, David S., Alan Randall, and John Raymond Stoll. "Valuing Increments and Decrements in Natural Resource Service Flows." *American Journal of Agricultural Economics* 62, no. 3 (1980): 478-488.

Brookshire, David S., Mark A. Thayer, John Tschirhart, and William D. Schulze. "A Tesonmental and Resource Economists. Stockholm, June, 1991.

Carson, Richard T., Robert C. Mitchell, W. Michael Hanemann, Raymond J. Kopp, Stanley Presser, and Paul A. Ruud. *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill.* Report to the Attorney General of the State of Alaska. Natural Resource Damage Assessment, Inc., November, 1992.

Carson, Richard T., Robert Cameron Mitchell, and Paul A. Ruud. "Valuing Air Quality Improvements: Simulating a Hedonic Pricing Equation in the Context of a Contingent Valuation Scenario." In *Visibility and Fine Particles*, edited by C.V. Mathai, 639-646. Pittsburgh: Air and Waste Management Association, 1990.

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

Cicchetti, Charles J., and V. Kerry Smith. *The Costs of Congestion: An Econometric Analysis of Wilderness Recreation.* Cambridge, MA: Ballinger, 1976.

Cocheba, Donald J., and William A. Langford. "Wildlife Valuation: The Collective Good Aspect of Hunting." *Land Economics* 54, no. 4 (1978): 490-504.

Cummings, Ronald G. Comment submitted to the Office of Environmental Project Review, Department of the Interior, in response to ANPR Concerning Type B Natural Resource Damage Assessments (54 FR 183, September 22, 1989). 1989.

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

Daniel, Terry C., Thomas C. Brown, David A. King, Merton T. Richards, and William P. Stewart. "Perceived Scenic Beauty and Contingent Valuation of Forest Campgrounds." *Forest Science* 35, no. 1 (1989): 76-90.

Daubert, John T., and Robert A. Young. "Recreational Demands for Maintaining Instream Flows: A Contingent Valuation Approach." *American Journal of Agricultural Economics* 63, no. 4 (1981): 666-676.

Desvousges, William H., Alicia R. Gable, Richard W. Dunford, and Sara P. Hudson. "Contingent Valuation: The Wrong Tool for Damage Assessment." *Choices*, no. 2 (1993): 9-11.

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

Desvousges, William H., F. Reed Johnson, Richard W. Dunford, Kevin J. Boyle, Sara P. Hudson, and K. Nicole Wilson. "Measuring Natural Resource Damages with Contingent Valuation: Tests of Validity and Reliability." In *Contingent Valuation: A Critical Assessment*, edited by Jerry A. Hausman. Amsterdam: North Holland Publishers, 1993.

Diamond, Peter A. "Testing the Internal Consistency of Contingent Valuation Surveys." Working paper 93-1, Department of Economics, Massachusetts Institute of Technology.

Diamond, Peter A., and Jerry A. Hausman. "Contingent Valuation: Is Some Number Better than No Number?" *Journal of Economic Perspectives* 8 (1994), 45-64.

Diamond, Peter A., Jerry A. Hausman, Gregory K. Leonard, and Mike A. Denning. "Does Contingent Valuation Measure Preferences? Experimental Evidence." In *Contingent Valuation: A Critical Assessment*, edited by Jerry A. Hausman. Amsterdam: North Holland Publishers, 1993.

Duffield, John W., and Chris Neher. *Montana Waterfowl Hunting: A Contingent Valuation Assessment of Economic Benefits and Hunter Attitudes*. Report to Montana Department of Fish, May, 1991.

Dunlap, Riley, E., and Rik Scarce. "Environmental Problems and Protection." *Public Opinion Quarterly* 55 (1991): 651-672.

Farkas, Steve, Robert Y. Shapiro, and Benjamin I. Page. *The Dynamics of Public Opinion and Policy*. Presented at American Association for Public Opinion Research Meeting, 1990.

Ferris, James M. "Demands for Public Spending: An Attitudinal Approach." *Public Choice* 40 (1983): 135-154.

Fischhoff, Baruch, and Lita Furby. "A Review and Critique of Tolley, Randall *et al.*: 'Establishing and Valuing the Effects of Improved Visibility in the Eastern United States'." Eugene Research Institute, Eugene, October, 1986.

------. "Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility." *Journal of Risk and Uncertainty* 1 (1988): 147-184.

Fischhoff, Baruch, M.J. Quadrel, M. Kamlet, R. Dawes, P. Fischbeck, S. Klepper, J. Leland, and G. Loewenstein. "Embedding Effects: Stimulus Representation and Response Modes." *Journal of Risk and Uncertainty* 6 (1993): 211-234.

Flores, Nicholas. "The Effects of Rationing and Valuation Elasticities." Unpublished paper. Department of Economics, University of California, San Diego, November 1994.

Gramlich, Frederick W. "The Demand for Clean Water: The Case of the Charles River." *National Tax Journal* 30, no. 2 (1977): 183-194.

Greenley, Douglas A., Richard G. Walsh, and Robert A. Young. "Option Value: Empirical Evidence from a Case Study of Recreation and Water Quality." *Quarterly Journal of Economics* 96, no. 4 (1981): 657-672.

Hanemann, W. Michael. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review* 81, no. 3 (1991): 635-647.

Hanley, Nicholas D., and Stephen Craig. "The Economic Value of Wilderness Areas: An Application to the Krutilla-Fisher Model to Scotland's' Flow Country." In *Environmental Policy and the Economy*, edited by F. Dietz, R. van der Ploeg and J. van der Straaten, 95-115. Amsterdam: North Holland Publishing Company, 1991.

Harrison, Glenn W. "Valuing Public Goods with the Contingent Valuation Method: A Critique of Kahneman and Knetsch." *Journal of Environmental Economics and Management* 23, no. 3 (1992): 248-257.

Hausman, Jerry A., ed. *Contingent Valuation: A Critical Assessment*. Contributions to Economic Analysis. Amsterdam: North Holland Publishers, 1993.

HBRS, Inc. "Pilot Test, Nonuse Value Study Glen Canyon Environmental Studies," Draft report, Madison, WI, July 1994.

Hoehn, John P. "Valuing the Multidimensional Impacts of Environmental Policy: Theory and Methods." *American Journal of Agricultural Economics* 73, no. 2 (1991): 289-299.

Hoevenagel, Ruud. *The Contingent Valuation Method: Scope and Validity*, Dissertation, Vrije Universiteit te Amsterdam, 1994.

Jakus, Paul M. *Valuing the Private and Public Dimensions of a Mixed Good: An Application to Pest Control*. Dissertation. North Carolina State University, 1992.

Johansson, Per-Olov. "Willingness to Pay Measures and Expectations: An Experiment." *Applied Economics* 22 (1990): 313-329.

Jones & Stokes Associates, Inc. *Environmental Impact Report for the Review of Mono Basin Water Rights of the City of Los Angeles*. Draft report to California State Water Resources Control Board. 1993.

Kahneman, Daniel. "Comments." In *Valuing Environmental Goods: An Assessment of the Contingent Valuation Method*, edited by Ronald G. Cummings, David S. Brookshire and William D. Schulze. Totowa, NJ: Rowman and Allanheld, 1986.

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

------. "Contingent Valuation and the Value of Public Goods: Reply." *Journal of Environmental Economics and Management* 22, no. 1 (1992b): 90-94.

Kaoru, Yoshiaki. "Differentiating Use and Nonuse Values for Coastal Pond Water Quality Improvements." *Environmental and Resource Economics* 3(May 1993): 487-494.

Kemp, Michael A., and Christopher Maxwell. "Exploring a Budget Context for Contingent Valuation Estimates." In *Contingent Valuation: A Critical Assessment*, edited by Jerry A. Hausman. Amsterdam: North Holland Publishers, 1993.

Knetsch, Jack L., and Robert K. Davis. "Comparisons of Methods for Resource Evaluation." In *Water Research*, edited by Allen V. Kneese and Stephen C. Smith. Baltimore, MD: Johns Hopkins University Press for Resources for the Future, 1966.

Krieger, Douglas J. *Economic Value of Environmental Risk Information: Theory and Application to Michigan Sport Fishery*. Dissertation. Michigan State University, 1994.

Krosnick, J.A., and L.R. Fabrigar. *Designing Questionnaires to Measure Attitudes: Insights from Cognitive and Social Psychology*. New York: Oxford University Press, Forthcoming.

Lin, C T.J. and J.W., Milon. "Contingent Valuation of Health Risk Reductions for Shellfish Products," Presented at U.S.D.A. NE-165 Workshop on Foodsafety and Nutrition, 1993.

Loomis, John B. "Balancing Public Trust Resources of Mono Lake and Los Angeles Water Rights: An Economic Approach." *Water Resources Research* 23, (1987): 1449-1456.

Loomis, John B., Michael Creel, and Timothy Park. "Comparing Benefit Estimates from Travel Cost and Contingent Valuation Using Confidence Intervals for Hicksian Welfare Measures." *Applied Economics* 23, no. 11 (1991): 1725-1731.

Loomis, John B., and Pierre duvair. "Evaluating the Effect of Alternative Risk Communication Devices on Willingness to Pay: Results from a Dichotomous Choice Contingent Valuation Experiment." *Land Economics* 69, no. 3 (1993): 287-298.

Loomis, John, Michael Lockwood, and Terry DeLacy. "Some Empirical Evidence on Embedding Effects in Contingent Valuation of Forest Protection." *Journal of Environmental and Economics Management* 25, no. 1 (1993): 45-55.

Madden, Paul. "A Generalization of Hicksian Substitutes and Complements with Application to Demand Rationing." *Econometrica* 59, no. 5 (1991): 1497-1508.

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

McClelland, Gary H., William D. Schulze, Donald Waldman, D. Schenk, J. Irwin, T. Stewart, L. Deck, and M. Thayer. *Valuing Eastern Visibility: A Field Test of the Contingent Valuation Method - Innovative Approaches for Valuing Perceived Environmental Quality*. Report to the U.S. Environmental Protection Agency Cooperative Agreement, 1991.

McConnell, Kenneth E. "Congestion and Willingness to Pay: A Study of Beach Use." *Land Economics* 53, no. 2 (1977): 186-195.

McDaniels, Timothy L. "Comparing Expressed and Revealed Preferences for Risk Reduction: Different Hazards and Question Frames." *Risk Analysis* 8 (1988): 593-604.

Milon, J. Walter, Eric M. Thunberg, Charles M. Adams, and C.T. Jordan Lin. *Recreational Anglers' Valuation of Near-Shore Marine Fisheries in Florida*. Report prepared for the Florida Marine Fisheries Commission, 1993.

Mitchell, Robert C., and Richard T. Carson. *Valuing Drinking Water Risk Reductions Using the Contingent Valuation Method: A Methodological Study of Risks from THM and Giardia*. Report to the U.S. Environmental Protection Agency, 1986.

------. *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington, DC: Resources for the Future, 1989.

------."Current Issues in the Design, Administration and Analysis of Contingent Valuation Surveys." In *Current Issues in Environmental Economics*, edited by P.O. Johansson, B. Kriström and K.G. Mäler. Manchester University Press, forthcoming.

Morin, Richard. "That Misleading Holocaust Survey: A Flawed Poll Overstated the Public's Skepticism." Washington Post, p. C3, March 20, 1994

Mueller, John E. *Policy and Opinion in the Gulf War*. Chicago: University of Chicago Press, 1994.

------. *War, Presidents, and Public Opinion*. Lanham: University Press of America, Inc., 1985.

National Opinion Research Center. *General Social Surveys, 1972-1991: Cumulative Codebook*. Chicago: NORC, University of Chicago, 1991.

Navrud, Ståle. "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, H. and E. van Ierland, Elsevier Science Publishers, 1989.

Neill, Helen Rosemary. *Valuing Public Goods with the Contingent Valuation Method: Three Essays*. Dissertation. Department of Economics, University of New Mexico, July, 1992.

Note. "Ask a Silly Question...: Contingent Valuation of Natural Resource Damages." *Harvard Law Review* 105, no. 8 (1992): 1981-2000.

Payne, Stanley L. *The Art of Asking Questions*. Princeton: Princeton University Press, 1951.

Propper, Carol. "Contingent Valuation of Time Spent on NHS Lists." *The Economics Journal* 100, no. 400 (1990): 193-199.

Randall, Alan, Berry C. Ives, and Clyde Eastman. "Bidding Games for Valuation of Aesthetic Environmental Improvements." *Journal of Environmental Economics and Management* 1 (1974): 132-149.

Randall, Alan, and W. Kriesel. "Evaluating National Policy Proposals by Contingent Valuation." In *Economic Valuation of Natural Resources: Issues, Theory, and Applications*, edited by R.L. Johnson and G.V. Johnson. Boulder: Westview Press, 1990.

Ready, Richard C. *The Value to Kentuckians of the Kentucky Equine Industry: A Contingent Valuation Study*. Preliminary report Lexington Chamber of Commerce, 1990.

Römer, Anselm U., and Werner W. Pommerehne. "Valuing Reductions of Public Risks: The Case of Hazardous Waste." In *Models and Measurement of Welfare and Inequality*, edited by Peter Eichhorn. Springer, Heidelberg, Forthcoming.

Rosenthal, Donald H., and Richard G. Walsh. "Hiking Values and the Recreation Opportunity Spectrum." *Forest Science* 32 (1986): 405-415.

Rowe, Robert D., Edward R. Morey, William D. Schulze, and William S. Breffle. *Compensable Natural Resource Damage Determination for the Upper Clark Fork River NPL Sites*. Report to the State of Montana, December, 1993.

Rowe, Robert D., William D. Schulze, W.D. Shaw, L.G. Chestnut, and D. Schenk. *Contingent Valuation of Natural Resource Damage Due to the Nestucca Oil Spill*. Final report to the British Columbia Ministry of Environment, June, 1991.

Samuelson, Paul A. "Alturism as a Problem Involving Group Versus Individual Selection in Economics and Biology." *American Economic Review*, 83, no. 2 (1993): 143-148.

Schkade, David A., and John W. Payne. "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, (1994): 88-109.

------. "Where Do the Numbers Come From? How People Respond to Contingent Valuation Questions." In *Contingent Valuation: A Critical Assessment*, edited by Jerry A. Hausman. Amsterdam: North Holland Publishers, 1993.

Schuman, Howard, and J. Scott. "Problems in the Use of Survey Questions to Measure Public Opinion." *Science* 236 (1987): 957-959.

Schuman, Howard, and Stanley Presser. *Questions and Answers in Attitude Surveys*. New York: Academic Press, Inc., 1981.

Smith, Tom W. "That Which We Call Welfare By Any Other Name Would Smell Sweeter." *Public Opinion Quarterly* 51 (1987): 75-83.

Smith, V. Kerry. "Comment: Arbitrary Values, Good Causes, and Premature Verdicts." *Journal of Environmental Economics and Management* 22, no. 1 (1992): 71-89.

Smith, V. Kerry, and Laura Osborne. "Are Contingent Valuation Estimates Economically Consistent? Visibility Changes at National Parks." Unpublished paper. Resource and Environmental Economics Program, North Carolina State University, January, 1993.

Smith, V. Kerry and Xiaolong Zhang. "Can Contingent Valuation Pass a Challenging Scope Test?: Visual Displays of Marine Debris," unpublished paper, Center for Environmental and Resource Economics, Duke University, January 1995.

Stevens, Thomas H., Jaime Echeverria, Ronald J. Glass, Tim Hager, and Thomas A. More. "Measuring the Existence Value of Wildlife: What Do CVM Estimates Really Show?" *Land Economics* 67 (1991): 390-400.

Schwarz, Norbert and Fritz Strack. "Context Effects in Attitude Surveys: Applying Cognitive Theory to Social Research." European Review of Social Psychology 2 (1991): 31-50.

Sudman, S., and N.M. Bradburn. *Asking Questions*. San Francisco: Jossey-Bass, 1983.

Sutherland, Ronald J., and Richard G. Walsh. "Effect of Distance on the Preservation of Water Quality." *Land Economics* 61, no. 3 (1985): 281-291.

Thayer, Mark "Contingent Valuation Techniques for Assessing Environmental Impacts: Further Evidence." *Journal of Environmental Economics and Management* 8 (1981): 27-44.

Thompson, Mark S. "Willingness to Pay and Accept Risks to Cure Chronic Disease." *American Journal of Public Health* 76, no. 4 (1986): 392-396.

Tobin, J., and H.S. Houthakker. "The Effects of Rationing on Demand Elasticities." *Review of Economic Studies* 18 (1951): 140-53.

Tolley, G.S., L. Babcock, M. Berger, A. Bilotti, G. Blomquist, M. Brien, R. Fabian, G. Fishelson, C. Kahn, and *et al. Valuation of Reductions in Human Health Symptoms and Risks: Volume II: Comparative Analysis of Approaches to Valuing Health Risks*. Final report to the U.S. Environmental Protection Agency, January, 1986.

Veisten, Knut, Hans Fredrik Hoen, Ståle Navrud, and Jon Strand. "Valuing Biodiversity in Norwegian Forests." Presented at Meeting of the European Association of Environmental and Resource Economists. Paris, June, 1993.

Walsh, Richard G., Donn M. Johnson, and John R. McKean. "Benefit Transfer of Outdoor Recreation Demand Studies: 1968-1988." *Water Resources Research* 28,(1992): 707-713.

Walsh, Richard G., Nicole P. Miller, and Lynde O. Gilliam. "Congestion and Willingness to Pay for Expansion of Skiing Capacity." *Land Economics* 59, no. 2 (1983): 195-210.

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

Whitehead, John C., and Glenn C. Blomquist. "Measuring Contingent Values for Wetlands: Effects of Information About Related Environmental Goods." *Water Resources Research* 27, no. 10 (1991): 2523-2531.

Wu, Pei-Ing. "Benefit Evaluation of Multidimensional Environmental Policy: An Application to Ohio's Big Darby Creek." Dissertation. *DAI* 52 (1991): 2228. Ohio State University.